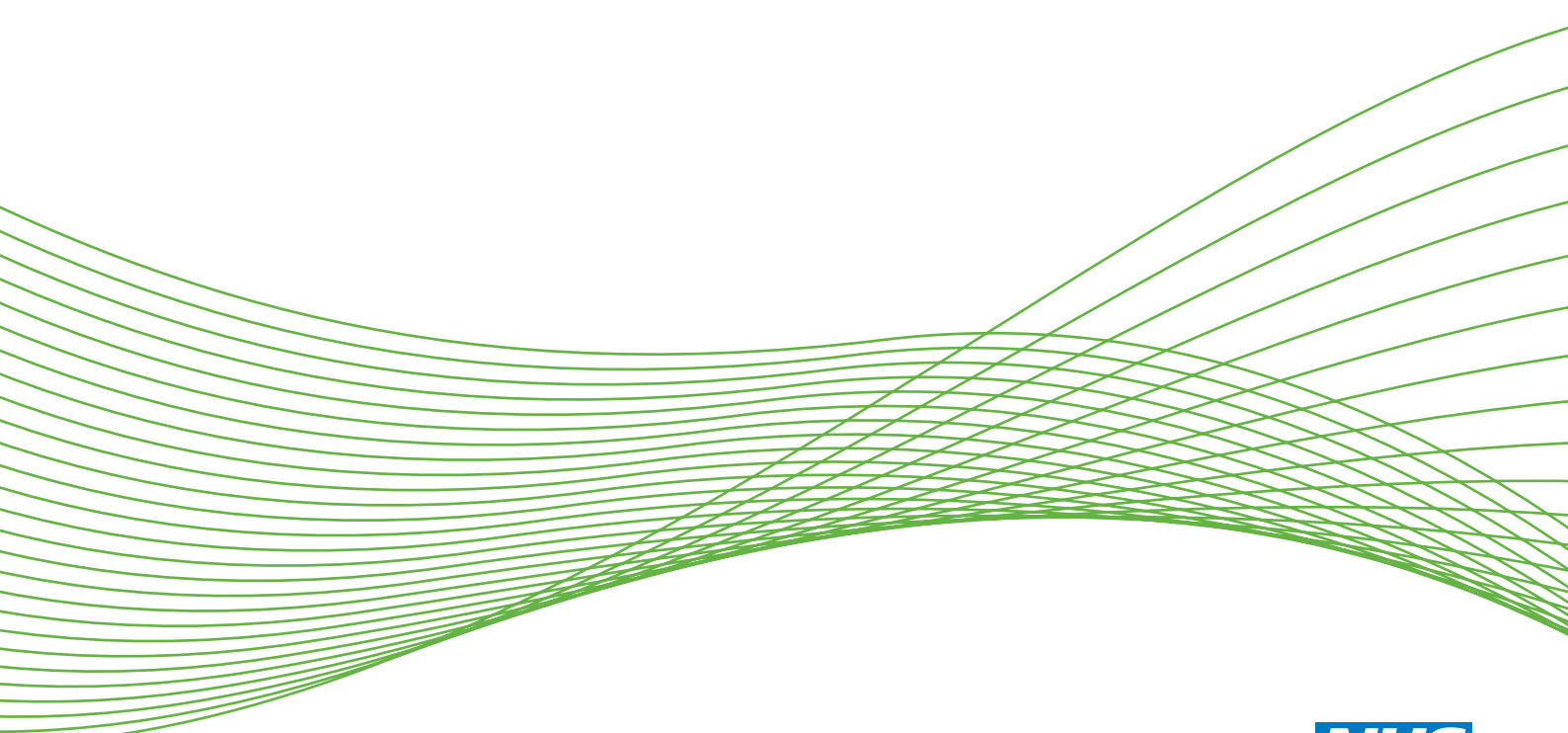


Demand management for planned care: a realist synthesis

Ray Pawson, Joanne Greenhalgh and Cathy Brennan



***National Institute for
Health Research***

Demand management for planned care: a realist synthesis

Ray Pawson,¹ Joanne Greenhalgh^{1*}
and Cathy Brennan²

¹School of Sociology and Social Policy, University of Leeds, Leeds, UK

²Leeds Institute of Health Sciences, University of Leeds, Leeds, UK

*Corresponding author

Declared competing interests of authors: none

Published January 2016

DOI: 10.3310/hsdr04020

This report should be referenced as follows:

Pawson R, Greenhalgh J, Brennan C. Demand management for planned care: a realist synthesis.
Health Serv Deliv Res 2016;**4**(2).

Health Services and Delivery Research

ISSN 2050-4349 (Print)

ISSN 2050-4357 (Online)

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: nihredit@southampton.ac.uk

The full HS&DR archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hsdr. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Health Services and Delivery Research* journal

Reports are published in *Health Services and Delivery Research* (HS&DR) if (1) they have resulted from work for the HS&DR programme or programmes which preceded the HS&DR programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

HS&DR programme

The Health Services and Delivery Research (HS&DR) programme, part of the National Institute for Health Research (NIHR), was established to fund a broad range of research. It combines the strengths and contributions of two previous NIHR research programmes: the Health Services Research (HSR) programme and the Service Delivery and Organisation (SDO) programme, which were merged in January 2012.

The HS&DR programme aims to produce rigorous and relevant evidence on the quality, access and organisation of health services including costs and outcomes, as well as research on implementation. The programme will enhance the strategic focus on research that matters to the NHS and is keen to support ambitious evaluative research to improve health services.

For more information about the HS&DR programme please visit the website: <http://www.nets.nihr.ac.uk/programmes/hsdr>

This report

The research reported in this issue of the journal was funded by the HS&DR programme or one of its preceding programmes as project number 11/1022/04. The contractual start date was in September 2012. The final report began editorial review in May 2014 and was accepted for publication in November 2014. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HS&DR editors and production house have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the final report document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded by the National Institute for Health Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, NETSCC, the HS&DR programme or the Department of Health. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, NETSCC, the HS&DR programme or the Department of Health.

© Queen's Printer and Controller of HMSO 2016. This work was produced by Pawson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).

Health Services and Delivery Research Editor-in-Chief

Professor Jo Rycroft-Malone Professor of Health Services and Implementation Research, Bangor University, UK

NIHR Journals Library Editor-in-Chief

Professor Tom Walley Director, NIHR Evaluation, Trials and Studies and Director of the HTA Programme, UK

NIHR Journals Library Editors

Professor Ken Stein Chair of HTA Editorial Board and Professor of Public Health, University of Exeter Medical School, UK

Professor Andree Le May Chair of NIHR Journals Library Editorial Group (EME, HS&DR, PGfAR, PHR journals)

Dr Martin Ashton-Key Consultant in Public Health Medicine/Consultant Advisor, NETSCC, UK

Professor Matthias Beck Chair in Public Sector Management and Subject Leader (Management Group), Queen's University Management School, Queen's University Belfast, UK

Professor Aileen Clarke Professor of Public Health and Health Services Research, Warwick Medical School, University of Warwick, UK

Dr Tessa Crilly Director, Crystal Blue Consulting Ltd, UK

Dr Peter Davidson Director of NETSCC, HTA, UK

Ms Tara Lamont Scientific Advisor, NETSCC, UK

Professor Elaine McColl Director, Newcastle Clinical Trials Unit, Institute of Health and Society, Newcastle University, UK

Professor William McGuire Professor of Child Health, Hull York Medical School, University of York, UK

Professor Geoffrey Meads Professor of Health Sciences Research, Health and Wellbeing Research and Development Group, University of Winchester, UK

Professor John Norrie Health Services Research Unit, University of Aberdeen, UK

Professor John Powell Consultant Clinical Adviser, National Institute for Health and Care Excellence (NICE), UK

Professor James Raftery Professor of Health Technology Assessment, Wessex Institute, Faculty of Medicine, University of Southampton, UK

Dr Rob Riemsma Reviews Manager, Kleijnen Systematic Reviews Ltd, UK

Professor Helen Roberts Professor of Child Health Research, UCL Institute of Child Health, UK

Professor Jonathan Ross Professor of Sexual Health and HIV, University Hospital Birmingham, UK

Professor Helen Snooks Professor of Health Services Research, Institute of Life Science, College of Medicine, Swansea University, UK

Professor Jim Thornton Professor of Obstetrics and Gynaecology, Faculty of Medicine and Health Sciences, University of Nottingham, UK

Please visit the website for a list of members of the NIHR Journals Library Board:
www.journalslibrary.nihr.ac.uk/about/editors

Editorial contact: nihredit@southampton.ac.uk

Abstract

Demand management for planned care: a realist synthesis

Ray Pawson,¹ Joanne Greenhalgh^{1*} and Cathy Brennan²

¹School of Sociology and Social Policy, University of Leeds, Leeds, UK

²Leeds Institute of Health Sciences, University of Leeds, Leeds, UK

*Corresponding author j.greenhalgh@leeds.ac.uk

Background: The task of matching fluctuating demand with available capacity is one of the basic challenges in all large-scale service industries. It is a particularly pressing concern in modern health-care systems, as increasing demand (ageing populations, availability of new treatments, increased patient knowledge, etc.) meets stagnating supply (capacity and funding restrictions on staff and services, etc.). As a consequence, a very large portfolio of demand management strategies has developed based on quite different assumptions about the source of the problem and about the means of its resolution.

Methods: This report presents a substantial review of the effectiveness of main strategies designed to alleviate demand pressures in the area of planned care. The study commences with an overview of the key ideas about the genesis of demand and capacity problems for health services. Many different diagnoses were uncovered: fluctuating demand meeting stationary capacity; turf protection between different providers; social rather than clinical pressures on referral decisions; self-propelling diagnostic cascades; supplier-induced demand; demographic pressures on treatment; and the informed patient and demand inflation. We then conducted a review of the key ideas (programme theories) underlying interventions designed to address demand imbalance. We discovered that there was no close alignment between purported problems and advocated solutions. Demand management interventions take their starting point in seeking reforms at the levels of strategic decision-making, organisational re-engineering, procedural modifications and behavioural change. In mapping the ideas for reform, we also noted a tendency for programme theories to become 'whole-system' models; over time policy-makers have advocated the need for concerted action on all of these fronts.

Findings: The remainder and core of the report contains a realist synthesis of the empirical evidence on the effectiveness on a spanning subset of four major demand management interventions: referral management centres (RMCs); using general practitioners with special interests (GPwSIs) at the interface between primary and secondary care; general practitioner (GP) direct access to clinical tests; and referral guidelines. In all cases we encountered a chequered pattern of success and failure. The primary literature is replete with accounts of unanticipated problems and unintended effects. These programmes 'work' only in highly circumscribed conditions. To give brief examples, we found that the success of RMCs depends crucially on the balance of control in their governance structures; GPwSIs influence demand only after close negotiations on an agreed and intermediate case mix; significant efficiencies are created by direct GP access to tests mainly when there is low diagnostic yield and high 'rule-out' rates; and referral guidelines are more likely to work when implemented by staff with responsibility for their creation.

Conclusions: The report concludes that there is no 'preferred intervention' that has the capacity to outperform all others. Instead, the review found many, diverse, hard-won, local and adaptive solutions. Whatever the starting point, success in demand management depends on synchronising a complex array of strategic, organisational, procedural and motivational changes. The final chapter offers practitioners some guidance on how they might 'think through' all of the interdependencies, which bring demand and capacity into equilibrium. A close analysis of the implementation of different configurations of demand management interventions in different local contexts using mixed methods would be valuable to understand the processes through which such interventions are tailored to local circumstances. There is also scope for further evidence synthesis. The substitution theory is ubiquitous in health and social care and a realist synthesis to compare the fortunes of different practitioners placed at different professional boundaries (e.g. nurses/doctors, dentists/dental care practitioners, radiologists/radiographers and so on) would be valuable to identify the contexts and mechanisms through which substitution, support or short-circuit occurs.

Funding: The National Institute for Health Research Health Services and Delivery Research programme.

Contents

List of tables	xi
List of figures	xiii
List of boxes	xv
List of abbreviations	xix
Plain English summary	xxi
Scientific summary	xxiii
Chapter 1 The challenge: reviewing the attempts to solve a ‘wicked problem’	1
Research method: realist synthesis	1
<i>Identifying the review questions</i>	2
<i>Searching for primary studies</i>	3
<i>Quality appraisal</i>	3
<i>Extracting the data</i>	4
<i>Synthesis</i>	4
Background: what is demand management?	5
Outline and design of the current review	8
<i>Part 1</i>	9
<i>Part 2</i>	9
<i>Appendices</i>	12
Part 1 Theory elicitation	13
Chapter 2 The runaway train: the multiple, intertwined causes of growth in demand for health care	15
Queuing theory and the question of time	16
Professional closure, informal control and turf protection	19
Micro-dynamics in the decision to refer	22
Self-propelling diagnostic cascades	24
Supplier-induced demand	27
Demographics and demand	31
The internet, the informed patient and demand inflation	35
Conclusion	39
Chapter 3 The policy response: charting the family of purported solutions	41
Introduction: eliciting and selecting the underlying programme theories	41
Taking stock: mapping the field of demand management interventions	42
Demand management interventions	45
<i>Strategic change</i>	45
<i>Administrative change</i>	47
<i>Role change</i>	50
<i>Procedural change</i>	53
<i>Motivational change</i>	60
Common denominator: a unifying programme theory?	60

Part 2 Theory testing	63
Chapter 4 Organisational change: referral management centres – can they control and shape demand?	67
Introduction	67
Referral management centres and triage: operational logistics	67
Programme theories: how are referral management centres intended to work?	68
Evidence review	70
Intended outcomes: patients are redirected to alternative services, referrals decrease, general practitioners change their referral behaviour	71
<i>White</i>	71
<i>Imison and Naylor</i>	72
<i>Offredy and colleagues</i>	73
<i>Moss and colleagues</i>	75
Unintended outcome 1: short-circuit	76
<i>General practitioners short-circuit the triage/referral management centre and refer straight to secondary care</i>	76
Unintended outcome 2: triage/referral management centres increase referral volumes	81
<i>Cox and colleagues</i>	81
<i>Imison and Naylor</i>	82
<i>Bungard and colleagues</i>	83
<i>Rymaszewski and colleagues</i>	84
<i>Maddison and colleagues</i>	84
<i>Raine and colleagues</i>	85
<i>Xiang and colleagues</i>	87
Chapter 5 Role change: general practitioners with a special interest – can they control and shape demand?	91
General practitioners with special interests: policy, promise and pitfalls	91
<i>Substitution or support</i>	93
<i>The anatomy of decision-making</i>	94
<i>Professional control and boundary work</i>	95
General practitioners with special interests evidence synthesis	96
<i>Outcome studies</i>	96
<i>Process and context (explanatory) studies</i>	103
Conclusion: emerging patterns	114
Chapter 6 Changing responsibilities: direct access to the results of clinical tests – can it control and shape demand?	117
Introduction	117
Programme theory	118
Impediments to procedural change	121
Evidence synthesis	121
<i>Typology of functions of diagnostic tests</i>	122
<i>Conclusion</i>	145
Chapter 7 Learned counsel: can guidelines control and shape demand?	147
Introduction: the guidelines industry	147
Approach 1: Cochrane orthodoxy – guidelines as interventions	148
Approach 2: thematic reviews – facilitators and barriers to guidelines	148
Approach 3: a rapid realist review	155
System strain 1: simple guidelines versus comorbid patients	157

System strain 2: the tension between (inter)national credibility and local control over guidelines	158
System strain 3: the tension between authoritative, population-based guidelines and discretionary, individual-based patient choice	160
System strain 4: tensions resulting from guideline oversupply – a new guideline can swamp a routine system	161
Conclusion	163
Chapter 8 Conclusions: facing the challenge of complexity	165
Lessons learned	165
Design principles in demand management	166
Group model building	172
Recommendations for future research	175
Chapter 9 'Thinking it through': prompts for practitioners	177
Prompt 1: what (exactly) is the problem?	177
Prompt 2: what are the options?	178
Prompt 3: will reorganisation work?	179
Prompt 4: can intermediaries do the trick?	180
Prompt 5: will direct access to tests results reduce excess referrals?	181
Prompt 6: will guidelines be followed?	182
Prompt 7: can productive change be accelerated?	183
Acknowledgements	185
References	187
Appendix 1 Search and selection strategies	205
Appendix 2 Reviewing the field	213
Appendix 3 Stakeholder and patient involvement	219
Appendix 4 Dissemination events and outputs	221

List of tables

TABLE 1 Ratio of rates of angiography, CABG and CARP in private vs. public hospitals	28
TABLE 2 Time spent on each activity over the three time periods (time recorded in 5-minute blocks)	108
TABLE 3 Details of savings from direct access to X-ray	117
TABLE 4 Endoscopic diagnosis in open-access cases and specialist referral cases (for the period November 1996 to December 1997)	126
TABLE 5 Referral rate according to condition	127
TABLE 6 Changes in GP diagnoses and diagnostic confidence	131
TABLE 7 General practitioner vs. orthopaedic referrals for MRI	133
TABLE 8 General practitioner vs. hospital outpatient referrals and GP patient management following MRI result	134
TABLE 9 Diagnosis of osteoporosis in referrals that met the referral criteria and those that did not meet the referral criteria	138
TABLE 10 Recommendations of DXA reports and GP action	139
TABLE 11 Management by GPs after open-access echocardiography	143
TABLE 12 Typological reviews: four accounts of facilitators and barriers to guideline implementation	150
TABLE 13 Gagliardi and colleagues' final framework of guideline implementability	153
TABLE 14 The typological review	213
TABLE 15 'Types' identified in demand management reviews	215
TABLE 16 Summary of findings: relocation of services to primary care	217

List of figures

FIGURE 1	Eliciting theory in reviews of 'focused' interventions vs. 'wicked' problems	16
FIGURE 2	Alternative strategies and probabilities of meeting fluctuating demand	17
FIGURE 3	Pooling strategy to smooth demand	18
FIGURE 4	A micro model of the referral decision	23
FIGURE 5	Rates of common surgical procedures among Medicare patients for 306 referral regions	29
FIGURE 6	Population projections by age group (US)	32
FIGURE 7	Office visits to physicians per year by age group (US)	32
FIGURE 8	Distribution of different approaches to referral management	43
FIGURE 9	Change objectives in demand management interventions	44
FIGURE 10	Prototypical structure of a RMC/referral management gateway	47
FIGURE 11	Interconnected levels of change	62
FIGURE 12	Referral management centres	63
FIGURE 13	General practitioners with special interests	63
FIGURE 14	General practitioner direct access to tests	63
FIGURE 15	Guidelines	64
FIGURE 16	Orthopaedic referral pathway: Goringhamshire	65
FIGURE 17	Existing elementary pathway and patient volumes	68
FIGURE 18	Revised 'ideal' pathway	69
FIGURE 19	Short-circuit	76
FIGURE 20	Increase in overall volume of referrals	81
FIGURE 21	Percentage change in volume of first attendances referred by GPs, all PCTs (excluding outliers) 2005–9	82
FIGURE 22	Distribution of GP practices by referral rate 2008–9, top 22 PCTs undertaking significant amount of referral management and against all practices	83
FIGURE 23	Primary care trusts with current and planned GPwSI services (%) by clinical area	92

FIGURE 24 Existing system and intended change	93
FIGURE 25 Intended and unintended decision-making structures	94
FIGURE 26 The DM _{2.5} decision hierarchy (ladder of decision-making)	95
FIGURE 27 The shaping of decision-making across professional boundaries	96
FIGURE 28 Direct access to tests	119
FIGURE 29 Pathology value chain	119
FIGURE 30 Birmingham South Central clinical pathway for headache	120
FIGURE 31 Birmingham South Central clinical pathway for 24-hour ECG	120
FIGURE 32 Patient outcomes following arrhythmia monitoring	141
FIGURE 33 A phases and interdependence model	156
FIGURE 34 Flow chart of inclusion of studies into review (RMC)	208
FIGURE 35 Flow chart of inclusion of studies into review (GPwSI)	210
FIGURE 36 Flow chart of studies included in the review (direct access to tests)	211

List of boxes

BOX 1 The rise in demand for health care	16
BOX 2 Reducing waiting in radiography	18
BOX 3 Summary 1	19
BOX 4 Summary 2	21
BOX 5 Summary 3	24
BOX 6 Common triggers for the cascade effect	25
BOX 7 Summary 4	27
BOX 8 Summary 5	31
BOX 9 Summary 6	35
BOX 10 A sequential model of an 'informed choice' consultation	38
BOX 11 Summary 7	39
BOX 12 Strategies for managing demand at the interface between primary and secondary care	43
BOX 13 Four models of demand management for access to specialist care	43
BOX 14 Typical sources on strategic change	46
BOX 15 Referral management centre programme theories	48
BOX 16 Referral management centre programme theories (negative)	49
BOX 17 Referral management centre programme theories (refined)	50
BOX 18 General practitioners with special interests' role as an intermediate tier of expertise	51
BOX 19 Paradoxes in the role of expert intermediaries	52
BOX 20 Supplementary changes to support role change	52
BOX 21 Feedback as shared wisdom	53
BOX 22 Resistance to peer review	54
BOX 23 Adapting the process of peer review	55
BOX 24 Guidelines provide standardisation around best practice	56

BOX 25 Guidelines as opposed to real clinical decisions	57
BOX 26 Supplementing guidelines	57
BOX 27 Simplifying and rationalising the diagnostic chain	58
BOX 28 Direct access: unforeseen difficulties and unintended consequences	59
BOX 29 'Modernising' direct access provision	59
BOX 30 Interim summary	80
BOX 31 Interim summary	88
BOX 32 Chapter summary	89
BOX 33 Royal College of General Practitioners' designation of GPwSI roles	92
BOX 34 Anticipated benefits of shifting care closer to home	93
BOX 35 Overall activity, referral rates, relative access and waiting times	98
BOX 36 Interim summary	102
BOX 37 Boundary relations and alliances	106
BOX 38 General practitioners with special interests' development and change management	107
BOX 39 The impact of system-wide contingencies on GPwSI services	113
BOX 40 Interim summary	129
BOX 41 Interim summary	136
BOX 42 Interim summary	145
BOX 43 Cochrane and colleagues' 'barrier groupings': categories (frequency)	151
BOX 44 Key design principles in demand and referral management	167
BOX 45 Vignette: supply-induced demand	168
BOX 46 Vignette: whole-system adaptation	168
BOX 47 Vignette: reconfiguring the administrative apparatus	169
BOX 48 Vignette: establishing the intermediate function	170
BOX 49 Vignette: differentiating patients according to test results	171
BOX 50 Vignette: the local adaptation and actualisation of guidelines into individual roles	171

BOX 51 Format description of improvement	174
BOX 52 Demand management options	179

List of abbreviations

AMBU	Abertawe Bro Morgannwg University	ICAT	independent clinical assessment and treatment centre
BMA	British Medical Association	LMC	local medical committee
BMD	bone mineral density	MHP	mental health professional
BMJ	<i>British Medical Journal</i>	MPTT	multiprofessional triage team
CAS	clinical assessment service	MRI	magnetic resonance imaging
CCG	Clinical Commissioning Group	NICE	National Institute for Health and Care Excellence
CIC	clinical interest company	NLIAH	National Leadership and Innovation Agency for Healthcare
CRG	Clinical Reference Group	PBC	practice-based commissioning
CT	computerised tomography	PCT	primary care trust
DAMASK	direct access to magnetic resonance imaging assessment for suspect knees	PICO	population, intervention, comparator, outcomes
DM	decision-maker	PwSI	practitioner with a special interest
DXA	dual X-ray absorptiometry	QOF	Quality and Outcomes Framework
EASE	Ensuring Access and Speedy Evaluation	QuANGO	quasi-autonomous non-governmental organisation
ECG	electrocardiogram	RCGP	Royal College of General Practitioners
ENT	ear, nose and throat	RCT	randomised controlled trial
GP	general practitioner	RMC	referral management centre
GPwSI	general practitioner with a special interest	RMS	referral management service
HIV	human immunodeficiency virus	SRIT	Service Redesign and Implementation Team
HSDR	Health Services and Delivery Research		

Plain English summary

Medicine is a victim of its own success. All advanced health systems face substantial increases in activity and costs with a seemingly unstoppable rise in demand for all aspects of care. A particular strain is often felt on the matter of referral management, where patients are relayed from one part of the system to another, often without due care being given to the availability of resources across the system.

The report reviews existing research on the effectiveness of the main interventions used to control demand and limit referrals for planned care. Four strategies are researched: referral management centres; using general practitioners with special interests at the interface between primary and secondary care; general practitioner direct access to clinical tests; and referral guidelines.

Using a mix of electronic and hand-searches, we identified a huge body of primary research on these measures. The assembled evidence amounts to a 'thwarted history' of demand control; there are no easily reproducible silver bullets available to tame the steady rise in referrals. In the main, the interventions reviewed were able to reform one aspect of a complex system without having the ability to co-ordinate change across the entire commonwealth of health-care responsibilities.

There are, however, pockets of progress. The most promising interventions are always local and adaptive, often comprising second- and third-generation attempts to co-ordinate fractious interest groups. The review is, thus, able to identify very specific conditions under which these interventions have been successful. It closes with suggestions on how such 'bottom-up' interventions can be quickened.

Scientific summary

Aims and objectives

This report attempts to meet the Health Services and Delivery Research commission to provide 'robust assessments of demand management interventions for elective care'. It does so by conducting a 'realist synthesis' of the primary research evaluating the wide range of strategies that have been devised to stabilise the threateningly high levels of demand for planned care that occur throughout modern health services.

This proved a challenging undertaking, given the extraordinary diversity of demand management activities. The task of the evaluator and reviewer is considerably simplified if the intervention under research is aimed at a well-defined problem, is implemented to a clear design and can be assessed on an agreed, measurable criterion. None of these desiderata applies in the case of demand management. The roots of the imbalance between capacity and demand are complex and intertwined. There are very few designated and independent 'demand management programmes'. Rather, attention to demand is part of the remit of specific agencies and one of the duties of particular post-holders. It is a routine aspect of the daily fabric of health management. Moreover, on outcomes, it transpires that there is no common comprehension of 'excessive' demand, with, for instance, the line between 'premature' and 'appropriate' referrals often being difficult to draw.

Anticipating these complexities our research began with two broad objectives. The first aim was to survey the landscape of activities that have been mounted in the name of demand management. Here, we sought to provide an overview of how the problems of excessive and inappropriate demand had been understood and also to provide a catalogue of the many and varied responses. The idea was to furnish the review with a 'menu' of potential causes and proposed solutions. The second and fundamental objective was to provide a robust assessment of the effectiveness of the various strategies and schemes mapped in phase one. Alongside the abundant variety and evident heterogeneity of these approaches it was also clear from our preliminary research that managing demand had proved an uphill struggle and that we would discover no 'best buy' interventions with the capacity to outperform all others. We thus interpreted our second objective in terms of the provision of an explanatory account of the complex medley of conditions that lead to successes and failures of the respective schemes.

Review strategy

Our method of collecting together and drawing lessons from primary research evidence is known as 'realist synthesis'. Realist synthesis is a theory-driven approach to evidence synthesis developed by one of the current authors. Realist synthesis finds use in complex interventions that are not easily reproducible and where there is considerable heterogeneity in both implementation and the contexts in which they are mounted. The focus of attention switches to programme theories, the ideas that drive interventions and the analytic approach is theory testing – discovering why and why not the programme theories come to fruition.

In the present instance the beginning logic was to provide a thorough review of how the problem of excessive demand had been diagnosed and to compare this with the compendium of proposed solutions. The basic motif is thus to discover how well the 'remedy' addresses the 'malady'. We know that demand management is a domain of partial solutions and this approach provides an explanatory focus pinpointing some of the unforeseen challenges and unintended outcomes.

Search strategy

Phase 1

Given that our focus was on underlying policy thinking, we directed initial attention to the 'ideas literature'. We searched for sources in the so-called grey literature (planning documents, guidance materials, discussion documents, proposals, rationales, policy expositions, professional journals and critical debate). We used simple search terms that were identified in our research brief and borrowed those terms and synonyms used in previous reviews of demand management, which identify the core approaches such as 'referral management centres', 'guidelines', 'feedback', 'general practitioners with special interests' (GPWSIs), 'direct access to test results', etc. We extracted the underlying programme theories, the log of potential problems and solutions, on the basis of a close reading of this documentation.

Phase 2

The second phase of the review focused on locating empirical studies that enabled us to test the intervention theories in practice. The primary materials of interest here are evaluative inquiries and so the stock-in-trade materials take the form of formal 'research reports' as well as papers and commentary from the many health-care journals. Theory testing can make use of findings that emerge from any form of quantitative or qualitative inquiry; there is no hierarchy of evidence and, thus, no search restrictions on that basis. Having identified a particular type of intervention (e.g. 'referral management centre'), electronic searches commenced utilising its key terms and synonyms and employing the standard databases for health service research. As our understanding grew of the flows and blockages associated with each approach (e.g. disputes over the control in such centres), we explored them further using iterative, 'snowballing' searches such as pursuing 'references of references'.

Synthesis method

Realist analysis has an explanatory role, focusing on the particular circumstances, respects and reasons why an intervention might work. An efficient way to expedite such analysis is to focus on the tensions between diagnosis and remedy – how well does demand control deal with the causes of demand inflation, with what unanticipated causes and unintended consequences. To this end we initially proposed investigation on four preliminary frictions. Is demand management able to:

- respond to different and sometimes conflicting motivations that prompt referral
- balance the varied and sometimes uneven expertise and mandates of the participants in referral chains
- promote accountability for cost-containment ambitions in NHS staff groups who traditionally lack such a remit
- regulate provision while responding to other initiatives, which provide patients with increased choice of provision?

In the course of the review we were able to extend and refine such questions. Early analysis revealed a core tension. The causes of demand and capacity problems are system wide: they are rooted in the perpetuation of historic organisational structures, the multiplication of treatment pathways within increasingly complex divisions of clinical labour, constant improvements in diagnosis and treatments, demographic change and increasing wisdom in the patient population, and so on. The policy responses, however, are invariably limited and tend to have more specific remits to generate improvements by remodelling organisational structures or by introducing new roles and procedures or by designing more exacting guidelines or by incentivising behaviour change. The end result is a patchwork of success and failure, which our synthesis attempts to map and explain.

Findings

Chapter 2 reviews the manifold interconnected processes that generate demand for health care. These explanations span a remarkable range of features, covering physician motivations, professional closure, demographic change, diagnostic improvements, supply-induced demand, the informed patient, etc. Our task was not to rank or adjudicate between these accounts but simply to provide a typology encompassing the wide range of demand pressures and to provide evidence showing that each one is substantial enough to command a policy response. It presented us with a daunting hypothesis, namely that multiple, intertwined problems are unlikely to yield to singular solutions, however well aimed. The demand for health care may be regarded as a swelling punch-bag. Landing a blow in respect of one problem may simply be absorbed as other vicissitudes gather.

Chapter 3 examines the programme theories that underpin the small army of interventions that have attempted to quell the inflation in demand for services. Potential solutions include referral management centres (RMCs), clinical assessment and triage services, service relocation, referrals to GPwSIs, financial incentives, audit and feedback, guidelines, queue sculpturing and behaviour change. We elicited the programme theories underlying each intervention. The scope of the intended solution varied substantially from the macro to the micro, with reforms being introduced, in turn, at strategic, administrative, role, procedural and motivational levels. We also observed that programme theories are not static. In the light of experience gained under critical scrutiny and via the trial and error of actual interventions, the core assumptions tended to modify, recommending the need for supplementary action. Invariably, the programme theories became 'whole system' models – suggesting the powerful hypothesis that sustainable change required the interweaving of the various macro, meso and micro mechanisms.

Chapter 4 examines the effectiveness of large-scale administrative reform in the guise of RMCs and other centralised triaging services. The programme theory posits that efficiencies to demand management are generated by rationalising decision-making along referral pathways between primary and secondary care. RMCs function to this end only in the presence of a cluster of supportive and somewhat rare conditions, for example (1) extensive collaboration between NHS managers and clinicians to agree on the governance structures; (2) continuity in the form of the recruitment of experienced local general practitioners (GPs) (rather than adjunct professionals) to help manage the triage; and (3) all protocols and guidelines that define the RMC's remit and functioning are codeveloped locally rather than imposed.

Chapter 5 examines the theory that establishing intermediate professional roles, such as the GPwSI, could manage demand in the system by siphoning off and dealing with an intermediate case mix appropriate to their medial experience and proficiency. This ambition fails to realise and the new role descends into an administrative support function when (1) consultants remain in relative control of referral decisions and the protocols that govern them and either (2) GPs retain referral habits by maintaining direct referrals to secondary care or (3) GPs use the new GPwSI pathway to offload many cases for purposes of patient reassurance. The new function comes to fruition only when there is protracted negotiation on (1) division of labour; (2) recruitment strategy; (3) case-mix; and (4) physical spacing.

Chapter 6 examines a procedural reform, namely the theory that providing GPs with direct access to diagnostic tests will enable them to distinguish between those patients who can be managed in primary care and those requiring referral to secondary care. Much depends on the nature of the test. GP direct access to tests designed to 'rule out' serious pathology or 'clinical indicator tests', designed to identify where patients were in a disease trajectory, led to greater efficiencies than GP direct access to tests designed to provide a 'differential diagnosis'. However, the improvements in patient flow resulting from direct access to rule out or clinical indicator tests are realised only when there is clear guidance from specialists indicating which patients should be referred and how they should be subsequently managed.

Chapter 7 examines the role of guidelines in moderating demand. The evidence revealed repeatedly an elementary design principle that they work more effectively if they are adapted to local circumstances. The process of adaptation, however, requires far more than the rewriting of guidance to a local rubric. The guidelines need to be mapped into local organisational structures. The people responsible for adapting the guidelines should be the same people who organise the implementation of the new guidelines. Without this level of accommodation, the documentation is likely to be added to ‘the guideline mountain in the corner of the clinic’.

Patient and stakeholder involvement

Throughout the review we sought advice on our theories and synthesis from a range of stakeholders, including those within our project team, individual and group meetings with key stakeholders within the local health economy, a specially convened stakeholder group meeting and regular meetings with a patient group. Their views helped us to develop and revise our programme theories, assess whether or not our findings were of use to their own decision-making and identify appropriate methods of dissemination.

Conclusions

There are no instant reforms that have the capacity to deal with the deep-seated and system-wide strain on capacity in health provision. There is no best-practice manual to be found, out there or in this report, to guide health-care personnel on when, where and how to make referrals. Instead, our review found many, diverse, hard-won, local and adaptive solutions. The key evidence in this review comes in the form of detailed expositions of the immense difficulties and occasional, bespoke successes in bringing into equilibrium the interlocking systems on which sustainable change depends.

Our conclusions (see *Chapter 8*) are thus presented as a small set of design principles, gathering together the configurations of ideas that apply in the most effective demand management interventions, pointing to the strains that have to be overcome and the interdependencies that have to be forged. The conclusions also include vignettes of some successful reforms – presented because they portray an understanding of the nature of system adaptation rather than as blueprints to be imitated indiscriminately. A final recommendation turns to the importance of system-wide collaboration in affecting sustainable change. ‘Group model building’ offers considerable potential in realising this goal.

Chapter 9 is aimed at policy-makers and practitioners who, perchance, may not have the time or inclination to wade through several hundred pages of detailed analysis. A constant theme of that analysis is that demand cannot be managed by rote. There are no silver bullets; there is no list of best practices to be imitated. There are, however, many instances in which capacity has been well managed and this has usually followed through trial and error as practitioners think through a succession of challenges that apply in their corner of the patient pathway. Different practitioners occupy quite different intersections in the system but what is transferable is the process of thinking thorough all of the conditions and caveats. Hence our final title: ‘Thinking it through: Prompts for practitioners’.

Funding

Funding for this study was provided by the Health Services and Delivery Research programme of the National Institute for Health Research.

Chapter 1 The challenge: reviewing the attempts to solve a 'wicked problem'

This chapter sets out our agenda. It covers the basic premise, the research strategy, the topic and the structure of this report. Above all, it is an attempt to measure up the considerable challenge we face in reviewing interventions that have sought to contain the burgeoning demand on planned secondary health care.

What follows is a 'review', a work of secondary analysis. This approach has become synonymous with the notion of evidence-based policy (evidence-based medicine, evidence-based management and so on). In pursuing it we concur wholeheartedly with the basic premise. Increasingly, the method of choice for tackling organisational problems and bringing about system improvement is the 'intervention' or 'programme'. Most interventions have a long history. They are tried and tried again and researched and researched again. Rather than adding unremittingly to the long list of primary inquiries, it seems sensible to work systematically through all of the existing research, with the idea of synthesising findings in the expectation of providing transferable lessons for future applications of that intervention.

This then is the common cause of all reviewers. Inevitably, given the vast repertoire of primary research methods, there is debate about the best method of drawing together all the extant material (for a wide-ranging appraisal see Suri and Clarke¹). Nowadays, secondary analysis is conducted under many different banners and it is possible to locate reviews that have employed systematic review, meta-analysis, meta-study, meta-ethnography, narrative review, meta-narrative synthesis, reviews, configurational reviews, Bayesian reviews, multimethods reviews, best evidence synthesis, critical interpretative reviews, realist synthesis, etc. There is so much activity at this secondary level that there are now tertiary 'reviews of reviews'.²

This is no place to argue methodological pros and cons. A potential point of consensus is that the very idea of a producing an 'aggregation', a 'pooling', a 'synthesis', an 'amalgam', a 'composite' of primary research can have different connotations and that this growing toolbox of synthesis methods allows the reviewer to ask a wide range of subtly different but equally valid questions. In the next section we outline our preferred strategy, pinpointing the specific questions that it is designed to ask.

Research method: realist synthesis

Our method of collecting together and drawing lessons from primary research evidence is known as 'realist synthesis'. It was developed by one of the current authors.³ A comprehensive overview on methodological guidance, publication standards, training materials and a bibliography of earlier realist reviews is available as part of a previous Health Services and Delivery Research (HSDR) report.^{4,5} Readers are referred to these sources for a detailed account of the method. In this report, we commence with a brief account of the core principles and of the basic steps of the realist approach, assigning the full details on the conduct of the complex search strategies to *Appendix 1*.

Realist synthesis avows membership of the 'theory-driven' school of evaluative inquiry. It differs from other forms of systematic review in spending a considerable amount of research time in attempting to tease out the ideas, the assumptions, the logic or the 'programme theories' that underpin the interventions under inquiry. All policy-making and all interventions begin with theory: accounts of the causes of the problem under scrutiny (diagnosis) and conjectures on what changes must be made to the system in order to alleviate that problem (remedy). This starting point establishes the fundamental question for the review. Did these theories come to fruition? Did the remedy meet the diagnosis? The underlying expectation, and one that holds especially true in the field of demand management, is that social and behavioural

interventions tend to meet with partial success. This state of affairs provides realist synthesis with its fundamental task, namely to chart the many contexts and conditions in which interventions work (or fail) and, above all, to explain the reasons for the inevitable mixed picture. Better explanations provide the ammunition to make better policy. They make provision for improved targeting and implementation of interventions.

The conduct of realist synthesis follows an orthodox path, starting with a research question and then finding the means to answer it by searching, identifying, appraising, extracting and combining together materials from the body of existing research. There are significant differences from other review methods, however, most notably in that this sequence is iterative – it is travelled repeatedly. The learning from one tranche of studies provides provisional explanations and new clues that that can be further refined by focusing and refocusing the searching for crucial primary materials. Learning accumulates as the review travels *around* the research cycle.

We now set out the basic sequence of steps in a realist synthesis.³ For each, we provide some forward glimpses into our review, illustrating its application in researching the field of demand management.

Identifying the review questions

The process of locating the proposed diagnoses and planned remedies that have gone into the making of an intervention is known as 'theory elicitation'. A search is undertaken to locate this material, which is normally discovered in the documentation produced in conjunction with the programme under investigation (plans, consultations, thought pieces, administrative accounts, legislation, critiques, etc.). Theory elicitation is an elusive business because the thinking behind interventions is never straightforward. Programme theories may be multiple; they change; they are assembled unevenly; they may be contested. Nevertheless, they prove a strategic point of origin for realist reviews, which then proceed to investigate whether or not, to what extent and why the programme theories have gone on to demonstrate their mettle in practice.

Already we see that an intensive search procedure precedes what is normally the very first stage in a review, namely the setting of the research questions. Advancing a pace or two into the evidence allows the reviewer to pose a better set of questions. In the present case, and prior to the research, we were aware, of course, of the very many schemes mounted in the name of demand management. A preliminary task is thus to sift and sort through this assemblage of schemes, classifying them by the underlying 'theory of change'. Varied objectives could be discerned, with different interventions seeking improvements at the strategic, administrative, functional, procedural and motivational levels.

We were aware, furthermore, of some of the major impediments to the effectiveness of many schemes. Prior to the review and in our initial research proposal we anticipated that reforms implemented to regulate demand might be snagged because of:

1. the different and sometimes contending motivations that prompt referral
2. the varied and sometimes uneven expertise and mandates of the participants in referral chains
3. the lack of any traditional remit for cost-containment ambitions in NHS staff groups
4. the difficulty in regulating provision while at the same time responding to other initiatives providing patients with an increased choice of provision.

Locating these issues in the 'grey literature' allowed us investigate these ideas using the very language of practitioners and in the contexts in which they had experienced them. This pointed us, for example, to very specific concerns about new, special-interest posts located between primary and secondary care. They had the potential to siphon off demand – but could they deal with issues 1 and 2? Could they bridge the disparate motivations and traditional mandates within the two sectors?

In the course of our preliminary search we thus adapted, extended and further specified our provisional hypotheses. Theory elicitation also alerts the reviewer to things they did not know at the outset of the research. A major issue we discovered over and over again in the critical debate about demand management was the idea of system complexity. Perhaps the key, emerging programme theory is the idea that multiple, synchronised improvements are necessary to promote sustainable change.

These preliminary investigations provide the eventual structure of our report. We identify four major subsets of demand management programmes and detailed subsets of research questions investigating the potential strains associated with each.

Searching for primary studies

Once there is a clearer, if broader, picture of what the programmes under investigation intend to do, realist synthesis now moves to gather evidence on whether or not they do so. This initiates another search of the literature to locate primary research studies that enable us to test the programme theories. Have the prior expectations proved justified? In this case the search is targeted at what might be considered as the more orthodox empirical literature, that is to say the papers, studies, reports and previous reviews that have undertaken an evaluation of the implementation and effectiveness of particular demand management interventions.

These searches utilise comprehensive databases such as MEDLINE, EMBASE and Google Scholar (<https://scholar.google.co.uk>; Google Inc., Mountain View, CA, USA). The search is organised by terms and synonyms describing the relevant interventions and, thus, for instance, might cover referral management 'centres' or 'gateways' or 'clinics' or 'systems'. Detailed flow charts describing the strategy pursued to underpin the report's major chapters are included in *Appendix 1*. More specific search terms were also determined by our prior investigations of programme theory. For instance, the avenue of inquiry on the use of intermediaries such as general practitioners with special interests (GPwSIs) turns on evidence about whether they provide 'substitution' or 'support' and these terms were used to locate further nuggets of evidence. In pursuing these more specific lines of investigation, realist synthesis turns increasingly to hand-searches and references of references.⁶ As noted above, such searches were sometimes renewed in the analytic stages of the review in order to further clarify outcome complexities.

What is achieved by this stage is a foregathering of the evidence matched to the structure of the inquiry devised at stage 1. All of this preliminary activity acts as necessary ground clearing to a review and, as is typical of realist synthesis, it took our exercise through its first third (in time and in the report). It presented us with a large 'menu' of potential lines of inquiry, which we then organised in the four 'empirical' chapters of the report.

Quality appraisal

Having uncovered a large number of potential primary sources to assist in our explanatory review, we now go on to the next stage, which is to weigh up the quality of the evidence provided by each study. Realist synthesis eschews the idea of a hierarchy of evidence. The interventions under review are complex systems embedded within complex systems. Delving into the assorted implementation process, contextual differences and outcome complexities requires us to call on the entire repertoire of social, behavioural and organisational investigations.

Standards still apply, of course, the first of which is that a primary study has to provide *relevant* evidence. This requirement commits the reviewers to a large amount of preliminary reading. To continue with a previous example, of the very many reports on GPwSIs, some explored closely the matter of whether the role evolved as support or substitution and how this went on to impact on patient flows. Other research, less relevant, considered the coverage of GPwSIs across clinical specialties and regions and was primarily concerned with the overall level of penetration of a national policy. Study-by-study decisions on relevance were made along these lines.

The second standard relates to *rigour*. The requirement here is for an investigation to be of sufficient standard *within type*, be it a qualitative interview, a process evaluation, a demographic analysis, a randomised controlled trial (RCT) and so on. For instance, studies that purported to provide testimony on GPwSIs' attitudes to their new roles varied from the anecdotal and partisan to those based on transparent and systematic research designs. Study-by-study decisions on rigour were made along these lines, a task requiring substantial methodological expertise on the part of the review team.

Extracting the data

Conventional reviews comb through all included studies, utilising a standard data extraction form mining exactly the same information (typically on treatment modality, population, effect sizes, etc.) from all studies. In realist synthesis the expectation is that each included study will address a different aspect of the programme theory. Broadly speaking: evidence on outcome patterns will be found in trials, survey research and administrative reports; evidence on implementation is found in process evaluations, qualitative interviews and personal testimony; and evidence on context is found in case studies and comparative inquiries. Primary studies vary in the coverage of this terrain. There is no expectation that any inquiry will be all-inclusive and the information extracted is determined by its theory testing potential.

The practical steps of eliciting the evidence begin with a rough annotation of the relevant text, testimony and tables for each included study (this process is aided given the preceding 'relevance test'). The materials highlighted are quite diverse. One arm of our inquiry tested whether or not providing general practitioners (GPs) with direct access to clinical tests generated demand efficiencies. We developed a hypothesis that gains would follow only according to the specific function of a test, which required the extraction of different bodies of evidence; for example, (1) on that function (from clinical documentation); (2) on the level of GP discretion in interpreting results (from administrative case studies); (3) on the level of reassurance provided by a test (from qualitative studies of physicians and patients); and (4) on subsequent patient referral routes and clinical outcomes (from before-and-after studies and RCTs).

Data elicitation is completed in the reportage of each study within the review. In realist synthesis, significant portions of the primary texts are propelled directly into the review. Looking ahead into this report, the reader will note that each of the included studies is covered in anything from half a side to half a dozen pages. The typical format is to justify the relevance of the study in relation to the theory under test by providing contextual details; to reproduce the pertinent evidence (of whatever type); to report on its quality; to record the inferences drawn by the original authors; and to draw these inferences into the overall explanation developed in our synthesis.

Synthesis

Bringing all this material together brings us to the purpose of a realist review, namely to provide a better understanding of how interventions really work. The aim of the exercise is to discover 'for whom', 'in what circumstances', 'in what respects', 'over what duration' and, above all, 'why' an intervention might work. And, although it is impossible to cover every single condition and caveat that might influence effectiveness, the parameters that are investigated afford a degree of generalisability to the review.

Realist *synthesis* refers to the idea of making progress in explanation. For instance, in reviewing interventions that seek to introduce an entirely new referral management centre (RMC) to control the referral process, we begin by discovering a study claiming that its effectiveness depends crucially on developing close collaboration between GPs and local consultants. This inference is hardened as further primary studies are uncovered providing evidence on the same proposition across a number of RMCs. This explanation, however, begs the question of how such collaboration is established and this directs the review to further studies and supplementary explanations about the joint authorship of referral guidelines, the recruitment of local, mutually respected professionals to manage the centres, the negotiation of clear time scales and logistical pathways for each referral, etc.

All of our evidence is presented in this sequential, accumulative manner. The developing explanations do not provide simple blueprints for 'best practice' that can be imitated blindly – all supply-and-demand pressure points are, to some extent, unique. The explanations do, however, provide decision aids on system adaptation. They guide the planner and practitioner on the many necessary conditions that lead to programme effectiveness. All of these exigencies require 'thinking through' to fit them to local conditions and it is this spirit that we present the conclusions to our synthesis.

Background: what is demand management?

We turn next to our substantive topic, namely to review interventions aimed at curbing demand for planned health care. The problem is immediately familiar to anyone with the slightest association with health service provision. Medicine has been a victim of its own success. All advanced health systems face substantial increases in activity and costs, with a seemingly unstoppable rise in demand for planned care. A particular strain is often felt on the matter of referral management, where the patient is relayed from one part of the system to another, often without due care being given to the balance of resources across the system. These are perennial problems and this is our playing field. As per expectations, putative solutions have been contemplated time and again, exploratory interventions have been implemented over and over and they have been researched again and again.

The task of reviewing demand management interventions, however, does have its particular challenges, which governed our approach and which, in turn, structure this report. The distinctive feature of our chosen 'intervention' might be best explained by reaching out to the much contemplated idea of 'wicked problems'. The basic notion is simple and sensible enough. Not all problems are the same, some are more intractable than others, and essential differences in terms of problem complexity should determine what we should expect by way of a solution and how we search for solutions.

Rittel is usually credited with introducing the distinction between 'tame' and 'wicked' problems into the planning literature and it is worth rehearsing the essential contrast as a prelude to our review.⁷ Tame problems, the argument goes, (1) have a well-defined and stable problem statement; (2) have an easily recognised stopping point when it is clear a solution is reached; (3) have a solution that can be objectively evaluated as right or wrong; and (4) belong to a class of similar problems that can all be solved in similar ways. Wicked problems lie at the opposite poles, where (1) the problem is ill structured, arising from sets of interlocking issues and constraints; (2) there is no overall resolution, only a series of partial and time-limited solutions; (3) there is no stopping point, with each partial solution creating a new system with its own strains; and (4) each manifestation of the problem is different, requiring the production of a revisable and versatile set of solutions. The last point should perhaps be emphasised. Wicked problems are neither unfathomable nor insoluble; rather, they require iterative, adaptive, ongoing responses.

Our thesis in this opening chapter is that the problem of demand management is much better understood as a 'wicked' rather than a 'tame' issue and that this has significant ramifications for the conduct of an evidential review. We are aware, in using this distinction, that it is in itself an oversimplification. The literature on systems theory and complexity science, originating in contribution by Rittel and Webber,⁷ has since become byzantine (for an elegant, modern summary see Byrne and Callaghan⁸). Each of the conceptual anchorages has become contested, with, for instance, wicked problems now being apportioned to somewhat different species as 'super wicked', 'chaotic' and 'messes'. Complexity theory itself has spawned many different subschools, each with its own way of describing complexity: general systems theory, morphological analysis, holism, chaos theory, actor–network theory, diffusion models, soft systems theory, agent-based modelling and so on.

Our starting claim avoids all such elaborations. It rests on a straightforward comparison that is difficult to contest – namely, that relative to most topics covered in health-care research the dilemmas of demand management typify all the features of a wicked problem. One characteristic in particular shines through.

As we shall see, every reform that is forwarded tends to expose new aspects of the problem, requiring further adjustment to the understanding of the initial problem. Our review thus imitates something that we discovered in the primary testimony, namely that an understanding of the true nature of supply and capacity imbalance develops alongside, rather than prior to, the application of potential remedies.

Coming to its specific features, we shall see as the review unfolds that demand management interventions are 'relatively complex' according to the following typology:

1. **Causal structure.** There are many notable interpretations of the causes of rising demand for health care, emanating from quite different sources: the inability to cope with backlogs arising from fluctuations in demand; the blockages in treatment pathways caused by professional closure and turf protection; the diverse motivations for referral that extend beyond formal clinical need; innovations in screening and testing producing self-propelling diagnostic cascades; innovations in treatments and services leading to supply-induced demand; changing patient demographics, particularly in age, driving up need for care and services; and the informed patient and the patient choice agenda increasing demand for all services. These causes, moreover, inter-relate, and their effects can be multiplicative.
2. **Contextual diversity.** As with its root causes, the locations of the potential demand management reforms are also widely dispersed. Demand has to be managed for every single condition from abdominal aortic aneurysm to zygote intrafallopian transfer. As one moves through the alphabet of disease, and from physical to mental health, and from pre-natal to end-of-life care, it becomes clear that each sphere has quite different supply and demand profiles. The institutional contexts responsible for managing these movements are also diverse. Modern referral services rarely run directly from the GP's surgery to the outpatient clinic – all manner of appointment, transfer and triaging services are positioned in between (choose-and-book systems, ambulatory monitoring devices, text reminders, community-based specialists, etc.). These managerial systems, moreover, are under the control of the wider apparatus of trusts, commissioning bodies and boards (national, regional and local). Quasi-autonomous non-governmental organisations (QuANGOs) such as the National Institute for Health and Care Excellence (NICE) and the Care Quality Commission as well as targets such as those contained in the Quality and Outcomes Frameworks (QOFs) all set parameters in which demand and capacity is constricted. All of these features are subject to periodic, top-down reorganisation, continually reconfiguring the scope for adjusting demand.
3. **Implementation diversity.** Referral chains themselves are long and complex, capable of providing potential pinch points at any point in the existing or newly modelled pathways. Again, in the modern era, the key players in referrals far exceed the traditional axis between the GPs and the specialist. 'Handovers' are also the business of practice managers, assistants, screening personnel, triagers, physiotherapists, radiographers, haematologists, geneticists, practitioners with special interests, appointment clerks and, of course, patients. The communications flowing across these channels are also diverse – referral letters, appointments, reminders, test results, patient records. The communication media are also manifold, with referrals being conducted on paper, by telephone, by text and on the internet. All of these implementation chains need to be reviewed with an eye on the 'weakest link' syndrome.
4. **Programme-theory diversity.** Perhaps because they echo the diverse range of causes, contexts and channels of demand, it turns out that interventions designed to contain demand not only are huge in number but also vary greatly in their basic ideas and programme theories. The procession of 'approaches' includes RMCs, administrative triage, joint working, service relocation, clinical triage, referrals to GPwSIs, financial incentives, audit and feedback, guidelines, queue sculpturing, behaviour change and so on. Another indication of complexity here is that the focus of the intended remedy varies from the macro to the micro, closing in progressively from shifts in overall strategy, to organisational remodelling, to change in individual's role, to procedural modifications, to motivational change. A further wicked twist here is that many interventions, intentionally or otherwise, end up pursuing change at all of these levels and are thus dependent for their success at achieving change across this entire piste.

5. **Outcome conundrums.** Although the aim of our class of interventions is seemingly clear, namely to control demand, the intended outcome is sometimes obscure and often contested. The moment when demand is considered 'unmet' is difficult to gauge, the politicised wrangling about the measurement of NHS waiting times being a prime example. Issues of quantity and quality also intermingle awkwardly. This is particularly evident in the underlying demand management aspiration concerning the 'appropriateness' of referrals. One of the causes of excessive demand is often considered to be the tendency of some practitioners to make improper or premature referrals. But the reasons for referral are themselves complex and judgements on appropriateness/inappropriateness is heavily contested. Indeed, this polarity is a two-way street for, as we shall see, it is quite possible that the goal of manufacturing better considered referrals can lead to an increase in demand.
6. **Intervention history.** These interventions are not mounted from scratch, that is to say they are never implemented in the absence of existing demand controls. As soon as any service is created (be it for the provision of housing, hamburgers, hunting horns or health care) it has to face the universal issue of managing shortfalls or oversupply in those services. All services wrestle with the permanent challenge of balancing demand with capacity. Accordingly, our inquiry presupposes that informal demand management regimes, in varying shapes and forms and operating at varying levels of success, are generally in place throughout a health service long before its practices become proceduralised and recognised formally as 'demand management programmes'. This previous history will always limit the effectiveness of 'new interventions'.
7. **Ever-present emergence.** The idea of emergence is perhaps the central proposition of complexity theory. It refers to the idea that novel structures are created by the interaction of the existing process within a complex system. It relates closely to the aforementioned wicked issue that some problems have no stopping point, with each partial solution creating a new system with its own strains. The concept of 'supply-induced demand' speaks to this very point. A new practitioner role or patient pathway may alleviate demand or referral pressures within some previous service arrangement – in time, however, it may become so popular that it becomes a pressure point of its own.

Hopefully, the above seven dimensions of complexity provide some indication that the problem under scrutiny is indeed wicked. We cease our initial exposition of intricacy at this point, reflecting briefly on the issue of how complexity impacts on the conduct of a systematic review. The implications are considerable and it is instructive to demonstrate this with a brief comparison of how topics are delineated in other modes of systematic review: 'Systematic reviews should set clear questions, the answers to which will provide meaningful information that can be used to guide decision-making. These should be stated clearly and precisely in the protocol'.⁹ Many reviews are traditionally conducted under the assumptions that they are dealing with stable systems – given populations (P), given interventions (I), given comparators (C) and given outcomes (O). When agreed operational definitions of this 'PICO' formula are in place, the review proceeds under the assumption that it is dealing with interventions that share these homogenous, reproducible and exact parameters.

To be sure, the more formal traditions of systematic review have also turned their attention to complexity, led in the health domain by the Medical Research Council's guidance on developing and evaluating complex interventions.¹⁰ This document, however, admits a rather restricted view of complexity into the evaluation palate – in recognising that programmes may have multiple components and multiple outcomes. Programme implementation is still seen as something to be controlled rather being subject to professional judgement, constant negotiation and perpetual improvement regimes. Heterogeneity is seen as a blight rather than an inevitability.¹¹

To make the point in a less partisan manner, it should be noted that the inclination to limit the remit of inquiry is indisputably *not* restricted to the statistical tradition. All reviews, whatever their methodological complexion, are assisted if their scope can be reduced. Scope control is clearly evident in existing reviews conducted under the realist banner. The study by Wong and colleagues contains a catalogue of realist health-care reviews conducted up to 2012, with such titles as 'Efficacy of school feeding programmes', 'Water, sanitation & hygiene interventions in reducing childhood diarrhoea', 'Role of district nurses in

palliative care provision' and 'Effectiveness of ban on smoking in cars carrying children'.⁴ Although the subject matters here are hardly 'tame', and although unforeseen complexity is discovered in each of these inquiries, it is a reasonable inference to suppose that the overall scope of each exercise is better circumscribed than in contemplating the full, sevenfold ferocity of entire 'demand management' regimes.

Outline and design of the current review

We close this introduction with our plans for squaring this circle and limiting the scope of our investigation. We also set out the chapter-by-chapter structure of the report. All realist reviews involve 'prioritisation', a decision on which of the myriad potential programmes theories should be put to investigation.³ Demand management reaches into every stitch of health service delivery: where should we begin to weave? In this respect it is appropriate to acknowledge some significant circumscription built in to the commissioning of this our review. Our remit from HSDR was to cover demand management for *planned care* rather than emergency provision. We were also commissioned to review referral management interventions, as they are located between *primary and secondary care*. At a stroke, this eliminates consideration of the formidable demand dilemmas involved at the very beginning and end of the care cycle. Intricate referral sensitivities permeate preliminary triage initiatives such as NHS Direct and the 111 Service and demographic change places an ever-increasing demand on palliative care. Although we are thankful that these interventions have been removed from our overflowing plate, it is the case that they too could benefit from rigorous secondary analysis.

So, how have we prioritised investigation of the residual, but still formidably complex, sectors of demand management? The most obvious strategy is to limit investigation to a subgroup of the familiar demand management interventions. Previous research has distinguished (at least in name) over a dozen different schemes and strategies. Lacking the resources to tackle them all we have focused on four: RMCs, GPwSIs, GP direct access to tests and referral guidelines. Our reasoning here is that these interventions span quite different theories of change – seeking to modify demand, respectively, by transforming organisations, inserting posts, swapping tasks and improving recommendations. We have attempted to build these subsections of the report in enough detail to supply some decision aids on the contexts and respects in which these different ideas work.

We have also striven to retain a 'complexity lens' throughout the report. Partly, this motivation emanates from broader ongoing changes in the focus of health service research.^{12,13} The latter group of authors begin a recent paper thus: 'Incremental approaches to introducing change in Canada's health systems have not sufficiently improved the quality of services and outcomes. Further progress requires 'large system transformation' considered to be the systematic effort to generate coordinated change across organisations sharing a common vision our goal'. Our findings point repeatedly to fact that demand management involves whole-system transformation and our review supplies a modest test of this ambitious thesis.

The main consideration for using a 'system complexity' approach, quite literally, is that it has been forced on us in conducting our review. Repeatedly, as we scrutinised the findings of a particular inquiry of a particular demand management initiative in a particular corner of the health service, we have emerged with the interpretation that what transpired was more a matter of the interdependence of the intervention within a wider range of determinants rather than what might be considered as the specific action of each programme. Repeatedly, we came across promising, partial and time-limited solutions that created a new system with its own strains. Over and again, we discover demand management has no stopping point. Rather than just report this as a nervous tic of a conclusion, we have sought to explain why it is the case and structured our report to this end. The remaining coverage is as follows.

Part 1

The runaway train: the multiple, intertwined causes of growth in demand for health care

Chapter 2 aims to 'size up' our wicked problem. Basically, the idea is to gain a sound understanding of the essential causes of demand/capacity imbalance – before we go on to study the research on its potential solutions. We undertake it here in pursuit of an idea that that the research team came to think of as the 'punch-bag hypothesis'. Interventions, via their embedded programme theories, tend to select out and attack specific interpretations of why the underlying problem arose in the first place. In the case of the disproportionate demand for health care we begin with multiple, competing accounts of its genesis. A meta-hypothesis thus lurks, namely that if one manages to land a blow on one of the obstacles to balancing demand and supply, it may simply be absorbed by all the other impediments stuffed in the punch-bag. Whether or not this conjecture turns out to be overly pessimistic will be determined in the remainder of the review.

The preliminary task here is to gauge the contents of the punch-bag, and this is also a task of research synthesis. The coverage here is designed to identify the main 'schools of thought' on the seemingly inexorable rise in the demand for health care. These explanations range across the social science disciplines, covering physician motivations, professional closure, demographic change, diagnostic improvements, supply-induced demand and so on. Our task is not to rank or adjudicate between these accounts but simply to provide a typology encompassing the wide range of demand pressures and to provide evidence showing that each one is substantial enough to command a policy response.

The policy response: charting the family of purported solutions

Chapter 3 aims to 'size up' the battery of potential solutions aimed at curbing demand for planned health care. As anticipated, they are many in number and include RMCs, clinical assessment and triage services, service relocation, referrals to GPwSIs, financial incentives, audit and feedback, guidelines, queue sculpting, behaviour change and so on. As per usual in realist synthesis, the aim here is to elicit the programme theories underlying each intervention, as exemplified in their supporting documentation. This task proved unusual and unusually instructive because of the intransigence of the demand problem and the longevity of attempted solutions.

The first notable feature of the programme theories is something we refer to as their 'ontological depth'. More simply, this refers to the scope of the intended solution, which varied substantially from the micro to the macro as one traversed the thinking behind the different interventions. The main engine of changed is construed, in turn, at strategic, administrative, role, procedural and motivational levels. What also could be clearly observed, with the passage of time, was a process of 'mission creep'. Programme theories are not static. In the light of experience gained under critical scrutiny and via the trial and error of actual interventions, the core assumptions tended to modify, recommending the need for supplementary action. Invariably, the programme theories became 'whole-system' models – suggesting that sustainable change required the interweaving of the various macro, meso and micro mechanisms. Combination punches become the order of the day.

Part 2

Many different health-care reforms turn on the idea of 'mandate' change, with the responsibility for decision-making changing hands among the ranks of health-care professionals, and this notion provides the linkage between *Chapters 4, 5* and *6*. The patient journey through initial consultation, diagnosis, testing, treatment, care and after-care is long and winding. The institutional and staffing structures originally designed to carry decision-making along this referral pathway have a tendency to become unbalanced under the perpetual rise in demand for health care. These incessant strains have been met by continual adjustments with the idea of introducing an improved division of labour in order to check and smooth the treatment tides. These new regimes differ in their latitude, some involving the implantation of new organisations, some introducing new staff positions and some simply changing responsibilities among

existing roles. Each chapter provides a review of the empirical evidence on effectiveness of the new mandates.

Organisational change: referral management centres – can they control and shape demand?

Referral management centres (and related services including 'gateways', 'triage centres' and 'single point of entry services') are intended to act as intermediaries between primary and secondary care. Their aim is to prevent unwarranted referrals from reaching secondary care, and to redirect them back to GPs or to redirect them to alternative community services. The empirical evidence in respect of these ambitions is mixed. A number of unintended consequences may occur. The co-option of clinicians into such centres may strengthen their power as services are managed according to clinical values. GPs may question the expertise of centre operatives and may bypass the new structures if they no longer retain decision-making powers in respect of patients with whom they are deeply familiar. Conversely, RMCs may also lower GPs threshold for referral and result in the offloading of routine work to lower-status professions. A number of characteristics can be identified in the more successful RMCs: all incumbents need to be engaged in their design; feedback to GPs about why their referrals have been blocked or redirected needs to be issued directly to the referring GP in a timely manner; and the local health economy should contain both the necessary skills and expertise and the capacity to manage the diverted referrals.

Role change: general practitioners with special interests – can they control and shape demand?

The development and formalisation of the innovative role of GPwSIs is usually traced back to the NHS Plan from 2000.¹⁴ These new positions marked a widely supported development away from the role of family practitioner. GPwSI-led services were charged with a number of functions including taking responsibility for an 'intermediate case-mix', with consequent reductions in hospital outpatient services and waiting lists as well as improved accessibility for patients. The evidence in respect of these ambitions is mixed. There is little solid evidence that GPwSIs have reduced the outpatient workload or outpatient waiting lists. There is, however, no evidence to suggest that clinical outcomes differ between GPwSIs and outpatient services, although there is a difference in their case mix. There is solid evidence to suggest that GPwSI services attract increasing referrals and offer shorter initial waiting times (which do, however, lengthen with the longevity of their services). There are mixed and indecisive data on cost savings. These irregular outcomes are rooted in the remarkably ragged professional remits of GPwSI posts and the uneven distribution of their posts in different corners of medicine. More successful outcomes ensue if (1) consultants retreat to the specialist role in the treatment of complex patients and become advisor to rather than supervisor of the system to identify them, (2) GPwSIs deal with an agreed and intermediate case mix and work in a (physical and social) space that is independent from the GPs and consultant surgeries and (3) GPs work to agreed referral protocols developed via GPwSI educational and management functions.

Procedural change: direct access to the results of clinical tests – can it control and shape demand?

General practitioner direct access to the results of clinical tests is the oldest intervention we examine in the plethora of demand management solutions; equity of access to pathology and radiography services between primary and secondary care physicians was a topic of much debate as early as the 1960s and 1970s. With the development of ever more sophisticated diagnostic tools and tests, choosing the right test at the right time and interpreting the results in the right way to guide subsequent management are key tasks in the patient diagnostic journey. Allowing direct access to diagnostic testing should enable the GP to distinguish those patients who can be managed in primary care and those who require the expertise of a specialist without the need for an initial referral to secondary care. Once again we find that the evidence is mixed; the nature of tests differs according to their function in patient diagnosis and there are noticeable differences between types of test. GP direct access to tests designed to 'rule out' serious pathology or 'clinical indicator tests', designed to identify where patients were in a disease trajectory, led to greater efficiencies in demand management than GP direct access to tests designed to provide a differential diagnosis, especially if each of the differential diagnoses required specialist referral. The evidence also

showed that specialists in secondary care rarely relinquished control over referral decisions, even with GP direct access to testing. Indeed, maximum utility in terms of demand management was realised when test results include clear guidance from specialists indicating appropriate management in the light of the results.

Learned counsel: can guidelines control and shape demand?

Regardless of health issue, health sector, patient condition or treatment modality, the chances are that provision is supported by 'a guideline' making professionally endorsed recommendations on best practice. And so it is with demand management, a major programme theory resting on the idea that demand can be curbed and sculptured with the appropriate formal guidance. This hypothesis has been met with a mountain of primary research and the main and undoubtedly sound conclusion from these inquiries is that guidelines in and of themselves have a limited role in addressing demand problems. The reason why guidelines falter is little to do with their content and format but is mostly due to complex decision structures in which they are embedded (and ignored).

Accordingly, the main body of research has amassed trying to identify the barriers and facilitators surrounding their implementation. Scores and scores of such impediments have been identified covering all manner of cognitive, attitudinal, patient, professional and resource constraints. Such has been the proliferation of these studies that, hitherto, the main task of review work has centred on the production of the definitive set barriers/facilitators producing a practical checklist of challenges that guideline producers and users will have to overcome. We argue that these overarching, itemised frameworks do not resolve into some sort of winning formula because the factors identified are always interdependent. Being alert to a 'systems' programme theory, our synthesis explains why some barriers are more intractable than others and why solutions always have emergent effects. Accordingly, we focus the review on some more specific remedies to some key 'system strains': (1) the tension in using simple guidelines for complex comorbidity; (2) the tension between (inter)national credibility of and local control over guidelines; (3) the tension between patient choice and top-down guidelines; and (4) the tensions involved when there are competing guidelines and contending targets.

Conclusions: facing the challenge of complexity

The evidence amassed in this report confirms our central hypothesis that demand management constitutes a wicked problem, which has defied clear and easily reproducible solutions. In scouring the evidence base we have unearthed an ever-growing list of contingencies that have to be overcome if such schemes are to succeed. We consider that the task of the impartial reviewer is to be the bearer of reasoned judgements, regardless of whether they constitute good or bad news. Most of the news in this review comes in the form of detailed expositions of the many difficulties and the rare gains in bringing into equilibrium the interlocking systems on which sustainable change depends. At this point it is worth recalling a general principle mentioned earlier, namely that the designation 'wicked problem' does not signify a hopeless, unfathomable, insoluble task. It means that solutions are adaptive, iterative, ongoing and, above all, local. We have indeed 'evidenced' this hypothesis – finding isolated islands of successful demand control amidst a rather storm-tossed sea.

Our conclusions are thus presented as a small set of principles, gathering together the configurations of ideas that apply in the most effective demand management interventions, pointing to the strains that have to be overcome and the interdependencies that have to be forged. We also reprise vignettes of successful 'case studies', showing how they exemplify the key design principles. System improvement is never a matter of blindly imitating 'best practice' but always an issue for understating the underlying mechanics of change.

Thinking it through: prompts for practitioners

Chapter 9 provides a small annex to our conclusions. It is aimed squarely at policy-makers, planners, managers and practitioners, providing them with some decisions aids to assist them in resolving demand pressures. It may be read independently from the rest of the report.

Appendices

Search and selection strategies

In *Appendix 1*, further details and flow charts are provided on the search strategy underpinning the review.

Reviewing the field

Appendix 2 provides a brief examination of previous attempts to review the field of demand management interventions in health-care provision. Given the gravity of the problem addressed, it is a topic that has attracted considerable attention in research review. Rather than engaging in a technical autopsy of each previous review, we have tried to assess them in terms of what is and what is not covered. And what is most often overlooked in the existing analysis is a whole-system perspective.

Stakeholder and patient involvement

Appendix 3 summarises the advice received in our meetings with project stakeholders.

Dissemination activities

In *Appendix 4* a table is provided detailing project dissemination activities.

Part 1 Theory elicitation

If you know the enemy and know yourself, you need not fear the result of a hundred battles. If you know yourself but not the enemy, for every victory gained you will also suffer a defeat. If you know neither the enemy nor yourself, you will succumb in every battle.

Sun Tzu, The Art of War, p. 84¹⁵

Part 1 provides the conceptual anchor to the study. Realist synthesis operates in the tradition of theory-driven evaluation. Programme theories, the ideas and suppositions underlying an intervention, provide the hypotheses to be investigated. Programme theories can be usefully divided into two types: (1) diagnosis of the problem and (2) the plans underlying the remedy. *Chapter 2* seeks to know the enemy: the causes of demand inflation in health care. *Chapter 3* pursues an understanding of remedies as they see themselves.

(Part 2 will assess more than 100 battles in order to see who is victorious and who succumbs.)

Chapter 2 The runaway train: the multiple, intertwined causes of growth in demand for health care

It is in vain to speak of cures, or think of remedies, until such time as we have considered of the causes . . . cures must be imperfect, lame and to no purpose, wherein the causes have not first been searched.

Robert Burton, The Anatomy of Melancholy, p. 52¹⁶

Because the problem of demand management is ubiquitous, it has attracted explanatory attention from a multiplicity of perspectives, both practical and philosophical, as well as from a range of disciplinary perspectives. There are numerous explanations of the notorious imbalance between demand and supply in health care; so extensive that we consider that they command a presence in our inquiry. *Figure 1* explains the function of this chapter within the overall report.

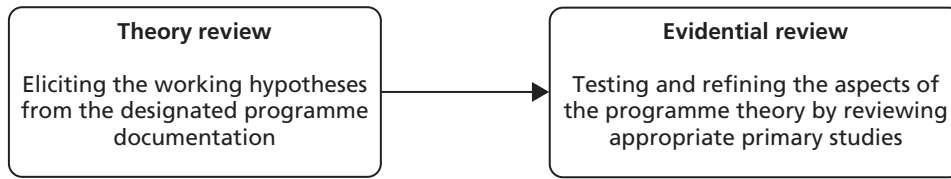
The top section of the figure depicts the standard routine of realist synthesis as applied in most reviews to date investigating confined interventions aimed at limited problems. Here, the existing literature is reviewed in two stages: (1) uncovering programme theories as expressed in the documentation underpinning the interventions under review and (2) searching for empirical studies that provide opportunities to test and refine the said theories. The lower section of the figure inserts an additional and prior search of the generic literature that is necessary as reviews take the leap from investigating well-delineated interventions to the present case in which we aim to evaluate whole suites of demand management programmes aiming at tackling a particularly longstanding, multifaceted and 'wicked' issue.

The entire history of programme evaluation tells us that interventions are never entirely and universally successful. Programme theories prefigure solutions that work only in certain circumstances and in particular respects – and it is the job of realist synthesis to articulate those contingencies. Such a prospectus presents a somewhat chilling challenge when it comes to the complexities of demand management. That is to say, the theories that are uncovered in the specific programmes (at stage 2) may well address only a limited range of theories (uncovered at stage 1) that explain the multifaceted roots of the underlying problem. A morose meta-hypothesis (the punch-bag theory) lurks, namely that we will be reviewing a domain of partial solutions – if one of the impediments to balancing demand and supply is solved, another might well pop up in its wake.

Undoubtedly this *is* too doleful a starting point, even if it is only hypothetical. Locating our review within this wider problematic does, however, justify beginning it with an overview of the broader literature on capacity problems – and this is the purpose of the remainder of this chapter. The coverage here, itself a review, is designed to achieve two aims. The first is to identify the main 'schools of thought', the key explanations for the seemingly inexorable rise in the demand for health care. The second is to assess the veracity and significance of each explanation, which we attempt by incorporating some crucial nuggets of evidence underlying each claim. Basically, the idea is to gain a sound understanding of the basic causes of demand/capacity imbalance – *before* we go on to study the research on its potential solutions. There is a considerable literature devoted to the issue, which we summarise and analyse into seven schools (*Box 1*).

Each section ends with a summary of the 'key issues to be taken forward', consisting of a preliminary attempt to 'size up' the extent of each of the presenting problems. The chapter ends with some overall reflection on the utility and the significance of the typology in *Box 1*.

Standard procedure



Revised procedure

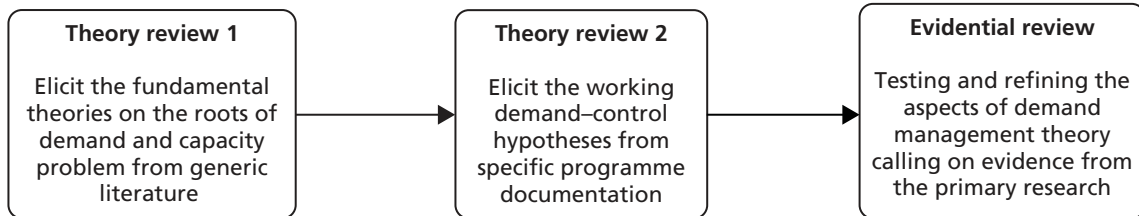


FIGURE 1 Eliciting theory in reviews of ‘focused’ interventions vs. ‘wicked’ problems.

BOX 1 The rise in demand for health care

Seven schools of thought

1. Queuing theory and the question of time.
2. Professional closure, informal control and turf protection.
3. Micro-dynamics in the decision to refer.
4. Self-propelling diagnostic cascades.
5. Supplier-induced demand.
6. Changing demographics and rising demand.
7. The internet, the informed patient and demand inflation.

Queuing theory and the question of time

Queuing theory is a formal branch of mathematics, which perceives ‘demand’ as a queue for services. And, as anyone who has queued will recognise, ‘time’ is the core concern – the demand management dilemma is understood as the task of bringing into balance these two sides of an equation:

1. ‘demand’ = the time *required* to complete a service/procedure in any period
2. ‘capacity’ = the time *available* to complete the service/procedure per period.

Fluctuating demand is seen as the underlying problem – it is irregularity of demand that creates queues. To use a contemporary example – planes arrive at Heathrow in uneven batches (seasonal and daily) and the immigration officials there to greet passengers sometimes sit twiddling thumbs and sometimes are over-run. How can the problem be managed?

Figure 2 (reproduced from Pandit and colleagues¹⁷) demonstrates the problem of setting the appropriate service level in the face of fluctuating demand. The three horizontal lines in *Figure 2* represent different ways in which capacity might be set and foretells of a delicate balancing act. Line 1 is set below the minimum weekly demand. Demand is never absorbed and backlog increases unceasingly. Line 2, set above maximum weekly demand, absorbs all demand, eliminates backlog but wastes considerable capacity. Line 3, the practical alternative, sometimes absorbs demand and sometimes increases backlog.

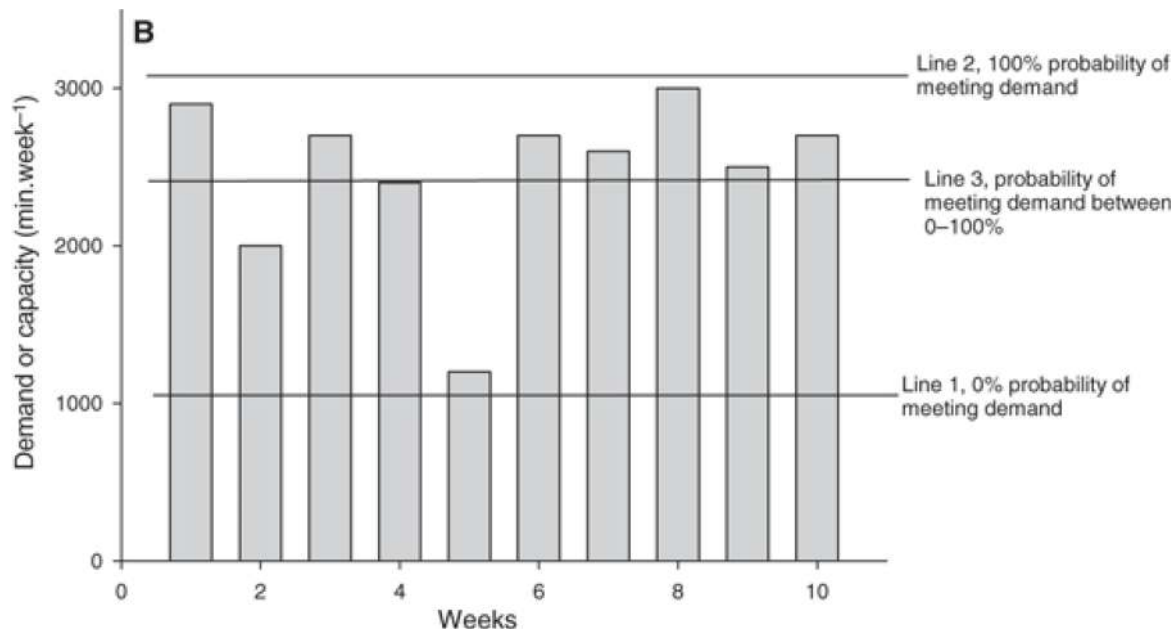


FIGURE 2 Alternative strategies and probabilities of meeting fluctuating demand. Reproduced with permission of John Wiley & Sons, from Pandit JJ, Pandit M, Reynard JM. Understanding waiting lists as the matching of surgical capacity to demand: are we wasting enough surgical time? *Anaesthesia*, vol. 65, pp. 625–40, 2010.¹⁷ © 2010 The Authors. Journal compilation © 2010 The Association of Anaesthetists of Great Britain and Ireland.

A basic and perfectly general insight can be drawn directly from the figure.

Spare capacity in one week cannot be carried forward to the next week. Time cannot be stored for future consumption or sale as can manufacture goods such as cars or washing machines. Furthermore it is not always apparent in any week that time is spare, since this is only known after the event. Therefore the occasions when demand outstrips capacity always contribute to backlog, but the occasions where capacity outstrips demand cannot always compensate and capacity is wasted

Reproduced with permission of John Wiley & Sons, from Pandit JS, Pandit M, Reynard JM, Understanding waiting lists as the matching of surgical capacity to demand: are we wasting enough surgical time? Anaesthesia vol. 65, pp. 625–40¹⁷

A potential paradox is raised, put mischievously in Pandit and colleagues' subtitle – 'are we wasting enough surgical time?'¹⁷ Some waste appears to be necessary to ensure that demand is met and this proposition provides us with one of the standard motifs of queuing theory, namely 'what are acceptable levels of waste?'

Queuing theory is able to make useful predictions, some counterintuitive, on queue dynamics under more complex conditions – what happens if you increase the number of servers? What happens if service levels also vary? What happens if you create a waiting list 'buffer' to even out demand? *Figure 3*, which is reproduced from Pandit and colleagues' figure 10, illustrates a standard response to fluctuating demand. The pictogram illustrates varying demand over a 4-week period. Rather than dealing directly with these demand contours, patients are placed on a pooled waiting list, which *if* operated flexibly and efficiently yields a constant demand in successive weeks. More sophisticated queuing models are able to incorporate and predict the smoothing effect of a range of further demand management strategies.

Critics have pointed out some limitations of queuing theory, namely that it suffers the characteristic limitations of a mathematical theory.¹⁸ It is defined by those inputs (number of arrivals, service requirements, number of servers, etc.) and outputs (probabilities of delay, queue length, mean occupancy, etc.) that can be modelled. Such models cannot incorporate all of the real world sources of variation and uncertainty, which reach right down into the idiosyncratic dynamics of every single consultation and up into contractual negotiations on service provision. Unwisely, some strict rules on service levels have been

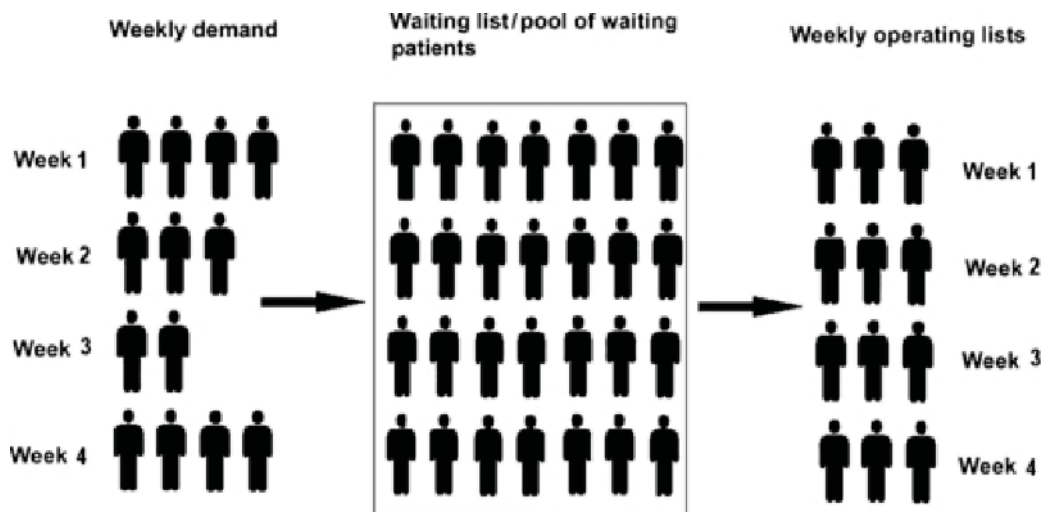


FIGURE 3 Pooling strategy to smooth demand. Republished with permission of John Wiley and Sons, from Pandit JJ, Pandit M, Reynard JM. Understanding waiting lists as the matching of surgical capacity to demand: are we wasting enough surgical time? *Anaesthesia*, vol. 65, pp. 625–40, 2010.¹⁷ © 2010 The Authors. Journal compilation © 2010 The Association of Anaesthetists of Great Britain and Ireland.

adapted on the basis of elementary queuing theory. Bain quotes one oversimplistic edict that was carried directly back into emergency medicine in Australia: 'Queuing theory developed by Erlang nearly 100 years ago tells us that systems are most efficient when they operate at 85% capacity.^[19] This applies to the queues at the local bank waiting for the teller or at a ticket booth at the MCG (Melbourne Cricket Ground)'.¹⁸

Practical resolutions to demand fluctuation involve close tracing of patient flows through every procedure and pause within a particular service. Ironing out demand generally involves close attention to each link in the chain (changing division of labour, changing appointment procedures and priorities, etc.). A glimpse of some specific strategies can be seen in the following 'case study' (Box 2) summarised from the NHS Institute for Innovation and Improvement.²⁰

Our purpose here is not to be drawn into the specifics of this case study, still less to comment on the effectiveness of the suggested improvements. The point is to establish that fluctuation of demand is all-pervasive and operates through all the micro process within a service. It is thus likely to remain pertinent to our review, with the following implications.

Box 3 provides a summary of this section.

BOX 2 Reducing waiting in radiography

1. Eliminating 'carve out'. Weekly slots had previously been divided between routine, inpatient and urgent requests without regard for demand.
2. Covering all sessions, rather than having unfilled slots.
3. Training radiographers to perform barium enemas to address the shortage of radiologists.
4. Introducing flexibility.
5. Introducing protocols for checking requests to avoid delay.
6. Giving individual appointment times rather than group booking times.
7. Booking only barium examinations onto lists, rather than booking many different examinations.
8. Introducing measures and targets where there had previously been none.

BOX 3 Summary 1

Queuing theory: basic issues to be retained

- 'Variation' in demand rather than just the 'level' of demand will itself produce a capacity problem. 'Time' is an important aspect of all health service provision and time cannot be stored.
- A complex system with many stages (i.e. referral systems from GPs to laboratory tests to special consultations to hospital admissions to operating theatre lists to postoperative recovery, etc.) will have 'internal queues', each with inherent variability. The characteristically long and back-and-forth chains of modern medicine will have additive or even multiplicative effects on backlog and wastage.
- Demand management interventions that operate by transforming organisational structures, professional functions or staff motivations are likely to retain a queuing footprint that still may lie at the root of the problem and still may require attention.

Professional closure, informal control and turf protection

From mathematics we move to social science. The sociology of the professions has always had a fascination with the medical establishment's ability to retain power, restrict entry and control clinical practices.^{21,22} The continued social transformation of the medical profession has been explored in terms of the clinicians' relationship with the state, with management, with patients, with the pharmaceutical industry and with evidence-based policy.^{23,24} Here, we want to trace the path of this literature into the daily routine of health care and into the matter of referrals. How might professional power affect the issue of handovers? Here, we report briefly on three of very many case studies that contain the typical features of 'turf protection' and show how these go on to create significant demand dilemmas.

Foote and colleagues' study of the management of ultrasound waiting lists examines a process beginning with the primary patient consultation with the referring GP, moving through radiological investigations and into the secondary care.^{25,26} The study begins with the basic predicament, namely that waiting lists are effectively formal institutionalised agreements between hospital departments and referring clinicians 'to attend to patients at a later date'.²⁵ Satisfaction with this concordat differs from stakeholder to stakeholder, and a range of strategies evolve to modify, manipulate or circumvent this bureaucratic pact. If they believed it was not in a patient's interest to wait at the back of the appointed queue, some GPs were quite prepared to load the system at other points. They made attempts to (1) bypass the ultrasound service by making hospital outpatient appointments, (2) send patients directly to the service in the hope of on-the-day cancellations, and (3) queue jump by forwarding urgent requests.

The real turf wars commence with the arrival of another set of stakeholders. Radiologists, sonographers and departmental managers held a different set of priorities: 'GPs and waiting patients were broadly interested in managing diagnostic uncertainty . . . In contrast, the ultrasounds service's account of the waiting list centred on difficulties in allocating scarce capacity to patients who were likely to have abnormal pathology in order to minimise the impact of waiting lists on ill patients'.²⁶ An 'all clear' diagnosis thus represented a positive outcome for GPs, while a congregation of 'all clears' constitute an unmerited demand for the radiological service.

Much of the immediate contestation between these professional groups occurred on the matter of prioritisation and handling of the 'urgent' requests. Radiologists and managers were fearful that the increasing number of urgent requests meant that the regular waiting list grew ever longer and the attendant publicity grew ever darker. As with most other forms of referral, bureaucratic hurdles were constructed to ration urgent requests. GPs were required to forward detailed, urgent requests only by fax/telephone and were heavily discouraged from not simply marking a referral form as urgent. But even under this system the majority of 'urgent' notifications fetched up with 'normal' scans. The result was a form of stand-off. Given the potential medico-legal consequences of declaring requests as inappropriate, 'radiologists rarely refused to scan GP requests and only passed judgement on the quality of the referral after the patient had been scanned'. That judgement was often brutal: 'The ultrasound service referred to such requests as "rubbish" ' though letting off steam in this manner did little to resolve the numerical log-jam.²⁶

A further internal fissure, with a considerable impact on demand from radiography, occurred between the professional groups responsible for its implementation: radiographers (clinicians) and sonographers (technicians). Ultrasound imaging occurs in real time, potential abnormalities being best uncovered in an intricate, immediate to-and-fro between the production of scans and their interpretation. Investigation and diagnosis are intertwined with a resultant challenge to the identities and the division of labour between the specialist and the technician. One way that the former group were able to maintain their expert role was through the widespread adoption of 'second-look stenography' or 'double scanning' – the radiologist rescans the patient to confirm the accuracy of the sonographer's scan. The demand dilemma is clear: 'Double scanning restricts session throughput (as patients may be scanned twice), creates session overruns (a single radiologist covers two ultrasound scanners) and makes session throughput vulnerable to the availability of a radiologist (as a radiologist must be present when patients are being scanned)'.²⁶

Abel and Thompson's qualitative study²⁷ sheds further light on the ways in which generalists may circumvent attempts by specialists to control their behaviour via guidelines. It also reveals the different perceptions of specialists and generalists regarding what constituted an 'appropriate' referral. The authors interviewed 15 GPs and 11 specialists in New Zealand to explore their use of surveillance guidelines for colorectal cancer. These guidelines were written by specialists and contain advice about when a patient should be referred to secondary care. The guidelines classify people into low-, medium- and high-risk categories and seek to control demand by stipulating that only high-risk groups should be referred. In and of themselves, the guidelines can be seen as attempts by specialists to exert control over the behaviour of generalists.

The authors concluded that specialists and generalists have different perceptions of risk by virtue of their clinical remit and relationship with the patient. Specialists, who are one step removed from the patient, managed risk 'scientifically', often based only on a referral letter and family history. They perceived that GPs referred people inappropriately, in ways that did not fit the guidelines, which they then assumed were 'too complex' for GPs. GPs felt that the formal guidelines often did not 'fit' the patient in front of them and made decisions based on their relationship with the patient and their clinical experience. They also took into account the patient's anxiety about their condition, rather than statistical calculations of risk. Once again, the turf wars resulted in a stand-off. GPs also embellished their referral letters to 'fit' with the guidelines in order to get their patient referred at an earlier date; specialists saw these embellishments as 'lying'.

Currie and colleagues' qualitative study of the development of genetics services run by GPwSIs provides further insight into the turf wars.²⁸ The focus in this case is on the rift that occurs between specialists and generalists. Twenty-four interviews were conducted with three key stakeholder groups (GPs, GPwSIs and specialist geneticists). The GPwSI role involves a reconfiguration of the standard health service division of labour. This new group of specialists take referrals from their fellow GPs, offer diagnostic and some treatment services and provide leadership in primary care in the reshaping of services around particular conditions and disease areas.

In the cases examined, GPwSI services that gained ready acceptance were those designed to *supplement* the work of consultant geneticists. Those that were set up to *substitute* the work of the geneticists encountered greatest opposition from the consultants. New functions such as 'raising awareness' in an 'educational role' prompted support; the development of new clinics within primary care did not. Consultants continued to hold sufficient power to constrain the development of GPwSI services through defining what constituted specialist knowledge and controlling access to training, education and support required to set up the new GPwSI services. GPwSI always had to establish consensus with the specialist regarding the purpose of service. Thus, whether or not GPwSI services were successful depended on whether existing inter- and intraprofessional relationships supported or constrained their development. Demand and duplication dilemmas continued and were sometimes exacerbated in new forms of informal control in which 'geneticists took on the role of appraiser, vetted all referrals to the GPwSI and filled a supervisory role'.²⁸

Although the case studies noted here refer to specific services dealing with specific conditions, the underlying dynamic undoubtedly will repeat itself in other parts of the health-care system. Note that our analysis in this section does not depend on the original research having tapped the precise extent of professional closure available to each subgroup. Accounts of baronial contestation will always be contested. The point is that, whatever the precise resultant, assertions of territorial expertise, in and of themselves, will shape patterns of supply and demand.

Box 4 provides a summary of this section.

BOX 4 Summary 2

Control strategies: basic issues to be retained

- Because of their diverse responsibilities, different professionals will hold contrasting constructions of what constitutes the demand management problem and, thus, different opinions on what strategies to follow in the face of continuing blockage in the system.
- Practical outcomes of this difference of opinion are likely to depend on the extent to which respective professional bodies hold controlling power over the constituent processes. Different coalitions of subprofessions are likely to form in the 'arms race' to retain control.
- Any fresh demand management intervention, be it the introduction of guidelines or the creation of new professional roles, should expect to operate in this contested terrain. The effectiveness of any innovation will, in part, depend on the process of addressing and harmonising contested professional interests and priorities.

Micro-dynamics in the decision to refer

Before corrections are applied to the referral process, it is wise to consider what we know about the inner workings of that process. Numerous studies have explored how generalists – usually GPs – make the decision to refer to specialist care. This work can be traced back to a classic 1981 paper on ‘referral thresholds’.²⁹ At issue is the identification of the exact point at which the GP decides that specialist help is required and, thus, referral is needed. There is a linkage to the above theories on turf wars, through the explorations of where a generalist’s remit ends and a specialist’s remit begins. Several studies have developed general frameworks describing spectrum of influences on GP’s referral decisions.^{30–32}

In this section we concentrate on studies in the micro-sociological or social psychological traditions, which attempt to model how GPs explore patients’ problems.³³ Often, this complexity is described using flow diagrams, such as *Figure 4*, of the interaction and decision-making process.³⁴ To illustrate these studies, we focus on three inquiries exploring how GPs make decisions to refer patients to psychiatric services. This is a domain with significant variation in referral rates stemming from the considerable uncertainty on whether or not, when and where a patient should be referred to secondary care.

Morgan tracked all new referrals to two outpatient psychiatric clinics over 6 months.³⁴ Of the 184 patients referred, 62 did not keep their appointment, 14 did not wish to be interviewed and 120 were interviewed to explore the history of their problems and the chronology of events before their referral. A total of 27 out of 31 referring GPs were interviewed to explore their views on the history and nature of the patients’ problems and how the referral came about. In addition, documentary evidence (referral letters, medical records) and clinical data recorded by the psychiatrist for each patient were scrutinised. In most cases, the decision to refer evolved out of a series of encounters between the GP and the patient and their family over a period of months, usually 3–5, but sometimes up to a year. This suggests, significantly, that referrals are best conceptualised as a process rather than as a single decision at a specific point in time.

Morgan identified three ‘patterns’ or types of referrals. The first category were patients experiencing problems that could clearly be labelled ‘psychiatric’ and beyond the remit of primary care – such as patients exhibiting violent, psychotic or suicidal tendencies. These patients were usually referred as soon as the problem was known. The second category, termed ‘elective’, was less clear-cut; in these cases, psychiatric disturbances were masked by physical symptoms, were transient or were accompanied by other events in the patient’s life. The referral was contingent on a series of events shown in *Figure 4*.³⁴ The GP first treated the physical symptoms and ruled out a physical cause for the problems, which were then recognised as ‘psychiatric’. The GP then provided reassurance and supporting treatment, which represented an attempt to ‘contain’ the symptoms and gave the GP time to ‘wait and see’. Persistent symptoms and pressure from relatives may undermine previous reassurance from the GP and lead to a referral, often in the context of difficult and unstable doctor–patient relations. In the third category, labelled ‘negotiated referrals’, the GP acted as an intermediary to arrange a psychiatric referral at the request of others – sometimes solicitors or hospital consultants but also patients and relatives with ‘ulterior motives’ (e.g. to gain better housing or sickness benefits, etc.).

The severity of the patient’s symptoms was judged a weak predictor of which category of referral came into operation. The authors conclude that the process of referral is structured by social as well as clinical events. GPs sometimes struggle to focus on symptoms rather than the patients underlying problems, which are not always seen as within their remit. Sometimes GPs do not feel they have the time or skills to address them. Furthermore, GPs find themselves managing not simply the patient’s condition but also their relationship with the patient, and they may choose to refer when this relationship breaks down.

In a related study, Sigel and Leiper interviewed a purposive sample of 10 GPs to explore how they defined and detected psychological problems, how they understood psychological therapies and their decision to refer a patient to psychological therapies.³³ They report that GPs saw ‘containment’ of mental health problems as a fundamental part of their management of patient’s mental health problems. Containment included helping patients through their problems, monitoring risk of self-harm and facilitating access to

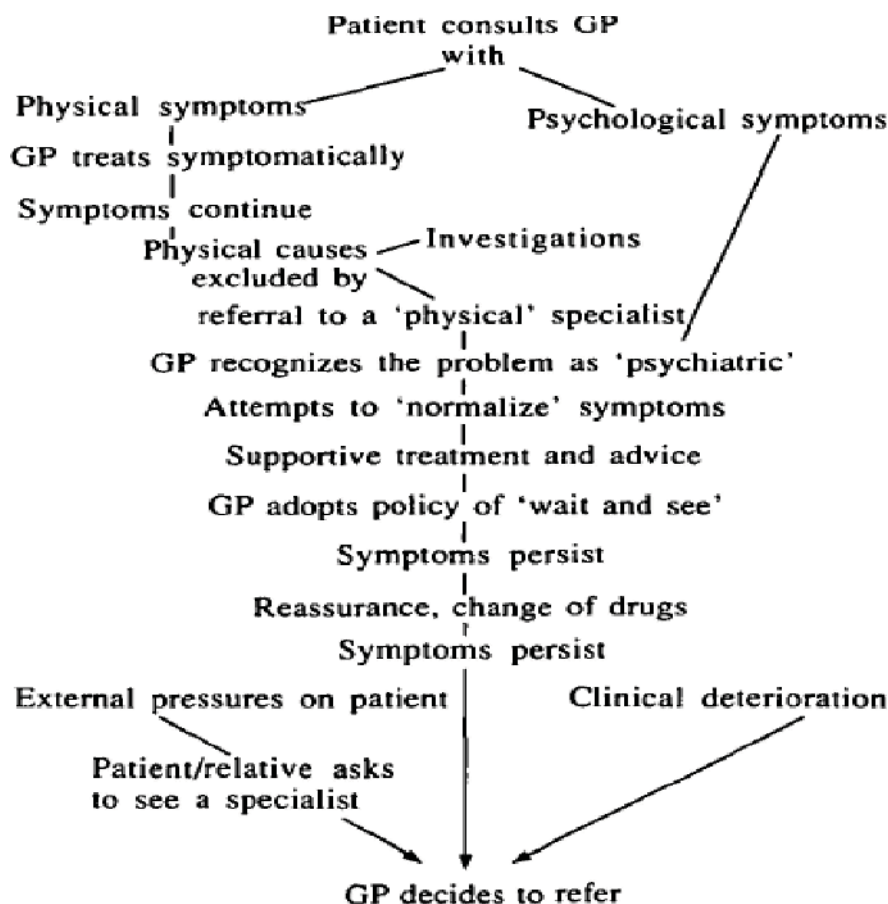


FIGURE 4 A micro model of the referral decision. Republished from Morgan D. Psychiatric cases: an ethnography of the referral process. *Psychological Medicine*, vol. 19, issue 3, 743–753, 1989, reproduced with permission from Cambridge University Press.³⁴

specialist services. Some elements of the findings are strikingly similar to Morgan's study.³⁴ GPs referred when they felt they had reached the limits of their capabilities for addressing a problem. This often occurred when the patient had not improved after the GP had attempted a course of treatment or the GP felt that they did not have the time to address the problem appropriately (similar to Morgan's 'elective' category). In other cases, GPs referred without trying to treat the problem, usually when they felt 'out of their depth' with problems such as psychosis (similar to Morgan's 'psychiatric' category).

In an earlier study, Nandy and colleagues interviewed 23 GPs about their management of minor mental illness.³⁵ They identified two key referral strategies used by the GPs, namely 'conduits and containers', which can be thought of as 'referral to' and 'referrals away'. Containment was adopted by GPs who saw minor mental illness as part of the remit of general practice, while the 'conduit' strategy was used by those who saw their role as diagnostic and that of others as being responsible for the management of patients. Similar to the previous studies,^{33,34} GPs tended to see the patient themselves and then refer when they felt they were not making progress. Containment behaviour was enhanced by an interest in minor illness and by confidence in dealing with mental health problems. 'Referrals to' describes referrals made for 'proactive' reasons, such as the skills of particular professional or because the patient desired to be referred. 'Referrals away' described referrals made for 'reactive' reasons and triggered by feelings of frustration, anger or irritation on the doctor's part where the GP felt that they needed help in managing the patient.

Although these studies have focused on referrals for psychological problems, it is likely that a number of the issues identified in these studies are mirrored in other domains where clinical uncertainty exists about referral.

Box 5 provides a summary of this section.

BOX 5 Summary 3

Decision to refer: basic issues to be retained

- The decision to refer is a process that unfolds over several encounters with the patient. Referrals are not determined by clinical factors alone, but also depend on social and emotional factors.
- GPs' perception of their clinical remit, skills and knowledge influence their perception of the point at which referral is necessary. These perceptions vary markedly according to the physician's confidence and interest in managing the medical condition at hand.
- At the same time as managing the condition, GPs also manage the doctor–patient relationship. The state of this relationship also influences the point at which referral takes place and whether it is a 'referral to' or a 'referral away'.
- Demand management interventions are often devised in the quest for the optimal identification of clinical needs. Their success may, however, be over-ridden by the empathetic, enduring, negotiated and iterative nature of referral decisions.

Self-propelling diagnostic cascades

Cascade effects refer to, 'a chain of events (which) tend to proceed with increasing momentum, so that the further it progresses the more difficult it is to stop'.³⁶ They occur throughout physical and social systems and even turn up in folklore. In one of Uncle Remus's tales, Br'er Rabbit becomes the victim of a cascade in the form of the 'Tar-Baby'.³⁷ In his ceaseless campaign to trap the wily rabbit, Br'er Fox places by the roadside a dummy child constructed of tar. The loquacious rabbit tries to engage with the uncommunicative baby, ending in a loss of temper during which he throws a punch. His fist becomes stuck, prompting even more punches and kicks, which eventually entrap him in a ball of tar. We omit the happy ending to the story, noting that in modern parlance, the 'tar-baby' effect refers to sticky situations that are aggravated by additional effort.

Mold and Stein introduced the term 'cascade effect' into the health-care literature via a treacly tale about a patient admitted for routine hernia repair.³⁶ His history of mild heart disease and anxiety about his precise coronary status led the surgeon to request a preoperative consultation. This was to include an exercise tolerance test, which it transpired could not be conducted for several hours. Meanwhile, the patient, waiting outside the test room, became agitated and suffered mild chest discomfort. Because of this condition, the cardiac physicians now responsible for his care advised against the exercise test and transferred the patient to the telemetry unit. The patient, now even more anxious, received medication and underwent further tests. Although his cardiac catheterisation eventually turned up with an acceptable result, the hernia repair operation was not performed because of the full theatre schedule. The result is a demand management problem – lost operating time cannot be recovered and the diagnostic mêlée delayed the procedure by 2 weeks.

The moral of the tale is simple. In becoming more thorough, the diagnostic pathway also becomes more technical, more formulaic and more time-consuming. This self-momentum may have unintended consequences – it is difficult to stop even when counterproductive. This scenario leads us to a more recent definition of the cascade effect, suggested by Deyo in what is widely regarded as the seminal inquiry: 'With regard to medical technology, the term refers to a chain of events initiated by an unnecessary test, an unexpected result, or patient or physician anxiety, which results in avoidable adverse effects and/or morbidity'.³⁸ Note that it is not diagnostic procedures of themselves that are deemed faulty or time-wasting; the problem is that they are often prompted too easily and once prompted they become self-propelling.

Following Deyo, research on the cascade effect has parcelled into two separate bodies of inquiry: one charting the psychological, institutional and cultural triggers that may prompt the overuse of screening and testing and the other describing the multiplicative, unintended chain of events that may follow from inserting new or additional diagnostic tests into the treatment pathway. Deyo's account of the 'triggering

events' that may fire misplaced effort is summarised in *Box 6*.³⁸ It is impossible to describe all of these scenarios in detail here; most can be deciphered from their titles. 'Defensive medicine' refers to the physician's inclination to request multiple tests in order to avoid potential legal liability for a missed diagnosis or treatment opportunity. 'Low tolerance of ambiguity' by patients or doctors refers to idea that, for a widening range of conditions, modern medicine offers scientifically warranted guidance on the best course of treatment. The corresponding prerequisite is the availability of the battery of tests able to identify each condition, clearly and unequivocally.

The utility of any classification system can always be questioned, of course and we note that some of these issues such as 'patient demand', 'defensive medicine', 'low tolerance of ambiguity' and 'overestimating benefits' enter the literature as generic causes of demand and capacity problems rather than specifically as diagnostic cascade phenomena (e.g. patients may attempt to 'demand' particular clinics, clinicians and booking times – and not just tests). Other items listed in Deyo's typology also seem to describe the consequence rather than the causes of diagnostic cascades.

Deyo's second task is to demonstrate how testing becomes self-propelling and for cascades to have unintended effects. He reviews a range of common tests [scans, foetal monitoring, angiography, spinal magnetic resonance imaging (MRI), pulmonary catheters] applied across a range of conditions. For each condition he describes the initial test, its complexities, the uncertainties of diagnosis and the chain of consequences that may follow. We quote his account of just one procedure:

Electronic fetal monitoring for women in labor is another technology that may lead to cascade effects. These devices monitor fetal heart rate, and certain patterns are associated with a greater likelihood of fetal distress. However, the risks of the test were not well considered before its adoption into routine care. Use of such monitors requires the mother to be relatively inactive in bed and this may increase anxiety levels. The combination of inactivity and anxiety may slow labor and lead to interventions to speed up labor (e.g., by artificial rupture of the membranes). When labor is accelerated, the pain of contractions increases and pain medication or epidural anesthesia may be requested by the patient. The loss of amniotic fluid may lead to higher pressures inside the baby's skull, which could lead to more abnormal readings on the fetal heart rate monitor. Use of epidural anesthesia may lower maternal blood pressure, similarly leading to more abnormal readings. Perhaps as a consequence of such events, Cesarean section rates are 40% higher when electronic monitoring is used rather than simple auscultation of the fetal heart rate.

Reproduced with permission of Annual Review of Public Health, Volume 23 © by Annual Reviews, <http://www.annualreviews.org>³⁸

What of the research quality of such evidence? It does, of course, capture the inexorable logic of a cascade. We note, moreover, that in the dozen years since the publication of Deyo's paper, technical advance will have added to rather than reduced the diagnostic permutations. But we also note that his conclusions on this as well as all other cascades are unremittingly negative and quite possibly selective.

BOX 6 Common triggers for the cascade effect

1. Shotgun testing.
2. Underestimating the likelihood of false-positive results.
3. Inappropriate screening.
4. Errors in data interpretation.
5. Overestimating benefits or underestimating risks.
6. Defensive medicine.
7. Patient demand.
8. Low tolerance of ambiguity by doctor or patient.
9. Desire to legitimise compensation claims.

Given Deyo's blanket conclusions and some confusion over its causal status, should the idea of the clinical cascade be taken forward into our analysis of demand dilemmas? There is clear evidence of a steady increase in the demand for and costs of diagnostic services. NHS Scotland report that:

[b]etween '2003/04 and 2006/07 the number of patients who had CT [computerised tomography], MRI and ultrasound tests (radiology tests) increased by 38 per cent to almost 736,000. The number of endoscopy and cystoscopy procedures increased by ten per cent over the same period; almost 168,000 procedures were carried out in 2006/07 . . . there were almost 73 million clinical chemistry tests, a 50 per cent increase since 2003/04.

Review of NHS Diagnostic Services, Audit Scotland, 2008, licensed under the Open Government Licence³⁹

In 2000, the Department of Health introduced a Cancer Equipment Programmes in England, spending '£407 million on new capital equipment, resulting in greater numbers of CT, MRI and linac machines to spur implementation of increased diagnostic and treatment capability' (quote © National Audit Office 2011, *Managing High Value Capital Equipment in the NHS in England*).⁴⁰

There is also good evidence that, despite increases in activity and equipment, demand is not always met. For instance, in the midst of the expansion reported above, a 'Scottish Government Diagnostic Collaborative Programme' was introduced to service delivery:

All have worked to identify any bottlenecks in their radiology and endoscopy services and taken action to improve them – for example, reducing the administrative tasks carried out by clinical staff to free up more time for clinical work. All of the sample boards have extended their working day for radiology, endoscopy and laboratory services or increased the number of sessions by working over lunchtimes.

Review of NHS Diagnostic Services, Audit Scotland, 2008, licensed under the Open Government Licence³⁹

These examples enable a clearer insight into the causal structure of the cascade effect. Ever-increasing diagnostic systems are set up in response to external demand from patients, doctors, outside agencies and financial incentives. The diagnostic system itself may then create new imbalances of capacity and demand.

What of Deyo's thesis on the adverse effects of the diagnostic cascade? A more balanced and potentially productive analytical tool may be found in another classic paper, *Diminishing returns on the road to diagnostic certainty*.⁴¹ Borrowed from economics, the 'law of diminishing returns' states that in all productive processes, adding more of one factor of production, while holding all others constant, will at some point yield lower per-unit returns. Applied to a diagnostic cascade, this leads to the question of whether or not the additional information given by an extra test leads to a genuine increase in diagnostic certainty at a reasonable cost. It asks of any test whether it involves a positive or negative return or, quite routinely, whether or not it is marginal? Difficult as it is to differentiate these outcomes, it is precisely this calculation that is involved in an evaluation of pre-operative laboratory testing,⁴² childhood cholesterol screening,⁴³ screening for Down syndrome and hypertension,⁴⁴ mammography,⁴⁵ etc.

Box 7 provides a summary of this section.

BOX 7 Summary 4

Clinical cascades: basic issues to be retained

- Modern medicine is defined by its machinery. Burgeoning demands for health care have led to parallel demands for an expansion in diagnostic testing, and the present-day apparatus used routinely in investigating, testing and screening is immense and intricate. The demand management problem is not only a matter of the efficient utilisation of such complex services but the potential for their spontaneous and institutionalised overuse.
- This precarious balance places a burden on demand management interventions. Demand control measures in other parts of the patient pathway (e.g. changing organisational structures, introducing new triaging roles, employing guidelines, etc.) may increase the demands on diagnostic services.
- Overuse of diagnostic services to adjudicate on referral decisions may suffer from the problem of diminishing returns.

Supplier-induced demand

It is often noted that the supply and consumption of health care bears little relation to the exchange of any other product. One of the key reasons why the health-care market is considered perverse is rooted in the phenomenon of supplier-induced demand or, in North American terminology, physician-induced demand. Supplier-induced demand occurs when there is asymmetry of influence and information between suppliers and consumers, allowing the former group to encourage the latter group into seeking more services than they would under conditions of fully informed (or perfect) competition. This underlying and deeply rooted asymmetry has been captured dramatically and in lay terms by Williams in the following description of the doctor–patient encounter, which:

ostensibly works on the principle that the doctor's role is to give the patient all the information the patient needs to make a decision, and the doctor should then implement that decision once the patient has made it. I am sure that the reader would find the above statement closer to his or her experience if the postulated roles of the doctor and patient were interchanged so that the sentence would then read: 'the patient's role is to give the doctor all the information the doctor needs in order to enable the doctor to make a decision, and the patient should then implement the decision once the doctor has made it.

Reproduced from Journal of Health Economics, 7, Williams A, Priority setting in public and private health care, 173–183, © 1988, with permission from Elsevier⁴⁶

The literature is crammed with examples of market distortions that stem from this potential deformation in the dynamics of demand. It should be said that supplier-induced demand is a contested theory, with claims and counter claims about both its source (what drives it?) and extent (how deep is the problem?). Before these complications are reviewed, some well-known studies providing basic examples of its presence are noted.

The most compelling evidence stems from studies in which physicians perform the same tasks under different payment regimes. Our first example is the dramatic result of a two-stage randomised trial.⁴⁷ Doctors working in a university hospital clinic were randomised to receive income in two modes: by salary or by fee-per-service. Patients attending the clinic were also randomly assigned, in this case to the different payment regimes. The key result was that for the same population of patients the fee-for-service doctors scheduled almost 30% more return visits that did those receiving a salary. Most of the discrepancy was attributable to a 50% increase in the scheduling of 'well child' visits, which the authors describe as having 'doubtful' medical worth.

Another significant supplier-induced demand outcome can be seen in the findings of a ‘natural experiment’ comparing the treatment of heart attack emergency admissions in public and private hospitals in acute care hospitals in Victoria, Australia.⁴⁸ The incentive structures are, of course, different. In the public sector costs are met by the state through taxation, no fee is passed over and there is no incentive for the doctor to promote services. In the private sector the patient covers payment through health insurance, full fees are earned and service promotion is the norm. The research compares usage rates of the (then) most expensive and up-to-date tests and treatments for acute myocardial infarction, namely (1) angiography tests (two forms), (2) coronary artery bypass graft surgery and (3) coronary artery revascularisation procedures. The social profiles and medical histories of patient groups entering the two systems were matched statistically and the treatments selected under the different regimes are reported in *Table 1*. The patterning is self-evident. Far greater usage of cutting-edge diagnosis and treatment is made in the private sphere, with some interesting variations in the mixed-mode central cells.

The authors of both studies^{47,48} deploy several interpretative caveats, most notably that the key inference is indubitably *not* about which patients receive the best care or best continuity of care. The crucial point is to understand who is driving the health-care choices. In the Australian case study, the patients are undergoing heart attacks and being treated in emergency conditions and, thus, unlikely to be exerting strong pressure on treatment modalities. In the absence of any convincing counter-explanations a powerful inference is thus drawn: ‘These patterns must therefore be driven by physicians’.⁴⁹

Another body of evidence for supplier-induced demand, incidentally the largest and most controversial, emanates from the ‘variations’ literature – the surprising but common observation that the use of services and expenditure on services varies significantly among apparently comparable populations and localities. The phenomenon is often traced back to a pre-war study which discovered a 10-fold difference in tonsillectomy rates between different UK school districts.⁵⁰ These variations were stubbornly persistent, with the Chief Medical Officer’s 1959 Annual Report continuing to note: ‘the tonsillectomy rate ranges from 0.5% in Merthyr Tydfil to 16.3% in Chester; in those aged 15 years, it ranges from 1.3% in Swansea to 36.5% in Kingston-upon-Hull’.⁵¹ Given demographic similarities between the latter two cities (e.g. dominantly working-class populations, estuary-based industries) it is very hard to see why patient demand would drive the vast procedural difference.

Since these early sightings, the phenomenon of variations in health-care practices has been documented at local, regional, national and international levels, the best-known database being the US *Dartmouth Atlas of Health Care* (www.dartmouthatlas.org/). One notable and repeated finding is that variation is strongly associated with the level of uncertainty about clinical outcomes. This pattern is illustrated in some data (*Figure 5*) reproduced from a *British Medical Journal* (BMJ) paper by Mulley.⁵²

TABLE 1 Ratio of rates of angiography, CABG and CARP in private vs. public hospitals

Patient modality	Rate ratio			
	Angiography	Angioplasty/stent	CABG	Any CARP
Public patients in public hospitals	1.00	1.00	1.00	1.00
Private patients in public hospitals	1.43	1.09	0.90	1.00
Private hospital patients	2.17	3.05	1.95	2.87

CABG, coronary artery bypass graft surgery; CARP, coronary artery revascularisation procedure.

Rates are for all Victorian residents aged 15–85 years admitted to Victorian acute care hospitals with acute myocardial infarction, July 1995 to December 1997, adjusted for age group and sex.

Adapted from Robertson IK, Richardson JR. Coronary angiography and coronary artery revascularisation rates in public and private hospital patients after acute myocardial infarction. *Med J Aust* 2000; 173(6): 291–295. © Copyright 2000 *The Medical Journal of Australia* – adapted with permission. *The Medical Journal of Australia* does not accept responsibility for any errors in adaptation.⁴⁸

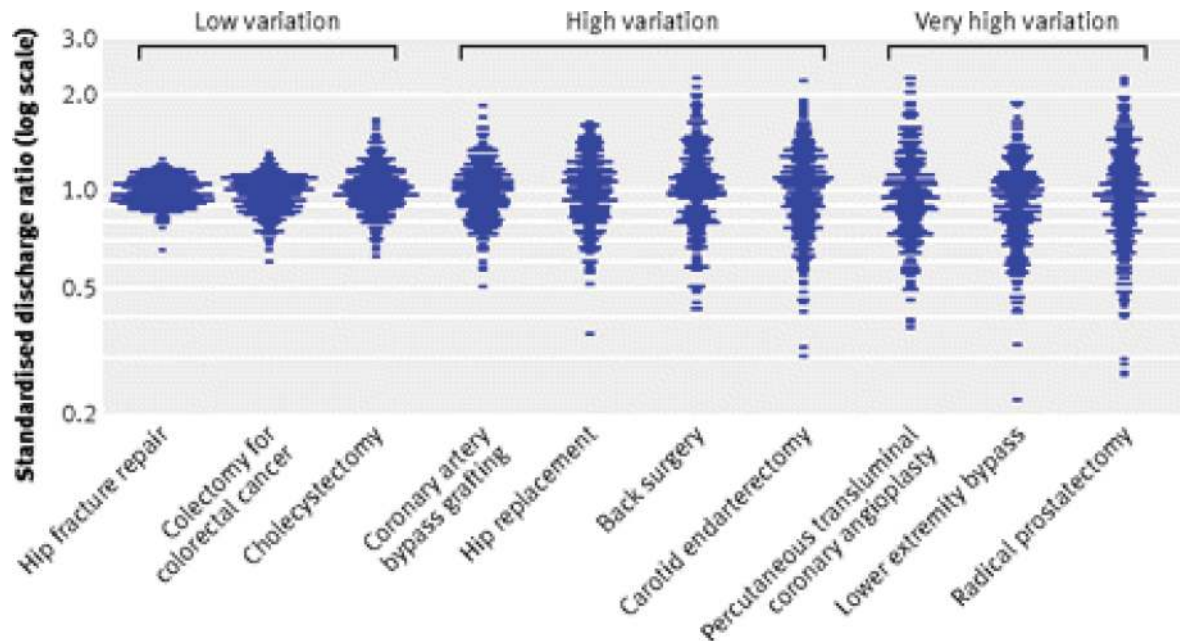


FIGURE 5 Rates of common surgical procedures among Medicare patients for 306 referral regions. Reproduced from *Inconvenient truths about supplier induced demand and unwarranted variation in medical practice*, Mulley A, *BMJ*, vol. 339, p. b4073, © 2009, with permission from BMJ Publishing Group Ltd.⁵²

Variations are plotted of the rates of common surgical procedures among Medicare patients across the 306 US referral regions. The rather tight cluster on the frequency of hip repair fracture indicates broadly similar rates for this activity across all US hospitals and regions. This contrasts with radical prostatectomy, one of the many procedures selected with 'very high variation' between surgeons, between hospitals and between districts. Again, the straightforward inference is that this pattern is unlikely to be driven by variations in patient demand, not many laypersons possessing knowledge of the vicissitudes of percutaneous transluminal coronary angioplasty. Rather, for an explanation of the disparity one needs to look across to the supply side. The working hypothesis is that the lack of an overall professional consensus about the effectiveness of particular treatments renders them open to parochial interpretation and implementation – and it is local custom, practice and vested interests that drive the high variation.

In the hope that these brief illustrations establish some of the empirical credentials for supplier-induced demand, we move to explanations for its ubiquity. There are a number of potential explanations, three of which are highlighted here: (1) technical innovation, (2) income maximisation and (3) professional uncertainty. As always in social science, these hypotheses both bestraddle each other and jockey for position. The common point of origin is that all are based on the idea of 'asymmetrical relations' between doctors and patients: the difference lies in the understanding of what it is that drives the physicians to promote supply in order to induce demand.

The first explanation perceives that demand is driven by technological innovation, to which medical care is particularly susceptible. Cain and Mittman trace an early example: 'In 1895, Roentgen discovered the X-rays during his study of light phenomena, and within six months the first diagnostic radiograph was performed' (quote ©1996–2015 California HealthCare Foundation).⁵³ Rogers has produced the best-known analysis of the diffusion of innovation and health-care organisations provide the prototypical example.⁵⁴ His first 'dynamic of diffusion' is the idea of 'relative advantage' – the more potential value or benefit of the innovation relative to current practice, the more likely it is to be adopted. Another key driver is 'observability' – the more obvious the evidence of improved experience and better outcomes the more likely is the adoption of new technology. Staying with X-rays, the relative gain is indeed palpable – being able to examine the functioning of internal organs without cutting the body. The potential outcome is that organisations failing to chase such technical gains will be left behind and this pressure perpetuates the asymmetrical understanding of treatment.

Although the quest for relative advantage is self-perpetuating this does not mean, of course, that technical capacity can be built willy-nilly. Demand leverage is not simply a matter of ‘if you build it, they will come’. Relative advantage is sustained and supported by marketing strategies: ‘build it and do everything in one’s power to make them come’. Any health-care organisation that has made substantial investment in equipment, scanning, testing and operating facilities will also pursue an engagement campaign. Two examples may be found in quite different health-care economies: the advice to new London polyclinics on how to ‘portray’ their relative advantages⁵⁵ and the ‘quick tips’ on technology diffusion prepared for the California Healthcare Foundation by the ‘Institute for the Future’.⁵³

The second explanation for supplier-induced demand rests on more brutal profit maximisation hypotheses and is often labelled the ‘target income hypothesis’: doctors are motivated to by a desired level of income and will modify diagnosis and treatment until patients respond in the right numbers at the right price. We have already provided some powerful randomised trial evidence for this at the level of the individual physician.⁴⁷ Perhaps the most notable institutional-level evidence for this theory comes in a US study of the relationship between physician financial incentives and rates of caesarean section delivery.⁵⁶ The paper exploits a historical change in the financial environment facing obstetrics and gynaecological services, namely the declining US fertility (births per 100 population), which fell by 13.5% between 1970 and 1982. They argue that such a significant reduction in demand increased income pressure and led these services to substitute normal childbirth with the more highly reimbursed alternative, namely caesarean delivery. Used in only 5.5% of births in 1970, caesarean delivery rose by over 240% over the subsequent 12 years, and became ‘the second most frequently performed major surgical procedure in the U.S., with a rate of 23.5 caesarean deliveries per 100 births’.⁵⁶

Clearly, the opportunities for the target income mechanism to come into play are limited by the reimbursement practices operating in particular health-care regimes, with fee-for-service and fee-for-target systems being the most open to excessive recommendations for expensive treatment. There are, however, apparent examples of this process in the ever-changing UK GP contract system. For instance, Hughes and Yule chart the effects of a highly specific change in the UK payment system for cervical cytology.⁵⁷ In the 1990s, per-item fees were replaced by target payments, with GPs being paid a graduated lump sum according to the proportion of cervical smears conducted among the eligible women in their practice. Their economic model suggests that the new contract accounts for ‘a dramatic rise close to 50% in smear tests performed by GPs’. The implications of such a shift remain contested. Is this a case of GPs manipulating demand by providing more information and encouragement to use the service? Or is this a case of meeting real population demand as captured in government targets? It is safe to conclude that there is wiggle room in many reimbursement regimes for physicians to recommend treatments that meet institutional targets and target incomes.

The third body of explanation for supplier-induced demand is characterised as the ‘professional uncertainty’ hypothesis. This model is driven by the data on local, regional and national variations in rates of treatment for particular conditions described earlier. Mulley provides a vivid explanation of the underlying and somewhat hidden mechanisms responsible for the disparity:

In the United States, striking differences exist among hospitals in the numbers of consultations and days in the hospital intensive care unit in the last six months that cannot be explained by differences in patients.

Mulley, 2009⁵²

and

Clinicians become accustomed to standard clinical policies that have been shaped over time by local capacity and are often surprised to learn that they are radically different from those in other regions. Though invisible, these local policies profoundly affect behaviour.

Mulley, 2009⁵²

He goes on to quote evidence from a study in which a representative sample of US clinicians, presented with standardised patient vignettes, make inconsistent treatment decisions – but ones which are entirely consistent with local service intensity.⁵⁸

This brings us to a very different explanation for the asymmetrical relations between doctors and patients, namely that it is 'Professional uncertainty rather than consensus about the scientific basis of clinical practise that is emerging as the dominating reality'.⁵⁹ As long as medical decision-making remains located in practitioner wisdom there will always be different preferences for diagnosis and treatment, and these preferences will become hardened unevenly in local institutional practices. The upshot is a different formulation of supplier-induced demand: 'With such a state of medical uncertainty it is clearly impossible for patients to evaluate what doctors cannot properly evaluate, and the abdication of consumer sovereignty represents a rational response to this'.⁴⁹

Box 8 provides a summary of this section.

Demographics and demand

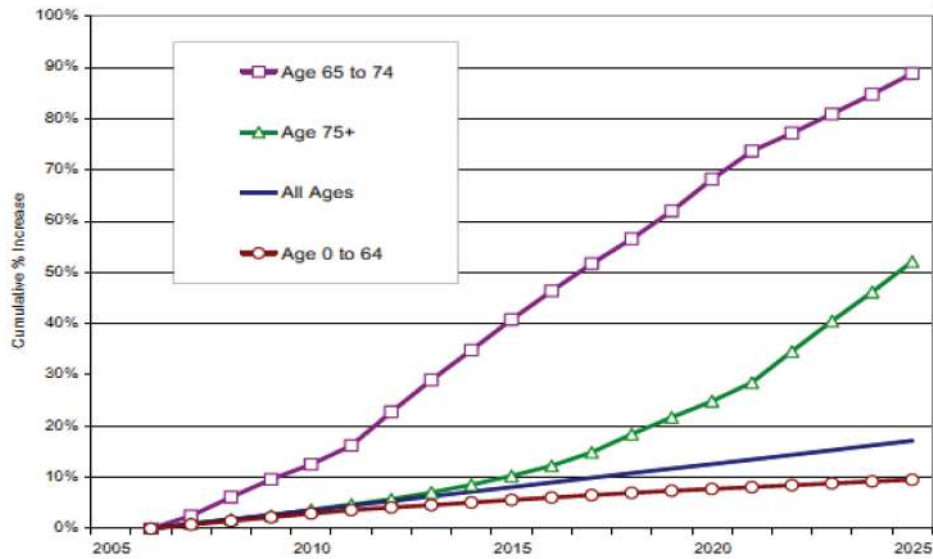
Demographic change is widely regarded as a fundamental force accelerating the demand for health care. Populations evolve and their changing composition affects overall demand as well as the need to service the particular requirements of different segments of the citizenry. Demographic change, as it impacts on health care, has been analysed by social class, race and ethnicity, locality and so on. The most profound transformation, and corresponding research effort, however, follows on the consequences of caring for an ageing population. We thus restrict this short summary to the issue of dotage and demand.

Populations are growing older, so much is certain, but people, including patients and physicians, are also busily redefining the process of ageing. Accordingly, there is a research literature on both the 'demographic time-bomb' and the 'red-herring' of ageing. Most proponents of the former, the so-called 'agequake' hypothesis, start with some basic trends and extrapolations. If one takes almost any modern society, it is clear that the baby boom has given way to grey growth. If we take the Scottish population as a random example and compare the magnitude of different age groups between 2000 to 2010, one notes a 7% reduction in the number of 0- to 15-year-olds, compared with a 13% jump in 60- to 74-year-olds and a similar spurt (14%) in the oldest segment (75 years and over). Other cohorts with longer life expectancy are building behind this advanced guard and if the trends are extrapolated it is plain to see, as in the US data in *Figure 6*, that 'seniors' are, and will continue to be, the fastest growing population sector.

BOX 8 Summary 5

Supplier-induced demand: basic issues to be retained

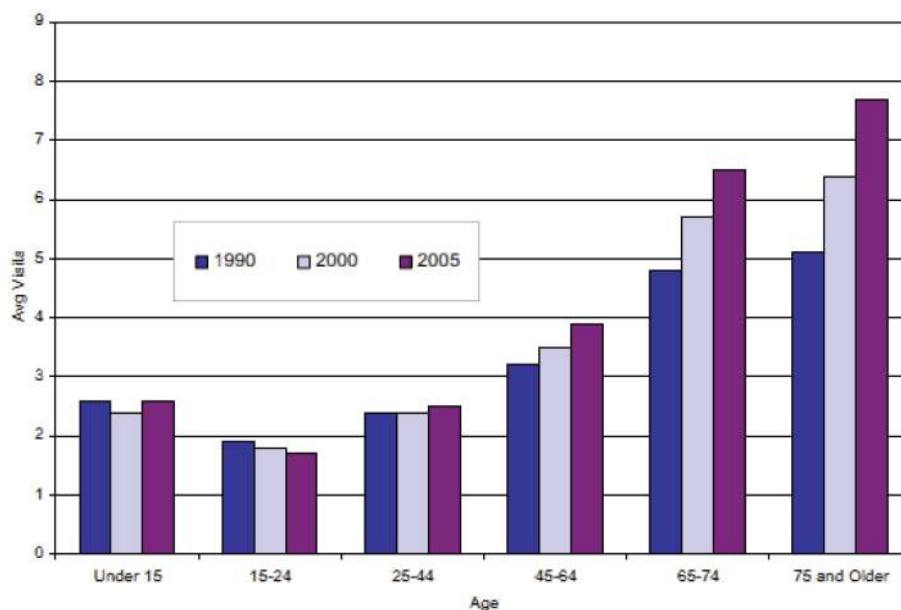
- Despite it being a contested concept, it is safe to conclude that, whether motivated by innovation, by avarice or simply by custom and practice, supplier-induced demand is a real and significant phenomenon. To be more precise, supplier-induced demand is at least three different phenomena and it should be recognised that, in an asymmetric market, 'technical innovation', 'income optimisation' and 'professional uncertainty' are all drivers of demand – but to quite different degrees across the diverse sectors of all health-care systems.
- The other pertinent feature of supplier-induced demand, in all its guises, is that these drivers of demand are perceived to reside in the health-care infrastructure. The root causes of demand imbalance lie in the very way medicine is practised and organised.
- This may prove a salutary lesson when it comes to evaluating interventions seeking to relieve demand pressure. Programmes that offer attractive new services with easy access may find that they become subject to supplier-induced demand. Programmes that offer solutions based on motivational, procedural and role change may find it difficult to act against wider and longstanding structural arrangements.



Source: Analysis of US Census Bureau Population Projections.

FIGURE 6 Population projections by age group (US). Reproduced from Dill and Salsberg⁶⁰ with permission © 2008 Association of American Medical Colleges.

The subsequent bulge in health-care demand, so the hypothesis goes, will occur because this particular population makes the greatest call on medical services. *Figure 7*, taken from the same US report, examines the number of ‘office visits’ to physicians across time for different age groups. At a glance, one sees that the oldest bums occupy considerably more seats and, significantly, that the rate of increase in consultations over the 15-year period is much more marked in these groups. By contrast, it seems that over the decades 15- to 24-year-olds in the USA have been less inclined to go to see the doctor. Treating such raw numbers as secular trends has led to economists at the Association of American Medical Colleges to conclude that, assuming that future growth follows historical trends of an increase in visits per person of 1.5–3%, the expected physician shortfall will almost double by 2025.⁶⁰



Source: Analysis of data from the National Ambulatory Medical Care Survey, National Hospital Ambulatory Medical Care Survey & the Nationwide Inpatient Sample.

FIGURE 7 Office visits to physicians per year by age group (US). Reproduced from Dill and Salsberg⁶⁰ with permission © 2008 Association of American Medical Colleges.

As well as such applications about 'manpower' requirements, this methodology is the basis of much cost estimation – discover the costs per capita of each age group and then extrapolate future costs allowing for a rising age profile and the uneven distribution of health spending per capita by age. This algorithm underpins many cost projections going back to the pioneering work in the UK by Abel-Smith and Titmuss, who predicted that NHS costs might increase, simply by dint of an ageing population, by 11.2% in the period 1951–71.⁶¹ A 1991 report on Canadian health spending, based on the same premise, follows the same pattern, forecasting that demographic shifts alone would generate a doubling of health-care spending in the following 40 years.⁶²

Since these early days, this method and its associated conclusions have been called into question. The first caveat, which required no major conceptual reorientation, acknowledged that these predicted increases should be expected to bite unevenly across different medical conditions. Put simply, demographic change should be expected to focus demand on particular services – and, therefore, a net X% increase in forward projections should not be met by X% increases in services across the board. A US study provides a comprehensive dissection of potential costs and savings by medical condition.⁶³ A handful of estimates illustrate the point: in the period 2005–15, age profiling alone should increase the call on services for respiratory disease by 10.2% and for circulatory disorders by 11.8% and, conversely, may also be expected to lower demand on pregnancy services by 2.6% and on human immunodeficiency virus (HIV) by 6.0%.

We now turn from caveat to critique. The 'demographic time-bomb' argument is regarded in some quarters as an 'illusion of necessity'.⁶⁴ If the health professions (doctors, hospitals, health services, insurance companies, drug companies, etc.) can attribute inexorably rising costs to an inevitable, uncontrollable externality such as the ageing population, then they, the establishment, are absolved of blame. More significantly, this political critique finds support in an empirical evidence base. From the mid-1980s 'researchers began to examine the possibility that a person's age might be a less reliable predictor of health care expenditure than their *proximity to death*' (emphasis added), because 'healthcare interventions are common in the last stages of life and . . . a high proportion of individuals who die are in hospital at that time'.⁶⁵

This hypothesis, that age is a demand 'red-herring', gained momentum in a study by Zweifel and colleagues.⁶⁶ This paper studied demand using longitudinal rather than cross-sectional data. We quote from the abstract:

The econometric analysis of health care expenditure (HCE) in the last eight quarters of life of individuals who died during the period 1983–1992 indicates that HCE depends on remaining lifetime but not on calendar age, at least beyond 65+. The positive relationship between age and HCE observed in cross-sectional data may be caused by the simple fact that at age 80, for example, there are many more individuals living in their last 2 years than at age 65. The limited impact of age on HCE suggests that population ageing may contribute much less to future growth of the health care sector than claimed by most observers.

Reproduced with permission from John Wiley & Sons, from Zweifel, P., Felder, S., & Meiers, M. Ageing and population health care expenditure: a red herring? Health Economics, vol. 8, pp. 485–96, 1999.⁶⁶ Copyright © John Wiley & Sons, Ltd

Since this time there have been many studies attempting to explore and further explain this pattern. Gray provides a useful overview of the longitudinal studies, from which we include two typical fragments of data (references not claimed for this paper).⁶⁵ An early study of Medicare expenditure in the USA discovered that members who died in 1984, while comprising 6% enrollees, accounted for 28% of total Medicare expenditure.⁶⁷ A Dutch study found a similar pattern, with hospital costs rising by 170% when moving from the second to the last year of life.⁶⁸ Studies of this ilk convince Gray to the extent that he declares that they 'show quite conclusively that time to death is a better predictor than age of health expenditure'.⁶⁵

Prediction is one thing; explanation is another. There is a need for precise understanding of how time-to-death shapes demand. A study by Levinsky and colleagues provides great insight by breaking down the aggregate bulge in final year of Medicare expenditure into figures on 'per head' costs.⁶⁹ Quoting Gray again:

during this last year of life health care expenditure per person actually declined with age, a pattern that was found in different geographical areas, for both sexes, for black and white beneficiaries, irrespective of degree of comorbidity, in hospices and hospitals, and regardless of the cause of death.

Gray, 2005⁶⁵

Gray explains that this was because the probability of being admitted to hospital or intensive care or of receiving interventions such as cardiac catheterisation, dialysis, ventilators or pulmonary artery monitors decreased with age.

Here is a rather subtle explanation, not always appreciated in the 'time-bomb versus red-herring' battle. The evidence reveals, once again, that proximity to death drives health-care demand but it also shows that age imprints itself on the subsequent supply of services. There is little evidence that demand at this point leads to treatment to prolong life at all costs. Other studies have found that costs peak in those aged 70–79 years but decline in the 'oldest olds'.⁷⁰

Further subtleties and conditionalities reassert themselves, moreover, if one examines studies that have pursued data that break down total health-care expenditure into different cost units for aged patients: 'ambulatory care', 'hospital care', 'drugs' and 'long-term care'. Werblow's investigation shows that the former three types of expenditure have very weak age correlations once time-to-death has been taken into account.⁷¹ Long-term care costs, for both nursing home and home care are, however, strongly related to the age of the recipient. Demand is met, it would appear, by a progressive shift from acute to long-term care.

So, where do we stand? As a matter of HSDR commissioning, we are absolved from tackling many of the subtleties of long-term and palliative care. Our brief is to study demand and capacity issues in the transfer from primary to secondary care. From this perspective the distinction between the 'age' of patients and their 'time to death' is perhaps oversubtle. There are a demand bulges associated with both, and the distinction between them is likely to be irrelevant to the physicians charged with making referrals and largely beyond the subtleties of programmes that attempt to control that demand. We conclude that age demographics are not so much a 'cause' of rising demand for health care as a 'site' or 'setting' or 'repository'. It is a 'structural condition', which acts as the carrier for more specific demand issues. And what the aged or those in their final years bring to the consulting room is, above all else, the challenge of comorbidity. It is the common factor in the studies above about what accounts for the high number of consultations as well as swelling treatment and care costs.

Typical and more precise evidence on the issue can be located in a study by Caughey and colleagues.⁷² It was based on the Australian Longitudinal Study of Ageing, which followed 2087 randomly selected men and women aged 65 years, correlating mortality with their clinical characteristics. Participants with '3–4' and '5-and-over' diseases had, respectively, a 25% and an 80% increased risk of mortality in comparison with those with no chronic disease, after adjusting for age, sex and residential status. Comorbidity is what killed them and what carried them into the health-care system.

Box 9 provides a summary of this section.

BOX 9 Summary 6

Demographics and demand: basic issues to be retained

- Much of the evidence utilised to referee demographic debates emanates from actuarial data and relates to population aggregates.⁷³ The demand and referral management interventions to be studied in the main body of our research are based, by contrast, on devising more immediate fixes to localised blockages.
- In this respect, the most interesting findings from this body of inquiry relate to the decisions that follow the population bulge, regardless of whether it is equated to age-of-bones or proximity-to-death. These show that population push is not and cannot be met mechanically by capacity increase. The vital responses are about which treatments are selected and, indeed, which are withheld.
- Demographic change is, thus, less of a driver of demand than a pressure point, one that accentuates the difficulties of making appropriate treatment and referral decisions. Comorbidities gather here, as do decisions about degrees of impairment and quality of life. The crucial issue for our review of demand management interventions concerns how well they are able to cope with demand from patients who enter the system frequently, presenting with multiple problems.

The internet, the informed patient and demand inflation

It is increasingly the case that patients have access to the internet and, at the same time, the amount of health information held online is increasing exponentially. There are bespoke, national sites specifically constructed for purveying advice, such as NHS Choices (www.nhs.uk) and Patient.co.uk (www.patient.co.uk). Evidence compilers such as NICE produce versions of all of their guidance for patients, carers and members of the public. Health-care charities such as Macmillan Cancer Support provide robust and timely cancer information for cancer patients, their families and carers. Once inside the system, the patient will be directed to condition-specific information as obtained through the websites of the Royal Colleges and other professional associations. There is also private provision of information should one choose to visit the like of the PruHealth official website (www.pruhealth.co.uk). Local, informal support groups crop up for most conditions, which can be accessed from bodies such as the Hysterectomy Association. Even the most routine activity of 'Googling' information will produce detailed results for almost any health problem.

As a result, patients' use of the internet as a source of information is also growing. One study indicates that in 1997, 41% of regular internet users in the USA reported having searched the internet for health information; this had risen to 50% by mid-2000 and then to 60% by the end of 2002.⁷⁴ Rates for the 'general population' are lower; a national survey of US citizens indicated that 31% had reported seeking health information online.⁷⁵ Younger women, people on high incomes and those who are better educated are more likely to search for health information on the internet.⁷⁴⁻⁷⁷ As we have already seen, the demographic footprint is never straightforward.

How might all this activity impact on demand for health services? At least four theories can be distinguished: that (1) health information acts as a substitute for visits to health-care professionals and reduces demand; (2) health information acts as a complement to information from doctors and means that time with the doctor is used more efficiently and, thus, reduces demand; (3) health information acts as a driver of demand, the newly informed patient being stimulated to seek health-care services; or (4) internet-informed patients 'time hog' medical services, leaving less time for other patients, and, thus, increase demand.

The first hypothesis, substitution theory, argues that people access health information as a surrogate for physical access to health care. Patients use health information on the internet instead of accessing doctors, or use it to decide whether or not they do really need to see a doctor. Both impulses, the argument goes, reduce demand for health care. A typical study is an evaluation of the NHS Choices website.⁷⁸ The authors

conducted two surveys: one of users of the NHS Choices website and another of people waiting to see their GP. The survey of 'Choices' users found that 59% reported using the website to inform their GP consultations. Of these, 33% reported that the website reduced their frequency of GP visits, while 3% reported that it increased their visits. Those who reported that the website reduced their GP visits tended to be young, female and infrequent users of primary care in good health. In other words, substitution hypothesis seem to apply more closely to people who do not make large demands on health services in the first place.

Furthermore, the survey of patients waiting to see their GP found that although 71% of patients used the internet to search for health information, only 8% reported having used the NHS Choices website. Of these, 18% reported that the online service reduced primary care consultations. Similarly, the majority fulfilling the substitution hypotheses were young, white women in good health who were infrequent users of primary care. The authors go on to make extravagant claims about potential financial savings to the NHS, using a dubious method of directly extrapolating from the survey sample into general population costs. More attention should have first been given to the response bias in 'pop-up questionnaires' and 'snapshot surveys'. We take a different key message from this survey: that health information may reduce demand for health care *among certain groups for certain conditions*. Accordingly, the net impact on the internet's capacity to reduce demand is conditioned by the relative contribution of these groups to the overall demand for health care. A subsidiary lesson is also drawn from the study: not all websites are the same in terms of traffic and style of usage, with the NHS Choices website being less visited than others.⁷⁸

Linked to substitution theories is the idea that health information acts as a complement to information from physicians and enables patients to utilise their time with the doctor more efficiently without having to be in permanent contact with the doctor. One study, through focus groups with patients with diabetes, heart disease and hepatitis C, found some evidence that patients used health information in this way: 'The Internet was perceived to be particularly useful for confirming and expanding on information received without "bothering" the doctor'.⁷⁹ Again, we note the methodological limitations of the study. The data are limited to people who were suffering from a chronic condition, all of whom required regular contact with the medical profession. The evidence is also 'prompted' rather than 'naturalistic'. Respondents were each allocated a computer with three internet interventions appropriate to their condition bookmarked, before they moved onto the focus group.

We move briefly to the rival set of theories, which suggests that the internet might act as a driver for demand for health care. There are a number of mechanisms through which this might occur. One is the 'hypochondriac' or 'cyberchondriac' effect, whereby otherwise healthy patients prone to hypochondria read about something on the internet, think that their minor symptoms represent a terrible disease and consult their doctor unnecessarily. A brief and distinctly anecdotal piece in *The Wall Street Journal* sets the scene.⁸⁰ It recounts the experience of Melissa Woycechowsky, who searched the internet to understand the cause of the numbness and tingling in her legs and spent many hours each day in a multiple sclerosis chatroom, becoming convinced that she had the condition. Several visits to the doctor and many hundreds of dollars eventually served to confirm that Woycechowsky suffered from hypochondria rather than multiple sclerosis.

A related hypothesis is that internet-informed patients may exacerbate or stimulate 'demand-induced supply' – as, for example, when patients with existing conditions read about a new drug on the internet and visit their clinician to ask for it, even though it may not be clinically appropriate to them. One study used surveys and econometric modelling to explore the effect of more informed patients on demand-induced supply.⁸¹ Do physicians accede to such requests? The survey data used in this study were taken from the public-use version of the Community Tracking Study Physician Survey – though the specific questions put to the medics are not reported. Multiple regression analysis is taken to show that a limited group of physicians are so influenced – according to their 'years of experience, race, international medical Graduate (IMG) status, Primary Care Physician (PCP) status, average weekly hours of work, ownership

status, practice type, sense of control over clinical decision making, levels of influence'.⁸¹ Again, methodological limitations are noted: general-purpose survey, self-report, absent questionnaire items and correlational evidence.

A yet further influence that might drive demand is a 'clarity over confusion' mechanism, whereby patients read the burgeoning materials on the internet but find the information confusing or conflicting or cannot see how the information relates to their specific woes and so consult their doctor for clarification. Two qualitative studies on small groups of self-selected physicians and patients found evidence that internet-related dialogue did feature regularly in consultations.^{79,82} It was, however, the medium of conversation rather than its inspiration. Patients found some information on the internet confusing or did not trust it and visited their doctors for clarification. These studies are primarily interested in the quality and consequences of these exchanges and provide little additional evidence on prevalence.

So, is there any numerical evidence that internet use does increase health-care utilisation? One study analysed data from the US Health Information National Trends Survey between 2003 and 2007 and found that, holding age, race, sex and education constant, 'e-health information seekers' had 1.2 more health professional visits per year than 'non health seekers'.⁷⁷ Further, frequency of usage also marks a difference in demand. Those who searched for health information once per week increased an individual's visits by 1.8 visits, compared with only 0.5 visits for infrequent e-health seekers. This study also examined the hypothesis that internet health information may complement rather than substitute visits to the doctor. The author tested a rather complex scenario, as follows. At onset of disease patients are less knowledgeable about their condition and thus would be more likely to visit their doctor following encounters with internet health information. As they become more familiar with their condition they may substitute the internet for health visits. However, if their condition deteriorates, they may look for more information over time and this may drive further visits. Suziedelyte tested this hypothetical sequence, restricting her analysis to those diagnosed with cancer and comparing the increase in the number of visits between those diagnosed 'less than 2', '3–5', '6–8' or 'more than 9 years' ago. Her results indicated a small variation in visits that did correspond to the U-shaped prediction.⁷⁷ There are signs of a pattern here but the progress of cancer is undoubtedly too complex to be tested using this data set. Some cancer 'sufferers' may be entirely cancer free 3 or so years after diagnosis while others may have experienced stasis, slow deterioration, recurrence or death. Disease trajectories do not march to a uniform tempo.

Fang and colleagues hypothesised that internet-informed patients may act as 'time hogs' and place additional demands on doctor's time, either through wishing to display their own knowledge to the doctor or by trying to secure more information from the doctor, which in turn leaves less time for other patients.⁸³ They tested this theory by applying an econometric model to the 2000–1 Community Tracking Study data.⁸⁴ This is a survey of US physicians' perceptions of their services, including questions on the percentage of patients they regarded as consumerist, on the quality of care provided and on time constraints. The authors found that physicians reporting more consumerist patients were substantially less likely to believe that they could provide high-quality care in the appropriate time periods. Again, although there is some support for the presence of internet-inspired time-hog, the study suffers clear methodological drawbacks (opaque questions, self-report by one party on another, correlational evidence).

In summary, there is little doubt that the internet-informed patient is real and little doubt that, collectively, these patients' new knowledge is affecting medical communication and, therefore, demand. However, the research paints a profoundly mixed picture of the impact of this transformation. Whether such information acts as a driver of demand or a substitute for services depends on (1) the quality of the information – poor-quality information is likely to increase demand while high-quality information may reduce it; (2) the characteristics of those absorbing the information and their existing rate of demand for health care; and (3) the point during the patient's journey in which the search for information takes place. The 'informed patient', it seems, faces both ways, sometimes elevating demand and sometimes stunting it.

These counterbalancing forces should not, however, be understood as a neutral causal affect. At quite another level, the 'informed patient' has become an institutional reality. Since 2006, patients in England requiring a referral to a specialist have been entitled to a choice of four or five providers. Since 2007, an attempt has been made to construct an 'extended choice network', enabling patients to choose treatment from any accredited NHS hospital. These reforms have been underpinned with the roll-out of 'Choose and Book', an e-booking software application enabling patients who need an outpatient appointment to choose the hospital, the date and the time of their appointment. Similar patient-centred reforms are in train in many other advanced health-care systems.⁸⁵

How might this change in institutional apparatus impact on demand management interventions? First of all we should make it clear that our remit does not include the evaluation of the extended choice agenda, although this has been included in previous studies.^{86,87} We do, however, discover patients and their volitions in many of the demand management subsystems already reviewed.

In terms of those causes located within the internal dynamics of the consultation, we do need to recognise the potential role of the patient within the growing culture of 'shared decision-making'. It is well known, of course, that many patients still prefer to delegate treatment and referral decisions to their patients, with one study demonstrating that the desire for involvement in decision-making varies by type, severity and longevity of a patient's condition.⁸⁸ Elwyn and colleagues, moreover, paint a picture of the complexities involved in trying to accommodate informed choice within treatment decisions in a model which we have summarised in *Box 10*.⁸⁹ The predicament for our review, therefore, is to recognise that some parts of the referral system are working to install these changes, while other elements are attempting to streamline consultations in the name of demand management.

Patients are also present at quite a different level in the service delivery apparatus of modern medicine. They are represented at committee level in all of the key decision-making bodies. Since January 2003, all NHS bodies have had a legal duty to involve and consult the public about the running of local health services. One example is the system of 74 Clinical Reference Groups (CRGs). CRGs are a primary source of clinical advice on the specifications, quality standards and commissioning policies for each specialist service, with each CRG appointing up to four patient and carer members.⁹⁰ In terms of our inquiry, we can also locate the patients' voice in the construction of several of the interventions set up to manage demand such as referral guidelines and RMCs. Accordingly, one of the many tasks for our review is to assess the significance of these contributions.

BOX 10 A sequential model of an 'informed choice' consultation

1. Opening up the possibility of, and opportunity for, shared decisions.
2. Exploring the patient's ideas and fears on their problem and possible treatments.
3. Portraying clinical equipoise, rather than indecision, during initial explanation of treatment options.
4. Identifying the preferred format for the patient to understand information on the various treatments.
5. Pausing the processes to check understanding and reactions to the data.
6. Rechecking the patient's acceptance of their role as joint decision-maker as choices harden.
7. Offering a cooling-off period and follow-up before arriving at the final decision.

Conclusion

These conclusions look forward to the main work of the project, which is to review the evidence on the effectiveness of the many schemes and strategies which have been devised to manage demand and make referral more effective. Our basic logic in this chapter is that knowledge of the symptoms comes before assessing the cure: we need to understand the nature of demand dilemmas before assessing any reforms. We should also remind readers that the above effort should be considered a rapid review, designed to carve out the main contours of demand and capacity imbalance rather than to inspect every local nuance. So, what have we learned?

The first lesson is that the roots of the problem are stunningly diverse. They range from the micro to the macro, from local to population levels, and from the immediate to the historic. Work routines and treatment pathways can fall in and out of rhythm. Power imbalances may sway services levels and slant rates of provision. Personal and emotional factors can shape the decision to refer. The technical apparatus of medical investigations can be self-propelling. Modern medicine and mercenary medics may induce demand. Demographic change can shape the profiling of demand. Populations grow more health literate and this can have a numerical effect on demand.

Box 11 provides a summary of this section.

The second conclusion is that these roots are bound to overlap and entangle in any particular corner of the health service. It is easy enough to imagine a scenario where all or most of these problems collide. An elderly patient under treatment may find himself in a queue-inducing population bulge, his comorbidities may prompt uncertainty and thus cascade testing, treatment decisions may be coloured by emotion and familiarity, his family may pester on the basis of cyber-health wisdom, his GPs may be confronting or indeed managing quotas and targets, his clinicians may have professional interests to defend and procedures to control, he may be from Hull (or Hell or Halifax) where a procedure is under- or overused or, indeed, he may become so agitated in working and waiting through this minefield that he develops some new disorder. Meeting this, or the endless equivalent configurations of pressure points, demonstrates the colossal difficulty of the task of demand management and affirms our initial contention that it may be considered one of the 'wicked problems' of contemporary public policy.⁷

BOX 11 Summary 7

(Internet)-informed patients and demand: key issues to be retained

- Thanks to the new mass and social media, masses of medical information are now readily accessible by patients. There is little doubt that the internet-informed patient is real and little doubt that, collectively, these patients' new knowledge is affecting medical communication and, therefore, demand. However, it may be said that the jury is still out on whether or not there is a direct causal link to the demand on services. There is empirical evidence to suggest that, in different circumstances, the informed patient may drive or may circumvent demand. Our review of the various interventions designed to contain demand cannot, therefore, anticipate a uniform reaction on the part of those on the receiving end.
- There is much clearer evidence on the penetration of patents into the institutional structure of medicine and of the growth of a culture of patient-centred care in all aspects of provision.
- Rather than simply taking on the task reviewing whether or not demand management programmes are able to steer well-informed and potentially stubborn patients, we should also anticipate the need for evidence on how and to what extent newly empowered patients steer the construction and implementation of demand management interventions.

Our third and final conclusion is to cast some critical self-reflection on our sevenfold typology of ‘causes of the rise in demand for health care’ in *Box 1*. We have just noted that these features may overlap and intersect, and this is the configuration we wish to stress when it comes to studying how we might review their potential resolutions. So, for emphasis, it should be made clear that *Box 1* is not simply a listing of ‘causes’ understood as direct pushes or elevators of demand. Nor should our brief evidence reviews be considered a matter of ascertaining that each push was of sufficient magnitude to consider it a worthy opponent in need of management. Instead, we have tried to spell out their precise aetiologies structures and discovered these are not the same across the seven features.

For instance, we have argued that demographic age is not in itself a generator of demand. ‘Proximity to death’ is regarded by some as a better indicator of increased call on services – but this gauge might too be erroneous as when services to the very old are withheld if considered fruitless and expensive. Thus, instead of ‘age’ being a direct cause, we have interpreted it as a ‘gathering point’ for comorbidity and consider this conjunction of clinical conditions to be a more direct press on demand for services. In the preceding section we argued that increasingly well-informed patients are not in themselves a direct cause of elevated demand. Health literacy may drive demand but may also apply a break. More fruitfully, we might consider that it is reforms in the overall management apparatus that increasingly recruit patients and that changing norms about their inclusion in treatment decisions are the crucial developments. Can these new privileges be accommodated in restrictive demand management regimes?

Another indication of causal complexity of the rising demand for health care is the recurrence of common issues across the typology. One feature occurs and recurs through these accounts: the issue of ‘clinical uncertainty’. It might be considered a key and eighth demand dilemma in *Box 1* but its capacity as a common contribution is what marks it out. The issue raised its head most clearly in item 5 within our discussion of supplier-induced demand. Recall Wennberg and colleagues’ study, which showed that procedures with the highest variation are those in which there is more uncertainty in the evidence base about the medical benefit of the procedure.⁵⁹ But uncertainty, it transpires, is everywhere. We noted in the section on the micro-dynamics in the decision to refer (item 3) that social and emotional factors often squirrel their way into the decision to refer. Partly this is because there is inherent uncertainty in the clinical consultation.⁹¹ Uncertainty is certainly a driver in the clinical cascade: the quest to bring finality to demanding diagnostic decisions, despite the law of diminishing returns. Uncertainty certainly underpins the turf wars, with generalists inflating demand by using referrals in seeking to reduce uncertainty and specialists seeking to control their services by reducing what they see as speculative referrals. Uncertainty certainly increases with comorbidity, in which treatment decisions are chosen on a selective rather than all-embracing basis. Uncertainty is certainly what drives patients to the internet, now established as a universal source of second opinion.

We conclude that our typology, imperfect and uncertain as it is, provides a useful entrée to systematic review. We hope to have captured some, if not all, of the obstacles curbing the demand for planned services, and to have understood something, if not everything, about how the drivers of demand intersect and combine together. Sun Tzu, the father of military strategists, once famously declaimed that ‘if you know your enemies and know yourself, you will not be imperiled in a hundred battles’.¹⁵ A careful appraisal of some the causes of health care demand may allow us to recognise the perils and the potential of some of the solutions.

Chapter 3 The policy response: charting the family of purported solutions

Text in this chapter has been reproduced from Pawson and colleagues under the terms of the Creative Commons Licence.⁹²

Introduction: eliciting and selecting the underlying programme theories

With this chapter we enter the first main phase of our review. All policy-making begins with theory: accounts of the causes of the problem under scrutiny (diagnosis) and conjectures on what changes must be made to the system in order to alleviate that problem (remedy). In realist terminology, these foundational ideas are known as 'programme theories'. The process of locating and articulating the proposed diagnoses and planned remedies that have gone into the making of an intervention is known as 'theory elicitation'.

Demand management is a perennial problem and so its principal ideas have been contemplated time and again. And, as we shall see, the result is a large family of interventions underpinned by a welter of interrelated ideas. The purpose of this chapter is to provide a thorough review and a detailed description of these intrinsic programme ideas. Note that the chapter travels no further than an exposition of these key concepts and conjectures. This is an ideas exercise; indeed, given the formidable ambitions for demand management, it is a 'big ideas' chapter. The assembled conjectures will go on to provide the foundation stone for further analytic work: they act as hypotheses to be tested. The basic objective of the upcoming phase of realist synthesis and subsequent chapters is to trawl the empirical evidence to gauge how, where, why and to what extent each of these theories has proved fruitful in practice. But that is for later.

Here, the key task is to hunt down the demand management programme theories. This was accomplished in the following stages, the chapter's sections mirroring this running order. Our preliminary task was to take stock of the field – to draw up an initial list of the interventions that have been applied in the name of demand and referral management. The purpose here is simply to scope and survey the field; we identified and classified the basic array of programmes without yet penetrating to their underlying ideas. Our attempt to map the coverage of the demand management interventions ends in a simple typology breaking down the programme theories according to what they perceived to be the main axis of change. Demand management schemes have been designed to operate at all levels from the macro to the micro. As already noted, these range from attempts to control demand by relocating the entire apparatus in RMCs to interventions seeking to change individual decision-making through financial incentives. To capture this range we made an initial sift of programme theories into five types: those seeking change at the (1) strategic, (2) administrative, (3) role, (4) procedural and (5) motivational levels.

For each class of programmes we then searched for sources that discussed and debated the underlying theories. As our focus is on underlying policy thinking, we directed our attention, self-evidently, to the 'ideas literature'. In other words we searched for sources in the so-called grey literature (planning documents, guidance materials, discussion documents, proposals, rationales, policy expositions, professional journals and critical debate). By and large our subject matter here is not located in the academic journals and we focused our search on documents and expositions produced by the health-care bodies, professional associations, institutes, agencies, services, QuANGOs, etc.

The next task was to extract, capture and compile the constituent programme theories. In essence we are collecting explanatory propositions. In a planning document an author might opine that intervention A works by doing B. In a further policy appraisal we might note claims that A will not work because of

stultifying force of C. Debate is often enjoined in further critical opinion as when it is hypothesised that A will work in respect D but not if the plan is to achieve E. There is no shortage of such material. Demand management has a considerable pedigree, and we expected and found and were able to catalogue a constant refrain of proposal, critique and counter-critique.

The next stage is known as theory selection. Thus far we had charted five broad theories about where change might manifest itself and scores and scores of subtheories about how these improvements might be achieved. It is impossible to review every idea that has ever dawned about an intervention: it is necessary to prioritise and seek out patterns within programme theories. Because we were investigating interventions with a considerable pedigree, one pattern became abundantly clear, namely the constant adaptation and revision of the original theories. We have thus presented many programme theories to capture this evolution, beginning by unearthing the founding ideas, and then charting some of the emerging critical material, and then moving to some of the prosed remedial improvements.

This idea of programme drift or implementation shift is routine in the policy literature. What we had not anticipated was that there might be an overall pattern to the transformation of ideas about demand management. Our inspection of the underlying programme theories reveals a trend for the incorporation of 'whole-system' thinking into demand management thinking. By this, we perceive a move away from the 'single lever' models which situate change at the strategic, administrative, role, procedural and motivational levels. This understanding mirrors much contemporary of thinking about change and innovation in organisations, which also lays stress on the idea that organisations are layered and that change is mediated by intercommunication between these layers. This idea, which we refer to as the 'depth ontology' of organisational change, provides us with a middle-range theory that finds its way into all subsequent stages of the review.

This brings us to the final section of the chapter. Remember that all of the above analysis lies in the domain of programme theory. To that extent it should be viewed as speculation and conjecture. We do not know as yet whether or not these policy revisions lead to improvement in demand outcomes. What can be said, however, is that that reformulation of the problem – namely, 'have the adaptations worked?' – captures the optimal strategy for reviewing this particular class of interventions. The concluding section considers how and how not to mount a review of a moving target. It outlines the strategy we use in the next phase of the review: what evidence will we need to test these adapting programme theories, where can it be located and what will be the shape of our emerging findings?

Taking stock: mapping the field of demand management interventions

Just as the causes of demand inflation are many, the efforts to stem the escalation are numerous. The purpose of this section is to gain an immediate measure of our review domain and to map the main types of interventions that have been implemented in the name of demand management. We are aided in this task by the production of several existing typologies and we commence by reproducing three of them as *Figure 8* and *Boxes 12* and *13*.⁹³⁻⁹⁵

These three typologies provide a blunt introduction to the complexity of the field. Favourably, we note many elements in common, the same interventions cropping up across the three classifications (guidelines, incentives, feedback, service regrouping, etc.). Contrarily, when it comes to the placement of the intervention types and their subspecies, the three figures diverge considerably. For instance, Coulter⁹⁴ regards 'financial incentives' as the umbrella term for different funding interventions. Imison and Naylor⁹³ perceive 'financial incentives' as one type of peer review and audit. In Roland and colleagues'⁹⁵ model 'financial incentives' become a form of professional behaviour change. The most telling pointer to the complexity of the field, however, is the upwards-pointing arrow in *Figure 8*, which is an indication of the 'different degree of intervention'. By this, the authors refer not just to the amount of effort and

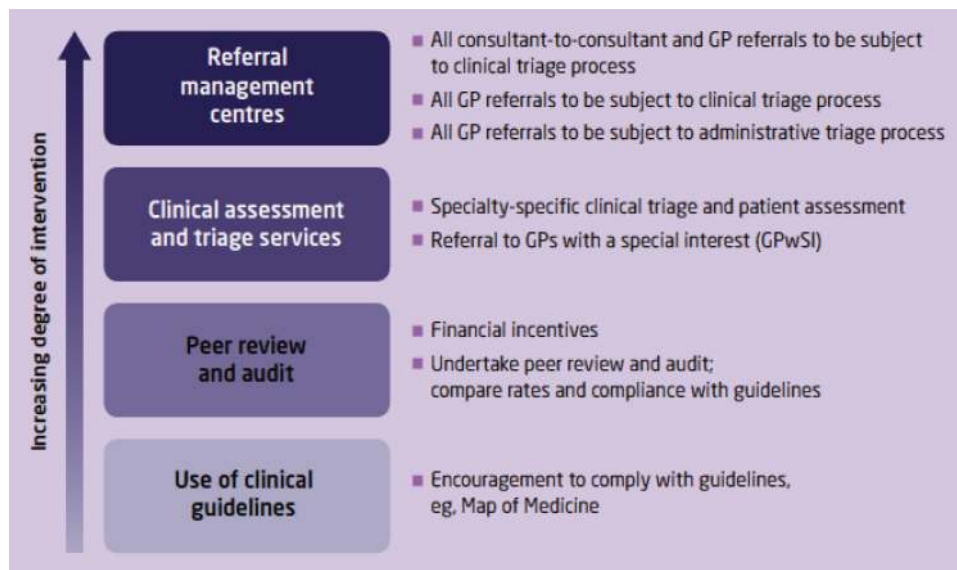


FIGURE 8 Distribution of different approaches to referral management. Reproduced with permission from The King's Fund from Imison C, Naylor C. Referral management: lessons for success. London: The King's Fund; 2010.⁹³ Hospital Episode Statistics data copyright © 2015. Reused with the permission of the Health and Social Care Information Centre. All rights reserved.

BOX 12 Strategies for managing demand at the interface between primary and secondary care

Information and audit: feedback of referral rates, measuring outcomes.

Decision support: guidelines, information for patients.

Financial incentives: practice based services, fundholding, primary care groups.

Reproduced from Managing demand at the interface between primary and secondary care, *BMJ*, Coulter A, vol. 316, p. 1974, 1998,⁹⁴ with permission from BMJ Publishing Group Ltd. Permission conveyed through Copyright Clearance Centre, Inc.

BOX 13 Four models of demand management for access to specialist care

Transfer: the substitution of services delivered by hospital clinicians for services delivered by primary care clinicians. This included: minor surgery, diabetes care, GPwSIs, discharge from outpatient follow-up, and direct access for GPs to hospital tests and services.

Relocation: shifting the venue of specialist care from outpatient clinics to primary care without changing the people who deliver the service. This included: shifted outpatient clinics, telemedicine (as a 'virtual' form of relocation), and attachment of specialists to primary care teams.

Liaison: joint working between specialists and primary care practitioners to provide care to individual patients. This included shared care and consultation liaison.

Professional behaviour change: interventions intended to change the referral behaviour of primary care practitioners, including referral guidelines, audit and feedback, and education and financial incentives.

From Roland *et al.* Reproduced with permission.⁹⁵

expenditure but to the fact that a composite programme, such as the introduction of a RMC, is likely to subsume many of the more targeted interventions. A RMC may well utilise guidelines, incentives, joint working, etc.

Without exploring other versions of these classifications (e.g. Pencheon⁹⁶) it can be seen that there are many equally logical ways of cutting the demand management intervention cake and that they are unlikely to match the complexity and variability of programmes 'on the ground'. Armed with this insight we sought to commence our exploration of programme theories with a relatively simple typology that distinguished the main 'axis of change' targeted by the programme: what is it that is supposed to change as a result of the intervention? Five different objectives are identified in *Figure 9*.

Our model responds to the idea, self-evident in the previous typologies, that change is envisaged at quite different levels from the macro to the micro. We make no claims for the originality of this idea, of course, the notion of 'vertical stratification' or 'depth ontology' being much utilised in the literature on organisational change.^{97,98} These terms refer to the idea that organisations are layered and that change is mediated by intercommunication between these layers. By 'strategic change' we refer to reorganisation of significant parts of the entire system; 'administrative change' anticipates the introduction of new teams or units to carry out specific tasks; 'role change' envisages a change in the division of labour between individuals or between teams; 'procedural change' focuses on changing the way tasks are performed; and 'motivational change' seeks to stimulate increased enthusiasm for performing a particular task in a particular way. As we are about to see, all of these objectives are envisaged in the entities we call demand management programmes and it is a simple enough task to fit the interventions compiled in *Figures 8–10* into the model in *Figure 9* (and vice versa). Our interest in *Figure 9* is to discover how the respective changes are understood and to trace the evolution of each programme theory.

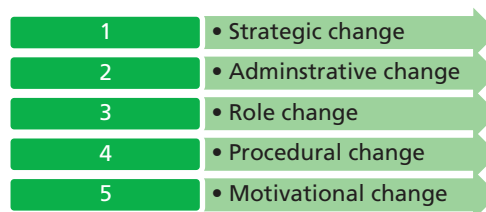


FIGURE 9 Change objectives in demand management interventions.

Demand management interventions

As noted, demand inflation and referral improvidence are perennial health service problems, which have gone on to inspire a medley of putative solutions. In this section we describe a spanning set of the major reforms. For each, we highlight typical examples of major UK NHS interventions, extracting brief illustrations of the core programme theory. Our immediate task is to show that the initial modes of change envisaged sit squarely within the stratified model provided in *Figure 9*. When planning change, programme architects grasp immediately for reforms of strategy or administration or roles or tasks or motivations.

However, we also take advantage of the longevity of demand management/referral management campaigns. A notable feature of our scrutiny of programme theory is that we came repeatedly on 'second thoughts'. Realist synthesis is a theory-driven approach and so engages in a systematic search to track down the assumptions, working ideas and plans that contribute to the building of any intervention. A typical feature of such programme architecture is that there is never just the single blueprint. Plans change. Programme designers adapt their ideas in the face of criticism and by dint of their own and others' experiences. Accordingly, the search for a programme theory rarely supplies an agreed model. More often it reveals a history of planning. In terms of long-running interventions, such as demand management/referral management, that history tends to build a picture of the adaptations that need to be made to foster successful change. Each of the following sections begins by unearthing the founding ideas, charting some of the emerging critical material, and then moving to some of the proposed remedial improvements.

Strategic change

The fundamental change envisaged in our first demand management intervention involves rethinking the full length of the referral chain. The big idea, clearly and unequivocally, is to improve the 'logistics' of the whole care pathway. This broad level of ambition fits with some of the grander, industrial-scale origins of the idea: 'lean thinking', 'queuing theory', 'continuous improvement', 'business process reengineering' and so on.^{19,99} In programme theory terms, the key change mechanism lies in the better co-ordination of practices and processes across complex organisations. In management parlance it about is getting the right provision in the right place, at the right time, in the right quantities. A useful pocket illustration of the core idea is located in the following passage:

... looking at the patient process and attempting to simplify and shorten it, identifying bottlenecks and pinch points for the individual process and then using the whole – hospital system perspective – to work out the best way of handing the flows into and through the process.

Reproduced under the terms of the Open Government Licence for public sector information v3.0¹⁰⁰

This strategy has a considerable history in the NHS management and improvement agencies and indeed in health-care systems worldwide. It is espoused in several major reports. Examples of typical documentation are provided in *Box 14*.

BOX 14 Typical sources on strategic change

NHS Modernisation Agency

Big Wizard (2002).¹⁰¹

Little Wizard (2002).¹⁰²

NHS Institute for Innovation and Improvement

*Improving Patient Flows in the NHS: Case Studies in Reducing Delay.*¹⁰³

*Reducing Delays in Patient Care – Seven Ways to No Delays.*¹⁰⁴

*Productive General Practice: Releasing Time.*¹⁰⁵

NHS Scotland: Scottish Executive Health Department

*Patient Pathway Management: Referral Facilitation.*¹⁰⁶

Health Service Executive Ireland

*Report and Recommendations on Patient Referrals from General Practice to Outpatient and Radiology Services.*¹⁰⁷

This strategic advice thus arrives with a certain professional authority, deriving from what Hage and Aiken describe as ‘centralization’.¹⁰⁸ Although cast as advice on what to do in the hurly-burly of demand decisions, the advice is pitched at a level of abstraction and within presentational formats associated with ‘co-operate strategy’. Two examples from *Little Wizard* give a good indication of the breadth of the programme theory:

Reduce hand-offs: most delay, error and duplication is at the point where responsibility for the patient is handed from one person, department or agency to another. Reducing the number of hand-offs can significantly reduce waits and delays.

Do as much as possible at the same time: many healthcare processes take a step-by step approach to performing tasks, especially where a number of specialties, departments or agencies are involved. If possible redesign the process to do all or some of the tasks simultaneously.

*Reproduced under the terms of the Open Government Licence for public sector information v1.0*¹⁰¹

In summary, the demand and referral challenge lies in the entirety of the patient processing system. Less piecemeal and more strategic thinking is needed. What has history made of these ideas? As recommendations come externally and from ‘on high’ from bodies such as the Modernisation Agency, one obvious concern is about whether or not such recommendations might be ignored in the ward and the waiting room. As is put in another of the agency’s pamphlets, ‘How can we avoid tokenism and ensure that any views will result in meaningful changes and service improvement?’¹⁰⁹ (© Crown Copyright 2005. Contains public sector information licensed under the Open Government Licence.)

Accordingly, the underlying theory comes in for revision. As a typical example of how it is supposed that logistical change will be delivered, we turn to a later document within the family. The NHS Scotland report on Patient Flow in Planned Care explains that ‘Improvements *cannot take place* without teams of people committed to improvement and with the knowledge required to change things’ [emphasis added].¹¹⁰

The report goes on to suggest the need for 'three generic team types that can be moulded and shaped depending on the improvement projects chosen'.¹¹⁰ The trio is designated 'Real Action Teams', 'Flow Improvement Teams' and 'Deep Analysis Teams'. Each of those teams is then kitted out with ideas in the form of 'toolkits', 'case studies' and 'educational forums'. We will not pause here to uncover the detailed suppositions behind these developments; we simply note that programme theory is on the move. In terms of their volume and prominence, programme theories that stress the need for strategic co-ordination are at the forefront of the demand management literature. On closer reading, and in terms of *Figure 9*, strategic change (level 1) appears to rest on a considerable degree of organisational reengineering (level 2).

Administrative change

The second proposed measure for improving demand management/referral management makes the case for transplanting one agency to do the work of another. The basic working hypothesis is that a fresh institution or agency, B, is better placed to carry out a particular function than the old regime, A. Specific responsibilities are handed over rather than the entire care process (which is the domain of level 1 change). Day-to-day improvements are deemed to flow by dint of better specialist management – the new body produces a better division of labour, better cost containment, better information flow, better training and so on. RMCs or referral management gateways are the perfect exemplification of this mode of change. A new body is transplanted into the midst of the care pathway to assume responsibility for core referral decisions. We borrow some promotional material from Fortis Healthcare,¹¹¹ which provides a vivid picture of the intended transplant and its interconnections (*Figure 10*).

Given the substantial reorganisation involved, it is perhaps not surprising that there are many different (and still evolving) subtypes. We provide further details on these variations on a theme in *Chapter 4*; here it suffices to point out that RMCs vary markedly in terms of the personnel responsible for making referral decisions. Much ink has been spilt describing the potential of these agencies in accompanying plans, proposals and position papers. In these narratives RMCs are envisaged to improve the management of demand by queue sculpturing of flows and blockages, with the better use of information for planning, by keeping a close eye on cost containment, with better support for patient choice and with the provision of education and training for key players in the system. Some brief and typical examples of the working ideas (programme theories) behind these claims can be seen in the following examples (*Box 15*).

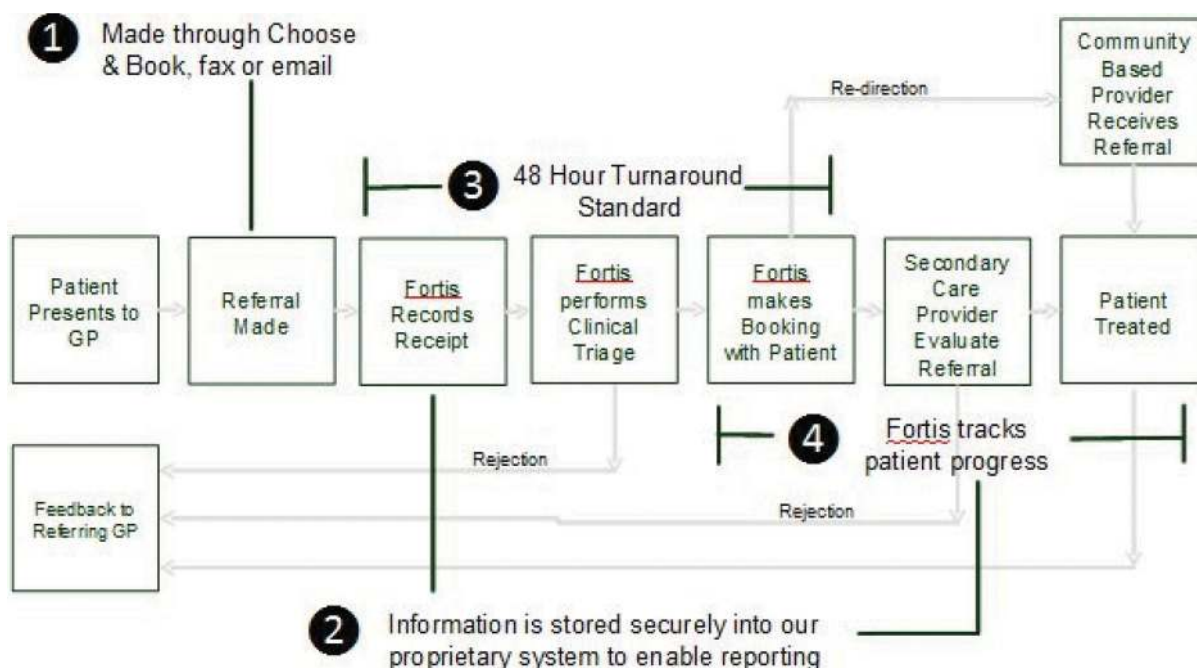


FIGURE 10 Prototypical structure of a RMC/referral management gateway. Reproduced with permission under the Open Government Licence 3.0.¹¹¹

BOX 15 Referral management centre programme theories

From Somerset Local Medical Committee's 2004 'position paper' on referral management centres

Not all referrals require a consultant opinion. Some can be dealt with by, for example, a specialist physiotherapist or a nurse practitioner. Yet others may be better handled by a GPwSI working for the PCT [primary care trust]. Clinical assessment of referrals should mean that the most appropriate person sees the patient – quickly, close to home, and at the lowest reasonable cost.

Reproduced with permission from Somerset Local Medical Committee¹¹²

From Billy Fashanu's comment on Elwyn and Davies' 2006 British Medical Journal paper

Better still, if the referral management system allows for closer access to hospital specialists guided by an agreed pathway, then the patient can be directed to the right services at the right time, hopefully in a more cost effective and clinically efficient manner (common evidence based pathways can be agreed by all clinician and issues of legality addressed as a group).

Reproduced with permission from Billy Fashanu¹¹³

From Somerset Local Medical Committee's 2004 'position paper' on referral management centres

Referral within the NHS has hitherto been chaotic. Without a checking system it has been impossible to make sure that a referral is despatched, that the hospital has received it, and that an appointment has been sent in response. An RMC can use simple IT to log and monitor all referrals. The patient will know that the referral has gone because RMC staff will contact them, and the PCT [primary care trust] will be able to check the progress of referral and initiate action if waiting list targets look like being exceeded.

Reproduced with permission from Somerset Local Medical Committee¹¹²

From Department of Health case study: Fortis Healthcare

Doctors get feedback about their referrals. If there is something missing, the doctor is asked for further detail. If the Gateway panel deem the referral unsatisfactory (ie, the referral is not reasonable), it is returned. This means that the doctors learn from the process and the quality of referrals improve. Monitoring referrals by type help the Gateway team to identify areas where GPs may require extra training.

Contains public sector information licensed under the Open Government Licence v1.0¹¹¹

Our purpose here is not to test the veracity of these claims but simply to identify the purported mechanisms for change. One sees immediately that RMCs are expected to work *through* a range of other processes: changing responsibilities, devising agreed pathways, improving information systems, employing feedback and training. In other words organisational change (level 2) actually works through role, procedural and motivational mechanisms (levels 3, 4 and 5).

Perusing exactly the same literature reveals another feature of these particular programme theories, namely the accumulation of critical voices about why their expectations may not come to pass (*Box 16*).

Again, our purpose is not to begin the process of gathering material to adjudicate between the various claims and counter claims here, but to show that working programme theories are always mobile. The first batch of theories envisages change in the decision-making apparatus to ensure that patients are directed to 'the right person in the right setting at the right time'. The second set of theories (some, interestingly, from exactly the same sources) foresees that some of the key personnel may doubt that the new incumbents can ever be that 'right person'. Further iterations, what is more, take us to a yet more discriminating 'compromise theory' and begin to include notions of trust, shared ownership and gradual development (*Box 17*).

This brief history of RMC provides salutary lessons for the reviewer. Programme theories are not respectful of programme labels, and nor are they likely to remain passively stable.

BOX 16 Referral management centre programme theories (negative)

Excerpt from Nowotny, 'Gatekeeper job is one for trained GP', *Pulse*, 10 August 2011

*Generally, the decision to refer a patient is a carefully considered one, the product of a GPs experience and training, and based on one or more face-to-face consultations. For an anonymous nurse working to a set protocol in a remote referral management centre to reject that carefully considered decision out of hand is professionally frustrating, to say the least. The primary concern, of course, is patient safety. Both GPs and consultants told *Pulse* they feared under-qualified medical staff were inappropriately rejecting referrals because they do not conform to inflexible algorithms.*

*Reproduced with permission from *Pulse* magazine¹¹⁴*

Somerset Local Medical Committee's 2004 'position paper' on referral management centres

GPs value their personal knowledge of a consultant's skills, practice, and personality when making a referral. This will be lost if referrals are directed generically or to a different provider. Patients also usually prefer to see the same specialist if they have previously been treated for a related condition.

Reproduced with permission from Somerset Local Medical Committee¹¹²

BOX 17 Referral management centre programme theories (refined)

Excerpt from Wootton and Whiting, 'Creating ownership of a referral gateway', Pulse, 3 September 2012

With the exception of one or two practices, the vast majority of GPs from the 103 practices in Manchester now use the gateway . . . a number of factors we believe have led to this almost universal acceptance. The first is that this was not something introduced suddenly and on a whim. It was developed over a long period of time and after plenty of discussion and collaboration with primary care through local meetings, the CCGs [Clinical Commissioning Groups] and the LMC [local medical committee]. We showed GPs it would be effective through the smaller pilot at the beginning. One very important factor in setting up the gateway was the use of local GPwSIs and senior, well-known GPs. It meant the other GPs respected the decisions made and the feedback they were receiving. Because the advice is provided in a very professional and supported way, the GPs saw it as more of a CPD [continuing professional development] tool than a criticism of their decisions.

Reproduced with permission from Pulse magazine¹¹⁵

Role change

This subset of theories refers to a less dramatic change in which systems and organisational structures remain intact but within which new roles are created and responsibilities are shifted. Decision-making is transferred into the hands of those who are considered better placed to interpret and act on referral options. The prime manifestation of this idea is the introduction of a new breed of practitioner known as a general practitioner with a special interest. This role change sees a blurring of the distinction drawn at the inception of the NHS between consultants as 'specialists' and GPs, ipso facto, as 'generalists'. There are several different models of GPwSI services.¹¹⁶ Some operate autonomously, some operate nearer the poles in either primary or secondary care and some function within new configurations such as RMCs. By this fact alone it becomes clear we are not dealing with a unitary programme theory.

Much of the remit of GPwSIs is concerned with demand management. Broadly speaking, they are supposed to improve patients' access to secondary care by managing less complex patients themselves or diverting them back into primary care. This in turn is deemed to shorten waiting lists and save costs. We undertook a search of the literature, again concentrating on official reports, positions papers, blogs and the professional press, seeking out the programme theories on *how* such a transformation is understood. There are many such hypotheses. GPwSIs are said to improve or 'stratify' the case mix arriving in secondary care; to handle a significant subset of patients themselves, thereby reducing handovers; to reduce costs associated with care in expensive hospital surroundings; to combine holistic patient care with specialist secondary knowledge; and to reskill and provide feedback to primary colleagues on the basis of the new role. We provide further details and analysis of the respective hypotheses in *Chapter 5*. Here, as previously, we chart the evolution of few key programme theories, starting with the core idea of GPwSIs as interlocutors or expert intermediaries (*Box 18*).

Awaiting in the critical literature are many confounding arguments that claim that, at best, GPwSIs will lack expertise and have no impact on referrals to secondary care or on waiting times and, at worse, will lengthen the patient pathway, increase overall demand for care and 'cream skin' more easily managed and higher tariff patients (*Box 19*).

One notes that under such common criticisms the GPwSI intervention begins to morph and is supplemented with other measures such as a case-mix control triage and a central booking system in the above testimony. With the further passage of time the GPwSI programme theory undergoes further adaptation, as in *Box 20*.

BOX 18 General practitioners with special interests' role as an intermediate tier of expertise

Perhaps the biggest advantage is that the GPwSI can provide an intermediate tier of expertise and advice to their primary care colleagues and alternative avenues for referral and access to specialist investigations.

*Reproduced with permission from the Royal College of General Practitioners from Gerada C, Wright N, and Keen J. The general practitioner with a special interest: new opportunities or the end of the generalist practitioner? The British Journal of General Practice: the Journal of the Royal College of General Practitioners 2002;**52**(483):796–8¹¹⁷*

Secondary care specialist time is best spent on managing less common or more severe respiratory illness. GPwSIs should enable the management of those common illnesses which, given appropriate resources, could be managed in the community . . . GPwSIs can be seen to be addressing problems at the interface between secondary and primary care particularly in relation to the knowledge gap between GPs and consultants. This gap . . . leads to the following problems: i) Inappropriate referrals: 40–80% of ENT [ear, nose and throat] referrals could potentially be dealt with by GPs, provided they acquire additional expertise, ii) Unnecessary appointments with hospital consultants: for some patients, tests may be necessary but treatment could then be carried out in primary care.

*Reproduced with permission from the Royal College of General Practitioners from Nocon A, Leese B. The role of UK general practitioners with special clinical interests: implications for policy and service delivery. The British Journal of General Practice: the Journal of the Royal College of General Practitioners 2004;**54**(498):50–6¹¹⁸*

The main aim is to stop the consultants seeing inappropriate referrals. At the end the day the orthopaedic surgeons should only see patients that need an operation and the rheumatologists should only see patients who need a second line agent and need be monitored by a practice nurse.

Honey et al.¹¹⁹

BOX 19 Paradoxes in the role of expert intermediaries

Other GPs may be reluctant to refer to colleagues, whom they perceive as generalists with much the same level of skills as their own, and there may well be a prolonged introductory period during which GPwSIs receive few referrals.¹¹⁸

Either they are doing what generalists should be doing anyway, in which case what is the point, or they are doing interventional procedures such as endoscopy that really should be done in a unit with supervision from experienced clinicians.¹²⁰

In assessing their value to the service as a whole, it is important to remember that even the best GPwSI will not have the skills or experience to deal with an unselected range of patients. Unless the case-mix is controlled by some form triage a proportion of patients will have to attend for two consultations: an unnecessary inconvenience which would actually increase the waiting list for some and add to the overall costs involved.

*Reproduced with permission from Robin Graham-Brown from Graham-Brown R. GPwSIs: a useful addition, at a cost, but not a substitute (rapid response). BMJ 2005;**331**:1441¹²¹*

In hard-pressed departments . . . the effect of creaming off the 'easy stuff' may simply be that the consultants can set more suitable (i.e. longer) appointment times for the connective tissue referrals, and thus there will be no impact on waiting times at all. Accordingly, although our GPwSI service has been well-advertised, we are at present running a common clinic booking system. Whether a GPwSI session is more useful than having extra consultant sessions may depend on the individuals involved; whether specialist services for small subgroups of GP practices in community settings can ever be a sensible use of scarce resources I doubt.

*Reproduced with permission from A Bamji from Bamji AN. Evaluating GPSIs (comment). BMJ 2003;**327**:460¹²²*

BOX 20 Supplementary changes to support role change

A minimum level of clinical experience, necessary equipment, and specialist support should be identified for all procedural or interventional services. A system of accreditation of competence to offer a pre-specified range of clinical services must be developed. General practitioners with special interests should not practice in isolation and should have easy access to advice, support, and professional development from local hospital specialists.

Reproduced from General practitioners with special interests: a potentially valuable asset, which requires evaluation, Rosen R, Stevens R, Jones RB, BMJ, vol. 327, pp. 460–2, 2003,¹²³ with permission from BMJ publishing Group Ltd

History, or to be specific, programme theory history, repeats itself in this example. Over time, basic ideas about role revision (level 3 change) are supplemented with the incorporation of guidelines (level 2) professional development and accreditation programmes (level 1) as well triage services (level 2) staffed by multidisciplinary teams.¹¹⁹

Procedural change

This subset of programme theories seeks to balance capacity and supply by acting more directly on referral decisions. All systems, organisations and roles remain in place, so that improvements are deemed to follow from giving individuals increased scope and/or guidance to reflect on and reshape their own decision-making. This domain has proved the busiest corner of policy-making in demand management. Accordingly, we split procedural change interventions into three subtypes: (1) peer review and feedback, (2) guidelines and (3) GP direct access to tests. Each one has attracted considerable debate on programme theory.

Peer review and feedback of referrals

Peer review of referrals operates, perforce, through 'peers', namely GPs, reviewing each other's referrals and giving feedback on their appropriateness. There is variation in how this is accomplished; whether it is undertaken prospectively, before the referral is made (in which case it is akin to a form of in-house triage), or retrospectively, after that referral has been made. There is also variation in who undertakes the review: GPs in the same practice, GPs from other practices or GPs together with consultants.

The core idea underlying peer review is that sharing knowledge and expertise collectively enables GPs to identify alternative management strategies for the patient, thus avoiding overreliance on secondary care. Shared wisdom will help GPs to make better referral decisions in the future. If GPs review their own referrals in relation to other practices, it will help them reflect on why those practices' rates might be higher or lower than theirs and take action to address this. There are additional benefits in the form of increased confidence to resist patient pressure for referral and having a greater tolerance of uncertainty and risk. Increasing contact with colleagues is said to also engender a culture of trust, which can support GPs to make difficult decisions and share the burden of blame should there be any adverse outcomes. *Box 21* provides first-hand testimony on these ideas.

BOX 21 Feedback as shared wisdom

This approach is not about 'rejecting' referrals, but about supporting GPs to make difficult decisions. General practice can be a very lonely life, and having the support of your colleagues and being able to call on expertise within primary care is important to the quality of care we provide. Through this approach, if the course of action is not to refer but manage the patient in general practice, if something goes wrong you can stand up and say: 'We spent time talking about the best thing we could do for you and several of us agreed this was the path to take'.

Wilczynski (reproduced with permission from Pulse magazine)¹²⁴

It is very much a developmental process, not a performance-management one. Practices and GPs can look at where they are against their colleagues. It is about getting people to reflect on their referrals – improving quality, not just reducing referrals. I think this is what every CCG [Clinical Commissioning Group] needs to do. All of us will be bottom of something so it is not about penalising. Its strength is in its simplicity.

Everington (reproduced with permission from Pulse magazine)¹²⁵

Cheek by jowl in the same literature, we identified numerous of pieces claiming that, although peer review may improve the knowledge, skills and confidence of GPs, it does not actually change their referral behaviour. GPs were able to forward 'special case exemptions' on the basis that collective referral data were out of date or inappropriate. They also offered alternative explanations as to why their referral rates are unusually high or low, which were out of their control. Typical examples are provided in *Box 22*.

What happens as a result is that 'peer-review programme theory' becomes the subject of debate. We encountered numerous claims and counter claims arguing that specific modes of conducting peer review are more likely to change referral behaviour than others. Some argued that prospective peer review, that is where referrals are reviewed before the referral decision is made, is more likely to change referral behaviour than retrospective peer review because it 'challenges behaviour in real time'. These forms of peer review are more like a 'triage function' and are, thus, a hybrid of administrative (level 2) and role change (level 3) theories. An additional administrative structure (review of referrals by a different GP or consultant) is placed on the patient pathway with the creation of additional roles. Other significant voices argue that impact of peer review can be strengthened by incentivising individual GPs (level 5 change). Thus, 2011 saw the inclusion of indicators relating to peer review of referrals into the QOF, which rewards GPs in the UK for implementing good practice. The seventh version rewards practices that participate in external peer review to compare secondary care outpatient data. Yet other revisions highlight the importance of dialogue with consultants (level 3) and protected administrative time (level 2) to engage in peer review as central to its success. *Box 23* provides a brief review of the medley of emerging propositions:

BOX 22 Resistance to peer review

Initially we worked as a clinical community on an incentive scheme based on peer review of referrals, which we believed would solve the problem. But while we identified learning opportunities, it seemed to do little to actually change behaviour. The main problem was that the data was always two or three months out of date.

Wootton and Whiting (reproduced with permission from Pulse magazine)¹⁵

Widespread scepticism about accuracy of their own and other practices' data seriously limited doctors' perception of the value of the feedback. The doctors in 89 out of 92 practices visited saw no need to alter their referral behaviour, and when rates were unusually high or low explanations were readily offered, which rarely included the doctors' own clinical behaviour.

*Reproduced from How valuable is feedback of information on hospital referral patterns?
de Marco P, Dain C, Lockwood T, Roland M. BMJ, vol. 307, pp. 1465–6, 1993
with permission from BMJ Publishing Group Ltd¹²⁶*

BOX 23 Adapting the process of peer review

The educational aspect of the prospective review approach is the most important factor in all of this. Even in the practice where referrals are looked at by partners before a final decision is made, there is a weekly meeting to discuss those referrals where an alternative option was recommended, with registrars presenting the case. We issue CPD [continuing professional development] certificates with these meetings, so instead of becoming a chore it is part of GPs' career development.

Wilczynski (reproduced with permission from Pulse magazine)¹²⁴

However, the [referral change] effect ceased as soon as the active intervention finished, indicating that active participation is key – reliance on a change of culture or referral behaviour to maintain reduced referral rates after a limited intervention cannot be supported from our experience. Furthermore, active participation of consultants appeared important.

Reproduced under the terms of the Creative Commons Attribution Licence from Evans and colleagues¹²⁷

The general practitioner participation was resourced to provide protected time for review and discussion in the practice. Each of these features of the project – continued active participation by GPs with resources and dialogue with consultants – appeared essential and required attention in planning further implementation of referral management strategies.

Reproduced under the terms of the Creative Commons Attribution Licence from Evans and colleagues¹²⁷

Guidelines on treatment and referral options

Clinical guidelines are methodically developed statements devised to help practitioners and patients decide on appropriate health care for particular clinical conditions and/or circumstances.¹²⁸ Under this intervention, all systems, organisations and roles remain in place – procedural improvements are deemed to follow from giving individuals guidance to reflect on and reshape their own decision-making. Guidelines apply to all aspects of health care, the principal application to demand management being the notion that they decrease unacceptable or unwanted variations in practice.¹²⁸ In *Chapter 2, Supplier-induced demand*, we have already seen evidence of the wide variations in clinical practice that apply between individual physicians, between practices, between regions and between nations.

Accordingly, the principal programme theory's guidelines work by 'standardising practice' – by imparting some systematisation in the decision-making process. The appeal to common practice is made under two further hypotheses, namely (1) on provenance (guidelines are produced by authoritative, multidisciplinary, professionally authorised, nationally representative groups) and (2) on the evidence base (guidelines reproduce expert, systematic, comprehensive, consensual and up-to-date recommendations). *Box 24* reproduces some typical presuppositions, varying somewhat in their assertiveness:

BOX 24 Guidelines provide standardisation around best practice

They offer explicit recommendations for clinicians, who are uncertain about how to proceed, overturn the beliefs of doctors accustomed to outdated practices, improve the consistency of care, and provide authoritative recommendations that reassure practitioners about the appropriateness of their treatment policies. Guidelines . . . clarify which interventions are of proved benefit and document the quality of the supporting data. They alert clinicians to interventions unsupported by good science, reinforce the importance and methods of critical appraisal, and call attention to ineffective, dangerous, and wasteful practices . . . Guidelines can also improve the consistency of care; studies around the world show that the frequency with which procedures are performed varies dramatically among doctors, specialties, and geographical regions, even after case mix is controlled for . . . Guidelines offer a remedy, making it more likely that patients will be cared for in the same manner regardless of where or by whom they are treated.

Reproduced from Clinical guidelines: potential benefits, limitations, and harms of clinical guidelines, Woolf SH, Grol R, Hutchinson A, Eccles M, Grimshaw J, BMJ, vol. 318, pp. 527–30, 1999 with permission from BMJ Publishing Group Ltd¹²⁹

NICE clinical guidelines are recommendations for the care of individuals by healthcare professionals, based on the best available evidence. They are also important for health service managers and commissioners. Guidelines can be used to develop standards to assess the practice of healthcare professionals, help in the education and training of healthcare professionals, help patients to make informed decisions, and improve communication between patients and healthcare professionals.

Reproduced with permission from NICE¹³⁰

Guidelines have a chequered history. We enter the debate after some have attempted to foreclose it. Imison and Naylor's King's Fund review concludes that there is little evidence to support passive use of clinical guidelines.⁹³ We return to this body of evidence in *Chapter 8*. Here, we follow what is a rather vituperative exchange on the basic principles, captured by the tragicomic title of a 2011 *Pulse* article, 'Protocols reduce general practice to Maoist collective'.¹³¹ Grave doubts about the ability of guidelines to shift existing patterns of decision-making are forwarded on at least three grounds: (1) one size cannot fit all, (2) clinical autonomy is threatened and (3) guidelines are not self-implementing and thus go unnoticed. *Box 25* voices some typical arguments.

These points have been recognised by friends and foes alike, resulting in major attempts to make guidance more user-friendly. There is widespread acceptance that the passive use of guidelines is impotent and a whole range of supplements are promoted on their active implementation. Readers will be unsurprised to learn that this takes this intervention far beyond its concern with the specifics of decision-making and into other realms of change such as incorporation in NHS governance strategies (level 1), adaptation to local needs (level 2) and the development of further procedures and protocols (level 4) to avoid the 'guideline bin'. Examples are provided in *Box 26*.

BOX 25 Guidelines as opposed to real clinical decisions

Guidelines that are inflexible can harm by leaving insufficient room for clinicians to tailor care to patients' personal circumstances and medical history. What is best for patients overall, as recommended in guidelines, may be inappropriate for individuals; blanket recommendations, rather than a menu of options or recommendations for shared decision making, ignore patients' preferences.

Reproduced from Clinical guidelines: potential benefits, limitations, and harms of clinical guidelines, Woolf SH, Grol R, Hutchinson A, Eccles M, Grimshaw J, BMJ, vol. 318, pp. 527–30, 1999 with permission from BMJ Publishing Group Ltd¹²⁹

GPs described a tension between their own experiences and the guideline recommendations . . . GPs saw consultations with real patients as more complicated than their portrayal in the guidelines. Guidelines were experienced as not flexible enough to take into account the complexity of individual circumstances, such as multiple diagnoses, . . . , and patient preference.¹³²

Conceptually, clinical guidelines and professional autonomy have a paradoxical relationship. Despite being the quintessence of medical knowledge at the corporate level, guidelines diminish the clinical autonomy of individual practitioners, and therefore threaten medicine's justification for its autonomy.¹³³

McLellan¹³⁴ states that guidelines do not implement themselves and many clinicians note that guidelines are often overlooked. In practice, guidelines are published, sent to professionals and supported by some lectures.

BOX 26 Supplementing guidelines

McLellan¹³⁴ states that national referral guidelines such as NICE guidelines can be controversial but if they can be developed with the participation of patients and professionals, they may provide better information, reduce treatment discrepancies and save the NHS money.

Uncertainty remains over the impact of guidance on reducing variation in referral rates from primary to secondary care, with the extent of its impact being dependent on the specific features of the guidance and the local cause of variation. An important and recurrent theme in the literature is a need to stimulate better joint working and dialogue between primary and secondary care. Referral guidelines should not, as has been cautioned, reduce the willingness of GPs to tolerate uncertainty and increase referrals to secondary care. Accordingly, NICE referral guidance should be used to encourage local health communities to discuss referral problems and develop local referral protocols.

Reproduced from NICE 'referral advice' recommendations database, Docherty MJ, Choudhury M, Littlejohns P, BMJ, vol. 342, p. d129, 2011, with permission from BMJ Publishing Group Ltd¹³⁵

Direct access to diagnostic tests

This is a third and most immediate example of procedural change. Decision-making is transferred into the hands of those who are considered better placed to interpret and act on referral options, one such example being providing GPs with direct access to diagnostic tests. The idea is to allow GPs to bypass the 'gatekeeper role' of the consultant in accessing tests results. The assumption is that, armed with immediate data, GPs will manage some patients directly within primary care and, thus, reduce the demand on the secondary system. Moreover, patients who do need to see a consultant will be able to access that care more quickly. Essentially, under this task reallocation model, GPs take on minor aspects of the consultant's brief. *Box 27* articulates the rationale.

Critics, however, have lined up to express concerns that GPs may then overutilise diagnostic tests and also to question whether or not GPs have the knowledge and skills to interpret the tests or to take the optimal action on the basis of the results.¹³⁸ *Box 28* provides a brief selection of examples.

The final quotation here heralds the now familiar process in which the initial programme theory is buttressed by a range of further stanchions that work across the service delivery infrastructure. We see that, in order to ensure that GPs use the test appropriately, the procedural change is overlaid with benchmarking, target setting and even the 'naming and shaming' of recalcitrant practices. Over time, these adaptations become part of the fabric of the intervention. *Box 29* illustrates some further examples in which (1) wholesale organisational change (level 2) and (2) feedback (level 3) and addition procedural changes (level 4) are utilised to manage direct access.

BOX 27 Simplifying and rationalising the diagnostic chain

Under conventional systems of care, outpatient clinics see patients referred by a GP for clinical assessment by a hospital specialist. Subsequent hospital visits are arranged to undertake any specialist diagnostic tests that may be required and to initiate treatment where necessary. In other words, the specialist in the outpatient clinic acts as a gatekeeper to other hospital resources. Allowing the GP to bypass this gatekeeper and gain 'direct access' to tests can enable GPs to make more efficient use of hospital resources . . . Direct access to diagnostic services should reduce outpatient attendance in that GPs may refer patients for diagnostic testing without prior consultant assessment. Waiting time from presentation to testing is accordingly reduced. If the patient can be managed by the GP without subsequent referral to a consultant, waiting time from presentation to treatment is also reduced and further outpatient attendance avoided.

Republished with permission of the Royal College of General Practitioners, from Direct access to diagnostic services, Sibbald B, British Journal of General Practice, 59(562), e144–5, 2009¹³⁶

GPs see only around 8 or 9 new patients with cancer each year. However, they see many more presenting with symptoms that could be cancer. And those symptoms are generally symptoms of many other diseases as well. This therefore makes it difficult for GPs to assess when it is appropriate to refer patients on to secondary care for investigation. GPs need easy access to the right diagnostic tests to help them to diagnose or exclude cancer earlier. We are committing additional funding over the next four years to enable GPs to have better access to selected diagnostic tests, along with funding for the additional costs of tests and treatment in secondary care.

Reproduced from the Department of Health under the terms of the Open Government Licence v3.0¹³⁷

BOX 28 Direct access: unforeseen difficulties and unintended consequences

A *Pulse* article features an interview with a GPwSI in musculoskeletal conditions who explains that a direct access scheme for MRI to investigate back or shoulder problems did not work because 'GPs had not been comfortable in interpreting the results'.¹³⁹

Direct access may increase demand for testing and lead to less appropriate referrals with a consequent reduction in diagnostic yield.¹³⁶

Exclusive: PCTs [primary care trusts] restrict access to diagnostic scans. 'GPs are facing heightened scrutiny over use of diagnostic scans, as PCTs scrap direct access to save costs, a GP investigation has found. Responses from 116 PCTs to a Freedom of Information request revealed that 28% of PCTs are benchmarking practices' referrals to diagnostic scans or plan to do so. In addition, a quarter of PCTs are supporting practices to reduce inappropriate access to scans. Others are taking more drastic action. Nine trusts have identified practices whose use of diagnostics is deemed excessive or under-used. The government prioritised GP access to diagnostic scans in its cancer strategy last year and PCTs were instructed to improve direct access as part of the 2011/12 NHS Operating Framework. But five PCTs have decommissioned some direct access schemes to key diagnostic scans, such as MRI, CT and ultrasound. One trust has implemented strict upper and lower target levels of use, and another is considering introducing limits'.

Reproduced with permission from GP magazine online/medeconomics¹³⁸

BOX 29 'Modernising' direct access provision

Electrocardiogram (ECG) with remote interpretation services. ECG is the basic tool for diagnosing heart problems. GPs are not always equipped to interpret ECG tests and consequently many refer their patients to secondary care for outpatient appointments or to diagnostic centres. Broomwell Healthwatch, a private provider based in the North West, offers GP practices, medical centres, and hospitals access to immediate, expert, clinician interpretation of ECG test results. This means ECG results can be interpreted immediately, and within the GP practice, thereby having great potential to save NHS money by reducing the need for outpatient referrals and improving patient care by enhancing patient convenience and reducing waiting times for diagnostic tests.

Reproduced from the Department of Health under the terms of the Open Government Licence v3.0¹³⁷

From the beginning of the scheme, in order to ensure that GP referrals were appropriate, a clinical triage function was established at the point of receipt of referral. This ensured that every test requested was appropriate, that the patient could be scanned safely and that the appointment for the test was planned for an appropriate setting. Three initiatives: referral guides, structured referral forms and feedback were used to improve the appropriateness of referrals for diagnostic tests from Primary Care.

Reproduced from the Department of Health under the terms of the Open Government Licence v3.0¹³⁷

Motivational change

The final theory of change (level 5) includes the practice of offering incentives to encourage GPs to reshape their referral behaviour. Change is located squarely in the hands of the individual practitioner; all system pathways, organisations, roles and routines remain in place. Improvements are deemed to follow from confronting individuals with financial carrot or stick to change and/or limit their referral practices. The intended change is behavioural and driven by personal interest rather than peer learning. Instrumental calculations are reckoned to insinuate change in the overall balance of referral decisions.

As noted above, indicators relating to the peer review of referrals and the development of care pathways were added to QOF indicators in 2011. In addition, some Clinical Commissioning Groups (CCGs)/primary care trusts (PCTs) have implemented 'referral incentive schemes' whereby GP practices receive payments for meeting targets to cut referral rates. A recent *Pulse* article entitled 'Five questions you should ask before offering referral incentives' suggested that 'Under a well-planned scheme, commissioners have the possibility of improving outcomes and to make care more cost effective, releasing money to be spent on other services'.¹⁴⁰

Once again, counter-theories await – postulating, in this instance, that referral incentives can encourage a blanket approach to all patients with a similar condition and, thus, fail to take account of individual patient differences and their referral needs. In this way they may introduce conflicts into the doctor–patient relationship, which doctors find unacceptable, particularly if decisions are felt to be motivated by cost rather than by clinical evidence.¹⁴⁰ Indeed, the proliferation of referral incentive schemes that match rewards to specific referral targets and therefore appear to reward arbitrary reductions in referral activity have been widely condemned by the British Medical Association (BMA).¹⁴¹ In their guidance to GPs and local medical committees (LMCs), the BMA raised significant concerns regarding the professional and ethical implications of target-driven referral incentive schemes where there is no concurrent assessment of the appropriateness of such reductions. In particular, they warn against pursuing the reduction of referrals at the expense of clinical appropriateness, as they fear that this may have a detrimental effect on the care provided to patients.¹⁴²

In order to be successful, revised programme theories advocate that incentives need to be targeted at outcomes that have a consensus regarding clinical importance.¹⁴³ As a *BMJ* article opines, GPs respond positively to incentives for peer review of referrals because 'it's a very good way of learning from one another'.¹⁴⁴ Here we see that motivational stimuli (level 5) appear to require the support of some mode of interpersonal adaptation (level 4) in order to maintain the delicate balance of quality and volume of referrals.

Common denominator: a unifying programme theory?

So what have we gleaned from this close inspection of the programme theories underpinning the many decades of efforts to contain demand for health care? Unsurprisingly, we learn that the initiatives vary massively in order to take care of considerable differences due to medical condition, due to the complexity of patient pathway and due to existing modes and levels of primary and secondary provision. The potential causes of demand inflation, discussed in the previous chapter, sit unevenly across these assorted clinical and service domains. Accordingly, we learn that interventions aiming to improve GPs' understanding of electrocardiogram (ECG) testing have little in common with those designed to improve referral letters or with those seeking to reorganise booking services, and so on.

We have, however, found a serviceable starting point to analyse these differences using a typology (see *Figure 9*), commonplace in the organisational studies literature, which differentiates the ambitions to produce change along a continuum ranging from the macro to the micro. Along with other authors we have been able to make a useful demarcation of demand management interventions into five main modes: strategic change, administrative change, role change, procedural change and motivational change.

However, by investigating these programme theories in a longer time frame, we notice a recurrent pattern under which they deform and reform. This should hardly come as a surprise. It is not the job of practitioners and managers to maintain interventions in their pristine purity. Their job is to make improvements work 'on the ground' and they work under the constant refrain of 'innovation' and 'modernisation'. Programme adaptations are thus myriad and tailored to local circumstances. But we have also been able to demonstrate that refinements are not simply piecemeal – administrative minds do tend to think alike. Accordingly, the prevailing motivation for the modifications has been to add supplementary procedures that operate across the levels of change (schematically one can say that 'level 3' plans are often buttressed by auxiliary ideas at levels '1', '2', '4' and '5', and so on).

It is useful to consolidate this proposition by returning to the generic literature on social and organisational change. We have already drawn on the idea of 'vertical stratification' to buttress our original fivefold typology (vertical stratification referring to the ontological levels in *Figure 9* rather than a hierarchy of 'seniority' that occurs in most organisations). This same literature also draws attention to the idea that change is mediated by the intercommunication or intersection between these layers. We provide a highly abbreviated tour of this recurring theme using two sources, one elderly and purpose-built, the other modern and metaphysical. Suitably ransacked, these authorities provide the basis for a simple model of the key, interrelated layers of change in health-care systems that will form the basis of our subsequent analysis.

A pioneering attempt to describe the corporate layer cake occurs in a book aptly titled *Social Change in Complex Organisations*.¹⁰⁸ Sitting at the conceptual apex is something these authors refer to as 'centralization' of the organisation. This is the platform for core decision-making; overarching strategies and systems are established here. Next comes 'formalization' – as the authors explain, 'organizations need daily guidelines for their operation; these guidelines are furnished by rules, the repository of past experiences'. At a level down, we have 'stratification', by which Hage and Aiken mean the division of labour between members and the degree of differentiation in their roles. The next two strata are labelled 'production' and 'efficiency'. These carry similar meaning referring to how organisational goals are embodied in everyday procedures and work routines. Finally in the model we reach the level of the individual member and the idea that change inheres in their 'morale', their levels of loyalty and their team spirit.

All of these layers (systems, rules, roles, routines, morale) wrap and unwrap as institutions evolve. And it is this image which provides Hage and Aitken's central thesis on the methodology for studying change: 'A temporal perspective, one in which the focus of attention is not only on the interrelationships of parts of an organisation, but more importantly on changes in these interrelationships is ultimately necessary for a more complete understanding of organizational behaviour' (p. 28).¹⁰⁸

A rather grander attempt to understand the anatomy of societal change can be found in 'realist' methodology, most especially in the works of Archer.^{145,146} Her theory of 'morphogenesis' attempts to answer the age old sociological chicken-or-egg question about what comes first in propelling social change – is it 'structure' or is it 'agency'? Put simply, her answer is 'chickenegg'. People's immediate actions are shaped within social structures in which they sit: communities, organisations, legal systems, power relationships, etc. However, in a longer time frame, these structures themselves change as a result of the activities and choices of the historically situated individuals who make them up. Social change, in short, occurs through a never-ending cycle: 'structural conditioning' shapes 'social interaction' which in turn shapes 'structural elaboration', and so on, and so on.

It takes more than structures and actors to make the Archerian world go round, of course, and in the detail of her work other layers of social reality are identified: different 'strata' that are also real and are also causally powerful. Thus, for instance, she makes the distinction between 'primary' and 'corporate' agents: individuals with momentary, specific leanings and people with roles and positions who punch more heavily in their world shaping power. Similarly, structures have 'first-' and 'second-' order shaping powers such as

that between a government and any administrative agency or QuANGO it may create. Further, there are distinctions between structures in terms of how they exert influence, through the 'material sphere' of power and resources versus the 'cultural sphere' of information and know-how.

Many other authors have tried to uncover these vital constituents of social change. We rehearse the myriad distinctions no further, however, for what is important is the overall recipe rather than the precise ingredients. Despite the difference in their longevity and ambition both models under discussion have a similar cadence. They stress that it is the interdependency of the strata which enables and constrains change. They agree that each stratum has causal potential and change can be instigated from any level. They posit that sustainable change requires harmonisation through the levels and that there is, of course, every reason to suspect that change provokes disharmony in institutions that are deeply stratified. Finally, these and similar models concur in supposing that change is something that no one steers. It happens perpetually and of its own accord whenever people and groups reflect on their own position.

This latter scenario raises an interesting conundrum: what happens if someone tries to steer change? What happens when there is a formal intervention seeking system improvement? The answer emerging from our longitudinal study of demand management programme theory is that, regardless of where it begins, the progress of any such intervention will also involve a considerable journey as all the other layers absorb the change in interlocking and temporally complex ways. A basic knowledge of system change tells us that what is provoked at one level may be prevented at another. Accordingly, researching service improvement really requires an evaluation of the entire perambulation of a programme theory around a system. With this in mind we submit that an unorthodox but equally useful enterprise for research synthesis is to explore the interconnectivity of different schemes through the different strata. This also supplies us with a broad hypothesis for reviewing specific interventions, namely that they are more likely to work if they gain traction at the different levels. An initial approximation of this conjecture strategy is provided by adapting *Figure 9* to *Figure 11*.

We close our review of demand management programme theories at this point and proceed to part 2 of the report, in which we review the available primary literature using the vantage point of *Figure 11*. Attempting to incorporate this 'whole-system' perspective with a systematic review places interesting demands on the exercise. From its origins in clinical research, systematic review began with the idea of examining all studies on one type of intervention. Increasingly, the situation where the reviewer examines many studies of several different families of interventions is becoming the norm. Now we are confronted with the additional challenge that each family should be considered a shifting lineage. As we noted in the introduction, we entered a busy domain with a colossal amount of primary research but also several existing reviews. It is useful to 'compare notes' on the most useful way of synthesising evidence on such a moving target. As the implications of such a comparison are largely methodological we locate these reflections in *Appendix 2* and more directly to an explanation of our chosen approach.

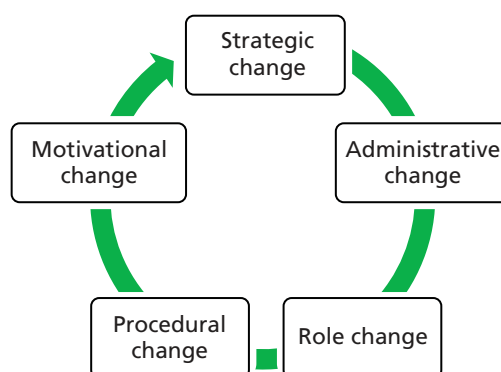


FIGURE 11 Interconnected levels of change.

Part 2 Theory testing

People in any organization are always attached to the obsolete – the things that should have worked but did not, the things that once were productive and no longer are.

Peter F Drucker, p. 66¹⁴⁷

This part carries the main empirical evidence in the review. Thus far, we have attempted to show how demand management interventions evolve under a continuous process of conjecture and critical reflections. Here, we want to scrutinise the evidence on whether or not these programme theories are met in practice. Have the expectations of the programme designers being matched with improved demand and capacity regimes? This prefatory section provides a simple pictorial reminder of the content of the four chapters, illustrating the diversity of the basic theories of change.

Chapter 4 focuses on RMCs, in which a new organisational structure takes over responsibility for the logistics of referral pathways (*Figure 12*).

Chapter 5 focuses on GPwSIs, in which entirely new roles are inserted into referral pathways (*Figure 13*).

Chapter 6 focuses on GPs' direct access to clinical tests, in which responsibility for individual referral and diagnostic decisions changes hands (*Figure 14*).

Chapter 7 moves from these varied attempts to 'shift the furniture' of service delivery to interventions which seek to plan and manage the decisions that make up the whole treatment pathway. Our focus here is on the effort to use formal guidelines in the attempt to curb demand and reduce inappropriate referrals (*Figure 15*).

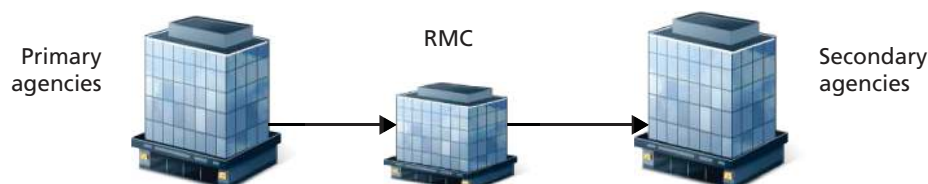


FIGURE 12 Referral management centres.

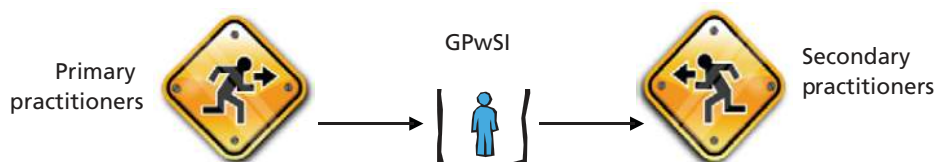


FIGURE 13 General practitioners with special interests.

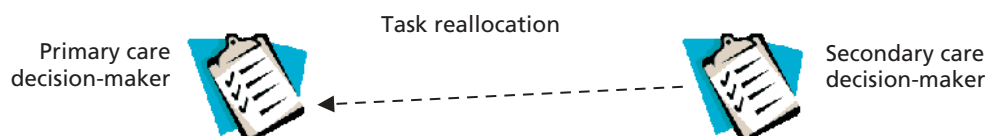


FIGURE 14 General practitioner direct access to tests.

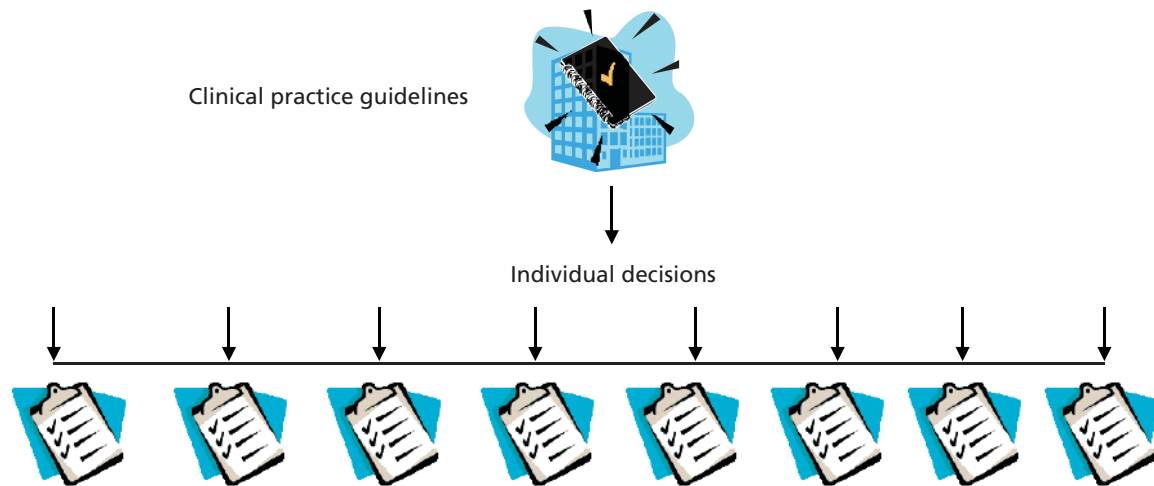


FIGURE 15 Guidelines.

While it is perfectly possible to differentiate these programme theories, we note once more that in the real world of service management they tend to run together. *Chapter 3* was devoted to detailing the interdependencies within the whole family of interventions. To capture the point in an instance, it is worth examining a real treatment pathway as depicted in a recent study on patient choice policies, which is reproduced as *Figure 16*.⁸⁷ It was produced as a case study within a project examining the impact of the patient choice policy on service delivery in the English region of 'Goringhamshire'. It demonstrates at a glance how any change to the system has to navigate and co-ordinate with a whole range of further decisions and decision-makers (DMs). In particular, it should be noted that two of our topics, 'GPwSIs' and 'Referral Facilitation Centres' (a cousin of the RMC), both make an appearance in *Figure 16*. 'Direct Access' and 'Guidelines' are not depicted directly but are ever present in such systems. The figure thus provides a useful reminder that all of the interventions reviewed here are mere portals in a wider system.

Our second preliminary observation is to note that the programme theories underlying these reforms have, of course, been contested. The notion of 'mandate change' is the core hypothesis across the interventions to be reviewed here. The division of labour into which occupations separate have deep institutional foundations and it is a matter of common sense, not to mention considerable academic study, to appreciate that these roots are not easily disturbed. We might anticipate, therefore, that this battle with the status quo will provide the vital evidence terrain for the review. In particular we can expect to examine:

- The extent to which functions actually change in the intended direction? Do the new roles involve significant substitution of decision-making duties or do they merely add additional layers of support?
- Who controls the change? Which professional boundaries come under pressure and what alliances form under the interventions?
- Revising decision-making structures can involve anything from strategic planning to prescriptive authority to clerical choices. Which decisions actually change hands?

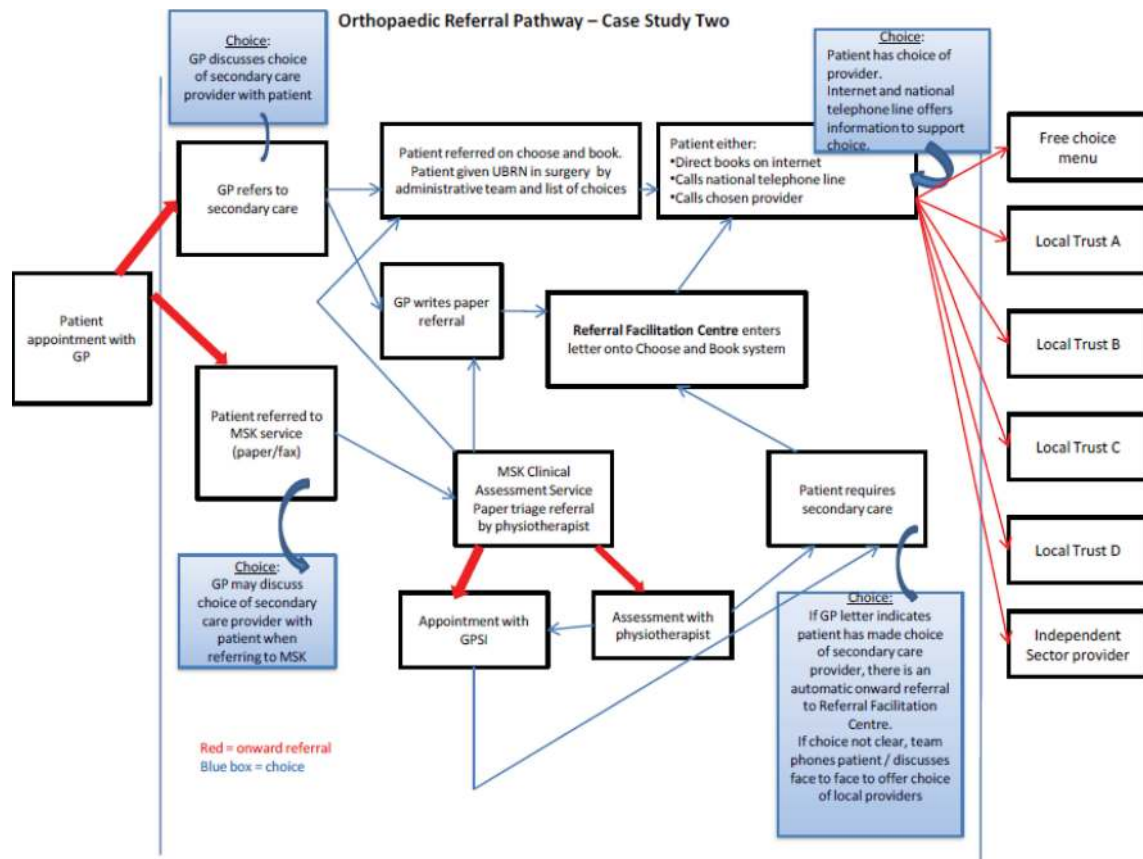


FIGURE 16 Orthopaedic referral pathway: Goringhamshire. Reproduced with permission from Peckham *et al.*⁸⁷

Chapter 4 Organisational change: referral management centres – can they control and shape demand?

Introduction

This chapter reviews the primary evidence on the effectiveness of RMCs. Their defining feature is a matter of the 'level' or 'scope' of the intended reform. RMCs operate within the organisational stratum and involve the insertion of a new institution, centre or unit into the patient pathway to take over aspects of the referral decision-making process.

The chapter begins with some terminological ground clearing and a brief history. RMCs are closely related to a whole family of similar gateways and assessment units and here we draw some simple boundaries around this phase of the review. Even within this relatively well-defined sphere we note heterogeneity. These new centres have been constructed with significantly different ambitions with respect to exactly which 'aspect' of the referral process they seek to control and we highlight some important contextual differences.

The bulk of the chapter is given over to reviewing the primary research on RMCs. The evidence base is not abundant and we comment on how we analyse a body of research mostly comprising descriptive case studies. The analysis follows the typical formula for realist synthesis. The basic RMC programme theory is elicited; the 'ideal outcomes' in terms of referral restructuring are formalised. The empirical evidence is then used to inspect departures from this basic model and two significant unintended consequences are unearthed in the primary literature. A summary explains the vital configuration of conditions that may contribute to the effective operation of RMCs.

Referral management centres and triage: operational logistics

Referral management centres do not appear explicitly in any policy documents emanating from the Department of Health, and nor do a closely related family of bodies variously labelled referral gateways, clinical assessment services (CASs), single point of access services, triage centres, and so forth. These units have largely been developed piecemeal by individual PCTs and later passed down to CCGs. A common ambition of improved steerage of patients through the referral pathway from primary to secondary care is clear to see. However, Imison and Naylor highlight that referral management schemes implemented by PCTs in their research had a subtly different range of different ambitions, with supporting the redirection of referrals to out of hospital assessment and treatment services and seeking reductions in the overall rate of referrals being the most common.⁹³ Davies and Elwyn¹⁴⁸ also note that RMCs vary in whether their remit is (1) to determine if patients could be diverted to a different service or specialist (such as a GPwSI) with greater capacity, (2) to ascertain whether or not the referral contains all the necessary information to permit further care or treatment to be undertaken or (3) to judge whether the referral was clinically necessary or the patient should undergo further investigation in primary care.

Referral management centres and triage operate as an intermediary between primary and secondary care. They are often set up to manage referrals for conditions where there are a number of different possible referral destinations and treatments, depending on the severity of the patient's condition. This proliferation of channels means that GPs may find it difficult to keep abreast of the different options and may also not know which service is most appropriate for the patient's clinical situation. RMCs may also offer a 'single

point of entry' for the different referral pathways and thus seek to take over decision-making from the GP about the severity of the patient's condition and who is the most appropriate professional to see that person. For example, GP referrals for mental health services are often channelled through a 'single point of access' service, which triages referrals and directs them to the most appropriate service for the person's needs. They may also offer 'direct access' to some services without having to see a consultant in secondary care. For example, in many musculoskeletal care services, in order to obtain a referral to a physiotherapist, it was necessary to be initially assessed by an orthopaedic surgeon. However, in many places across the country, services have been reorganised so that the physiotherapist is now the first person to see the patient. They treat those patients within their specialist remit and refer those that require surgery on to the orthopaedic surgeon. Although all these services operate with slightly different desired outcomes in mind (improved access, reduced waiting times), a key underlying aim is to ensure patients see the right person, at the right time, in the right place.

Accurate data on the prevalence of RMCs are difficult to obtain. In 2007, Coleman and colleagues conducted a survey of 152 PCTs to explore developments in practice-based commissioning (PBC), of whom 101 (66%) responded.¹⁴⁹ Of these, 26 PCTs (36%) had put in place new referral management arrangements and 31 (43%) were developing them. Of the 26 who had referral management procedures in place, only three were using units formally recognised as RMCs. Imison and Naylor's⁹³ analysis of Choose and Book data from 2009 (which is not representative of all referrals as a significant percentage of GPs still use paper referrals) suggests that 91% of PCTs have triage or assessment services that divert referrals that historically would have gone to secondary care.

Programme theories: how are referral management centres intended to work?

The perceived root of the demand management problem is the generic tendency in the existing patient pathway for unnecessarily large groups of patients to be referred directly from primary to secondary care (*Figure 17*). RMCs and triage seek to modify the patient pathway by acting as a block both to prevent inappropriate referrals and to triage referrals and sort them into different case-mix categories (e.g. 'less complex' to 'intermediate' to 'complex' patients). The underlying programme theory is that this enables patients to be redirected to the most appropriate services with the necessary skill set to meet the patients' needs, such as those run by GPWSEs or other Allied Health Professionals. Consequently, this means that patients are diverted away from consultants in secondary care. The desired outcome, as depicted in *Figure 18*, is that some patients (e.g. less complex) can be referred back to primary care for the GP to manage, the 'intermediate' patients can be directed to GPWSE or other Allied Health Professionals in alternative community services and the more complex patients are referred on to the consultant. Referrals for procedures on the 'effective use of resources list' are blocked. These are procedures that are deemed to be cost-ineffective, according to systematic reviews of the evidence, and which many CCGs have decided that they will not fund.

Note further that this model assumes a deficiency in the existing system. Triage in its various forms already occurs informally and unevenly in the interface between primary and secondary care. RMCs are perceived as improving rather than introducing the triage process. They are perceived as an attempt to ensure that triage happens, as GPs or consultants are assumed not to be undertaking this process effectively (as depicted in *Figure 17*). They are also understood to have an educational function, being seen as a means of providing GPs with feedback on their referrals, with the idea that GPs will gradually modify their referral behaviour. Some direct testimony on these various claims has been reported in *Chapter 3*.

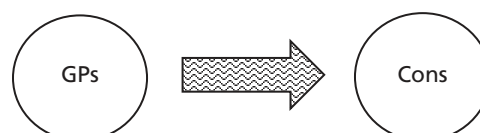


FIGURE 17 Existing elementary pathway and patient volumes. Cons, consultants.

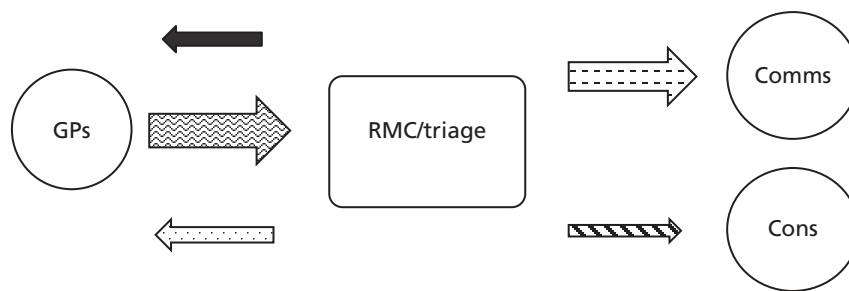


FIGURE 18 Revised 'ideal' pathway. The size of the arrows indicates approximate volumes; the pattern indicates case complexity. Wavy line, existing patient volumes; spots, less complex patients who can be redirected to the GP; horizontal stripes, intermediate patients who can be redirected to community services; diagonal stripes, complex patients for the consultant; solid black, referrals on the 'effective use of resources list'/require more tests or information. Comms, community services including GPwSIs; cons, consultants; RMC, referral management centre (and synonyms).

Figure 18 and the accompanying key describe the desired outcomes of RMCs in more detail. If this redistribution of case mix works, the ideal is that fewer patients see the consultant and, consequently, referrals to secondary care go down, without increasing the overall volume of referrals. This also may reduce waiting times to see the consultant. Those who may otherwise have been referred to secondary care are sent to alternative community services. Through having some referrals rejected and receiving feedback, GPs assimilate what constitutes an 'appropriate referral' and modify their referral behaviour in the future.

As is self-evident in Figure 18, these are ambitious reforms requiring a high degree of co-ordination of a range of different practitioners at the interface with different institutions. With ambition comes contention and realist synthesis seeks to examine how and to what extent plans come to fruition. One way of achieving this aim is to consider the programme theory alongside the critical ideas that have been raised in the face of such reforms. The empirical literature can then be scrutinised in an attempt to adjudicate between the contending claims.

One critical voice argues that RMCs can be seen as part of the infrastructure through which attempts are made to impose evidence-based practice on clinicians. Harrison and Wood coined the term 'scientific bureaucratic medicine' to describe the policies, institutions and organisations responsible for driving forward the implementation (or imposition) of evidence-based practice.¹⁵⁰ They argue that evidence-based practice is a means of managing demand in the NHS through limiting the availability of interventions to those which, according to 'scientific' evidence from RCTs, are cost-effective. Thus, 'scientific' knowledge about the effectiveness of interventions is summarised and codified into protocols and guidelines (the 'bureaucratic element') which guide the user on particular forms of action. In RMCs, decisions about the appropriateness or otherwise of a referral are commonly made through the application of clinical guidelines or referral criteria to referral letters. RMCs also act as a means of preventing referrals for procedures that have been shown, through evidence-based research, to be of 'limited clinical value' and have thus been placed on the 'effective use of resources list'. These are procedures that PCTs and later CCGs have decided they will not fund because the evidence suggests that they do not provide sufficient value for money.

The co-option of clinicians to work within RMCs could constitute a strengthening of the power of the medical profession, as they may impose their values on how decision-making about referrals is made and how guidelines are applied. However, it may also represent a weakening of medical power if these clinicians take on the values of NHS managers to effectively police the activity of rank and file of GPs. Which of these outcomes emerge is likely to depend on how GPs, consultants and GPwSIs are involved in the ways in which the RMC is set up and the control they have over the means through which decision-making process about referrals are made. It is on this matter we need the decisive evidence.

Figure 18 also represents a considerable degree of 'boundary work' on the part of different professional groups. By dint of their specialised knowledge and control over accreditation, professional groups gain power that allows them to define turfs, control jurisdictions and exclude outside groups from participating in a rigidly defined range of specialised activities. From this perspective professions can be seen 'as convex bodies with secure heartlands deep behind the boundary territories'.^{151,152} It follows that understanding a particular activity, such as the right to make referrals, is first of all matter of seeking out the controlling power base. But these controls and their boundaries are not cast in stone. As Abbott puts it, 'Professions only come into existence when social actors contest and create boundaries. A profession is a set of turf battles that are later yoked into a single defensible position in the system of professions'.¹⁵² This alerts us to the need for our review to locate empirical inquiry that has examined boundary walls and boundary work.

In this light, GPs have expressed a number of objections about RMCs. We focus on two key criticisms here. First, GPs express concerns regarding the clinical governance of RMCs, specifically whether or not non-medical personnel have the skills and expertise to make triage decisions about patients. RMCs require GPs to refer patients to them in order to divert patients to alternative services. If GPs hold concerns about the triage process, they may not refer patients to RMCs and find ways of avoiding or circumnavigating them. Second, if GPs do refer to an RMC, they may not take the time to consider the options for the best management options for the patient. Over time, this may lead to a deskilling of GPs as they lose experience in selecting referral routes for certain groups of patients. A further possibility, especially where alternative community services are set up to manage patients, is that GPs may offload unwanted patients to these services. Hughes notes how professional groups with a high standing retain the desirable and interesting work and delegate the more unpleasant, uninteresting or 'dirty work' to those with lower standing.¹⁵³ This may particularly apply if RMCs or triage are seen as a route to offloading less complex or less interesting work to other Allied Health Professionals with less power or lower standing than GPs.

One final observation to note in *Figure 18* is the assumption that both RMCs and triage services have the option of redirecting patients to alternative community services. It goes without saying that this would not be possible if such alternative services did not exist. The devil is clearly in the detail. In some areas, alternative services, including those run by GPwSIs or Allied Health Professionals, have been set up alongside and in conjunction with RMCs. In other areas, GPwSIs and independent clinical assessment and treatment centres (ICATs, which are run by the private sector) pre-existed the RMC, with the RMC being set up in order to make better use of these services. The precise pre-existing configuration of the local health economy is, thus, likely to influence the outcomes of any attempts to redirect patients within it.

We now move to review the evidence on RMCs and triage to test this bundle of competing theories. The synthesis is structured in recognition of three core aims and their potential intended and unintended outcomes.

Evidence review

Most of the evidence evaluating the fortunes of RMCs in reducing and/or redirecting referrals away from secondary care comes from relatively small 'case studies'. Although there is considerable discussion in academic journals, much of the research literature emanates from 'professional publications' and the 'grey literature'. One reason for this particular profile is that RMCs and triage have never carried the status of a 'formal' programme initiated and backed up by a succession of central policy edicts. Instead they represent ad-hoc, if much borrowed, responses to a more general policy imperative to reduce the number of referrals from primary to secondary care. As such, they have been devised, implemented and evaluated separately and locally, with relatively few attempts to 'compare and contrast'.

Such an evidence base meets with some consternation in the methodological literature. For obvious reasons, this is a RCT-free zone, making a necessity of the idea of secondary analysis of mixed-methods primary data. More problematic is the usage of case studies that verge on promotional material. Especially if the evaluation comes from in-house, there is a well-recognised tendency for the production of 'good news' stories. Our response, following Pawson,¹⁵⁴ is led by the idea that there can be useful evidence nuggets in studies that are otherwise indifferent (or worse). Accordingly, we do not simply import the material from these case studies: we interpret it. In the material below, the approach is most clearly evident in the discussion of the Manchester RMC in we combine the insight from several different sources that trace its history.

Intended outcomes: patients are redirected to alternative services, referrals decrease, general practitioners change their referral behaviour

White¹⁵⁵

This *Pulse* article from 8 October 2012 described the construction of a referral management service (RMS) in Kernow, Cornwall. This was developed by a team of local GPs who promoted the idea to fellow local GPs and secured the buy-in of all 58 practices. A small GP community interest company was set up, with each practice holding a single £1 share. The RMS was funded and 'owned' by the PCT but the clinical direction was set by the clinical interest company (CIC) board consisting of five locality GP leads, a LMC representative and the PCT's director of commissioning.

In terms of its operation, call handlers checked whether or not the referrals complied with local guidelines, and for those that did they booked referrals and offered patients a choice of location. Where referrals fell outside local guidance, were on the 'limited clinical benefit' list, might require upgrading to the 2-week rule or had information missing, they were returned to the GP.

To support the functioning of the RMS, referral guidelines were jointly written between the CIC board and local secondary care consultants. It was 'hoped that this would give GPs confidence to try to manage the patient themselves or undertake investigations before referring the patients'. They report that consultants were initially concerned that the RMS would 'restrict their workflow' but the CIC explained that they were trying to 'free up their time so they could do their job properly and help them reposition their services in the face of competition from other providers'.¹⁵⁵

The article indicated that before the introduction of the RMS, referrals were 'rising by 5% per year'; in the year following the introduction of the RMS, the authors report referrals 'have declined by about 5%'. The authors attribute this to 'GPs looking more closely in-house at what needs to be referred . . . not the RMS rejecting hundreds of referrals'. The report also claims that:

Consultants [indicate] referral quality is . . . higher, patients are arriving with the right test already done and conversion rates . . . are higher. For one . . . surgeon the pay off was immediate. Usually 14% of patients would be booked straight to operation – in her first fully sifted clinic, bookings rose to about 80%.

White¹⁵⁵

This RMS was owned by a consortium of local clinicians and the PCT. In terms of the theory under consideration, this joint ownership appears guidelines on which referral decisions are made by both administrative staff and GP sifters were jointly agreed by GPs and consultants as the RMS was being set up. This configuration is said to have resulted in a significant degree of clinical engagement with the RMS. It is also said to be the mechanism that leads to the reported reduction of referrals (though the reviewer, of course, has no grounds on which to check their reliability and validity).

It is also interesting to note a further mechanism through which stakeholders perceived the RMS to work – not purely through the imposition of guidelines to ‘check’ that correct decisions are made or block referrals but instead via education and feedback to GPs so that they reflected and changed their practice. This simple testimony provides a thin ‘first coating’ of evidence that clinical engagement, gained through giving clinicians significant powers to control how RMCs function, is contributing to the successful implementation of RMCs.

*Imison and Naylor*⁹³

Some of this text is reproduced from The King’s Fund from Imison C, Naylor C. *Referral Management: Lessons For Success*. London: The King’s Fund; 2010.⁹³

This study consisted of a literature review, an examination of routinely collected data on referrals, qualitative interviews with stakeholders in 21 PCTs and in-depth study of four case studies using different referral management approaches.

Two case studies from this report highlight the importance of clinical engagement but also clinical governance arrangements as being central to the success of RMCs. The report achieves a rare contrast between two different RMCs, one led by GPs (site A) and one led by the PCT (site B). Site A had ‘very robust governance’ and respondents raised ‘few concerns about clinical risk’. The RMC itself was a not-for-profit organisation owned by local practices and employees. In this case, the remit of the GP triagers was to ascertain whether or not the referral was urgent and needed upgrading under the 2-week wait rule, and also to check the referrals for completeness and appropriateness. Missing information was sought from the practice and referrals could be returned to the referring practice if they did not contain sufficient information or if they did not fit local or national guidance pathways. The clinical leaders felt that ‘risks of triaging patient referrals in the absence of full information were managed by the use of clinical pathways, knowledge gained through experience of being a triager, the availability of peer support and the availability of a robust tracking system for referrals’. Both triage GPs and local GPs highlighted a key mechanism through which they felt triaged worked: that knowing a peer would review your referral motivated GPs to improve the quality of their referral letter:

I think instantly they get a lot better straight away . . . because you really have the feeling that one of your peers is going to be reading the letter you write

Triage GP, site A

Personally, I think very clearly now about what I am referring for . . . what background information does the triager need to know?

Local GP, site A

Respondents also reported that in order to work effectively, clinical engagement with RMCs was essential. The working hypothesis here is that where clinicians have control over the design and governance arrangements for a RMC, they are more likely to cede any control over referral decisions to the RMC, as they can trust that the RMC is operating safely. This case study site saw a number of improvements in the quality of GP referral letters over time: the recording of previous medical history increase from 61% to 81%, the recording of medication history increased from 65% to 75%, the recording of blood pressure increased from 36% to 54% and the recording of body mass index increased from 32% to 48%. However, this site did not see a reduction in referral volumes.

In contrast, the RMC in site B was led by the PCT, and Imison and Naylor’s respondents raised many concerns about clinical governance, particularly about the conduct of clinical triage in the absence of the patient and all the necessary clinical information. In this site, referrals were reviewed by a GP assessor who

either rejected or accepted the referrals. Rejected referrals were returned to practice with an explanation of the reasons why. Accepted referrals were sent to the appointments booking team. However, one of the GP triagers in this site noted the variability in practice between the triagers and questioned the ability of some of the triagers to do their job:

Some assessors were very keen on rejections and some did very few. Some did so few one had to question their ability to do the job . . . it gets back to individuals making decisions on other people's decisions

GP triager, case study B

The booking team then attempted to contact patients (up to three times) to arrange an appointment; if the patient could not be contacted by telephone, a letter was sent to the patient asking them to book an appointment. If the patient did not contact the RMC within a further week, the referral was removed from the system. However, a PCT employee in this site commented that between 30–35% of patients were not able to be contacted by telephone owing to the high turnover of patients moving house, resulting in addresses being out of date. However, in this site, GPs still retained clinical responsibility for the referral until it reached its destination, leaving GPs in a potential risky situation where they had no means of keeping track of the referral. One of the GP triagers commented:

I think it works best when it is clinically-led. I think when it is introduced by management with the sole purpose of trying to reduce referral numbers it doesn't work. It has to be done as a process that is going to improve patient quality and outcome

GP triager, case study site B

This lends some support to the theory under consideration, that clinical engagement and leadership is important to secure the safety and thus success of RMCs. The authors report that this site 'continues to experience high rates of referral and ongoing growth in the number of referrals despite the introduction of the centre'.

Once again the anecdotal nature of this evidence, as with much case study material, can be called into question. But, once again, it provides some useful building blocks about the consequence of different governance arrangements.

Offredy and colleagues¹⁵⁶

This case study provided an insightful description of the process of the development of a CAS in Harrow. Such studies are useful in review work because they follow the history of the service and they track lessons learned – including, in this instance, a description of a pilot study which was not successful. Our interest is in the nature of the adaptations and improvements made rather than in the declarations of 'success' by the local PCT.

The initial driver for the service was the long waiting times for neurology appointments and analysis of the case mix of patients indicated that many could have been treated more quickly and appropriately in primary care by GPs with expertise in headache. A 'Service Redesign and Implementation Team' (SRIT) was set up, which consisted of PCT staff and clinicians to oversee the process of service redesign. The authors describe it is 'fortunate' that there was a GP in Harrow with expertise in headache who was willing to develop a community service. Initial discussions were held with the relevant GP and local hospital consultants. In addition, the PCT made available greater access to CT scanning for GPs in order to reduce waiting times for this procedure (a form of direct access to testing discussed in detail in *Chapter 7*). The team developed detailed referral protocols (a form of guidelines discussed in detail in *Chapter 8*). These covered (1) sequencing to ensure that GPs performed CT investigations prior to referral, (2) identification to distinguish patients GPs could manage themselves and those who needed to be referred to secondary care and (3) guidance to GPs in how to manage patients prior to referral.

However, the resulting headache service was located in the local hospital under the day-to-day management of a consultant neurologist (rather than in the community, as originally envisaged). A GPwSI with an interest in headache worked alongside the consultant and the consultant triaged referrals that could be seen by the GPwSI. This arrangement was piloted for 3 months. The authors conclude that ‘the result of the decision to locate the service in the local hospital was an under-utilisation of the GPwSI skills as only a limited number of cases were being referred to him’ and the PCT perceived that the service was ‘not cost effective and was still based in the hospital as opposed to being in the community’.

Here we see unfolding a rather typical example of organisation change (recall *Figure 11*). Roles were changed (the attempted substitution of GPwSIs for consultants and allowing GPs direct access to tests) and procedural change was instigated (expecting GPs to follow guidelines and referral protocols). There was also availability in the local infrastructure to support the service – a GPwSI with an interest in headache was available to support the service. However, the consultant conducted the triage and, thus, remained in control of decisions about which patients could be seen by the GPwSI. Consequently, the expected change in referral volumes depicted in *Figure 18* in this chapter failed to materialise.

The SRIT remained keen to find another way of improving services. They undertook an exercise to identify clinical expertise to support community services alongside a process to identify which specialties would benefit from community services. They identified cardiology and dermatology as services that had a high local incidence and long waiting lists. In addition, owing to advances in technology it was possible to carry out some cardiology investigations in the community. The SRIT decided that RMCs should be introduced only ‘where skills were available to ensure a practitioner with specialist skills undertook triage of referrals and there was an appropriate alternative to a hospital appointment.’ They undertook an intensive, five-step process to design and implement the service:

1. Process mapping – to map the patient journey and identify bottlenecks and unnecessary journeys.
2. Involvement of key stakeholders – patient pathways were redeveloped based on the evidence-based literature and with the involvement of key stakeholders including GPs, nurses, secondary care clinicians, support staff, PCT staff and patients. Personal visits were made to all GP practices.
3. Redesign of the service so that GPwSIs now undertook triage in a community setting and were supported by consultant mentorship. This was enabled by the fact that in Harrow there are ‘high quality GPs, many of whom have additional qualifications in the relevant specialties on offer in the redesigned services’. They further add later in the paper that a range of other professionals are also involved in triaging including ophthalmologists and specialist nurses.
4. Introducing the service – publicised the service to patients and the public through printing leaflets about the CAS in different languages that were distributed to health-care facilities in Harrow, it was advertised in local newspapers, in patient forum newsletters and on the PCTs website.
5. Implementation – monitored impact and extended the number of specialties involved; they extended it to other services only where they could appropriately divert patients away from secondary care.

The authors also indicate that there were also ‘disincentives for general practices to send less than 90% of their referrals through the CAS’. In their conclusions the authors proclaim that ‘the key to its success has been the whole-system approach to its implementation, in that stakeholders have been involved with the service redesign work’.

This paper is essentially a ‘good-news story’ and, as noted, reviewers need to approach such material with caution. The report omits, for instance, any insight into the views of consultants or GPs on the success or otherwise of CAS. Nevertheless, the paper has face validity; it carries explanations for both failure and success. It supports and is supported by the theory emerging in *Chapter 3* that in order to be successful, organisational change needs to be harmonised across the strategic, administrative, role, procedural and motivational levels. Any one of these phases that is out of equilibrium may cause the intended reforms to fail. In this case the control of the triage arrangements was the sticking point. Left to themselves, existing power dynamics may be reinstated. The solution here rests on an evolving compromise. Although the power to triage patients eventually resided with GPwSIs within CAS, GPs and consultants were engaged in

developing guidelines to determine the basis on which this triage took place. It seems that achieving the 'right' or 'optimal' balance of power between different clinical and managerial interests that determines the success of a RMC.

*Moss and colleagues*¹⁵⁷

This is a small case study describing the implementation of a RMS in Abertawe Bro Morgannwg University Health Board (ABMU). The health board provides services for the residents of Swansea, Neath Port Talbot and Bridgend, but also purchases services from other NHS Wales providers, including Cardiff and Vale and Hywel Dda health boards. ABMU set up the service in order to 'actively manage the referral of residents to ABMU-provided services and away from traditional referral routes into neighbouring health boards'.

The article reports that the service was set up with the BMA principles for the operation of RMCs in mind. These principles include:

- The purpose of RMCs should be to improve the patient pathway, not to lengthen it.
- Active consultation with clinical colleagues in primary and secondary care to agree the process through which the RMC would operate.
- Supporting referrals to specific consultants where clinically indicated.
- Review of referrals by clinically qualified colleagues.
- Robust clinical governance procedures.
- A RMS policy that was circulated to all relevant clinical and non-clinical colleagues across ABMU and external health boards.

The authors describe the steps taken to set up the service. The lead planning officers met with GPs to develop criteria to determine which referrals should be sent to the RMS and which would be excluded, and these were then endorsed by the ABMU clinical directors and the LMC. The lead planning officers worked with ABMU clinical directors to agree the process through which referrals would be reviewed and a lead clinician was identified in each acute directorate of ABMU to take this forward. GPs and ABMU agreed the time scales under which referrals would be reviewed and the mechanism through which feedback to the referrer would be given. Patient information leaflets were developed in partnership with a patient experience involvement group and were circulated to the ABMU GP practices. Meetings were held with external health boards to inform them about the RMS and the operating procedures for the RMS were circulated to each GP practice and ABMU clinical directors.

The RMS became operational on 24 November 2010, and between this date and 31 March 2012 the RMS received 502 referrals for ABMU residents, which were directed to Cardiff and Vale or Hywel Dda health boards, and of which 327 were authorised and 175 were redirected. The authors estimated that redirecting these referrals saved the health board £250,000 through either full or marginal cost recovery. The number of referrals for ABMU residents directed to Cardiff and Vale Health Board reduced by 19% from 2010–11 to 2011–12 and the corresponding reduction of referrals directed to Hywel Dda over the same period was 49%. At the same time, there was an increase in GP referrals for ABMU residents for ABMU services, which the authors interpret as evidence that feedback from the RMS the increased awareness of local GPs about the availability of local service options within ABMU.

This is a further small-scale before-and-after study and caution must be exercised in interpreting and generalising the results. The net impact of the RMS on the overall volume of referrals is not clear. Nevertheless, the RMS appears to have been successful in redirecting referrals and in raising the awareness of local GPs about the availability of local services. It once again lends further support to the theory that clinical engagement and leadership in determining the operational logistics of RMCs are important to the success of RMCs. It also supports the hypothesis that feedback about the availability of local services provided by RMCs to GPs can lead to GPs changing their referral behaviour to make more use of local services.

Interim summary

These few case studies describing ‘successful’ RMCs highlight a number of contextual factors that appear to be important to their success. Clinical leadership and engagement in the development of RMCs appear to be important in order to ensure robust governance procedures to oversee the process through which referrals are reviewed in the absence of full clinical information. The use of experienced clinicians to do this appears to be important, not only in conducting the triage, but also in providing feedback to GPs. The peer feedback function of RMCs was seen in a number of these case studies to be a key mechanism through which GPs changed their referral behaviour, rather than the blocking of referrals.

Unintended outcome 1: short-circuit

General practitioners short-circuit the triage/referral management centre and refer straight to secondary care

Much is to be learnt about the importance of clinical engagement in the success of RMCs by exploring what happens when it is not there. We found a number of examples in which GPs did not engage with the triage/RMC process at all and simply referred directly the consultant, thus short-circuiting the RMC (Figure 19).

CRG Research/Cardiff University¹⁵⁸

This report provides a number of case studies of RMCs, one of which briefly describes the RMC set up in Hounslow. The focus of the Hounslow RMC was to reduce the quantity of referrals to secondary care through triaging referrals against agreed referral guidelines and returning to practices those referrals that do not meet the guidelines. The RMC intended to enlist GPwSIs and physiotherapists to undertake the triaging. However, the PCT changed their mind about this practice because GP practices had expressed concern about the use of triaging by GPwSIs and physiotherapists and, consequently, were refusing to use the RMC. In addition, a local shortage of GPwSIs to undertake the triaging in the local health economy brought further challenges to the successful implementation of the RMC. In this set of circumstances GPs appear to choose not to refer to the RMC, resulting in short-circuiting rather than a remodelling of support and substitution.

The authors of the report conclude that when referral management is understood as resting wholly within the remit of professional management, there is a substantial risk that usual practice will find a way to bypass the procedures introduced. However, when clinicians are involved in the decision-making to set up a RMC and see referral management as a process that can support or improve their decisions, simplify the pathway and make the patient’s journey more efficient, referral management will be supported by clinicians. This is another small-scale study with few reliable outcome data. It does reveal another small nugget of evidence adding further support to the developing theory, namely that if clinicians are expected

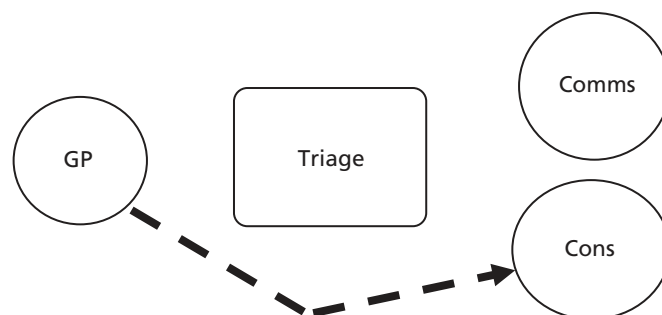


FIGURE 19 Short-circuit. Comms, community services including GPwSIs; cons, consultants.

to behave and take on the values of managers and have no power to shape the nature of RMCs, they will find ways of circumnavigating them. RMCs under bureaucratic control will trigger short circuit rather than support or substitution. If RMCs are shaped in a way that represents professional values and agreed outcomes, clinicians will engage with them.

Papers on the Manchester Referral Management Centre

- CRG Research/Cardiff University, Evaluation of Referral Management Pilots in Wales, Final Report.¹⁵⁸
- NHS Manchester, Agenda item 15, paper Number L, Individual funding requests (Non-commissioned Procedures) and Referral Gateway Initiative, NHS Manchester Board Meeting, 7 July 2010.¹⁵⁹
- Introducing a GP-led referral gateway, *Pulse*, 18 May 2011.¹⁶⁰
- Gateways using nurses to screen referrals, *Pulse*, 10 August 2011.¹⁶¹
- GPs axe remote referral gateway run by private firm, *Pulse*, 2 March 2012.¹⁶²
- Referrals gateway worked well, *Pulse*, 20 March 2012.¹⁶³
- Creating ownership of a referral gateway, *Pulse*, 3 September 2012.¹¹⁵
- Harmoni website NHS Manchester Referral Management Centre.¹⁶⁴

Further evidence about the importance of clinical engagement of local clinicians to the success of RMCs also comes from the evolution of the RMC set up in Manchester. Evidence about the development of the RMC in Manchester has been collected from a number of sources including the above report from NLIH on referral management pilots, a number of *Pulse* articles, websites from the company Harmoni and NHS Manchester and papers from NHS Manchester Board meeting from July 2010 (listed above). The result is a useful example of a blow-by-blow history of the service, which traces the rise and fall of different approaches to referral management in one location. Such historical detail is rarely reported in journal papers, which largely focus on the immediate impact of programmes, rather than providing longitudinal and contextual detail on their implementation. With its emphasis on the fickle fortunes of programme theories, realist synthesis provides a useful means of corraling together different accounts of the same programme. The customary caveat must also be added. This small section follows 'the story' of only one RMC, and caution must be exercised in generalising these findings.

The NLIH report indicates that in 2006, the initial focus on referral management in Manchester was referral tracking and information gathering about the volume of referrals to inform planning to meet capacity by the PCT.¹⁵⁸ This took place with 11 RMCs, which covered 14 PCTs. No referral diversion was taking place, although some RMCs had also introduced referral letter templates in order to gain rapid access to chest-pain clinics. These did not relate to or require compliance with clinical criteria or guidelines.

The report goes on to describes that, outside the daily remit of RMCs, a number of 'Tier 2' services had developed whereby GPWSIs assessed some categories of referrals, such as those in areas of low urgency referral impact [ear, nose and throat (ENT) and dermatology] to determine whether or not they could be managed by an alternative provider, such as a GPWSI in a relevant field. GPs, however, did not receive feedback on these referrals. In addition, a number of ICATs had been developed to provide a diagnosis in advance of attendance in outpatients, where, normally, diagnosis would occur within secondary care followed by a further appointment for treatment. The ICATs were run by a multidisciplinary team who confirmed a diagnosis and then directed patients to a range of providers including those in the private sector. Here we see how services within a health economy are often set up in a patchwork manner and there may not be co-ordination in how the services are intended to complement each other.

Two *Pulse* articles^{115,160} and information from the Harmoni website and NHS Manchester's website then chart the further history of RMC activities in Manchester. By 2009, the 14 PCTs had merged into three PBC consortia. The three PBC consortia in Manchester had a 7-year contract with the ICATs described above, which cost £7M per year, but it is described in the sources as being 'drastically underused'. As this was a block contract (the PCTs had to pay for it whether or not it was used), the PBC consortia were effectively wasting capital while hospital outpatient referrals were rising unchecked. The PBC consortia's goal was to

'reduce hospital outpatient activity levels by using alternative services'¹⁶⁰ – attempting to shift demand away from the overused hospitals towards the underused ICATs. The commissioners initially tried to address this via a peer review of referral scheme by feeding back hospital referral data to GP practices but 'while we identified learning opportunities, it seemed to do little to actually change referral behaviour' and 'generally outpatient referral activity hadn't reduced'.¹¹⁵ Lack of significant change here is attributed to the data being 2–3 months out of date.

Accordingly, the three PBC consortia tried a different tack. Local clinicians in South Manchester developed a referral gateway, which was piloted in four practices in South Manchester. Two senior GPs triaged referrals and a central booking team managed data flow and the Choose and Book function. Initial outcome data 'showed great potential to affect outpatient activity'.¹⁶⁰ The implication of these reports are that the initial success of the pilot lay in the fact that local GPs, administrative staff and GPwSIs were involved in setting up and running the RMC. All three PBC consortia then agreed to develop a city-wide referral gateway. Each PBC consortium voted to use part of the local incentive scheme funding to pay for the gateway.¹¹⁵

The three PBC consortia then met to agree how the referral gateway would work – these included how to give advice back to GPs, governance arrangements and who would triage referrals – Allied Health Professionals and consultants were considered. The PBC consortia took their plans to the LMC (a body that represents the views of GPs to the PBC consortium/PCT) who rejected them the first time. The LMC was 'not convinced that the scheme would cut referrals massively or save any money. It also felt very strongly that this could not be mandated and GPs should be free to make direct referrals if they chose to do so'.¹¹⁵ All three boards wanted local Manchester GPs to provide triage support in order to gain the good will and engagement of practices at the initial stage.¹⁶⁰ Owing to 'problems with internal staffing', the PBC consortia enlisted a private company Harmoni to run the gateway. The clinical triage work was subcontracted to the local out-of-hours provider, 'with the proviso that Manchester specialists did the triaging'.¹⁶⁰

Wootton and Whiting¹¹⁵ explain in their article how the referral gateway operated. GP referrals were initially checked by the RMC/triage team for completeness and to determine that the referral was not for a 'non-commissioned' procedure; that is, procedures that were not funded by NHS Manchester unless clinically appropriate. These included procedures such as alternative or complementary medicines, cosmetic surgery not related to breast reconstruction or trauma and tattoo removal. If data were missing or the referral was for a non-commissioned procedure, an electronic advice note was returned to the practice. The referral then underwent clinical triage by a team of GP triagers who determined whether the referral could continue, be diverted to an alternative service or returned to the GP with advice on management and further tests. A paper regarding the referral gateway tabled an NHS Manchester Board meeting on 7 July 2010¹⁵⁹ expressed the hope that feedback from the gateway would enable GPs to understand which procedures were on the non-commissioned list and, therefore, would reduce their referrals for these procedures in the future which, in turn, would result in a decline in the number of referrals returned to GPs by the referral gateway.

The referral gateway cost £220,000 to set up.¹⁶⁴ In terms of referral outcomes, it was a success. After 6 months, there was 'a 1.2% overall reduction in outpatient activity instead of the forecast 3.8% growth'.¹⁶⁰ After 18 months there was a '12–15% reduction in outpatient referrals and an increase in ICAT usage from 42% to 79%'.¹¹⁵ Wootton and Whiting also report that during the 18-month pilot period, the gateway produced and estimated cost saving of £4–6M.¹¹⁵ The Harmoni website reports that the RMC returned 7% of referrals to GPs.¹⁶⁴

However, other articles reveal that the RMC underwent further changes in its configuration, despite its apparent success in curbing referrals. A *Pulse* article on 2 March 2012 reported that the Manchester CCGs had opted to bring the referral gateway 'in-house' and had not renewed their contract with the private provider Harmoni.¹⁶² Two main reasons cited are costs and that the referral gateway did not have the required local knowledge of Manchester services.¹⁶² Two main reasons cited are costs and that the referral gateway did not have the required local knowledge of Manchester services.¹⁶² The secretary of the Manchester LMC and board member of the CCG said outsourcing was expensive and some minor issues occurred owing to lack of local scheme knowledge and tier two services. Most local GPs are happy back in-house with local people involved.¹⁶² A NHS Manchester spokesperson was cited as saying the decision was based on cost and also a need to employ local people and help the local economy.¹⁶² However, to counteract this implied claim that part of the reason why the CCG did not renew its contract with Harmoni was because it did not employ local people, yet another *Pulse* article asserted that the location of the referral gateway in Southampton had 'no impact' on its functioning.¹⁶³

Different accounts of this period in the development of the Manchester referral gateway make different claims about who undertook the triage and where this took place. It is worth dwelling on the ambiguity here as it relates to general concerns expressed by GPs about the way RMCs are run. These were well rehearsed in *Chapter 3*, but to reiterate, GPs have argued that 'For an anonymous nurse working to a set protocol in a remote RMC to reject that carefully considered decision out of hand is professionally frustrating, to say the least' and that 'Nurses assess patients according to rigid criteria and do not have the experience to make flexible decisions in the same way a doctor can'.¹¹⁴ We can surmise that the growing complexity of the system, led to concerns about who exactly was undertaking the triage. This might have influenced GPs' support of the gateway (and perhaps their willingness to refer to it) and might have influenced the decision to bring the gateway back in-house.

For example, the paper prepared for the NHS Manchester board meeting on 7 July 2010¹⁵⁹ indicates that although the initial checking of procedures would be undertaken by Harmoni, based in Southampton, the clinical triage would be carried out remotely by experience local GPwSIs working in defined areas. Indeed, a *Pulse* article from 2012 declaring the success of the Manchester RMC argued that 'One very important factor in setting up the gateway was the use of local GPwSI and senior, well-known GPs. It meant the other GPs respected the decisions made and the feedback they were receiving'.¹¹⁵ Although this might have been the aim, an article in *Pulse* based on a Freedom of Information request discovered that 'two podiatric nurses and a non-medical prescriber with a diploma in podiatric medicine' were used to triage GP referrals for vascular surgeons.¹⁶¹

This case study illustrates how local approaches to referral management change are often in partial flux. To summarise, we see that policy-makers adopt new solutions when initial approaches fail (e.g. peer review) and that they attempt to roll out successful initiatives (the pilot RMC set up by clinicians) on a larger scale. However, in doing so, changes and adaptations to the original approaches are made; in this case a lack of local staffing infrastructure meant that the service was no longer run by local staff, and so the private company Harmoni, based in Southampton, ran the referral gateway. The local involvement of GPs, GPwSIs and administrative staff had contributed to the success of the initial pilot in terms of both clinical engagement and familiarity with the local availability of services. Scaling up meant that these two essential features were lost. The LMC also saw the RMC as an attempt to curb the clinical autonomy of GPs and wished to defend their right to refer directly to secondary care. Although the scheme successfully reduced referrals, after 18 months the configuration of the gateway was changed again to bring the gateway back 'in-house.' According to some reports, this was to tackle the problem that those undertaking the triage lacked local knowledge about available services.

The above exposition probably reflects the limit of secondary analysis in piecing together a reliable account the history of one RMC. For instance, one member of our advisory group with knowledge of this particular RMC also suggested, contrary to some of the published accounts, that a further reason why the RMC had to be reconfigured was a lack of engagement by local GPs because they did not know the GPs who were

conducting the triage and questioned their capability for doing so. The pocket history, however, does provide value in pointing to the systemic character of organisational change. It illustrates something we characterised earlier as the ‘punch-bag’ problem of demand management: solving one problem at one level has a tendency to create an unintended outcome at another.

Box 30 provides an interim summary.

BOX 30 Interim summary

Comparing the case studies reported to this stage reveals some incertitude about outcomes but also something of the pattern to confusion. In order to be ‘successful’, RMCs require collaboration between NHS managers and clinicians to agree on the purpose of the RMC and its remit and functioning. In the positive accounts described above, clinicians were able to exert considerable control over the development of the guidelines that determined how referrals were dealt with and the governance structures for ensuring this was done safely. In terms of the theory under consideration, RMCs were successful because this level of clinical engagement allowed clinicians to maintain their power over how the RMC is run. A key element of this was the use of experience GPs to conduct the triage. In turn, the perception that the RMC is being managed according to clinical standards of governance leads to more widespread support and trust in the RMC on the part of local GPs and consultants. Finally, the mechanism through which the RMC was deemed to work was not simply through the blocking or redirection of referrals but through GPs receiving feedback about their referrals and learning about what constitutes an appropriate and the availability of local services. This led to GPs reflecting on their own practice changing their own referral behaviour. The role of peers in delivering this feedback was seen as crucial to its success.

Where RMCs were not ‘successful’, there is evidence that this was due to them being perceived as attempts by the NHS managers to curb clinical autonomy. In terms of the underlying theory on scientific bureaucratic management,¹⁵⁰ a RMC can be viewed, and was viewed in some quarters here, as an attempt to enforce managerial values on the medical profession. This suggests that if clinicians are not able to exert significant control over the remit and functioning of RMCs, they will not garner the trust of local clinicians. A further key issue in this case study appeared to be whether or not local GPs were involved in the triage process. This may have been important in ensuring that, and fostering trust that, the triage process was being conducted safely. Further, RMCs were not successful where it was perceived that a single interest dominated the triage decision-making – recall the CAS in Harrow, where the initial pilot study was not successful because the consultant maintained control over the triage process and few patients were referred on to the GPwSI.

Unintended outcome 2: triage/referral management centres increase referral volumes

A second pattern discernible in another series of case study reports is that RMCs/triage may have no effect on, or even increase rather than decrease, referral volumes. A function of RMCs and triage is to divert referrals to alternative community services, such as GPwSIs or those run by Allied Health Professionals. However, these services may simply provide additional capacity, rather than reducing referrals.

The lack of change or increase in the volume of referrals can be a consequence of patients who would not normally have been referred finding their way into the system, as GPs pass on 'dirty work' to other professionals. Alternatively, it may simply be a result of previously unmet demand that can now be addressed as access to the service has been improved. The result is that there is no change to the case mix of patients reaching secondary care. This is depicted in *Figure 20*, in which the total number of referrals entering the system increases, as does the number reaching alternative community services. The total number of referrals to consultants may decrease. A number of studies in our review show how this pattern of outcomes may arise. Here, we review a number of studies that explore this outcome in more detail.

Cox and colleagues¹⁶⁵

This study addressed the overall question of whether or not RMCs can curb the rise in outpatient attendance attributable to GP referrals. It is a retrospective time series analysis of the directly aged standardised monthly rate of first outpatient attendances (as a proxy for referral rates) of 85 practices with 714,000 patients in north Norfolk between 2006 and June 2011. These 85 practices formed five groups in order to manage referrals: three groups set up RMCs, while two groups engaged in peer review of referrals. The attendance rate ratio was calculated as a secondary outcome by dividing each group's attendance rate by the equivalent England attendance rate for every month.

The study found no statistically significant changes in the rate of outpatient attendances or the attendance rate ratio for two of the groups using RMCs and an increase in both measures for one group using RMCs, which also happened to be the group with the largest budget. As this is an observational study, it is not possible to know whether or not outpatient attendance rates would have increased in the absence of the RMCs. Follow-up periods for the initiatives varied between 9 months and 21 months; for some of the initiatives, it is arguably too early to capture their impact. 'First outpatient attendance rates' may also constitute an imperfect measure as they can also be influenced by consultant to consultant and tertiary referrals, which were not subject the influence of RMCs. Nevertheless, this is the only peer-reviewed published study of the impact of RMCs on outpatient attendance rates. In contrast to the single case studies cited earlier, which claimed a reduction in referral rates, this study shows a much less positive and more mixed picture. Within these practices in these time intervals, RMCs, at best, have no impact on referral rates and, at worst, may increase them.

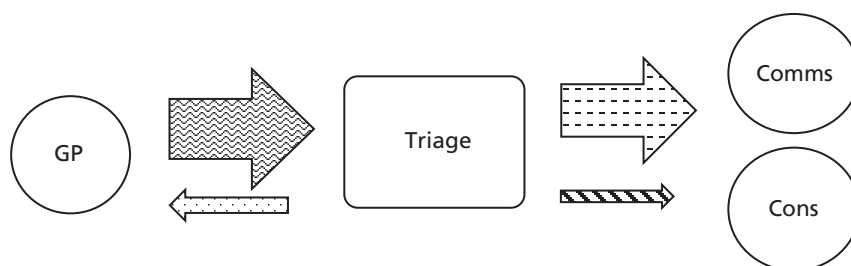


FIGURE 20 Increase in overall volume of referrals. The size of the arrows indicates approximate volumes; the pattern indicates case complexity. Wavy line: existing patient volumes; spots: less complex patients who can be redirected to the GP; horizontal stripes: intermediate patients who can be redirected to community services; diagonal stripes: complex patients for the consultant. Comms, community services including GPwSIs; cons, consultants.

Imison and Naylor⁹³

Some of this text is reproduced from The King's Fund from Imison C, Naylor C. *Referral Management: Lessons For Success*. London: The King's Fund; 2010.⁹³

Here, we draw on the authors' quantitative analysis of referral rates across PCTs and interviews with key respondents to give some insight into the extent to which RMCs in particular had achieved their outcomes and also some possible unintended outcomes of these and triage services. They compared changes in referral rates and referral volumes using Choose and Book data from 2005 to 2009 between the 22 PCTs with the highest proportion of referral activity going through some form of referral management activity (at least 30% of Choose and Book referrals in these practices went through triage or to a GPwSI) and PCTs nationally. This analysis indicated that there were no significant differences between the 22 PCTs undertaking a substantial amount of referral management activity and PCTs nationally. The authors report that 'PCTs that had introduced full referral management centres were picked out, but were no more or less successful than others'.

They also used Hospital Episode Statistics data to analyse percentage change in volume of first attendances between 2005 and 2009 and found 'no discernible difference between the pattern of activity for PCTs with active referral management and those without (*Figure 21*).

They also examined the distribution of GP practices by referral rate in 2008–9 and found no significant difference in the distribution of referral rates at practice level between PCTs undertaking referral management and all practices (*Figure 22*).

Qualitative data from interviews with stakeholders from PCTs indicated that participants felt that referral management schemes generally had been effective in diverting activity into alternative out-of-hospital services but were far less effective in reducing the overall demand and improving the overall quality of the referral process. Some reported that reducing waiting times and increased access to services was encouraging GPs to refer more and that referral to triage services might mean GPs were less likely and less needful of 'thinking for themselves' and more likely to refer onwards than to manage patients themselves. There was also a belief that new out-of-hospital triage services had increased awareness of specific conditions and, thus, identified unmet needs in the patient population (a phenomenon discussed earlier as supplier-induced demand). For those PCTs where referral management schemes were not mandatory, respondents believed that the 'worst' or 'highest referring' GPs could refer outside the schemes and directly to the trust, evidence of the unintended consequence discussed in the previous section (see *Unintended outcome 1: short-circuit*).

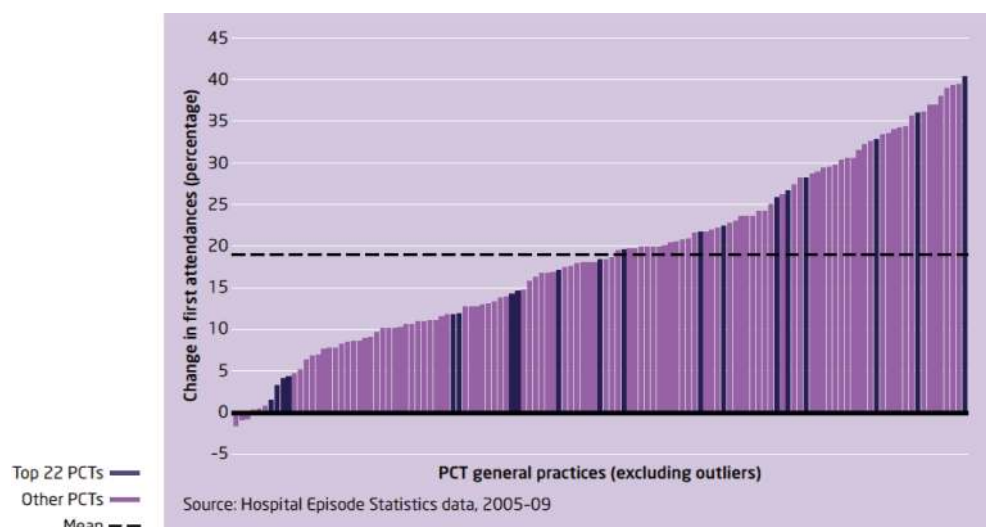


FIGURE 21 Percentage change in volume of first attendances referred by GPs, all PCTs (excluding outliers) 2005–9. Reproduced with permission from The King's Fund from Imison C, Naylor C. *Referral management: lessons for success*. London: The King's Fund; 2010.⁹³ Hospital Episode Statistics data copyright © 2015. Reused with the permission of the Health and Social Care Information Centre. All rights reserved.

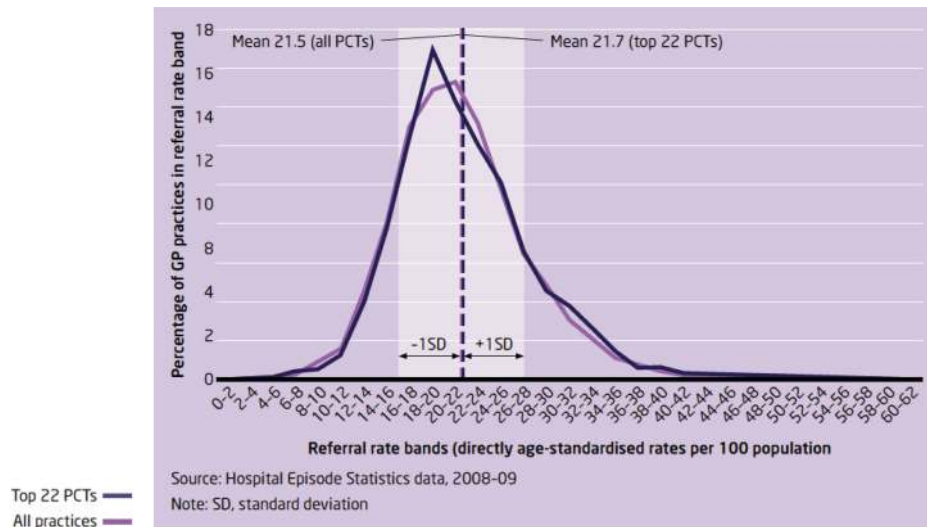


FIGURE 22 Distribution of GP practices by referral rate 2008–9, top 22 PCTs undertaking significant amount of referral management and against all practices. Reproduced with permission from The King's Fund from Imison C, Naylor C. Referral management: lessons for success. London: The King's Fund; 2010.⁹³ Hospital Episode Statistics data copyright © 2015. Reused with the permission of the Health and Social Care Information Centre. All rights reserved.

Bungard and colleagues¹⁶⁶

This was a before-and-after evaluation of a 'triage service' for patients referred to cardiology. Properly speaking, it lies at the boundary of this chapter and *Chapter 7*, being a mixture of triage and direct access to tests. Patients were referred to a triage team consisting of a cardiologist, a nurse practitioner and a doctoral-trained pharmacist (Dpharm). When patients were referred, the EASE (Ensuring Access and Speedy Evaluation) team followed an algorithm that used the patient's pre-consultation symptoms to determine which tests should be ordered and carried out in advance of the consultation, all according to the Guidelines of American College of Cardiologist. The nurse practitioner used information from the patients' pre-consultation symptoms and any test results available in order to triage patients into four categories: (1) to be seen by the emergency department, (2) urgent – seen by EASE within 1 week, (3) stable – seen by EASE within 4–6 weeks and (4) elective – referred back to GP. When patients were seen by the EASE team, it was hoped that all relevant tests would have been already carried out by the nurse practitioner and the doctoral-trained pharmacist, who collated this information for the consultant so that the consultant could focus their history-taking and examination on the most pertinent aspects of care, giving more time to discuss diagnosis and treatment plans with the patient. The nurse practitioner and the doctoral-trained pharmacist then co-ordinated any follow-up procedures or tests. The intervention ran alongside the conventional route for referring directly to a cardiologist, but the team chose to compare speed of access to a consultation with a historical control of patients referred to cardiology in the year before EASE was implemented. The main outcome measures were time to initial consultation and time to receive diagnosis.

Following the introduction of EASE, the total volume of referrals to cardiology services increased by 170%, from 1512 to 2574. The service did not significantly reduce the number of referrals via the conventional route but simply provided additional capacity. The mean wait time reduced from 71 days prior the introduction of EASE to 33 days for the EASE group. The volume of referrals referred to EASE also increased over time, an increase of 50% in 2005 relative to 2004 and an increase of 19% in 2006 relative to 2005. After being seen by EASE, the vast majority of patients were referred back to the GP (98%); 20% had a follow-up with the cardiologist.

This study suggested that triage and direct access to tests may speed up waiting times but may not reduce the total volume of referrals or reduce the volume of referrals to secondary care. Triage as operationalised here merely provides additional capacity. This study, of course, does not permit us to identify whether or

not that additional capacity was being used appropriately. It does indicate that the new service did not substitute for the old one; it merely provided additional layers of support.

Furthermore, we learn that the additional DMs in the team, nurse practitioners and pharmacists, followed an algorithm and carried out tests to support the work of the consultant. This suggests another conjecture much in evidence in chapters to come, namely the ‘dirty work’ hypothesis, whereby more simple tasks such as ordering and carrying out tests are performed by lower-skilled professionals, while responsibility for the interpretation and decision-making on the basis of those tests are retained by the more highly skilled professional. However, the authors pick up on a rival interpretation, namely that this freed up time for the cardiologist to spend longer with the patient, making more effective use of the cardiologist’s time. Both explanations are consistent with the idea that triage may increase patient volumes.

*Rymaszewski and colleagues*¹⁶⁷

This was a small before-and-after evaluation of the introduction of a nurse-led triage service and team-based approach to the management of patients with musculoskeletal conditions in one hospital in Glasgow. Prior to the introduction of the team approach, patients were referred directly to the orthopaedic surgeon, who would treat those patients requiring surgery but refer those who did not require surgery on to other specialties (e.g. physiotherapy, podiatry or rheumatology). Under the new service reconfiguration, a range of enhanced services were introduced, including specialist physiotherapist and a rheumatology specialist nurse working alongside the consultant rheumatologist to examine and investigate patients. A specialist nurse triaged referrals according to a series of evidence-based protocols and redirected patients to the specialist physiotherapist, the newly configured rheumatology service, podiatry or an orthopaedic surgeon. Waiting times to be seen reduced from 182 days in 1995 to 90 days in 1999. However, the number of referrals rose from 2226 referrals in 1995 to 4391 referrals in 1999. The conversion rate to orthopaedic surgery improved from 20% (556 out of 2779) in 1995 to 38% in 1999 (242 out of 618), suggesting that the triage process was successful in diverting patients away from consultant care and increasing the appropriateness of referrals to consultants.

This study suggested that improving access to services by reducing waiting times results in an increase in the total number of referrals. The triage process successfully diverted patients away from consultants and did improve the appropriateness of referrals for surgery; however, many more patients were seen by other Allied Health Professionals than had previously been the case.

*Maddison and colleagues*¹⁶⁸

This study described a clinical assessment and triage service set up in North West Wales Trust, designed to form a ‘single point of entry’ for musculoskeletal services. Musculoskeletal services were provided by four secondary departments (orthopaedics, rheumatology, pain management and therapy services) each of which had a long waiting time. GPs are reported to have perceived the services to be inadequate for conditions such as back pain and ‘saw little point referring to the service’.

A central triage service was set up to direct patients to the most appropriate service based on information contained in a generic referral letter pro forma. The services were managed by a senior manager and a ‘co-ordinator’ oversaw the referral process and to set up the service on the patient information management system. Alongside the triage system, new services were also set up, including a back-pain pathway led by extended scope GPs, and three community-based musculoskeletal clinics run by GPwSIs and extended scope physiotherapists. Workshops were held with local GPs to gain their input into the development of referral guidelines. The outcomes measured were waiting times, number of referrals and rate of conversion from orthopaedic clinics to operating lists as a measure of appropriateness. These outcomes were assessed on a quarterly basis between October 2001 and July 2003. The TEAMs service was implemented in April 2002 and, thus, the authors compared these outcomes before and after the service was introduced. As ever there are methodological limitations. For instance, no attempt was made to adjust for seasonal or case-mix variation in the volume of referrals.

Following the introduction of the TEAMS, there was an increase of 116% in the total number of referrals for musculoskeletal problems to North West Wales Trust. The number of orthopaedic referrals was 'slightly' reduced (we are not presented actual figures, just a graph). There was a sharp decrease in waiting times for pain management, rheumatology services and back pain from May 2002 to September 2003. There was no decrease in waiting times for orthopaedic referrals to hand clinics and for all other orthopaedic referrals, the decrease took longer to materialise and the drop was not as steep as for other services. The conversion rate for orthopaedic referrals remained unchanged at 37%. The community musculoskeletal clinics were 'well received' by GPs and the short waiting time of 4–6 weeks 'put them in demand'. Patients also liked the service, with 75% of patients being 'completely satisfied' with the service.

Other data reveal that fewer than 10% of patients were referred on or followed up in these community services. 'Referring on' and 'follow-up' are very different indicators of success for the service and, analytically speaking, should not be fused. If a patient is referred on from a community clinic, this suggests that they could not be managed in the clinic and needed to be seen by secondary care, indicating that triage has not placed them in the correct category (they are 'more complex' than the triage suggests). Even assuming that all 10% of referred patients required onward referral suggests that triage was reasonably accurate in classifying these patients as less complex. However, if only 10% or fewer of patients required follow-up, this suggests that the vast majority of patients were seen in one consultation and their problem was easily dealt with in that time. This could indicate that the problem was less complex and one might question if they might have been effectively managed by a GP. We can hypothesise that these are the sorts of patients that GPs may not have previously referred if the waiting list was too long. In their discussion, the authors note that 'a large proportion of the increase burden of referrals has fallen on physiotherapy services. It is apparent, however, that many referrals to physiotherapy are inappropriate'.

The paper reads as, and is perhaps intended as, a 'good-news story' by virtue of the impact of the service on waiting times and patient satisfaction. However, overall, these findings suggest that the triage service had only a minimal effect on diverting orthopaedic referrals away from secondary care. Referrals to orthopaedics reduced only slightly, the conversion rate (an indication of appropriateness) remained unchanged and waiting times took longer to decrease and did not drop to such a large degree as other musculoskeletal services. Instead, there are indications that the new arrangements simply created new, sometimes inappropriate, demand for musculoskeletal services provided by GPwSIs and Allied Health Professionals, mostly physiotherapists. In terms of the theory under test, it resulted in 'substitution' but in an unintended way: the alternative services were taking on the work of GPs *and* the consultants rather than of *just* the consultants. There are aspects of the 'dirty work' hypothesis here, whereby GPs transferred responsibility for the care of patients with minor musculoskeletal problems to lower-skilled professionals. Arguably, in the long run, if GPs do offload dirty work to other DMs, this may result in a deskilling of GPs, as they gain less experience of managing these problems. We found some evidence to suggest that this may occur in a qualitative study of GP views to a service with a similar underlying programme theory of referral redirection: a single-point-of-entry service for mental health, reviewed below.

Raine and colleagues¹⁶⁹

This was a qualitative study of GP views of a single point of access to mental health care. The study consisted of six focus groups of 54 GPs and 15 mental health professionals (MHPs); three groups consisted only of GPs and another group was a mix of GPs and MHPs. These groups were purposively selected from a wider study of 16 focus groups, though the basis of this sampling is not detailed. The focus of the analysis of the six focus groups in this paper is to explore 'the attitudes of English GPs and MHPs to single point of access for mental health referral'. The paper thus reports on only a portion of a larger study and it is possible that the authors did not probe and explore the respondents' views on triage in sufficient detail. Nevertheless, the study provides a useful exploration of the theory that RMCs and triage may deskill GPs.

Raine and colleagues began with the rationale for the introduction of single-point-of-entry/triage services – that of supporting GPs to find their way through the maze of mental health services on offer. GPs and MHPs in this study accepted that the rationale for a single point of access was that GPs could not be

expected to know which specialist therapy would be appropriate. Indeed, some GPs were in positive support of this referral route as they felt that they did lack this relatively esoteric knowledge. The authors reported that 'MHPs liked it because they felt that the choice of therapy demands expert patient assessment' (implying that GPs were not experts).

However, some GPs questioned if MHPs would reach consensus on the best way to manage these types of illnesses – raising questions about whether or not those working in the service would be any better than the GP at making referral decisions:

Would they agree amongst themselves as to the best way of approaching some of these chronic illnesses? I am not sure they would.

The study also provided further insight as to why GPs may be reluctant to refer to the service, because it does not allow them to select which professional they may wish their patient to see. GPs build up knowledge about the custom and practice of different consultants and this allows them to match these practices to the needs of the patient. Not being able to undertake this matching process leads GPs to feel that they are giving patients a poor service:

I would like to make the choice . . . in the best interests for my patient. If I don't know whether they are going to see Dr A or Dr B, then I would feel that I'm giving my patient a worse service.

For the rest of secondary referrals we know the patient will do better with this consultant because he is more direct, with this one because he is much more reflective.

In addition, the anonymous nature of the single point of access made it difficult to explain its purpose to patients.

When the patient asks 'why are you sending me there and what is going to happen to me?' if you have no idea, that could be a bit of a problem

Furthermore, the study identifies two possible mechanisms through which a single-point-of-entry gateway may deskill GPs in making referral decisions. The first was the lack of feedback or advice on patient management following referral to the service and a lack of dialogue between professionals.

You get 'this patient completed a course of assessment'; you don't get their opinion on where to go from there.

Second, referral to a single-point assessment service eliminated the need to choose an appropriate therapy for patients and it was felt that this perpetuated GPs' lack of knowledge about mental health interventions:

The single point of entry encourages you to be quite lazy . . . and not really to apply your mind as to what sort of therapy might be useful.

These two quotations speak directly to the 'dirty work' hypothesis and the deskilling of GPs. Referring to a RMC or single-point-of-entry gateway may deskill GPs as it does not support them to manage the patient themselves; they can instead 'off load' patients somewhere else. Furthermore, without any feedback on how to manage patients, GPs do not learn how to manage patients more effectively.

The issue of the extent to which the feedback provided to GPs by a RMC is successful in changing GP behaviour is explored in a recent study, reviewed next.¹⁷⁰

*Xiang and colleagues*¹⁷⁰

This study focused squarely on one particular aspiration of RMCs, that of improving the quality of referral letters through feedback to GPs. Xiang and colleagues assessed a single RMC set up in 2008 and examined its impact on the quality of GP referral letters. The programme theory is that the RMCs scrutinise referrals letters and return those that contain inadequate information to GPs with feedback on how they can improve. In turn, GPs reflect on this feedback and subsequently improve the quality of referral letters. In this RMC, the triaging was performed by a GP and the feedback was given by the triaging GP to the referring GP over the telephone. However, alongside this, the RMCs remit also involved review of referral letters to assess whether or not they adhere to local management pathways and to direct referrals to an appropriate provider. The 'discussion' section of the paper hints that it was also in the RMC's remit to block referrals. Thus, the triaging GP makes judgements on the referring GP's decision-making, a form of substitution.

This study assessed the quality of the referral letters in a 60-day period 1 month after the RMC became operational and again during a 30-day period 7 months later. In each time period, 300 letters were randomly selected from those reaching the RMC and their quality assessed using two sets of criteria. The process of judging the quality of referral letters highlights the difficulties of defining this construct and the context-dependent nature of referral letters. The first set of criteria were those agreed by the local PBC board to monitor the quality of referral letters; four items were considered to be 'core' including information about the patient's blood pressure, body mass index, past medical history and medication history. However, in applying these criteria, the researchers considered that not all items were 'essential' to the referral. Consequently, they adjusted the way in which quality was judged, allowing the option for the researchers to signal where missing data were 'not relevant' to the referral.

The second set of criteria were devised by Grol and colleagues and consisted of two components: provision of clinical information, judged as including at least four of five pieces of information, and clarity of the reason for referral, judged as including at least one specific request for the specialist.¹⁷¹ However, in applying these criteria, the researchers found that requests were often implied, rather than explicitly stated. Again, the researchers adjusted the ways in which they judged the quality of referral letters to make a distinction between 'implicit' and 'explicit' requests.

Here, we see that even the researchers used a subjective over-ride to make sense of 'objective' indicators of referral quality when interpreting referral letters. This begs the question of whether or not the triaging GPs also experienced difficulties in judging the quality of GP referral letters. The paper provides no information about the ways in which triaging GPs approached this task; however, some findings from the study allude to the difficulties they may have encountered. At time 2, the researchers judged that only 38.9% of referral letters contained all four criteria set by the CCG, with over 70% of letters recording past medical history and medication history. At least 25% of letters were deemed by the researchers to be missing core information relevant to the referral request. Only 43.5% of referral letters met the Grol criteria at time 1 and 39.6% of letters met them at time 2. However, only 7% of letters received peer-to-peer feedback. This may suggest a mismatch between the triaging GPs' judgement of referral letter quality and that of the researchers. It may also suggest that triaging GPs did not always give feedback when referral letter quality was poor; the researchers note anecdotal comments from triage staff that they were reluctant to give feedback when they recognised that the letter was written by a locum, as they knew there was little chance of being able to contact that doctor.

In terms of outcomes, the study found that there was a statistically significant increase in the percentage of referral letters containing information about the patient's past medical history and medication but there was no similar increase in letters containing information on patient blood pressure or body mass index. At baseline, over 70% of letters contained information on the former two criteria, whereas around 50% contained information on the latter, suggesting that the intervention further improved what GPs were already doing well but did not change what they were not doing well. There was no statistically significant difference in the percentage of letters meeting the Grol criteria between baseline and follow-up. Although not a statistically significant difference, 39.6% of referral letters at baseline contained an explicit request for help, compared with only 30.4% at follow-up, suggesting that referral letters became less rather than more explicit following the introduction of the RMC. The researchers interpret this as a possible 'unintended consequence' of the RMC as GPs attempted to minimise the risk that their referral was deflected or refused by the RMC.

The *Box 31* provides an interim summary and *Box 32* provides a summary of the chapter as a whole.

BOX 31 Interim summary

These studies suggest that if additional services are set up alongside referral gateway or triage services, they may improve access to such services by reducing waiting times. However, they may not reduce the volume of referrals to secondary care and may increase the overall volume of referrals in the system. This is because they may be dealing with unmet need; as access to the service becomes easier, GPs may be more likely to refer directly to, for example, physiotherapy, whereas they may have previously told patients there was nothing they could do if waiting times were too long. However, this may also represent a lowering of GP thresholds for referral and GPs passing on 'dirty work' to lower-skilled professionals. There is also some evidence that this may in turn deskill GPs as they reduce their experience of managing these patients and they do not have to think about how this might be done if they can simply pass the patient on elsewhere. The lack of feedback on patient management from triage services may further compound this problem. In this situation, alternative services take on the work of both GPs and consultants.

BOX 32 Chapter summary

Referral management centres, triage and single-point-of-entry services act as intermediaries between primary and secondary care. Their aim is to prevent referrals from reaching secondary care that require further investigations to be done in primary care, that do not meet current guidelines or that are for procedures that are not locally funded due to a lack of evidence on clinical effectiveness. They are also designed to triage patients and redirect them to alternative community services (such as those run by GPwSIs, Allied Health Professionals or the private sector), which better match the patient's needs or have greater capacity. Through providing feedback to GPs on the appropriateness of their referral, the expectation is that GPs will reflect on and modify their referral behaviour. The intended outcome of RMCs, triage and single point of entry is that both overall referral volumes and referrals reaching secondary care will decrease.

Referral management centres are part of the infrastructure that attempts to impose or infuse managerial values on clinicians and the medical profession in particular. The critical literature suggests a number of unintended consequences may occur. The co-option of clinicians into the services may strengthen their power as services are managed according to clinical values. However, attempts to impose managerial values on clinicians may lead to concerns about the clinical governance of RMCs and questions about the skills and experience of those working within RMCs. This, in turn, may result in GPs circumnavigating RMCs. If effective feedback is not given to referring GPs, RMCs and single-point-of-entry services may deskill GPs as they are no longer required to engage in decisions about the most appropriate referral pathway for certain patients. They may also lower GPs' threshold for referral and result in GPs offloading 'dirty work' to lower-status professions.

The quality of the empirical evidence evaluating RMCs, triage and single-point-of-entry services is variable. There are few high-quality empirical studies of RMCs and a small number of case studies. These studies present a mixed pattern of outcomes. They suggest that if clinicians are not involved in the development of RMCs or decisions about the means through which blocking or redirection is achieved, this leads to concerns about clinical governance, resulting in clinicians circumnavigating RMCs. GPs may also short-circuit RMCs if they have concerns about the skills and experience of those conducting the triage. RMC and triage services may have no impact on referrals or may even increase referral volumes to secondary care. This is because alternative community services are often set up in addition to secondary care, making access easier. Consequently, GPs may lower their referral thresholds and refer patients to triage or single-point-of-entry services who previously they may not have referred to secondary care. We also found some evidence to suggest that GPs may offload less complex or less interesting patients to services run by Allied Health Professionals. We also found some evidence to suggest that RMCs and single-point-of-entry services do not always provide effective feedback to GPs. By referring to a single-point-of-entry service or RMC, GPs are no longer required to engage in decisions about the appropriate referral pathway for a patient, resulting in a deskilling of GPs.

Our review identified the resources and ideas need to support the successful implementation of RMCs, triage and single-point-of-entry services. These pointers for success have emerged from a small number of case studies and so caution is needed. However, the case studies highlighted remarkably similar messages about what makes for a successful RMC. RMCs are more likely to be successful if all those whose work is likely to be influenced by the introduction of a RMC (e.g. GPs, consultants and NHS managers, other Allied Health Professionals) are engaged in the set-up of the RMC and, most crucially, in the revision of care pathways and the development of guidelines that determine how referrals are blocked and redirected. Feedback to GPs about why their referrals have been blocked or redirected needs to be done directly to the referring GP in a timely manner. Indeed, GPs reflecting on feedback and changing their referral behaviour was seen as a key mechanism through which RMCs may deliver a change in referral volumes, rather than through blocking referrals. It is essential that the local health economy contains both the necessary skills and expertise and the capacity to manage the diverted referrals. Finally, patients need to be informed about the presence of RMCs and triage and the impact it may have on their care.

Chapter 5 Role change: general practitioners with a special interest – can they control and shape demand?

General practitioners with special interests have already featured as bystanders in the review of RMCs. They move to centre stage in this chapter, which examines their new role in modifying demand for secondary care. Role changes have occurred in every corner of health provision throughout the world and under quite different tags: ‘physician extenders’, ‘triggers’, ‘support assistants’, ‘specialist practitioners’, ‘allied health assistants’ and so on. In the UK, the phrase ‘with special interests’ has often been used to demarcate these emerging roles. A Royal College of General Practitioners (RCGP) 2006 information sheet lists five such roles:¹⁷²

- GPwSIs
- nurses with special interests
- Allied Health Professionals with special interests
- pharmacists with special interests
- practice managers with special interests.

The chapter begins with a brief introduction to policy impetus and programme theories underlying the introduction of the GPwSI role. It then considers some of the counter theories suggesting limited leverage for incumbents of these new positions. These critical hypotheses are already familiar and concern the extent of actual role change (substitution or support) and the possibility of ‘turf war’ protection to maintain existing lines of control. The main body of the chapter is the evidence review, which attempts to adjudicate between these various hypotheses. The review is structured into two parts, beginning with outcome studies and moving to those that have concentrated on the implementation processes associated with the new role. Analysis is presented in an interim summary and a final conclusion.

General practitioners with special interests: policy, promise and pitfalls

The development and formalisation of the innovative role of ‘Specialist GPs’ is usually traced back to the 2000 NHS Plan,¹⁴ which called for a thousand new appointments. The initial designation itself was subject to a degree of inter-professional wrangling and has since been replaced with the more precise ‘general practitioner with a special interest’. The significance of the new role was further underpinned in a parallel Department of Health policy of organisational decentralisation promoted in *Shifting the Balance of Power*,¹⁷³ which recommended increases in provision in the community and the local health-care economy. Two further 2002 documents,^{174,175} produced jointly by the Department of Health and the RCGP, marked the professionalisation of this new function, setting expectations for duties, recruitment, activities, competencies, facilities, accountability and payments. A rather wide range of GPwSI functions was envisaged (*Box 33*), listed in a RCGP information sheet (updated 2006) as follows:

Another crucial facet of the GPwSI initiative is the very wide clinical terrain in which these individuals have been expected to operate. These are listed below in a table reproduced from a ‘snapshot’ in a 2006 RCGP information sheet that lists current and planned services (*Figure 23*).¹⁷⁷ More recent figures on the extent of GPwSI provision do not appear to be available (a request to the RCGP information service on this matter drew a blank).

BOX 33 Royal College of General Practitioners' designation of GPwSI roles¹⁷²

- Specific clinical areas.
- Education (undergraduate teaching, vocational training and postgraduate education).
- Leadership in service provision and representative organisations.
- Research and academic general practice.
- Quality assurance (including examiners, mentors, appraisers and assessors).
- Management (for example in primary care organisations or in deaneries).
- Public health.

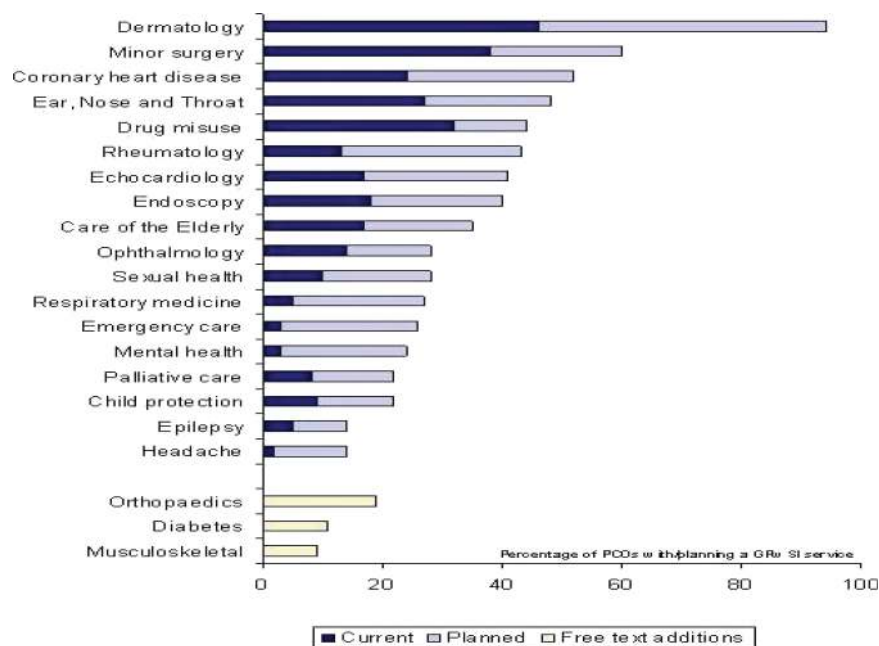


FIGURE 23 Primary care trusts with current and planned GPwSI services (%) by clinical area. Reproduced from Pinnock and colleagues¹⁷⁶ under the Creative Commons Attribution Licence Attribution 2.0.

Another notable variation in the development of GPwSI provision was spatial: where do these practitioners tend to operate? Locality is not insignificant as the ability to professionalise is partly territorial, whether one substitutes or supports depends on whether one's potential partners are at arm's length or cheek by jowl. Rosen and colleagues report on a variety of operational bases: (1) GPwSIs providing services in outpatient departments, (2) those offering supplementary procedure in community settings and (3) those developing services and education in management centres.¹⁷⁷

These initial snapshots of duties, clinical specialisms and locations already carry some status as 'evidence'. These bureaucratic domains have quite different patient pathways, staffing structures, organisational logistics and decision dynamics, and these contingencies alone carry fair warning against the notion that GPwSI interventions could and would be rolled out in similar fashion. One size is very unlikely to fit all. We will return to evidence on the further implementation of the scheme in the ensuing years, and complete these introductory remarks simply by noting key criteria on which the scheme has become evaluated:

'The stated aim of government in formalising GPwSI services was to shift the balance between primary and secondary care for reasons of lowering waiting lists, increasing cost effectiveness and patient accessibility'.¹⁷⁸ These key policy ambitions with regard to demand management are articulated further (especially in points 1, 2 and 3) in the following summary (Box 34), paraphrased from the Department of Health policy document.¹⁷⁹

BOX 34 Anticipated benefits of shifting care closer to home

1. Timely: reducing waits and delays and contributing to the delivery of the 18-week Public Service Agreement target, especially for diagnostics and outpatient care where long waits have an impact on the NHS's ability to meet this target.
2. Efficient: streamlining patient pathways by removing unnecessary steps (e.g. stopping follow-ups when not needed, direct placement on waiting lists, using referral protocols). Improving care through better use of skill mix (e.g. follow-up by community physiotherapy after joint replacement) and preventing the need for more costly interventions.
3. Effective: providing services that are clinically appropriate and based on evidence of proven benefit; refraining from providing services to those not likely to benefit (i.e. avoiding under- and overuse).
4. Equitable: providing consistent care regardless of demographic, socioeconomic or geographical status.
5. Patient-centred: providing care that is respectful of, responsive to and guided by patient preferences, needs and values. Ways in which this can be achieved include improved access, enabling patients to be treated closer to home without having to attend hospital and offering patients a choice of provider which includes the local specialist services from PwSIs.
6. Safe: avoiding harm to patients.

PwSI, practitioner with a special interest.

Reproduced from the Department of Health under the terms of the Open Government Licence v3.0.¹⁷⁹

Programme theories articulate the rationale behind an intervention: the reasons why it *should* work. All of the above claims envisage GPwSIs as having the capacity to redirect patient flows, change the skill mix, moderate service overuse and underuse, and so on. Whether or not these functions come to pass depends on the pre-existing systems into which they are dispatched, and the critical literature also spells out reasons why complex services are reluctant to change. Here we articulate three potential pitfalls to stand beside the policy promise, leaving the main body of the review to investigate how the balance falls in practice.

Substitution or support

Figure 24 presents a highly simplified diagram that begins to explain the intended purpose of role substitution (DM in this and subsequent figures refers to a 'decision-maker'). A new post (or service) DM_{2.5} is introduced, drawing some of the decision-making mandate from the previous incumbents DM₂ and DM₃. The responsibilities associated with the new roles and the qualities of their incumbents are assumed to be the underlying mechanisms that will generate a more efficient harnessing and smoothing of demand. Our investigation can thus be thought of as an exploration of the functions and functioning of 'DM_{2.5}s' in their GPwSI guise.

The figure (upper section) represents the patient pathway and the decision-making responsible for referrals along that chain. DM₂ represents more junior and more generalist DMs; DM₃s are more specialist and

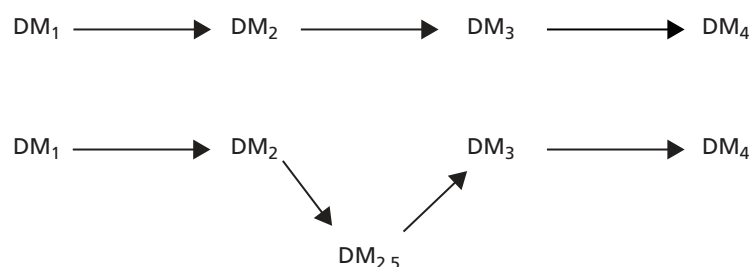


FIGURE 24 Existing system and intended change.

senior. The basic assumption (lower section) is that $DM_{2.5}$ s can take on those tasks that by a process of accretion have fallen on DM_3 s but which do not really require high and specialist levels of expertise. A newly trained and accredited $DM_{2.5}$ can thus take over this portion of decision-making, with benefits ensuing because of task appropriateness, economies of scale, remuneration efficiencies and a consistency of approach. A subsidiary theory, much in play in the introduction of the GPwSI role, is that a $DM_{2.5}$, often with a similar background as a DM_2 , will be in a better position to understand and to translate between the concerns of the generalist and the specialist.

A basic familiarity with the critical literature also provides a pointer to the main area of contestation, namely whether the new roles fetch up in substitution or support? A new role is devised but does it take on substantial decision-making duties or does it remain subordinate to the previous decision-making channels? As in the previous chapter, we will use the ‘substitution or support’ motif in assessing the impact of role change. Dozens of primary studies make reference to this benchmark in relation to quite different substantive domains (e.g. Zwijnenberg and Bours¹⁸⁰) and there is a clear imperative to investigate this distinction in our GPwSI review.

Figure 25 provides a visual reference to this key benchmark, though note that we add a third potential outcome: ‘short-circuit’. Here we refer to another unintended consequence of the new role. Existing referral pathways and systems are rarely singular, especially under the current movements towards ‘informed patient’ and consumer choice policies such as ‘Choose and Book’, which are intended to widen rather than tailor patient pathways. In some cases these alternative routes (both formal and informal) enable DMs to bypass the new role altogether and this further eventuality can also be monitored in our review.

The anatomy of decision-making

Many different modifications are intended to follow with the introduction of GPwSIs, but one key assumption is about the handover of decision-making. The ‘decision-making’ literature is voluminous: conceptual distinctions are drawn between rational and irrational choice, bounded rational choice, interpretative repertoires, tacit and explicit decision-making, evaluation, judgement, discretion, intuition and so on. For a more concrete illustration it is useful to examine the everyday stock of decisions made in the work environment. Thompson and Dowding, for example, produce a 10-fold typology covering the typical array of decision type in the routine duties of nursing.¹⁸¹ Nurses’ decisions involve (1) diagnostic judgements on a patient, (2) gauging whether or not they are stable, (3) evaluating for any deterioration, (4) predicting upcoming problems, (5) choosing between interventions, (6) targeting who will receive the most benefit, (7) choosing the optimal timing of a procedure, (8) determining which patients and their families need support, (9) deciding how to reassure patients and (10) organising handover to other staff.

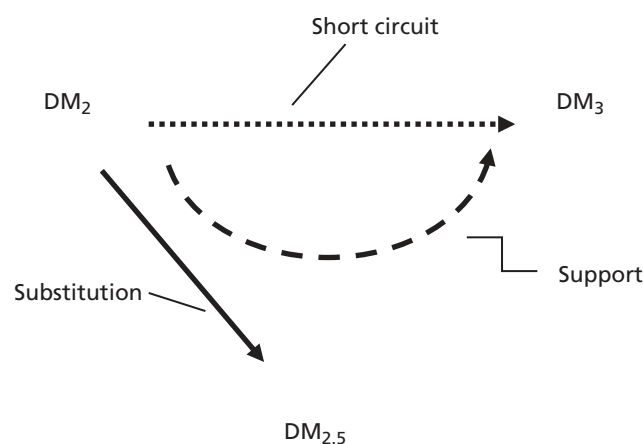


FIGURE 25 Intended and unintended decision-making structures. Substitution: the new $DM_{2.5}$ role is established with real decision-making powers independent from DM_2 and DM_3 . Support: the new $DM_{2.5}$ role is established, which provides intermediary support but remains under the control of DM_2 and DM_3 . Short-circuit: the new $DM_{2.5}$ role short-circuited with DM_2 utilising alternative pathways to refer patients to DM_3 .

All of these distinctions, theoretical and practical, point to the fact that decisions are never singular, and they suggest that decision-making has its own internal architecture. In trying to assess the extent of change introduced with the introduction of GPwSIs it is crucial to understand more precisely what kinds of decisions are changed (or are reluctant to change) as these new roles and regimes are introduced. Clearly, we cannot operate with the close focus of Thompson and Dowding's model, which uses fine-grained and role-specific distinctions. We require a more abstract classification system that establishes some core differences in decision-making responsibilities. *Figure 26* is a simple hierarchical model that we will use in gauging the ambit of role change in the interventions under review. Decision-making ranges from the strategic to the routine. Role-change interventions that achieve 'substitution' involve climbing such a ladder of decision-making, whereas others will end in 'support' and would only cover the lower-order functions. These lower-order functions carry a name in the literature namely the 'dirty-work' hypothesis¹⁸² and a crucial point of investigation in what follows is to examine to what extent GPwSIs have to settle for enacting rather than making decisions.

Professional control and boundary work

Much individual and institutional capital goes into development and maintenance of existing role hierarchies and changing the health-care division of labour will always be politically contested – medicine often being regarded as the archetype of profession power. The introduction of GPwSI services involves adjustments to existing decision-making structures and this in turn invariably involves the attempt to manipulate professional boundaries. Much prior scholarship has gone into what it is that make such boundaries porous or unbridgeable and we need to apply these theories in understanding the elbow-room provided to GPwSIs.

Studies of how professions come into being and are able to exert control over significant national agencies and agendas are one of the mainstays of the social science literature. Here we call mainly on a basic structuration theory, usually traced back to the work of Abbott in two influential contributions.^{153,154} By dint of their specialised knowledge and control over accreditation, professional groups gain power that allows them to define turfs, control jurisdictions and exclude outside groups from participating in a rigidly defined range of specialised activities. From this perspective professions can be seen 'as convex bodies with secure heartlands deep behind the boundary territories'.¹⁵⁴ It follows that understanding a particular activity, such as the right to prescribe medicines, the right to make referrals and so on, is first of all matter of seeking out the controlling power base.

But these controls and their boundaries are not cast in stone. As Abbott puts it, 'Professions only come into existence when social actors contest and create boundaries. A profession is a set of turf battles that are later yoked into a single defensible position in the system of professions'.¹⁵⁴ This alerts us to the need for our review to locate empirical inquiry that has examined boundary work as well as boundary walls. We depict this line of inquiry in *Figure 27*. As is evident in the *Care Closer to Home* 'manifesto', role-change policies are intent on

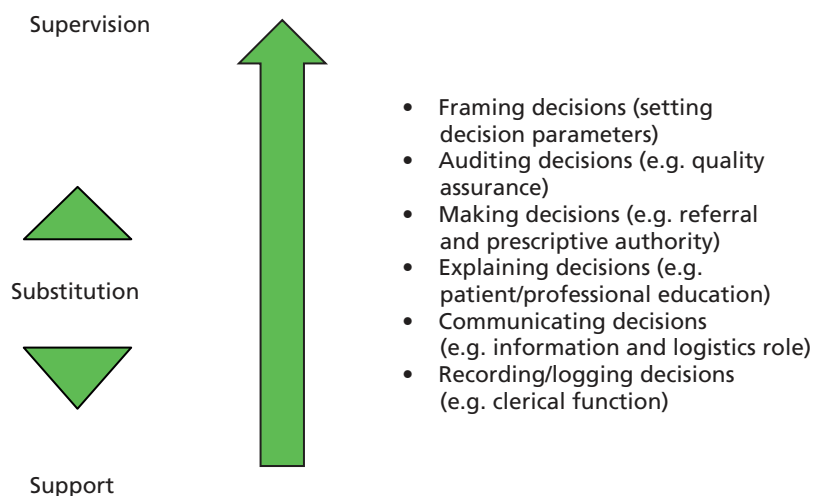


FIGURE 26 The DM_{2.5} decision hierarchy (ladder of decision-making).

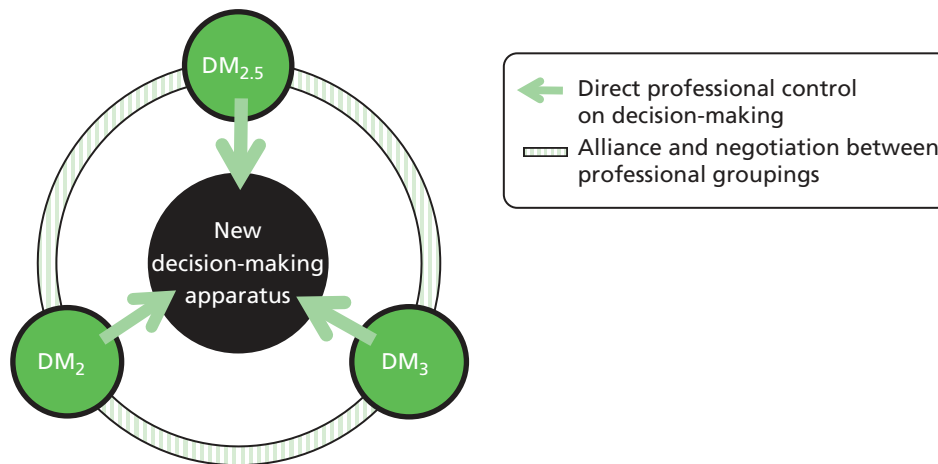


FIGURE 27 The shaping of decision-making across professional boundaries.

creating a new decision-making apparatus (depicted in the centre of the diagram). Whether the new structure materialises, whether there is a real change in decision-making responsibilities, depends in part on the dynamics of professional closure (depicted in the periphery of the figure). The solid arrows represent pre-existing lines of professional control – here we examine the staying power of direct, authoritative and established lines control over a particular decision. The striped connectors represent the boundary work that occurs under ‘reprofessionalisation’ – how are alliances drawn and redrawn with the introduction of new roles? The role of alliance-making, as we shall see, turns out to be crucial in the development of the GPwSI role.

General practitioners with special interests evidence synthesis

We now move to the main body of the GPwSI review, which unfolds as follows. We begin by analysing the main ‘outcome studies’ covering materials on referral volumes and pathways, accessibility, waiting times and patient profiles associated with the introduction of GPwSI services. There have been surprisingly few such investigations, perhaps because the dislocated and amoebic nature of the GPwSI role makes the isolation of cause and effect patterns particularly difficult to pin down. The studies that do exist reveal a distinctly assorted and irregular picture in relation to changes in demand-and-referral regimes.

We then proceed to examine primary materials that may help to explain the mixed messages from the outcome literature. These commence under a subheading ‘process and context studies’. Our strategy here is to use the studies to interrogate the hypothesis just raised on substation or support, on decision dynamics and on professional alliances. We also encounter a substantial amount of evidence on the impact of reforms in wider NHS policy, and system complexity and heterogeneity.

The primary inquiries in both sections are presented and analysed study by study. We make brief references on the rationale for study selection and on their research quality but concentrate on the detailed findings emerging from each piece of research. Each primary study, of course, has its own set of concerns. We extract from each study only those items that contribute to our conceptual framework (ignoring, for instance, material on GPwSI remuneration and field training). Some studies have findings that reach across our analytic framework but, with a few exceptions, we stick to the one-study-at-a-time format, which we consider helpful in building the explanatory narrative.

Outcome studies

There are a number of studies which *estimate* the potential clinical utility of a GPwSI service: what proportion of patients sent to outpatients *could be* handled by a GPwSI? This is a fundamental question about capacity and the largely positive findings from these assessments featured prominently in much of the early-stage policy promotion for such a service.

Gilbert and colleagues¹⁸³

All GP referral letters to the respiratory medicine department of a teaching hospital, apart from urgent cancer referrals, were identified over two 2-week periods. All patient and practice identifications were removed. Two GPs and one consultant respiratory physician assessed each of the anonymised referral letters to determine the patient's suitability to be seen in a GPwSI respiratory clinic, 'assuming such a clinic had a predetermined range of investigative facilities'. The study thus assesses the extent of agreement between DM₂s and DM₃s on referral appropriateness and thus provides a simple test of whether or not there is scope for reform.

Out of 96 referrals covering a wide range of respiratory conditions apart from lung cancer, 22 (23%) were considered by all assessors to be suitable for a GPwSI clinic, and there was full agreement that 40 referrals (42%) were unsuitable. The other 34 referrals (35%) had varying degrees of agreement on suitability. The largest groups of patient referrals considered suitable for a GPwSI clinic were those with chronic obstructive pulmonary disease or cough as the main presenting clinical problem. The commonest groups considered unsuitable were referrals of patients with an abnormal chest radiograph, haemoptysis or possible interstitial lung disease.

The authors reach the following conclusions: 'This small study has shown that at least a fifth of GP referrals to a hospital Respiratory Medicine clinic could be seen in a suitably resourced GPwSI clinic, with consequent reductions in hospital outpatient waiting lists and improved accessibility for patients'. The operative word here, of course, is 'could'. This is a simulation study on the problem rather than the remedy and it requires testing against 'actual' rather than 'assumed' GPwSI provision. In this type of 'paper exercise', GPs' and consultants' views have parity, and we have yet to discover whether or not this holds in their real-world interactions. As noted, however, the study demonstrates the clear potential for reform and, indeed, it was widely cited in the expansion of GPwSI policy along with another inquiry by Sanderson and colleagues.¹⁸⁴ Using a similar method, this latter study of GPwSIs working in ENT found that 30–40% of patients referred to secondary care *could* have been seen by a GPwSI instead.

Rosen and colleagues¹⁷⁷

We turn to studies of 'actual' provision with this study, the first major evaluation of GPwSI services. It remains by far the most comprehensive inquiry, running to 300 pages and examining evidence on activity and referral rates, case mixes, clinical quality and safety, change management, demand management, stakeholder assessments, establishment and running costs and patient opinion. Four sites were studied, three implementing GPwSI dermatology clinics and the fourth providing a musculoskeletal service. Detailed descriptions of the set-up phases are provided and these supply a first illustration of a common feature of GPwSI interventions. Even though the services under research relate to a narrow clinical domain, the schemes on the ground are implemented in a remarkably heterogeneous fashion, with differences in instigation (funding from external *Action On* awards or teaching hospitals or PCTs), in personnel and their experience (GPwSIs recruited from GP principals, locums and hospital clinical assistants), in contracts (full- and part-time, lone work and team work), in catchment (14–62 GP practices), in service location (within outpatients, GP surgeries or bespoke sites) and so on.

This study provides the closest available examination of the changes in referral patterns consequent on the new services. The basis of the design is complex (a stratified, comparative cohort method) but has to be of this level of sophistication to monitor and establish differences in referral activity. In three of the sites different clusters of general practices were identified, some of which could *only* refer patients to hospital outpatients and some of which could refer to *either* the hospital consultants or the GPwSI clinic. This allows for a mapping of patient flows with and without a GPwSI services (cohorts) as well as situations when both services can be chosen (comparative). There is also careful consideration of sample size, time periods and the statistical power to detect significant changes. The subsequent findings are somewhat difficult to summarise; *Box 35* makes an attempt to do so.

BOX 35 Overall activity, referral rates, relative access and waiting times

- **Volume of clinical activity:** following the introduction of the GPwSI services there was no change in overall outpatient hospital activity (number of new patients seen in a given period) in the dermatology sites. However, there was a 12% increase in volumes for musculoskeletal outpatient clinics. Activity levels at all GPwSI clinics increased over time with steadily rising appointments.
- **Referral rates:** the study was unable to detect significant changes in hospital referral rates following the introduction of the GPwSI services in any of the sites studied. Overall (i.e. combined) referral rates from GPs to hospital clinics and to the new GPwSI clinics increased in three out of the four sites.
- **Relative access:** the study compared GP practices with single 'access' (to outpatient clinics) and dual access (to both outpatients and GP). In one site both single access and dual access practices increased their hospital referrals. In another site both types decreased these referrals.
- **Waiting times:** on balance there was no change in waiting times in hospital outpatient clinics (increased in two sites, decreased in the other two). The waiting times in two sites were significantly shorter in GPwSI clinics than for hospital clinics. However, GPwSI waiting times tended to increase over time as the service became recognised and established.

The overall conclusion, reflecting these unanticipated and rather febrile findings, is that 'the association between introduction of GPwSI clinics and hospital referral rates was variable and unpredictable'.¹⁷⁷ It is the sheer inconsistency of outcome patterns that is perhaps the most apparent and revealing finding. The key policy expectation that GPwSIs might lighten the outpatient load does not seem to have materialised. There are also significant indications in the overall activity data that GPwSI schemes sometimes absorb and sometimes generate demand – it appears that GPwSI services can be substitutive, supportive or quite supplemental.

Rosen and colleagues' evaluation also includes considerable and valuable 'explanatory data', which may account for these perverse effects. We return to this material in subsequent sections and move on with the business of reviewing other outcome studies.

Rogers and colleagues¹⁸⁵

This is an evaluation of a multiprofessional triage team (MPTT) in a community-based musculoskeletal clinic. The team in question was led by GPwSIs and supported by physiotherapists. The 'audit' was conducted by a team of orthopaedic surgeons. The conclusion, even in the modulated tone of academic journals, is condemnatory: 'Time delays, patient confusion regarding professional roles and diagnostic indecision are significant problems for patients referred to hospital clinics from MPTT clinics. This risks sub-optimal patient care and may lead to future medico-legal implications'.

Patients judged by their GP to have lower-limb problems requiring further investigation had a choice of whether to be referred to a MPTT clinic or directly to the consultant orthopaedic surgeon in secondary care. A total of 191 patients were referred to the GPwSI-led clinic and 145 patients were referred directly to the orthopaedic surgeon. Of the 191 patients referred to the GPwSI clinic, 132 (69%) were subsequently referred on to the secondary care orthopaedic clinic. Of these 132, 76 patients (56%) were seen by a GPwSI and 56 (44%) were seen by a physiotherapist.

Waiting times differed considerably. The cohort of 132 patients seen in the MPTT clinic waited a mean of 52.6 days following their GP referral and, subsequently, were seen in the secondary clinic 88.4 days later. This was statistically significantly longer than mean of 62.4 days for those patients who were directly referred to orthopaedic clinic. Patients were asked who they had seen in the MPTT clinic and '84% incorrectly identified the healthcare professional they had seen' – 36% thought that they had seen an orthopaedic consultant, 42% thought that they had seen a GP, 20% thought that they had seen a

physiotherapist and 2% thought that they had seen a nurse. Diagnostic accuracy is also audited and reported as follows: of these 132 MPTT patients referred to secondary care, 47% were given a diagnosis that agreed with that made by the orthopaedic surgeon, 31% had no diagnosis and 22% had an 'incorrect diagnosis'.

Several of these inferences are questionable. There is no information on how many of those patients who presented to their GP with lower-limb problems were not referred on at all. The inability of patients to identify the 'rank' of a presenting professional is well known. Much depends on how the 'choice' was presented to in the first place and how the research question on identity was posed. Patients presented with a list of options are more likely to tick a (possibly wrong) answer even if they did not really discover who they saw, whereas patients offered open-ended recall are more likely to leave the question blank if they are unsure (there appear to be no 'don't knows' in these data). The statistically significant findings on length of wait compare a two-stage with a one-step process, with obvious consequences. They ignore the 30% of patients dealt with entirely and slightly more speedily at the MPTT.

This study does provide some tentative evidence of a potential unintended consequence of GPwSIs. Rather than streamlining the patient pathway, this study suggests that GPwSIs may lengthen it if a significant proportion of patients referred to them subsequently require referral to secondary care. In this study, there was no initial triage process and GPs and patients were allowed to choose whether to be referred to a GPwSI or a consultant. It points to the importance of the initial triage process in ensuring that the appropriate 'intermediate' case mix of patients is referred to GPwSI, rather than the more 'complex' patients who require consultant care. The findings on 'incorrect diagnosis' are somewhat tempered in the body of the report, with the acknowledgement that diagnostic accuracy is somewhat subjective and varies from condition to condition. These data, properly presented as 'diagnostic difference', together with the findings on patients who were referred without diagnosis, do signal significant disparity in GPwSI and consultant motivations. We encounter this perspectival difference in many other primary studies still to be reviewed. The key task of the synthesis is to explain it and not, as in this paper, to treat it in a politicised fashion with dire and unproven inferences about patient confusion and medico-legal ramifications.

Data from RCTs are often relied on for providing the soundest indication of outcome differences resulting from the introduction of a new intervention. As we have already noted, the GPwSI services that have been introduced are associated with significant heterogeneity and constant adaptation. Creating the investigative conditions for a clean comparison between a 'GPwSI experimental group' and a 'previous treatment control' is, thus, perilous and we uncovered only one study making the attempt.

Salisbury and colleagues¹⁸⁶

Patients with skin problems were randomly assigned to a new, suburban GPwSI dermatology service or a hospital dermatology outpatient clinic. Only those with non-urgent problems (49%) were included in the experiment (i.e. those falling squarely within what was considered to be the GPwSI remit). Outcome difference were assessed by (1) disease quality of life index, (2) patients perceptions of access to service, (3) patient satisfaction with consultation, (4) preferences for site of care, (5) proportion of failed appointments and (6) waiting times for first appointment. The majority of these data were compiled from a patient questionnaire and the remainder from their records.

No difference was found in terms of the clinical outcomes – though it should be noted that these were based on patient-reported 'quality of life' and 'perceived improvement' rather than on direct clinical measures. Respondents' perceptions of access (location, travel, transport and parking) favoured the GPwSI service (access scores 76.1 vs. 60.5). Respondents reported little difference in terms of satisfaction with the consultation other than viewing the amount of 'time spent' with the GPwSI and 'the facilities' more favourably. Fewer GPwSI patients failed to meet first appointment (6% vs. 11%), although missed appointments equalised in subsequent follow-up. Waiting times for first appointments were significantly shorter for GPwSI patients than for outpatients (72 days vs. 113 days).

The study is considered by the authors to 'provide support for the effectiveness of accessibility and acceptability of general practitioners with special interests'. We would qualify this interpretation by noting that this is very much a patient perception study. Regardless of specific outcome studied, results are heavily weighted towards and investigated through the patient perspective. The overall balance of findings might be attributed to the well-known preference of patients to be treated in community outreach locations and by professionals with whom they are more familiar. Indeed, before the trial '61% expressed a preference to be seen by a GPwSI'.

The other caveat to note is that the referral dynamics under investigation are significantly governed by research decisions on assignment to the two trial arms (outpatients vs. GPwSI). These were set at a 2 : 1 ratio in the expectation that this would ensure that the GPwSI service was working at a 'reasonable level of capacity'. In other words, research design decisions fix referral flows that might or might not apply to the actual proportions handled in any real interplay between the services. Methodological decisions determine accessibility rather than investigate its evolution. It is the volatility of referral flows and the capacity of GPwSI services to sift and handle them that is the real issue and these are better investigated in the to-and-fro of actual service delivery (cf. the Rosen and colleagues' study¹⁷⁷).

We located three studies with a specific concern for the economic impact of the new GPwSI services, from which rather different conclusions emerge: Coast and colleagues,¹⁸⁷ Jones¹⁸⁸ and Ridsdale and colleagues.¹⁸⁹

Coast and colleagues¹⁸⁷ study compares costs in the two arms of a randomised trial comparing patients receiving 'care by a GPwSI service' or by 'usual outpatient care' for non-urgent dermatology treatment (it uses the same service analysed in the Salisbury RCT, described above). 'Costs incurred by the NHS for a general practitioner with special interest service for non-urgent skin problems were about 75% higher than those for care provided in a hospital outpatient clinic.' Although the number of consultations was slightly higher among patients receiving care from a GPwSI, the major contribution to the increased costs in the service was the higher unit costs associated with consultations rather than with hospital specialist consultations. The main reason for these higher costs is that 'the patients attending the general practitioner with special interest service always see the relatively costly general practitioner with special interests, whereas outpatients might see the relatively costly consultant but may also see one of the consultant team (registrar or clinical assistant) who is less costly'.¹⁸⁷

Jones's¹⁸⁸ study examined prescribing costs for respiratory drugs in practices with GPwSIs in asthma compared with average costs for their Family Health Services Authority. The GPwSIs appear eager to climb the ladder of decision-making responsibilities illustrated in *Figure 26*. Perhaps unsurprisingly, given their special interest, the average costs for their prescriptions were significantly higher than the Family Health Services Authority norm (an increase of £195 per 100 patients per quarter).¹⁸⁸

These two studies give a significant indication, in their respective and highly specific domains, that the GPwSI services may prove more costly than 'usual provision'. The typical 'limitations' of economic evaluation apply and are acknowledged, most particularly on the matter of what to include in the baselines and whether or not this enables a fair comparison. From the point of view of this study the main limitation is that the GPwSI services are simply regarded as alternative forms of 'treatment' for the same condition. This interpretation, which is a function of cost comparison methodology, perhaps misses the essential point of the GPwSI role, which is to reform referral pathways making them more appropriate for *different* sets of patients. It also fails to take into account the full range of activities of GPwSIs [see the RCGP 'job description' in *Box 33* and the Wilson and Cox study,¹⁹⁰ in *Process and context (explanatory) studies*, in which an 'activity breakdown' indicates the direct treatment function is a relatively small aspect of the new role]. The 'take-home' message from these early studies is thus tightly bounded and foreseeable, namely that the establishment of a new service will attract increased professional changes and prescription costs, especially at 'set-up'.

The Ridsdale and colleagues' study might be considered a fairer and certainly a fuller assessment of the cost of a new GPwSI service.¹⁸⁹ Estimated cost of running a GPwSI headache service are compiled, covering sessional, management, administrative, reception, overhead and contingency costs, as well as charges for conducting and receiving supervision. Estimates are derived from the hourly costs of each activity and the number of sessions delivered. 'The cost per hour was . . . £272. Assuming 30 minutes for a first appointment and 15 minutes for a follow up, [costs were] £136 and £68, respectively . . . the cost of neurology appointments nationally (derived from 2006/2007) HNS references costs are £203 for first appointments and £121 for follow ups'.¹⁸⁹ On this basis the authors conclude that 'GPwSI services can satisfy the needs of patients with similar headache impact at cost that are lower than those for secondary care'.¹⁸⁹

Quite typically in economic evaluation, different conclusions are drawn according to the items included in the baseline. Ridsdale and colleagues' study is outwardly more comprehensive in this respect but the great absence here is any detail on the crucial comparative baseline measure used in the 'NHS reference costs'.¹⁸⁹ It is not possible to discern if these are exactly the same as those used in the assessment the GPwSI service. The eternal 'like with like' problem is further exacerbated had other potential costs on estates, training, patient expenditures, onward treatment and external disinvestments had been included in the analysis (see Kernick¹⁹¹) for a consideration of the multitudinous options). We conclude that the economic analysis of GPwSI services is somewhat inconclusive and this assessment is likely to continue to apply given the heterogeneity, complexity and adaptability of these services.

In investigating outcomes, one of our intentions was to locate studies that quantified the different decision patterns made by GPwSI (DM_{2.5S}) as opposed to GPs (DM_{2S}) and consultants (DM_{3S}), with aim of mapping any differences in terms of the hierarchical model in *Figure 26*. We uncovered several studies that describe the various decision repertoires (which we will review later) but noted a lack of appetite to pin numbers on the patterns. The following study provides an exception to this rule.

Elliot and Kernick¹⁹²

Fourteen GPs with a special interest in headache recorded consecutive consultations over a 4-month period, logging whether or not they had referred patients for radiological tests and their motivations for doing so. The data reveal that 'headache GPwSI' use these investigations at a level (30.16%) that is above those reported for GPs who have direct access to neuroradiological services (between 1.3% and 5.3%) and below those reported for neurologists (up to 60%). Other significant findings are (1) the main reason for investigation reported by the headache GPwSI was for 'patient reassurance', a reason which falls outside the British Association for the Study of Headache (BASH) Guidelines and (2) the 'yields of significant findings' in these tests are low but nevertheless broadly similar across all three groupings.

The paper thus provides another useful indication that GPwSIs are prepared to climb the ladder of decision-making responsibilities (see *Figure 26*) and it is this very opportunity that excites the main concern of the authors in their opprobrium for the 'increasing and unnecessary pressure on neuroradiological services caused by the widening of access to tests'. We explore the issue of patient reassurance and GP direct access to tests in more detail in *Chapter 6*; in terms of our analysis on role change and the respective positioning of DM₂, DM_{2.5} and DM₃, the example provides a compelling example of role strain. The GPwSIs clearly have a considerable appetite in calling for tests and are able to do so with the same clinical significance as senior groups. But they also appear to persist with the traditional family GPs interest in bolstering the patient, while also being met by admonition from clinicians who consider this motivation to reassure 'difficult to justify on clinical grounds, particularly against a background of limited healthcare resources and a very low rate of significant finding where there is no clinical suspicion'.¹⁹²

Although it is somewhat beyond the purview of our theoretical framework, there is another phalanx of outcome data that warrants a brief mention at this point. One outcome that is quantifiable and indeed much quantified in respect of GPwSI services is the associated degree of 'patient satisfaction'.

Many studies, including ones reviewed to this point, have monitored, usually through surveys and questionnaires, levels of patient approval of all aspects of GPwSI provision.

To make the point we briefly reprise three further findings from studies already discussed. Ridsdale's comparison of GPwSI and consultant headache services utilised a 19-item questionnaire on rate of satisfaction with services (everything from 'ease of access' to 'would you go back'). On every item the GPwSI service was preferred.¹⁸⁹ Such consistency of opinion might be considered a 'halo effect', the crucial point, however, being about the unequivocal placement of the corona rather than its size. The evidence on patient perspectives on GPwSI service from Salisbury's study has already quoted in terms of heightened satisfaction on 'access' and 'short waiting times'.¹⁸⁶

Once again the Rosen study provides the most rigorous evidence.¹⁷⁷ It uses the same comparison of patients attending GPwSI and consultant services. The samples are larger than the other surveys (1082 GPwSI vs. 1111 consultant appointments, with a response rate around the norm at approximately 50%). There was no difference in reported clinical 'quality of life' between groups and no difference in 'confidence in the doctor' but on most of the accessibility measures ('duration of wait', 'reasonableness of wait', 'parking', 'clinic hard to find', 'adequacy of time spent') the GPwSI services are favoured.

Unlike any of the other outcome measures under review, there is relatively little equivocation to be discovered here – patients appreciate ingress to and approachability of the new providers. In a direct sense this evidence is somewhat distant from our concern about whether or not GPwSI services impact of referral rates. Indirectly, however, it may have some explanatory value. From our examination of the 'harder data' on patient flows (referral patterns, costs and waiting times), it is apparent that GPwSI services have both absorbed and generated demand. One potential explanation for the latter lurks under the notion of 'supply induced demand'. That thesis is at its most powerful if a newly introduced service is accessible and familiar to patients – and that proposition is clearly exemplified in the reported satisfaction data. This particular pocket of evidence thus begins to construct an explanation for one aspect of the changes in referral rates, a conjecture that we will investigate with other types of primary materials gathered in the next sections.

Box 36 provides an interim summary.

BOX 36 Interim summary

In the above section we have attempted to summarise those studies aspiring to paint a quantitative picture on the changes in activity levels and referral patterns following the introduction of GPwSI services. There is little solid evidence that GPwSIs have reduced the outpatient workload or outpatient waiting lists. There are a few more positive but decidedly more anecdotal claims on the former. There is no evidence to suggest that clinical outcomes differ between GPwSIs and outpatients, although there is a difference in their case mix. There is solid evidence to suggest that GPwSI services attract increasing referrals and offer shorter initial waiting times (which may lengthen with the longevity of the service). There are mixed and indecisive data on cost savings. There is good evidence of GPwSIs' willingness to enlarge their decision-making roles. There is some initial evidence suggesting the phenomenon of GPwSI 'supply induced demand'. Although this summary is drawn from the best available evidence, a more general caveat might be applied to each and every proposition. There is huge disparity in the way that GPwSI services and GPwSI roles have been implemented; no study could possibly cover every permutation, rendering these results time and place specific. It does not follow, however, that the findings are mere happenstance because, as we are about to see, it is perfectly possible to develop an explanatory account.

Process and context (explanatory) studies

In this section we seek out studies that will help explain the lacklustre results of the arrival of GPwSIs on activity levels and referral rates. We analyse them according to the conceptual framework described in the chapter introduction, thus uncovering rich empirical detail on whether or not GPwSI interventions involved new decision mandates and role change, and how these in turn are influenced by professional boundary rivalries, wider policy changes and whole-system complexity. Again and perforce, primary studies tend not to fall neatly into these compartments, but, once again, we continue with the study-by-study presentation, with the occasional need to reprise the more comprehensive inquiries.

We begin with papers on boundary work: the question of how pre-existing inter- and intraprofessional relationships impinge on the development of the GPwSI role. There are several studies of this ilk, most notably a body of work carried out by Currie and colleagues²⁸ on the introduction in 2006–7 of pilot GPwSI services in genetics. We report on two characteristic outputs from this work and reprise another section from the Rosen report on the same topic before moving through the remainder of our analytic framework.

Currie and colleagues²⁸

Currie and colleagues'²⁸ qualitative study of the development of genetics services run by GPwSIs provides an in-depth test of the boundary theory. The research focus is on the relationships (sometimes smooth and sometimes choppy) that occur between GPs, GPwSIs and consultants. Case studies of four different pilot sites were undertaken based on thematic, semistructured interviews of each stakeholder group, numbering 36 in all. Further emergent themes were revealed in the analysis of transcripts and the conclusions were supported by a respondent validation exercise.

The remit of these pilot sites followed our standard 'programme theory'. The GPwSI role was intended to reconfigure the standard health service division of labour. This new special interest group was expected to lighten the load on the consultant genetics service by taking some referrals from their fellow GPs and offering some diagnostic and treatment services. It was also expected that they would provide leadership and education in the reshaping of primary care services in a clinical area in which GPs traditionally lacked experience.

A notable feature of these particular schemes was that the planning was 'funding initiated'. GPwSIs had responsibility for designing the schemes – following the receipt of bespoke funding and sometimes without full consultation with all parties. Accordingly, they were implemented with four subtly different visions and Currie's study has the great opportunity of following which perspective came to pass. Across the four sites, GPwSI services that gained ready acceptance were those designed to either complement or supplement the work of consultant geneticists. Those that were set up to substitute the work of the geneticists encountered greatest opposition from the consultants. New functions such as 'raising awareness' in an 'educational role' prompted support, as did providing 'psychological support' to patients undergoing genetic testing. The case in which the GPwSI sought to develop a new genetics clinic within primary care did not. Clinicians continued to hold sufficient power to constrain the development of GPwSI services by rather classic means – by defining what constituted specialist knowledge and controlling access to training, education and support required to set up the new GPwSI services. GPwSIs always had to engage in alliance building, establishing and re-establishing consensus with the specialist regarding the purpose of their services.

In terms of *Figures 25 and 27*, DM₃ (the consultant) retained control over the sorts of decisions and tasks that DM_{2,5} (the GPwSI) was able to undertake. In terms of *Figure 24* the GPwSI (DM_{2,5}) was largely limited to explaining decisions (i.e. professional education and patient support). Where GPwSIs did undertake decision-making, namely in the management of low-risk patients, the consultant determined which patients they saw. Calling on their professional expertise, geneticists retained control over the guidelines for deciding patient allocation and the development of protocols about matching patient complexity with available resources. Demand and duplication dilemmas continued and were sometimes exacerbated in new

forms of informal control. In the most contested case, the ‘geneticists took on the role of appraiser, vetted all referrals to the GPwSI and filled a supervisory role’.

Interestingly, and perhaps surprisingly in this particular study, the ‘effectiveness of GPwSIs in genetics appeared stymied because genetics was perceived as peripheral to mainstream GP work’. One significant aspect of the ‘mandate change theory’ relates to GPs (the DM₂s). The GPwSI education function is intended to (1) extend GPs’ knowledge base and (2) allow for some career progression as more GPs seek to take on such roles. We will return to these subhypotheses in more detail, noting here that the evidence from this genetics service might prove an exception to these rules. Two quotations, one from a GPwSI and one from a GP, indicate some lack of mutual appreciation at this particular boundary:

Genetics is an unknown. GPs don’t realise the importance of it, but also don’t know where to go for genetics advice or they don’t think anything can be done. It’s not part of GPs’ day to day thinking unfortunately.

GPwSI

Genetics is within the smaller print of medicine compared to what we see day in and day out. There will be some people who would think the GP should know more about genetics, but then they don’t understand the breadth of the workload with which we work, the more important disease areas with which we work with.

GP

Quotes reproduced with permission from Sage Publications from Currie G, Finn R and Martin G (2009) Professional competition and modernising the clinical workforce in the NHS, Work, Employment and Society, 23(2):267–84²⁸

Currie and colleagues’ overall conclusion was that, owing to these boundary manoeuvres, referral routes to the genetics service were largely unchanged and the education function was not always appreciated.²⁸ In the few instances where substitution occurred, GPwSIs managed heavily vetted, low-risk patients while consultants dealt with complex, high-risk patients. Alas, as in many qualitative studies, there are no firm data to confirm these specific manoeuvres but we have already seen quantitative evidence from elsewhere to suggest that this might be a familiar general pattern.

Martin and colleagues¹⁹³

We turn to another study relating the same ‘Genetics GPwSI’ scheme. Although it employs the same methodology, it adds explanation in that it follows boundary work through the entire history in which the new division of labour was ‘opened, negotiated and reclosed’. This is a difficult sequence to review succinctly but we attempt to do via testimony recorded by the authors.

The opening salvos are fired by the specialist consultants:

[After the GPwSI was funded] he . . . asked how we felt this should happen, and it was . . . only then . . . his ambitions for this became clear. There was quite a lot of conflict . . . about what he wanted to do and what we felt it would be appropriate for him to do clinically.

Geneticist, site D

We didn’t see any . . . advantage in having [her] . . . as a GP doing lots of genetics clinics. [. . .] the sensible thing was for the first bit of her work to . . . establish whether or not there was a need for her.

Geneticist, site A

The GPwSIs are first to blink:

They felt slightly threatened by what I was trying to do in terms of might I be competing for patients, trying to create an alternative service, which has never been my intention. [. . .] As months go by we can see what our different roles are and how we complement each other.

GPwSI, site D

I've had a number of sessions with [the geneticist quoted above] and we decided I actually . . . to look . . . at whether to set up clinics or not . . . because I don't want to be doing the same work as a genetic counsellor. [. . .] So . . . at the moment, I'm concentrating on education.

GPwSI, site A

Inevitably, compromises are reached. The ensuing and limited changes in responsibilities are mapped in the Currie and colleagues' paper;²⁸ the Martin and colleagues' paper¹⁹³ teases out the reasoning:

The gap between primary and tertiary care is enormous. It's difficult for me to go and speak to an audience of GPs. They speak a different language, GPs are different types of doctors and what they want to know isn't always apparent [to us].

Area Geneticist

Consequently, the educational work of GPwSIs was largely welcomed by specialists and the supportive clinical role was gradually developed by accommodating GPwSIs:

There's a preconception that genetics is difficult and it's a specialised, rarefied subject, and it's not. Part of what [the GPwSI has] done is demystify, simplify and she's provided access into genetics for people.

Geneticist, site A

. . . things with guidelines, I'm happy to sort. [. . .] Any queries I'll pass on to the consultant. I'd still receive a copy, but even if it was something very simple, I might give some advice first off, but anything I'm not so sure of, I always discuss with the consultant.

GPwSI, site A

Martin and colleagues' conclusion of that the eventual change under the GPwSI genetics programme is closer to 'reproduction' of the existing division of labour rather than 'reform' and much closer to support than substitution:

This model of work represented a rather different construction of what constituted legitimate GPwSI knowledge. Rather than involving the 'upwards' extension of knowledge into the realm of esoteric clinical genetics, the key to the viability of an educational role – and more importantly, to its acceptability to geneticists – was that it drew upon a central component of the GP identity. Rather than encroaching on genetics expertise, the knowledge base to be exploited was the GPwSI's knowledge of the pressures of everyday primary-care practice.

Reprinted from Social Science and Medicine, Vol. 68, Martin GP, Currie G, Finn R, Reconfiguring or reproducing intra-professional boundaries? Specialist expertise, generalist knowledge and the 'modernisation' of the medical profession, pp. 1191–98, copyright 2009, with permission from Elsevier¹⁹³

Rosen and colleagues¹⁷⁷

Another section of the comprehensive Rosen and colleagues' study provides further analysis of the interprofessional negotiations which take place on referral routes, protocols, divisions of labour and case mixes as they become established in new GPwSI services. This part of the study was interview based, utilising a rather different group of stakeholders – 11 PCT managers, nine GPwSIs, three physiotherapists, one hospital manager and one clinical administrator. We begin by reproducing (*Box 37*) three key paragraphs from the summary of this phase of the inquiry. The first two make for familiar reading following the studies of Currie and colleagues.²⁸ All of the propositions therein are amply supported by detailed testimony from key informants. Having been raised previously, the additional evidence is not reproduced here. The third paragraph makes a significant and fresh point on change management, which we review in more depth.

As we have seen the development of GPwSI services involves steering around the agendas and sensitivities of different stakeholder groups. Rosen and colleagues' study¹⁷⁷ pays close attention to these 'relational' aspects of the services (or what we have termed alliance building in *Figure 29*). The evaluation, recall, was of four sites – three implementing GPwSI dermatology clinics and the fourth providing a musculoskeletal service – each with quite different development trajectories (*Box 38*).

These case studies were carried out at the development stages of the various GPwSI services and no doubt further perturbations occurred on their way to 'maturity'. They demonstrate amply the importance of boundary work and alliance building in the success of such schemes. They also illustrate that there is likely to be no single origin, route or terminus for success.

BOX 37 Boundary relations and alliances**Rosen and colleagues' study¹⁷⁷**

Different stakeholders were pursuing different objectives through GPwSI clinics. PCTs were particularly interested in understanding and taking control of specialist referrals. For GPwSIs the clinics offered intellectual stimulation, a more varied work life, a sense of personal development and a sense of clinical autonomy and control over their specialist work. Consultant views varied depending on their support for or scepticism about the service but several expressed concern about the threat they posed to the 'consultant specialist' identity and the ability of consultants retain control over patients referred by GPs. They were enthusiastic about the possibility of being able to focus on more complex patients but also concerned about GPwSI competence.

The relationships between GPwSIs, PCT staff and hospital specialists varied from site to site and were a key determinant of the acceptability of the service. These ranged from near universal support among all key groups through to outright hostility. The former situation resulted in widespread enthusiasm for the service with strong GP–consultant relationships and easy access for support and supervision. The latter resulted in tensions, lack of access for training and support and many operational problems for the services. Generally, consultant acceptance of GPwSI services was shaped by their level of control over triage and patient selection and their knowledge of GPwSI skills.

Consultant acceptance of the GPwSI services was the key determinant of the level of access for support, supervision, ongoing training, etc. Without consultant acceptance, GPwSIs could end up working in isolation and with insufficient clinical back-up.

Extract reproduced from Rosen and colleagues with permission.¹⁷⁷

BOX 38 General practitioners with special interests' development and change management**Rosen and colleagues' study¹⁷⁷**

Scheme 1 was set up with support from all stakeholders, consultants contributing to the initial training of GPwSIs. 'We knew if we tried to do it as a thing from the centre out with a diktat, "this is what we want to do!" it was doomed to fail'. Some operational frictions followed, especially between the PCT and the consultants. Compromises were reached, such as the GPwSI holding clinics in both the hospital and in the community.

Scheme 2 was set up as a joint initiative between the PCT and the hospital following external funding. It was initially supported by one dermatology consultant but not by the other four, who felt 'detached' from the service. The referral protocol was developed largely by the PCT working with a local GP and the prospective GPwSI. The GPwSIs (2) had short-term contracts and were appointed from another area and found their role 'unsettling'.

Scheme 3 was initiated by an enthusiastic consultant with no initial involvement from the PCT. The consultant approached the prospective GPwSI and gained support and funding from hospital management. Another consultant completely withdrew his support for the GPwSI service and his referrals continued to operate externally to the new triage arrangements. This dual arrangement seemed to find acceptability.

Scheme 4 was part of an extensive and locality-wide redesign initiated by two PCTs working collaboratively. It centred on the introduction of a central triage service run by physiotherapists and supported by the GPwSI. There had already been a limited GPwSI-style (if not GPwSI-badged) clinic for specific conditions (hand problems) in one of the areas and this was a basis for recruitment. Consultants opposed the scheme roundly; none identified any positive aspects of the new service. The former 'GPwSI' felt that the redesign was 'hasty'. Management toughened. Stalemate ensued.

Extract reproduced from Rosen and colleagues with permission.¹⁷⁷

Wilson and Cox¹⁹⁰

This study is a mixed-method evaluation of the first 2 years of a GPwSI post, using qualitative interviews, a satisfaction survey and an activity log. The stakeholder interviews have wide coverage, involving not only the GPwSI post-holder but also Macmillan nurses and managers, district nurses, hospice staff, PCT managers, palliative care specialists and GPs. The particular GPwSI post provides a reminder of the diverse range of their functions, being a region-based role in a service providing palliative care for a 650,000 population over a region of hundreds of square miles. The remit was to provide 'better focussed clinical care for patients requiring specialist palliative care' than the existing services: the post was filled by one GPwSI on a part-time basis.

The research speaks to several of the issues outlined in our explanatory theory. There is an 'activity log' of routine work activities (reproduced as *Table 2*) that throws light on whether the post involves substitution or support. The short answer seems to be both. There a steadily growing workload on taking referrals, case review and work directly with patients and their relatives. There is also a good deal of information provision to patients and fellow DMs as well as time allotted to receiving training. The main activity, however, seems to revolve around administration, meetings with staff and service development, which the authors find questionable. Arguably it is neither substitution nor support but simply the assembly of bureaucratic apparatus to consolidate the new role.

TABLE 2 Time spent on each activity over the three time periods (time recorded in 5-minute blocks)

Activity	April 2004 to October 2004: period 1	November 2004 to May 2005: period 2	June 2005 to December 2005: period 3
Administration related to the GPSI role (including telephone calls and e-mail not related to ongoing cases, preparing for meetings, writing case notes)	37 hours 15 minutes	34 hours 20 minutes	21 hours 40 minutes
Administration relating to the evaluation project	36 hours 5 minutes	5 hours 45 minutes	35 minutes
Meetings with staff	22 hours 15 minutes	24 hours 10 minutes	30 hours 50 minutes
Service development (e.g. involvement in GSF, OOH)	34 hours	8 hours 25 minutes	23 hours 5 minutes
Case review/advice	0	16 hours 55 minutes	6 hours 45 minutes
Taking referrals	45 minutes	50 minutes	1 hour 15 minutes
Patient visits	17 hours 15 minutes	20 hours	24 hours 25 minutes
Bereavement visits	0	40 minutes	3 hours 30 minutes
Follow-up with patient	40 minutes	55 minutes	2 hours 25 minutes
Patient update (to other health-care professionals involved in the case)	30 minutes	7 hours 40 minutes	6 hours 40 minutes
Providing GPSI information (responding to requests)	50 minutes	50 minutes	30 minutes
Post promotion (including speaking at conferences)	49 hours 25 minutes	54 hours	3 hours
Training (for the GPSI)	44 hours	0	0
Supervision (of the GPSI)	27 hours	19 hours	10 hours 20 minutes
Shadowing (of other staff)	27 hours 30 minutes	16 hours 30 minutes	0
Attending conferences (for the GPSI's development, including travel)	40 hours	39 hours 30 minutes	37 hours 30 minutes
Other (e.g. the GP bursary scheme)	2 hours	0	5 hours 55 minutes

GPSI, GP with a special interest; GSF, Good Standards Framework; OOH, out of hours.

Table reproduced from Wilson E and Cox K. Community palliative care development: evaluating the role and impact of a general practitioner with a special interest in palliative care. *Palliative Medicine* (vol. 21, issue 6) 527–35, copyright © 2007 by (Sage Publications).¹⁹⁰ Reprinted by permission of Sage.

There is also useful material on decision mechanics and professional boundary work. The new guidance introduced a further bureaucratic layer resulting in lower than expected referrals to the GPwSI (new DM_{2.5}). The problem was that the system already had an existing support network of district and Macmillan nurses (old DM_{2.5S}), who used to be able to refer to the palliative care physician (DM₃) with only notional support from the GP (DM₂) ('they didn't have to lift a finger'). Under the new bureaucratic system, if the support teams thought that a referral to the GPwSI was needed they had to go through the GP (DM₂) and 'get him to do all the paper work'. This was eventually remedied with the revision of guidelines back to the status quo.

Moreover there was a belief, especially from Macmillan nurses, that they had already developed high levels of professional skill in palliative decision-making and that the GPwSI role added little. Antagonism at this boundary was exacerbated as a result of capacity issues with the part-time and geographically isolated GPwSI. His lack of time, resources and feedback facilities and the dispersed community needs meant that the system tended to 'short-circuit' back to the existing services (thus completing our trio of role scenarios). The ensuing role strain was met with a compromise. The GPwSI was increasingly used as a source of

'second opinion'. Formal referral re-routing was seen by Macmillan nurses as less important, as it provided a check 'on whether there was anything else we could be doing in addition to what we were doing at the time'.

This study describes the set-up phase of the new inquiry and, arguably, some of the tensions above may be considered the teething troubles of a highly specialised service. The authors, however, prefer a 'system complexity' explanation, in which the utility of any 'new' team player is dependent on existing staffing, expertise and service capacities. Their evidence, they submit, calls on a prior question – the obligation 'to assess whether a specialist input is actually needed and what it is needed for'.

Pickard¹⁷⁸

This paper speaks directly to many of our explanatory theories. In the words of the author, coverage relates to 'rationale behind the establishment of GPwSIs; defining the field of expertise; negotiating boundaries with other professions; issues of professional control and autonomy and intra-organisational issues'. The paper is part of a more general 'qualitative evaluation' of *Care Closer to Home*, examining GPwSI services being established in dermatology, ENT, gynaecology, general surgery, orthopaedics and urology. It follows the typical research design of interviews with key stakeholders, including consultants, managers, GPs, GPwSIs and practitioners with a special interest (PwSIs). Given the breadth of coverage, there is a lack of empirical detail at several points, the author moving directly to an interpretation of the data rather than providing the 'thick description' associated with much qualitative research.

Within the wider remit, two 'case study vignettes' are presented. The first describes a relatively autonomous GPwSI-led day surgery service, directed by two GPwSIs and two salaried GPs, one of whom had been a staff-grade hospital doctor. The service had existed for 10 years and had incorporated the GPwSI initiative rather than being its product. The second was a new GPwSI/consultant collaboration who, with a nurse practitioner, provided gynaecological services in a community diagnostic and treatment centre. The lone GPwSI had recently been accredited in gynaecology under the national scheme and by the local PCT and was trained by the consultant in specific procedures. Once again we find more evidence demonstrating that GPwSI services are particularly diverse in their origins and organisational structure. The study continues to chart these differences showing how they seed contestation on decision-making mandates, referral routing personnel and profession building.

In terms of decision-making, Pickard¹⁷⁸ notes a now familiar battle over claims to expertise and the ability to diagnose. Consultants opine of their own profession, 'diagnosis is the most fundamental aspect of the specialist role' and 'they are just ten-times more expert than GPs', whereas, in staking their own claims, GPwSIs contested certain aspects of the consultant's role, claiming competence in 'anything not done under general anaesthesia'. This frontier was sometimes quite unyielding, with some consultants continuing to 'review' their GPwSI's diagnosis remotely.

Pickard¹⁷⁸ also outlines some differences in referral patterns, under the question: did rank and file GPs retain traditional control to send their patients for secondary care opinion or were they obliged to follow the new pathway to GPwSIs? She notes an initial reluctance and a preference to short-circuit referrals directly along the time-honoured path to the consultant. Over time the GPwSI substitution function becomes established (1) with considerable alliance building from GPwSIs (publicising 'evidence' on service quality and patient satisfaction), (2) by the setting of a lower tariff than acute hospitals and (3) because of strong support for the GPwSI routing from PCTs.

The main contribution of the paper is a consideration of the wider professional terrain of the GPwSI. This analysis is supported by a longitudinal reading of the policy frameworks that underpinned the development of GPwSIs. This shows that GPwSIs occupy a position that, save for the gendered terminology, is perfectly labelled as a 'no man's land'. We have listed a number of key policy documents that proved decisive in the development of the role. The key imperative of a wave of initiatives is to reorganise by moving away from a 'consultant aristocracy' in order to establish a system of 'countervailing

powers' between professional subgroups, managers and consumers. Increasing emphasis over the years on local or community service delivery itself disturbs the professional boundaries in which GPwSIs work and their opportunities for establishing their new roles. For instance, as part of the Department of Health policy *Implementing Care Closer to Home* in 2007, a number of key management 'domains' were redrawn in a way that shifted professional alliances. Part 3 of that document sought to locate the activities of GPwSI and pharmacists with special interests within a new 'service delivery framework'.

The consequence is that GPwSIs sit between GPs and consultants (and their respective Royal Colleges). Their duties are carved out somewhere between individual ambition, professional self-regulation and administrative standardisation. Control over their role also lies somewhere between the national delivery frameworks and the requirements of managers according to local service requirements. The result is that the professional status of GPwSIs remains extremely fluid. For instance, in Pickard's¹⁷⁸ fieldwork she encounters some GPwSIs who continue to work part-time, retaining a GP role. Others became full-time GPwSIs with roles that went beyond diagnosis and referrals, becoming 'champions' (with a remit to promote, educate and support other GPwSIs).

There was also corresponding uncertainty in the training function. Some consultants were reluctant to engage in training (seeing it as a 'threat to their own position'). Some consultants took local GP protégés under their wing. Some qualified GPwSIs took responsibility for training further GPs into the GPwSI role. In one of the case studies, 'PwSIs working under the direction of a consultant were carrying out the same procedures as GPwSIs on another site'. Furthermore, the accreditation function was also unclear and changed over time. Initially, it was a method of formalising existing partnerships that 'had developed through consultant input'. With the publication of the Department of Health policy in 2007, accreditation fell more squarely in the hands of the PCT – although, once again, they had no formal jurisdiction over the consultants and the Royal Colleges.

Finally, Pickard¹⁷⁸ unearths another newly developing 'no man's land'. Emerging service requirements and frameworks often do not correspond to the clinical subdivisions in use within the disciplinary field of medicine. GPwSIs increasingly work in fields defined according to management criteria. 'Thus the areas of "drug misuse", "epilepsy", "headache", "endoscopy" and "child protection" took their place alongside "dermatology", "rheumatology", "ENT" and "orthopaedics"'. As a result, not only were GPwSIs obliged to fit into contingent and possibly shifting specialist areas but they also had the potential to compete with other providers, including both private providers of the same services and those from other professional groupings such as nurses and therapists who were often working in the community.

All of this is rather well summarised in Pickard's conclusion: 'The [GPwSI] professional field is thus contingent, managerially led and insecure'.¹⁷⁸

We follow this sweeping analysis of policy change with a much more detailed examination of a key constituent. One of the broader institutional impulses that may be significant in sustaining the GPwSI role is the changing nature of GP practices in the UK. This transformation has had considerable institutional backing, being the subject of a major discussion paper issued by the RCGP of the BMA.¹⁹⁴ Well-documented changes in practice size, budget control, purchasing power and contractual authority have had considerable implications for the GP roles, creating a much more hierarchical internal division of labour. The studies that follow fit GPwSIs into the picture.

Lester and colleagues¹⁹⁵

Lester and colleagues¹⁹⁵ summarise the key dynamic:

Greater specialisation appears to be occurring in primary care, with GP principals actively "specialising in generalism" or adopting new roles and identities, through vertical substitution, for example as GPs with special interests. Salaried GPs, in contrast, appear to be adopting the left-over or discarded jobs, mopping up the less complex and perhaps less professionally satisfying or challenging patients, echoing Hughes work on the division of labour based on 'dirty work'

Reproduced with permission from the RCGP from Lester and colleagues¹⁹⁵

The reference here is to Hughes's 1958 thesis¹⁵³ that those with high professional standing retain the more desirable work, delegating the less pleasant or stigmatising work to others with less standing.

The study comprises detailed qualitative interviews with two emerging groups: practice partners and their salaried junior colleagues. Two pairs of quotations from the respective parties mark an increasing awareness of role differentiation:

I get the messy things and I rather like mess. Part of the philosophy is we try to allow chaos to tumble and that we systemise it and we organise it and then we handle it.

GP principal

I don't see much minor illness anymore and most consultations are either psychological or quite long and complicated.

GP principal

Most of the time I see new problems . . . they are coming in with new problems, but I do see chronic follow-up problems when the regular GP is not there.

Salaried GP

We just have to get them all in and have their blood pressure checks and send them all off again. Yeah they all just have to whiz in and whiz out.

Salaried GP

Reproduced with permission from the RCGP from Lester and colleagues¹⁹⁵

The main concern of the study is about the increasing concern about the future and overall function of primary care. For this review the key element is the widespread recognition among GP principals that their future lies in a more specialised role (even if it is captured by the oxymoronic term 'specialist in generalism'). The first two quotations above are from partners who are not GPwSIs (other respondents held such posts). What is notable, nonetheless, is that both of them describe their roles in terms akin to those laid down in policy documents that established the GPwSI role. Whatever turbulences GPwSIs encounter in implementing and developing their new role, it would seem that there is a considerable following wind intent on creating and establishing that role within their own profession.

There are many further studies, reporting in the GPwSIs' own voice, which describe the developing internal divisions that are overtaking the traditional GP role. Two are covered briefly here: Newman and colleagues¹⁹⁶ and Honey and colleagues.¹¹⁹

The Australian paper uses qualitative interviews to unpick the many influences that might lead a GP to take up a specialist interest in human immunodeficiency virus (HIV).¹⁹⁶ Despite the area being underserved and politically isolated, there are still perceived advantages (clinical, professional and political) for the committed in becoming a community-based HIV medication provider:

To be a competent prescriber . . . one needs a fair bit of work to keep up with developments . . . you'd sort of want a big cohort of patients to make it worth your while [GPs by contrast] Their books are closed. They can't see any more patients. They want to be able to see patients with their coughs and colds and flus and kidney stones and ingrown hairs, and all the other things.

Reproduced from Newman C and colleagues¹⁹⁶ with permission from Taylor and Francis Ltd (www.informaworld.com)

The Honey and colleagues' study in Bradford provides a detailed account of the perception of two GPwSIs in homage to the new role:¹¹⁹

My work is more fulfilling now because I feel I can help patients more. As a generalist the patients who come to see me either don't need to see a doctor and we have to pat them on the back and send them out the door – which can be difficult – or they have got something wrong with them but we haven't got the time or the facilities to deal with them. So we have to pack them off somewhere else, like the hospital, to see a specialist. Very often that process is problematic because there are often waits and problems with communication. So now in a small area – urology – there are a group of patients who walk through the door on an average morning who have a problem that I can start to deal with and even if the urological problem they are bringing is beyond me e.g. a patient who needs a prostate operation or the removal of a kidney . . . I am working very closely with a team of clinicians who are able to do them and I know what checks to do first and I know what to tell the patient. So I feel I can help those patients.

Honey S, Small N, Ali S, General practitioners with special interests: the potential benefits and possible risks for primary care, Primary Health Care Research & Development, vol. 6, issue 1, pp. 5–16, 2005, reproduced with permission from Cambridge University Press¹¹⁹

Jones and Bartholomew¹⁹⁷

This very early paper provides another indication of the disintegration of the traditional role of the family practitioner and telling details of the momentum towards special interest status. It consists of survey of randomly selected GPs ($n = 931$, response rate 43%) mounted shortly after the 2000 NHS Plan.¹⁴ Data reveal that 70% of respondents claim a special interest and 38% of them report undertaking clinical sessions in areas of particular interest. 'Even if none of the non-respondents undertake clinical sessions this means that around 16% (approximately 4000) of all GPs in England do one or more sessions in their areas of special interest.'¹⁹⁷ The date here is crucial. Substantially more GPs than envisaged in the GPwSI plan (recall that it was 1000) were already providing clinical specialist sessions. Relevant postgraduate qualifications were held by 71% of respondents and 82% had undertaken continued medical education in relation to a specialism. Free-text responses confirmed the wish to add to their professional life. The study provides another indication of the unstoppable bandwagon on role change. GPwSI interventions climb on board a train that has already left the station.

Although each and every one of the studies so far has suggested that the adoption of the GPwSI role involves the incorporation of a complex delivery system with an already complex delivery system, we have yet to encounter a study that devotes itself to describing the multiple contingent relationships in the system (recall the 'Goringhamshire' orthopaedic referral pathway in *Figure 16*). Almost inevitably, the study that comes closest to this objective is the Rosen evaluation.

Rosen and colleagues¹⁷⁷

Chapter 8 of the study includes explanatory data that help explain the limited impact of the GPwSI services on referral capacity. This follows, argue the authors, because of a whole range of 'system wide' impediments. Introducing a new service at point D requires the harmonisation of facilities and capacity at points A, B, C, E, F, G, etc. *Box 39* provides a brief list of the logistical difficulties encountered in the case study sites.

In terms of the distinction between substitution, support and short-circuit, we see a mixture of all three operating simultaneously in these interventions. GPwSIs (DM_{2,5}) aspired to and, in a few cases, achieved a fully substitutive role in their implementation of an extended, see-and-treat service. The outpatient consultants (DM₃), however, saw the new services essentially as support owing to the demonstrably small and simpler set of patients referred to the GPwSIs. In one instance *they* referred their own routine patients in the reverse direction to GPwSI clinics (in order to 'kick start' the new service). Some GPs (DM₂) saw the GPwSI service as a reservoir of support absorbing minor consultations and treatments that they might otherwise have to deliver. Finally, there is the matter of control of the referral route. Patients were

BOX 39 The impact of system-wide contingencies on GPwSI services

Staffing levels. Changes in decision-making responsibilities need to be matched with the reshaping of capacity at in different parts of the referral pathways. High staff turnover in both outpatient and GPwSI services limited the capacity to make planned changes.

Capacity limitations. The relatively small size of the newly introduced GPwSI service automatically limits impact on overall referral rates (one dermatology GPwSI clinic saw only 6% of the total area caseload).

Time pressures. GPwSIs, under pressure to introduce patient-centred care, and lacking the full range of administrative support, held longer consultations (from 15 to 40 minutes) than in the outpatient clinics (10 minutes). Establishing these see-and-treat responsibilities reduced overall throughput in some cases.

After care. Routine follow-up (following treatment) is made more complex and difficult to manage – aftercare being assigned to additional routes within the new system – via outpatients or via GPwSIs or via GPs [or by the newly emergent brand of practitioners with special interests (PWSIs) or by reception and office services].

Alternative routing. The lack of clarity in expectations about the new musculoskeletal service led to an increase in external sources of referral to outpatient clinics (from accident and emergency, walk-in centres and other consultants), which increased the pressure on them.

Lowering thresholds. Respondents to Rosen and colleagues' inquiry acknowledge that one potential effect of the new GPwSI services was the lowering of GP's 'referral thresholds', leading to an aggregate increase in referrals throughout the system. Minor conditions that might previously not have been referred found their way to the GPwSI.

Unmet demand. More generally, against the background of long waiting lists, it was sometimes understood that the new system was a way of meeting rather than curbing previously unmet demand.

Payment systems. Sometimes these were revised and became out of kilter with the intended effect of reducing referrals. GPs who would have been previously paid for minor surgery were more inclined to 'send it off'.

Interpersonal changes. Referral systems take on their own custom and practice and hard won person-to-person preferences on referrals are lost under a more anonymous system.

sometimes referred directly to GPwSIs by the GP. In other models there was prior triage (either centrally from the PCT or by a clinician) with the GPwSI referrals clearly consigned to support mode. As for short-circuit, we have already noted the unintended consequence of increases in alternative routes to the outpatient clinics – using accident and emergency, walk-in centres and consultant-to-consultant exchange.

The authors refer to very many of the above logistical and capacity issues as 'local contingencies'. They bring heterogeneity and unpredictability to the systems studied. Is there any reason to believe that these unforeseen conditions would apply to other and perhaps more mature GPwSI interventions? Are they merely teething troubles and domestic difficulties? Although there is good reason to suppose that, over time, these interventions would reach a different equilibrium, there is no reason to suppose that that equilibrium would be final and no reason to suppose, moreover, that any subsequent system would be that envisaged in the initial policy proclamations. What is clear from the 'explanatory data' above is that the installation of a GPwSI involves far more than adding another step or quality filter in the decision chain. How the new stage operates will always be constrained by staffing levels, capacity limitations, time allocations, after care procedures, available alternatives, decision thresholds, unmet backlogs and so on. These processes will never be identical from location to location but their presence will always be felt.

Conclusion: emerging patterns

Putting the outcome and process studies together reveals something of a service delivery melee – but there is some order within that complexity. Our first conclusion takes us back the matter of what exactly is being evaluated in this research on the GPwSI role. GPwSI services do not represent an orthodox ‘planned programme’. There was never any point in the above history when a distinct ‘GPwSI intervention’ was specified and designed. As with much policy-making, it all began with a vision (thou shalt create a thousand GPwSIs), a rough specification of their role and the introduction of accreditation and training paraphernalia.

Research synthesis, of course, is much simpler if there is a designated series of programmes implemented to an agreed blueprint. GPwSI services, by contrast, originate at the confluence of two grand policy imperatives. The first is a breaking of the ranks of GPs. For many years there has been restlessness with the GP ‘generalist’ role and a fragmentation into an internal hierarchy beginning with the distinction between GP principals and salaried GPs. As noted in several papers above, this move to increasing specialisation began long before the RCGP call for a thousand GPwSI posts. A parallel and more momentous policy shift occurred under the broad rubric of ‘care closer to home’. In the same time period, fundamental changes in health-care management through PCTs and CCGs sought to redraw professional boundaries and power structures. PBC allowed for a more local approach to initiate and support services, moving them to locations ‘nearer to the patient’.

The resultant vector was the GPwSI, a notional function with uncertain powers. As we have seen, GPwSI posts have been deposited unevenly across every corner of medicine, with a professional remit ranging across clinical practice, education, leadership, management, research, quality assurance and public health. GPwSIs can be considered quite properly as a portmanteau profession launched, from the point of view of the reviewer, in a perfect storm of heterogeneity. Against this background it is not possible to issue an overall verdict on their contribution in solving demand and referral overload.

It is possible, however, to discern within the primary literature certain configurations into which the GPwSI role has settled. Several of the primary studies above describe a ‘default’ arrangement in which referral patterns remain relatively unchanged owing to residual pressures in the system. This occurs when (1) the consultant (DM_3) remains in relative control of referrals and the protocols that govern them, and when (2) the GPwSI ($DM_{2.5}$) settles for a role that concentrates on the support functions in managing the subsidiary apparatus and focuses on prescribed subset of ‘simpler’ patients, and when (3) the GP (DM_2) retains referral habits either by short-circuiting the process with direct referrals to secondary care or by concentrating on patient reassurance regardless of demand implications. Under this scenario there are no gains in demand management.

Such a pattern contrasts with a second configuration, often won under many years of adaptation and compromise, in which the division of power, motivation and expertise is more perfectly spaced, and in which (1) the consultant (DM_3) retreats to the specialist role in the treatment of complex patients and become advisor to rather than supervisor of the system to identify them, and (2) the GPwSI ($DM_{2.5}$) deals with an agreed and intermediate case mix and works in a (physical and social) space that is independent from the GPs and consultant surgeries, and (3) the GP (DM_2) works to protocols delivered via the GPwSI educational and management functions. This scenario represents a system-wide reconfiguration and is most likely to occur under the remit of the triaging and gateway systems reported on in the previous chapter. Referral outcomes are similar to best practice in RMCs; that is to say referral accuracy may improve, waiting times may shorten and patient satisfaction may increase, but the overall level of referral activity is likely to increase. Improvements at many points across a system tend to prompt supply-induced demand across the board.

The question arises about the best way of getting from scenario 1 to scenario 2. In many of the cases reviewed, this journey has involved a reflective process of 'muddling through', with one after another of the systems strains being identified and dealt with. *Box 18* provides an agenda for such an exercise as perceived through the eyes of GPwSI services in the Rosen and colleagues¹⁷⁷ study. We borrow a term here from Charles Lindblom's classic article 'The science of muddling through'.¹⁹⁸ Instead of comprehensive analysis of every policy option, a much more constrained process of 'successive limited comparison' is really how policies are developed, insists Lindblom. While we accept this non-pejorative assessment, it still begs the question of whether or not GPwSI services could be developed in a less piecemeal fashion.

We close with a tentative solution being developed to handle these situations in which there is a grouping of semi-independent bodies, each with different agendas and capabilities, which are set some goal that involves collaboration. The interdependencies in these networks are so complex that it is impossible to design and control them centrally. Rather than dealing with a thousand post-hoc improvements, it is sometimes possible to incorporate 'group model building' into the development stage of interventions. Group model building is a form of decision-making that involves diverse stakeholder groups working with a facilitation team to solve a focused problem within a complex system. The classic components involve testing different scenarios and strategic options, modelling their outcomes (intended and unintended) from different perspectives and testing to find points of compromise between different 'scripts'. Recent examples in devising new systems of obstetric care¹⁹⁹ and responses to domestic violence²⁰⁰ have been published. This approach captures perfectly the underlying dilemma of introducing new roles, such as GPwSIs, into complex referral systems. Whether or not it is able to shift solutions from 'ex post' to 'ex ante' remains to be seen.

Chapter 6 Changing responsibilities: direct access to the results of clinical tests – can it control and shape demand?

Introduction

We turn in this chapter to a review of our third form of mandate change, in which systems and organisational structures remain intact but the responsibilities and remits assigned to a given role are shifted. An example in this family is GP direct access to diagnostic tests and services.

The notion of providing GPs with direct access to clinical tests has been around for half a century. We open the chapter with a very early sighting in a study by Clarke and Rickards.²⁰¹ They audited a 3-month period of GP open access to an X-ray service from a Sheffield practice with 4000 patients (*Table 3*). During the 3 months, 29 patients used the direct access service and subsequently only five (17%) were referred to hospital or outpatient services. The authors calculated a cost saving to the NHS of £108.13s.9d.

Diagnostic tests such as medical imaging are a key feature of modern medicine; it has been suggested that 75–80% of medical decisions are based on a test result.²⁰² Many diagnostic tests and services such as imaging and endoscopies have conventionally been situated in secondary care and often the specialist in the outpatient clinic is the gatekeeper to these tests. Traditionally, the services of radiology and other diagnostic services have tended to be available only through hospital infrastructure. Indeed, a special article in the *Lancet* in 1962 suggested that there needed to be a physical redesign of hospitals to bring diagnostic services to the front and thereby enable speedy access for patients from outside the hospital.²⁰³ Direct access to tests is, therefore, an extension of GP responsibilities into a remit that previously required specialist referral.

General practitioner direct access to tests has a relatively long history within the sphere of demand management interventions. A review of open-access pathology services as long ago as 1973 suggested that 'use of pathology should be a significant feature of modern general practice' and that the policy of the then Department of Health and Social Security was that GPs should have free and open access to the pathology services provided in hospitals.²⁰⁴ A joint working party of the Royal College of Radiologists and the RCGP suggested in 1981 that direct access to radiological services was essential for primary care physicians. The report outlined that GPs should have a similar right of access as a consultant, although provision should be determined locally bearing in mind resources and ensuring that proper indications for testing are observed.

TABLE 3 Details of savings from direct access to X-ray

Details of saving		
Attendance saved	Cost per item	Saving
18 first attendances (non-teaching hospital)	£3.4s.8d.	£58.4s.0d.
One first attendance (teaching hospital)	£3.14s.5d.	£3.14s.5d.
14 second attendances (non-teaching hospital)	£1.1s.11d.	£15.6s.11d.
1 × 1 week's inpatient care (non-teaching hospital)	£31.8s.5d.	£31.8s.5d.
Total saving in 3 months		£108.13s.9d.

Reprinted from *The Lancet*, Vol. 286, Clarke AH, Rickards DF. Economics of open access to diagnostic services, pp. 336–7, © 1965, with permission from Elsevier.²⁰¹

More recently, the Department of Health's 2008 policy document *Care Closer to Home*²⁰⁵ advocated radical service redesign, with an emphasis on devolving aspects of a patient's care pathway from secondary to primary care. The coalition government's 2010 White Paper *Equity and Excellence: Liberating the NHS*²⁰⁶ aimed to shift decision-making closer to the individual patient and devolved power and responsibility for commissioning services to local consortia of GP practices and included specific targets for introducing patient choice for diagnostic testing by 2011.

This chapter first outlines the programme theories that underpin how GP direct access to tests is purported to work in managing demand for specialist services. It then offers some potential impediments to enacting procedural change as envisaged through the programme theories. The now familiar 'turf war' issues arise here in battles over claims to expertise and into whose remit the 'dirty work' of organising necessary diagnostic tests falls. The main body of the chapter is the evidence review, which is organised around a typology of the functions of diagnostic tests.

Programme theory

The principal idea behind allowing direct access to tests is to enable GPs to carry out tests required for the ruling in or out of diagnoses for certain conditions, or the initiation/maintenance of treatment, without the need to referral to secondary care. The underlying theory is about creating a more efficient division of labour, whereby the GP can use the test results to decide whether the patient requires management or treatment by a consultant or can be managed in primary care. There are a multiplicity of tasks and procedures to be implemented and managed along the patient pathway from primary to secondary care. The basic demand management programme theory here is about ensuring that the right person conducts the right task in the right sequence at the right price. The basic premise is that a GP is able to select an appropriate test given the patient's presenting symptoms and to interpret the results accurately in order to make further management decisions without an initial referral to a specialist.

Diagnostic tests are routinely used in modern medicine to assist in diagnosis and establish appropriate management plans. Without direct access to diagnostic tests, a GP must make a referral to a specialist in secondary care for a patient to undergo the test. If the test excludes any serious disease pathology, or indicates a condition that can be managed in primary care, the patient is then discharged back into the ongoing care of the referring GP. With GP direct access to tests the specialist referral can be avoided in some patients and for others the results of the test can ensure the referral is to the most appropriate consultant. The premise is that a better, more efficient division of labour is achieved by shifting the point at which some decisions are made to earlier in the patient pathway, in primary as opposed to secondary care.

General practitioner direct access to diagnostic tests is thought to improve demand management by:

1. Preventing inappropriate referrals: patients are often referred to secondary care for a diagnosis as the GP is uncertain if their symptoms are indicative of serious pathology. Direct access to tests can reduce the number of referrals, as it allows the GP to make the diagnosis/establish a treatment plan and then manage less complex patients in primary care. Only those patients with serious pathology, those requiring surgical intervention and those with complex management issues need to be referred.
2. Earlier diagnosis and reduced waiting times: a number of patients will require referral to secondary care subsequent to the test. A further benefit of GP direct access is that waiting times to the diagnostic test can be reduced as the need for a specialist appointment to obtain the test is bypassed. Furthermore, treatment can be initiated sooner, even if that treatment needs to be managed by a specialist, as the results of the test will be available for the initial consultation.
3. Reducing costs: there are expected cost savings from reducing the number of referrals to secondary care and in reducing the number of consultations necessary if referral is required.

The adjustment may be relatively minor, involving the customisation of responsibilities for tasks rather than the creation of new roles or organisational units. The basic rationale is depicted in *Figure 28* and a value chain for the role of diagnostic tests is shown in *Figure 29*. However, further examples of specific patient-care pathways utilising direct access given in *Figures 30* and *31* demonstrate how complex direct access to testing is in reality, with a myriad of diagnostic and treatment decisions to be made along the pathway.

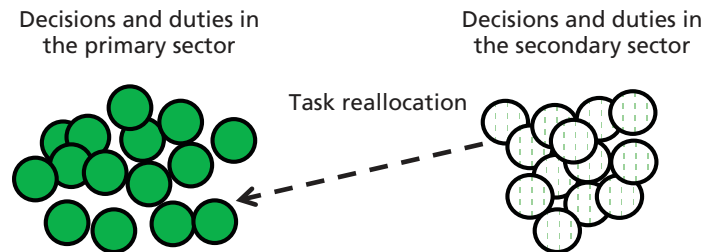


FIGURE 28 Direct access to tests.

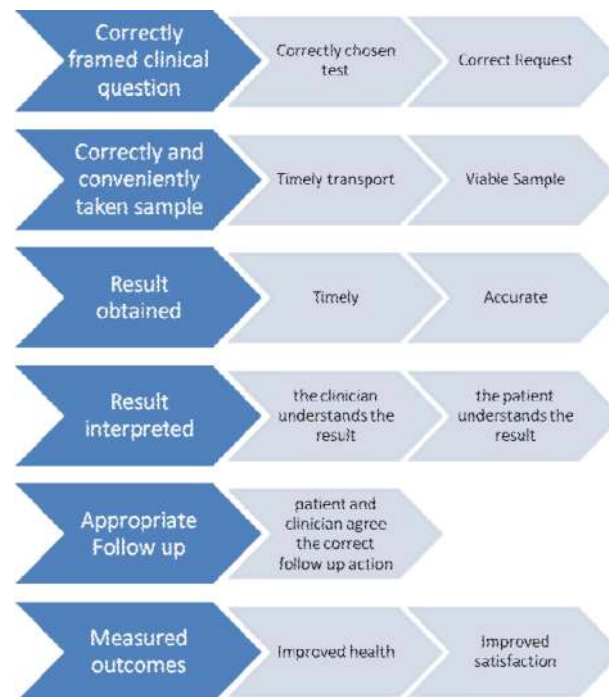


FIGURE 29 Pathology value chain. This information is licensed under the terms of the Open Government Licence (www.nationalarchives.gov.uk/doc/open-government-licence/version/3/) with permission from David Laszlo Partnership from David Laszlo Partnership and the Department of Health Aligning Diagnostics, London, Department of Health.²⁰⁷

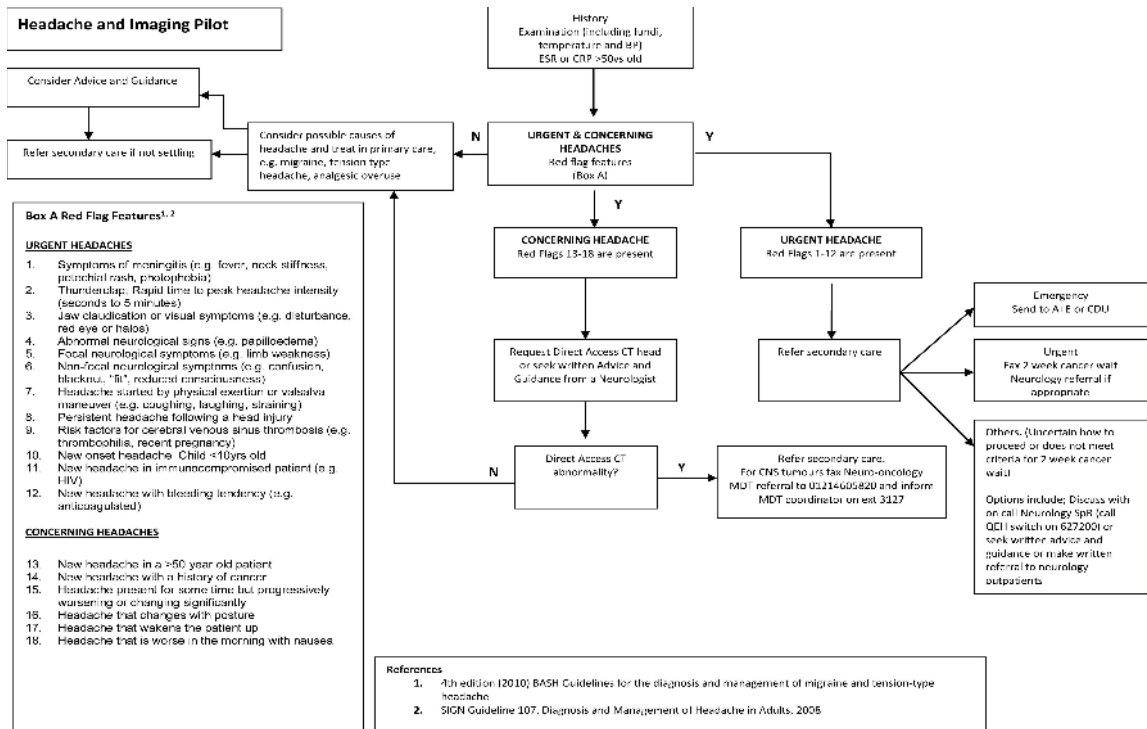


FIGURE 30 Birmingham South Central clinical pathway for headache. Reproduced with permission from Birmingham South Central CCG.²⁰⁸

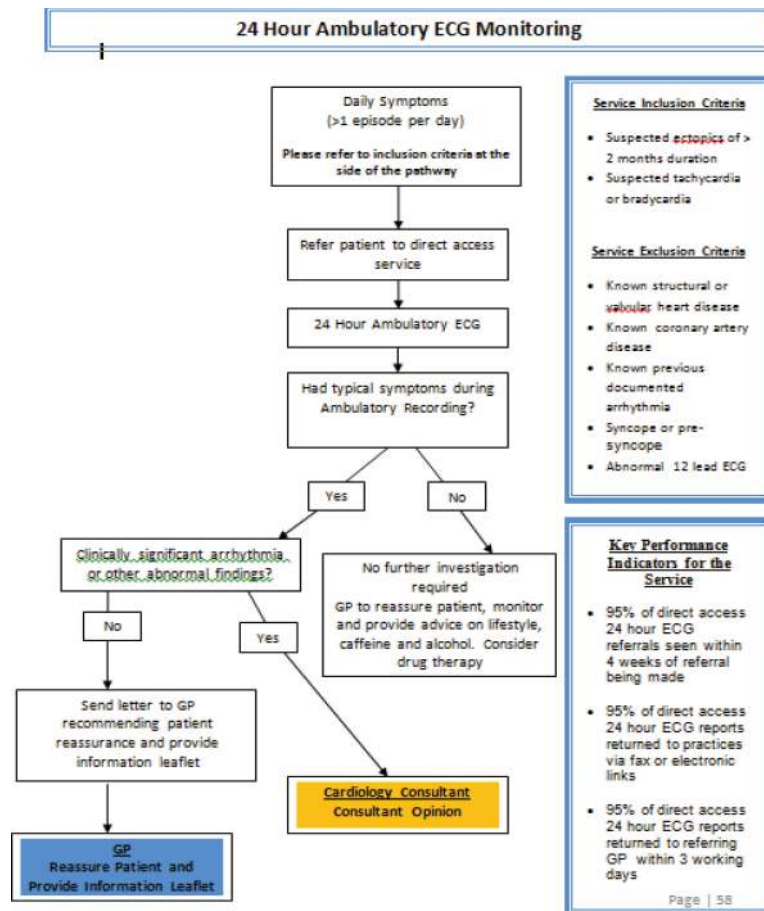


FIGURE 31 Birmingham South Central clinical pathway for 24-hour ECG. Reproduced with permission from Birmingham South Central CCG.²⁰⁸

Impediments to procedural change

Inevitably, in practice there are barriers to such shifts in decision-making responsibilities. The basic programme theory rests on the assumption that the GP is best placed and has the appropriate skills to decide when a test is necessary and to make appropriate management decisions on the basis of the results.

However, some argue that GPs do not have the appropriate skills to make informed decisions regarding the *need for* a test and an unintended consequence is that some patients will receive tests inappropriately and demand on diagnostic services will, therefore, increase.²⁰⁹ Once again, we may expect 'turf wars' at play here; specialists lay claim to the expertise and clinical skills that can diagnose without excessive testing.

Perceived expertise is also a potential issue in the *interpretation of* the results of a test. Turf wars abound here, too, in that the specialist may again lay claim to the expertise necessary to make appropriate management decisions rather than relinquish control of this to a generalist. The specialist, therefore, controls access to diagnostics and guards their professional boundaries closely. In some cases it may be that the GPs themselves feel that the interpretation of the results lies outside the limits of their expertise and an unintended consequence is then that direct access is bypassed and referral to secondary care remains the pathway of choice.

As in the previous chapters, there are also potential impediments concerned with the nature of the duties which change hands. There are some tests that can be perceived as 'dirty work'. No one camp wishes to lay claim to the work as within their specific boundaries, although the test remains a necessary stage in the patient care pathway. For such unwanted labour, an unintended consequence may well be a bypassing of direct access by the GP and referral to secondary care. However, specialists, in turn, may block referrals in the absence of the necessary tests and the patient wait may be extended. As per usual, we enter the review against a background of contested ideas and seek evidence to shed light on the hypotheses presented.

Evidence synthesis

An important proviso in reviewing the evidence on direct access to tests is that 'all tests are not created equal'. Notwithstanding the differences between imaging, pathology, endoscopy, etc., the function of a particular test in a particular context differs significantly. We begin the evidence synthesis with an outline of typologies of the functions of tests and organise the subsequent review according to one typology, as this provides more explanatory power than reviewing the evidence by test or by condition.

In discussing the types of tests available, a common approach is to distinguish between the technologies or sample source underpinning a test, for example endoscopies versus CT scanning or blood tests versus urine tests. A more useful typology in the sphere of demand management is one which distinguishes between the function of a test. For example, the AHRQ Methods Guide for Medical Test Reviews differentiates three types: those that screen for the likelihood of a disorder; those that diagnose current disease; and those that monitor progress and/or assess response to treatment.²¹⁰ The Department of Health strategic guidance for aligning diagnostics outlines five functions of tests: screening for clinical risk; providing reassurance; supporting diagnosis; monitoring treatment effectiveness; and guiding the selection and/or titration of therapy. Such a typology can, therefore, make distinction between a CT scan for chronic headache whose aim is to reassure an anxious patient that they do not have serious pathology, and a CT scan in a patient with focal neurological symptoms where there is suspicion of cerebrovascular insult.

We present here an adapted typology that we then use to structure the review of the evidence. In this typology we make a distinction between diagnostic tests that identify the rare cases of serious pathology (ruling-out tests) and those that aim to provide a definitive diagnosis where one of a number of positive results are expected. Both screening tests and those that monitor treatment may be regarded as clinical indicator tests.

Typology of functions of diagnostic tests

1. A ruling-out test: these are tests where a low diagnostic yield is expected but a positive result suggests serious pathology; the expectation is that most patients would receive a negative result. The rule-out test is necessary to provide reassurance that serious pathology is not present.
2. A differential diagnosis test: these tests would generally have a high diagnostic yield if employed appropriately; the expectation is that most patients will receive a positive result as a test outcome. However, the diagnostic outcome may be one of a number and the actual diagnosis would drive subsequent management.
3. A clinical indicator test: these tests serve to identify at which point on a clinical trajectory the patient is and the results would guide the initiation of treatment or a change of management.

Ruling-out tests

One might expect that direct access to ruling-out tests is a prime candidate for a successful demand management intervention. In these cases a negative finding is just as important as a positive result. The key expectation in terms of demand efficiencies is that the majority of patients will have a negative scan and, therefore, a referral to secondary care can be avoided. This scenario, moreover, provides the patient with considerable reassurance. The literature gives us two examples in this modality: imaging for patients presenting with chronic headache and endoscopy for patients presenting with dyspepsia. In both cases, the test is used to rule out serious pathology, such as a brain tumour in the case of headache and gastroesophageal cancer or peptic ulcer disease in the case of dyspepsia.

As we will see, these referral efficiencies do not always materialise. Some evidence indicates that a number of patients with negative results still receive a referral to see a specialist. There is also debate about whether or not GPs have the skill to interpret such tests and whether or not they go on to request further tests that the experienced consultant might regard as invasive and unnecessary. We explore these unintended consequences and counter theories via a series of primary inquiries.

***Simpson and colleagues*²¹¹**

This paper reports the results of a database review and GP questionnaire of direct-access CT for chronic headache to assess rate of significant pathology, outcomes, cost benefit and GP satisfaction. All patients who had been referred from primary care with a headache for longer than 3 months for a brain CT between 1999 and 2007 were identified from a database. The referring GP was sent a questionnaire along with the original radiology report to assess impact on treatment and referral and GP satisfaction with the service. It is unclear from the paper if the review was prospective or retrospective; it may be that GPs were asked to make comment on referrals from up to 8 years ago. All scans that reported abnormal findings were reviewed by a consultant neurologist who classified the findings into those that were likely to be incidental to the presenting symptoms and those that were potentially causative of the presenting symptoms. The subsequent management of all those with potentially causative abnormalities was reviewed.

During the 8-year study period, 4404 direct-access CT scans of the brain were performed. Of these, 461 (10.5%) had abnormalities reported with these thought to be potentially causative in 60 (1.4%). The authors report subsequent resource utilisation only on the 60 patients where potentially causative abnormalities were found and an additional 17 patients who had normal scans but returned for further imaging. It is not clear if the remaining patients were then managed in primary care. However, later in the paper, as part of the economic estimation, the authors cite that 14% of the 4404 patients went on to have a consultant outpatient appointment; this would amount to 616 patients, suggesting that some patients with normal scans still had a specialist referral.

There was a poor response rate to the GP questionnaire: 986 out of 2998 usable questionnaires were returned (32%). When asked what the management would have been without direct-access CT, 90% of the GPs would have referred to a specialist clinic with the majority of these referrals to either neurology (45%) or general medicine (38%). The authors thus cite an 83% reduction in the total number of referrals for specialist opinion. This, however, was calculated from the poorly responded questionnaires, so it is by no means certain that this figure would apply to those referrals where the GP did not respond.

The authors used these data coupled with audit data of imaging referrals from a headache clinic to suggest that direct access demonstrated a cost saving of £86,681.81 across the 4404 patients.

This study suggests direct access to CT scanning may reduce referrals to secondary care by ruling out serious pathology at an earlier stage in the patient pathway. However, there is no direct comparison with actual behaviour in the absence of direct access and the findings are based on hypothetical, post hoc, GP-reported assessments of management that may have been gathered some years after the actual test. Therefore, the reported effect should be treated with caution.

In this study, it appeared that the only task transferred to the GP was the decision to offer the test. All scans were reported by a consultant neuroradiologist and so the pre-existing professional boundaries between specialist and GP remain intact. CT scans are often ordered to provide patient reassurance and the majority of scans in this population are expected to be normal. However, in this study, some patients with normal scans still had a subsequent specialist appointment, suggesting that perhaps the scan did not always provide reassurance. It was unclear from this study if these referrals were made at the request of the GP, the specialist consultant or the patient, and so we do not know who needed the reassurance.

*Taylor and colleagues*²¹²

This paper was a before-and-after comparison of GP direct access to brain MRI for patients with chronic headache. Outcomes for the first 100 direct investigations were compared with the last 100 referrals to neurology for the condition before the introduction of the direct-access service pilot. It should be noted here that the two groups of patients might have been quite different and this, rather than the presence or absence of direct access to brain MRI, could have affected how they were managed. Indeed, they differed in terms of mean age and gender split.

The authors reported that:

The main concern of neurologists prior to initiation of this scheme was that incidental findings would be investigated unnecessarily, with increased cost and possibly increased morbidity. This was thought to be due to the limited exposure of GPs to neuroradiological reporting. To address this risk, a standardized reporting style was introduced, which has been adhered to in practice

Reprinted from Clinical Radiology, vol. 67, Taylor TR, Evangelou N, Porter H, Lenthall R, Primary care direct access MRI for the investigation of chronic headache, pp. 24–7, 2012,²¹² with permission from Elsevier

Here, we see that neurologists express concerns about the ability of GPs to interpret tests, as GPs do not have a neurologist's specialist experience. As result of this lack of experience, neurologists perceive that GPs may then order further, potentially invasive, investigations unnecessarily. To avoid this potential unintended outcome, in this study, all reports sent to the GP were expected to have a clear summary of the findings and guidance with regard to how to act on the results. In terms of the theory under test, the GPs were provided with support in the form of a standardised report to assist them with interpreting the test results.

The GP direct-access cohort had an onward outpatient clinic referral rate of 17%. This was twice as high as that advised in the reports accompanying the scan results; however, the referrals were not all to neurology clinics and included, for example, referrals for sinus problems or to a general headache clinic rather than to a neurologist. Furthermore, in one case, where the report gave advice that onward referral was indicated, no subsequent referral or appointment could be identified.

The authors conducted a crude cost-effectiveness analysis that suggested GP direct access resulted in a 46% cost reduction, compared with referral to a neurology clinic. This was mainly due to the cost saving of an outpatient appointment in 83% of the cohort, as it was assumed that all those referred would still have a MRI. However, this may not be a correct assumption, as there are no data to establish how many referrals to the neurology clinic undergo a MRI.

This study suggests that direct access to MRI for chronic headache seemed to reduce the need for referral; however, the results of a before-and-after study should be treated with some caution. In terms of the theory under consideration, neurologists perceived that GPs did not have the necessary experience to interpret the test accurately and may unnecessarily investigate incidental findings. To avoid this potential intended outcome, the neurologists retained some control over which patients were referred and how the results were interpreted by developing a 'clearly defined referral pathway' and including a guidance statement with the results. In this study, these strategies appeared to be moderately successful in these aims.

*Thomas and colleagues*²¹³

This paper reported on a prospective observational study of primary care access to head CT for chronic headache. All GPs in an area of Scotland were given open access to head CT, subject to specific referral criteria:

- aged 18 years or over
- normal neurological examination
- symptoms consistent with chronic, unchanging headache.

All referrals over a 1-year period were included and a follow-up questionnaire was sent to the referring GP with the scan report to ascertain the impact of the result on subsequent referral decisions. Outpatient clinic databases were checked for 12 months to record any subsequent referrals.

During the study period, 232 patients were referred for brain CT out of an estimated 18,702 headache consultations, giving a referral rate to open access of 1.24%. Of these 232 referrals, 215 were scanned (11 patients cancelled/did not attend appointment, and six were not performed by radiologist as forms were incomplete/the patient did not fulfil the study criteria). Of the 215 scanned patients, 30 were referred to a neurology clinic because of headache [note that almost half of these ($n = 12$, 40%) had the referral made at the same time as the scan request].

Completed questionnaires were returned for 189 (88%) of the referring GPs and reported that the scan had stopped a referral to secondary care in 167 (88%) of the patients. As the scan caused a referral in 10 (5%) of patients, the authors give an estimated predicted reduction in clinic referrals to be 83% if GPs had open access to CT scans for chronic headache.

The authors noted that a number of referrals were made at the same time as the request for a scan and suggested that in cases where the GP is unsure of the diagnosis or in need of advice about management then direct access to CT will not alter referring behaviour. However, for many patients a scan is simply about reassurance, and in these cases a referral to secondary care is not always necessary:

The primary reason for arranging neuroimaging must remain its clinical indication, but the results of this study suggest that open access scanning may give GPs more confidence in managing the small minority of patients for whom a normal scan will provide reassurance and prevent secondary referral.

In terms of the theory under test, the low referral rate for open-access imaging per headache consultations found in this study would suggest that GPs are not overusing the direct-access service. It may be that the clear guidance on who to refer and the gatekeeping role of the radiologist ensured that use was controlled. The results suggest that clear referral criteria may help GPs to decide which patients are

appropriate for referral to direct-access CT scanning and, as a result, the test results can provide the necessary reassurance to the patient. Consequently, in these circumstances, direct access can reduce the number of referrals to secondary care. However, short-circuit may still occur if either the GP or the patient is not reassured in the absence of a specialist opinion.

Benamore and colleagues²¹⁴

This study did not exclusively examine the effectiveness of a rule-out test as the focus is on access to CT of the brain regardless of the function of the scan in the diagnostic decision pathway. However, in the study results there are indications that point to unintended consequences even when the results rule out significant pathology.

The paper reported on a retrospective scan and case note review of all primary care direct-access referrals for head CT scans to a local hospital radiology service over 8 years (1995–2003). All scan reports were reviewed and classified as normal, incidental or significant findings. All reports were cross-referenced for referral to secondary care and the case notes of those with incidental or significant findings were reviewed for secondary care attendance and outcome.

Over the time period, 1645 scans were performed, although only 1403 were available for review (85%). Headaches were the commonest reason for scan ($n = 654$, 46.61%) and, in keeping with the test performing a rule-out function in this context, the least likely to have significant findings on scan (18/654, 0.02%). However, when the referral rates were examined the scan result was a poor discriminator of who received a subsequent referral; not all those with significant findings were referred (13/18, 72.2%, were referred), a significant proportion of patients with normal results still received a referral (78/513, 15.2%), and a similar proportion of those with incidental findings were referred (22/123, 17.9%).

It is difficult to delineate how direct access might impact on referrals from this study; there is no assessment of expected referral patterns in the absence of direct access or what one would expect to happen given the scan results. In this study, it was unclear if the referring GP was given any supplementary detail except the scan results. The categorisation of significant, incidental or normal was not performed as a routine part of the result reporting, suggesting the GPs were left to interpret the scan results themselves. The authors do suggest that referrals may be generated by findings that a specialist would classify as incidental.

The authors noted that for some presenting symptoms (notably headache) a scan is primarily for reassurance against serious pathology as the expected diagnostic yield is low. They suggested that 'It is reasonable to assume that the initial CT, generated by primary care, failed to provide the desired reassurance and potentially generated referrals from incidental findings'.

In terms of the theory under test, this study suggests that when GPs are not provided with guidance on how to interpret scan results, the meaning of the scan results are not always clear to the GP and, consequently, the scan itself may not provide the necessary reassurance to either the patient or the GP. This may then lead to GPs failing to refer patients with significant findings or referring those with incidental findings.

Wong and colleagues²¹⁵

This paper reported on the evaluation of an open-access endoscopy service for patients presenting with dyspepsia in Hong Kong. Primary care physicians could use the open-access service but could also still refer direct to the specialist clinic. An extra morning session was established each week for open-access requests so that waiting times could be compared with the usual route of specialist referral. Waiting times and endoscopy results were compared for the two groups.

During the 3-year study period (1996–9), 978 referrals were made to the open-access service. The mean waiting time from referral to test was 6 weeks. For those referred to a specialist, the mean time to first consultation was 17.5 weeks and a further 4.5 weeks' wait subsequently for the test.

The results of all patients during an 11-month period were assessed and there were no significant differences in diagnostic yield between the open-access tests and those from specialist referral except for an increased incidence of oesophageal varices (a known complication in patients with liver disease) in the specialist group, which may be explained by the preference to refer patients with liver disease to a specialist rather than go through direct access to the test. *Table 4* from the paper is reproduced here as illustration.

The authors examined 166 of the referrals in detail (they did not state how they chose these). Of these 166 referrals, 126 attended for endoscopy as scheduled and 107 of these (88%) responded to a telephone interview at least 8 weeks after their endoscopy. All patients reported that they had received a rapid and definitive diagnosis from the endoscopy. Twelve per cent of the interviewees reported receiving a referral to a specialist clinic or an admission to hospital for management following their scan report and a further 12% were followed up by their family physician. The remaining 76% reported no further contact with services following the results of their test. The authors suggest that open-access endoscopy 'gives a rapid and definitive diagnosis, which satisfies both patients and family physicians'.

Of interest is the 24% default rate, with 40 patients not attending for the direct-access test. When the patients who did not attend were contacted, the reasons given were:

- good response to empirical treatment since the last doctor consultation (40%)
- symptoms got worse and attended private hospitals or private doctors immediately (32.5%)
- went abroad (12.5%)
- a fear of endoscopy (2.5%)
- other personal factors (12.5%).

Direct access in this instance reduced the referrals to secondary care and reduced the time patients waited to receive an initial result. The authors noted that this is not merely about diagnostic yield as 'a negative endoscopic finding is as important as a positive one in the management strategy of dyspepsia'. In terms of the theory under test, the results of the endoscopies in this study suggests that the case mix of patients referred via the two routes was not significantly different; however, many more patients were referred directly to a specialist than to direct access. Furthermore, even when patients were referred to the direct-access clinic, one-quarter did not turn up for their appointment and one-third of these patients sought advice from a specialist, thus short-circuiting direct access to the test. This suggests that GPs, patients or both would prefer access to a specialist rather than GP direct access to a test.

TABLE 4 Endoscopic diagnosis in open-access cases and specialist referral cases (for the period November 1996 to December 1997)

	Open access (n = 367)	Specialist (n = 967)	p-value
Number of normal results (%)	297 (81)	756 (78)	NS
Number of peptic ulcers ^a (%)	66 (18)	162 (17)	NS
Number of gastric cancers (%)	2 (0.5)	1 (0.1)	NS
Number of others ^b (%)	2 (0.5)	48 (5)	< 0.005

a Peptic ulcers included duodenal and gastric ulcers.

b 'Others' included esophagitis, esophageal varices and other minor abnormalities.

Reproduced with permission from John Wiley and Sons, from Wong BC, Chan CK, Wong KW, Wong WM, Yuen MF, Lai KC, Hu WH, Lau GK, Lai CL and Lam SK. Evaluation of a new referral system for the management of dyspepsia in Hong Kong: role of open-access upper endoscopy, *Journal of Gastroenterology & Hepatology*, 15(11), 1251–6, 2000.²¹⁵
© 2000 Blackwell Science Asia Pty Ltd.

Todd and colleagues²¹⁶

This paper reported on a case note review of direct-access referrals for endoscopy for the 12 months of 1995 in one endoscopy unit in Scotland. The authors identified all diagnosed cases of Barrett's oesophagus, gastric ulcer and colonic polyps from the endoscopy reports and reviewed the case notes to ascertain GP referral behaviour. These conditions were chosen as specialist assessment and treatment may be required.

During the study period, 1451 procedures were performed under direct access and 89 of these were diagnosed with one of the identified conditions. The referral rates for all conditions were ascertained.

It is not possible to glean from this paper what impact direct access to endoscopy had on overall demand for secondary care; only the referral behaviour following a positive diagnosis on the test was investigated. The results showed that not all patients with positive diagnoses receiving a referral (*Table 5*).

The paper did not explore the reasons behind the differences in referral behaviour for the identified conditions. The authors did suggest that the results are reassuring as 'the overwhelming majority of patients with adenomatous colonic polyps are referred to a specialist'. However, they also highlighted that 'a substantial minority of patients with Barrett's oesophagus and gastric ulcers are not referred to a specialist', even despite local policy in place for the management of ulcers which suggests a specialist referral.

No advice was given to the GP about appropriate management of the patient following the endoscopy results and the results suggest that the test result on its own is not sufficient for GPs to manage patients appropriately. The authors suggest that:

It may be that general practitioners preconceived ideas of the importance of a condition affect their decision to refer more than the existence of guidelines. We believe that the existence of guidelines alone is not enough, education is a vital partner to ensure good practice.

In terms of the theory under test, this study suggests that direct access to a test does not necessarily allow a GP to identify those patients for whom a referral is appropriate. Even where there are guidelines in place regarding management of particular conditions, if the results are received without advice then GPs may not always follow the appropriate guidelines.

TABLE 5 Referral rate according to condition

Referral rate according to condition				
Condition	Number of cases	Number of cases referred	Referral rate	<i>p</i> -value ^a (chi-squared)
Barrett's oesophagus (all types)	25	14	56%	<0.001
Barrett's oesophagus (with intestinal metaplasia)	18	10	56%	0.08
Barrett's oesophagus (with gastric metaplasia)	7	4	57%	0.25
Gastric ulcer	18	10	56%	0.08
Non-adenomatous polyps	24	3	12.5%	<0.001
Adenomatous polyps	17	15	88%	

^a As compared with adenomatous polyps.

Reproduced with permission from Sage Publications from Todd JA, Zubir MA, Goudie BM, Johnston DA, *Scottish Medical Journal*, 45, pp. 49–50, copyright © 2000 by (Sage Publications). Reprinted by Permission of Sage.²¹⁶

Morgan and colleagues³¹

This study did not explore direct access to tests but can help to shed some light on why some patients may receive a referral even when a test has ruled out serious pathology. The paper reported on a qualitative interview study exploring GP beliefs about patients presenting with headache and the reasons for their referral behaviour. GPs were purposively sampled according to referral behaviour in the preceding year (from no referrals for investigation of headache to three or more referrals). Data were collected from 20 GPs and analysed using a framework approach. The authors reported that recruitment was stopped at 20 GPs as data saturation was reached.

Quotes from the GP participants suggested that some patients would be reassured only if they were seen by a specialist:

... ultimately they say that's all very interesting and thank you very much for the time you've taken, but I still want to see someone ...

GP0604

However, results from the study also suggested that some GPs may be better at instilling confidence in a patient:

I think part of my ability to help people is [a] projection of my confidence, personality, holistic approach, whatever ... to give them confidence in me so when I say: 'it's fine' they believe me.

GP1704

In addition, it may be that some GPs have a lower threshold for managing clinical uncertainty and may make referrals to manage their own anxiety.

I'm, I would say, cautious I suppose. I tend to err on the side of caution and refer more ... that's my nature.

GP0607

... I might refer people just for my own anxiety and fears of missing something, even if they're not anxious. So, really, I suppose it depends on lots of factors.

GP1803

This study suggested that referral despite a negative test result may be explained by patient pressure and the perceived confidence in the opinion of the GP. In these instances a negative result from a ruling-out test may not be sufficient to provide the reassurance needed to satisfy the patient that there is nothing seriously wrong. It may be that for some patients only a specialist can provide the necessary reassurance.

The study also suggested that GPs may vary in their tolerance of diagnostic uncertainty. In these tests a negative result can rule out serious pathology; however, this can mean that presenting symptoms remain unexplained and GPs may refer for a specialist opinion to ensure that they have not missed an important diagnosis.

Ridsdale and colleagues²¹⁷

This was a cohort study by the same research group as that in the above study; it compared 488 patients who consulted their GP for headache over a period of 7 weeks but were not referred and a cohort of patients consulting their GP for headache over 1 year who were referred. The two groups completed questionnaires measuring the disability experienced due to headache, anxiety, depression, illness perceptions and frequency of consultation in the past 3 months.

The authors found that both groups experienced severe disability and that headache significantly affected their functioning. There was no significant difference between the two groups on measures of disability, anxiety, depression or satisfaction with care. Referred patients had consulted more frequently in the 3 months prior to their appointment and were more likely to link more of their symptoms to headaches. They were also more likely to have stronger emotional representations of their headaches and to worry more and were made more anxious by their headaches. The authors argued that referral was not related to the clinical severity of headaches but was associated with higher consultation frequency and patients' anxiety and concern about their headache symptoms.

This study did not explore direct access to testing but the results suggested that patient anxiety does influence referral decisions and this may help to explain why some patients with a negative result on a ruling-out test still receive a referral. *Box 40* provides an interim summary.

BOX 40 Interim summary

A 'rule-out test', such as CT/MRI scanning for headache, is used to rule out the possibility that the patient has a serious pathology. A negative result is expected for the majority of patients and it is often ordered to provide patient reassurance. The theory is that GPs can decide when and for whom this test is appropriate and can interpret the test results in order to reassure the patient that they do not have a serious pathology and confirm that specialist referral is not, therefore, necessary.

In terms of the theories under test, the evidence clearly shows that direct access to rule-out tests can reduce referrals to secondary care. Such tests allow GPs to identify patients without serious pathology that do not require referral to a specialist and in the majority of cases a referral is thus averted. There was no evidence of overuse of rule out tests in GP decision-making. However, there were some inconsistencies where a minority of patients with a negative test result still received a referral to a specialist and some patients for whom the results indicated a referral was warranted did not receive a referral.

The evidence further suggests that the utility of rule-out tests in demand management can be maximised if results are accompanied by guidance from a specialist on interpreting the results and clinical advice on management given the results. Where these guidelines are not present, GPs may investigate incidental findings unnecessarily or fail to refer when necessary. In terms of the 'turf war theory' under test, defined referral criteria and guidance on management attached to the test results ensure specialist control over referral and management decisions, which, in turn, ensures that the appropriate patients are referred to secondary care.

However, some studies showed that even when a scan suggested no serious pathology, a small percentage of patients were still referred to the specialist. Even with clear guidance, some patients received a referral to a specialist following a negative result on a ruling-out test. In cases such as these it is likely that a negative result does not provide the necessary reassurance to the patient and/or the GP.

Differential diagnosis tests

The principal demand reduction theory underlying direct access to differential diagnostic tests is similar to that of the rule-out test. The idea is that through bypassing the need to see a consultant, patients can have speedier access to the test. The results of the test then enable GPs to distinguish between those who can be managed in primary care and those who require an onward referral to secondary care, thus reducing the number of referrals to secondary care.

Differential diagnosis tests sit in the same place as a rule-out tests in some typologies, as they are both diagnostic tests. However, there are important differences between 'ruling-out tests' and 'differential diagnosis tests', which have significant consequences for their impact on demand management. Compared with 'ruling out tests', the expected diagnostic yield in a differential diagnosis test is higher, as the test is performed in circumstances in which a clinical history is suggestive of significant pathology and the test serves to confirm (rather than rule out) this suspicion and/or to distinguish between different definitive diagnoses. Accordingly, direct access to a differential diagnosis test may not reduce referrals to secondary care if most patients undergoing the test turn out to require a referral. In this situation, direct access can reduce waiting time from presentation at the GP to diagnosis and treatment and reduce the number of specialist appointments necessary as the results are to hand at the first outpatient appointment.

As ever, these chains of events are the subjects of debate. One potential unintended consequence may occur if there are not agreed indications between the specialist and the GP for the test in the diagnostic pathway. In this context, direct access to a differential diagnosis test may increase demand unnecessarily for the test and result in a patient receiving an avoidable test.

We move to consider what the evidence has to say about various theories and counter theories with the examples of primary research on MRI for suspected internal derangements of the knee and MRI for the investigation of back pathology: Brealey and colleagues²¹⁸ and three DAMASK (Direct Access to Magnetic resonance imaging Assessment for Suspect Knees) studies.²¹⁹⁻²²¹

These papers reported on the same pragmatic multicentre randomised trial of GP direct access to MRI versus referral to orthopaedic surgeon for patients presenting with suspected internal derangement of the knee. All patients presenting to practices where referral to secondary care was considered were randomised within practice to either the intervention (direct-access MRI) or an orthopaedic surgeon appointment. In the intervention arm, the MRI was used by the referring GP to inform diagnosis and plan subsequent management. For those who were sent for MRI, a provisional referral to the orthopaedic service was also made to ensure that the waiting time from the GP consultation to the orthopaedic appointment was similar for both arms.

To note here is that this design is not pragmatic; it may not mirror actual practice, as it is unlikely that GPs would directly access differential diagnostic tests *and* refer to a specialist at the same time. Indeed, the study design, by making a provisional referral for all those in the direct access arm, implicitly acknowledges one of the potential unintended consequences of direct access, that patients may wait longer to see a specialist, as the GP may wait for the results to be available before making a referral.

Included in the trial was an educational intervention to support the dissemination of the clinical guidelines with the aim of improving GPs' knowledge of MRI use. GPs were incentivised to attend training, although attendance was not a condition of participation in the trial. A further educational message was attached to the radiologists report to remind the GP that subsequent management should depend on both the MRI results and the clinical findings. Thus, both referral and management guidance were provided to GPs.

Only 44% (285 out of 647) of practices approached to take part accepted the invitation. This low participation rate perhaps suggests that the majority of GPs preferred to make their own decisions about which referral pathway they chose for patients, rather than have those dictated by the process of randomisation. Those participating were likely to be further from a hospital. Over a 2-year period

553 patients who were eligible and gave consent were recruited from 163 practices, 279 were randomised to receive the intervention (MRI) and 274 were randomised to the control (referral to orthopaedics). Follow-up was 85% at 24 months.

The authors reported minimal difference between the groups in terms of orthopaedic consultations; 230 (82%) patients in the MRI group and 236 (86%) in the orthopaedic group. However, this adds little explanatory value as the control arm was orthopaedic referral where the majority of the 14% who did not have an appointment were patients who failed to attend following allocation. The authors noted in the discussion that although GPs had the option to cancel the orthopaedic appointment on the basis of the MRI results, they rarely did so with only 27/279 (10%) appointments cancelled on the basis of the scan results. In terms of the theory under test, the design of the study meant that it was difficult to draw conclusions about whether or not direct access to MRI reduced referrals to secondary care.

The referring GP was asked to record the main diagnosis and how confident they were in this diagnosis at baseline and these questions were repeated on receipt of the radiologists report in the MRI arm and the letter from the orthopaedic specialist in the referral arm. *Table 2* from the paper is reproduced here (*Table 6*) to show that there was no difference in the likelihood of change in diagnosis between the two arms, although GPs reported an increased confidence in their diagnosis in the MRI arm. The study was not designed to shed light on why this might be. We could hypothesise that GPs in the MRI arm were perhaps more likely to attempt a definitive diagnosis prior to the sending patients to the MRI as they would have to interpret the MRI scan and make subsequent decision, rather than the consultant. Receiving confirmation about their initial diagnosis was perhaps more likely to increase their confidence. In contrast, perhaps GPs in the consultant arm felt less pressure to make a definitive baseline diagnosis as they expected this to be made by the consultant.

This study showed that direct access in this context had little impact on referral decisions, as the majority of patients received a referral following their MRI. It also pointed to some important differences in subsequent treatment, which give us some interesting evidence towards exploring the unintended consequences of direct access. There was a difference in the number of patients having an arthroscopy (a form of keyhole surgery used to both diagnose and treat joint problems): 40.1% in the direct-access group compared with 28.1% in the orthopaedic group. Interestingly, this was classified as a diagnostic arthroscopy in 39.5% of the patients in the orthopaedic arm, compared with 18.5% in the direct access arm. It is not reported in the papers how many patients in the orthopaedic referral also had a MRI scan; however, the economic

TABLE 6 Changes in GP diagnoses and diagnostic confidence

	Original randomisation		Between-group change ^a
	MRI referral	Orthopaedic referral	
Diagnosis altered, <i>n</i> (%)			
Yes	170/279 (61)	165/266 (62)	$\chi^2 = 0.07, p = 0.79$
No	109/279 (39)	101/266 (38)	
Change in diagnostic confidence, <i>n</i> (%)			
Increased	168/264 (64)	80/249 (32)	$\chi^2 = 51.43, p < 0.001$
No effect	77/264 (29)	129/249 (52)	
Decreased	19/264 (7)	40/249 (16)	
Within-group change ^b	$p < 0.001$	$p = 0.006$	

a The χ^2 test.

b The Wilcoxon-signed rank test for within-group change in diagnostic confidence between trial entry and follow-up.

Reproduced with permission from the RCGP, from DAMASK.²¹⁹

evaluation gave a much lower mean resource use cost for MRI in the orthopaedic referral group – £128.96 versus £326.79 in the direct-access arm – which suggests that a significant number of patients referred direct to orthopaedics did not have a MRI scan as part of their diagnostic testing.

This suggests that a MRI in the case of suspected internal derangement of the knee is not necessarily the diagnostic test of choice by the specialist, which may explain why more arthroscopies in this arm were classified as diagnostic. It also shows that patients referred directly to a specialist had fewer arthroscopies and fewer MRI scans than those in the direct-access group. In terms of the theory under test, this implies that specialists are more likely to have the clinical skills necessary to reach a differential diagnosis without the need for scans or tests and that referral first to a specialist may prevent the use of unnecessary scans. In contrast, a clinical pathway where a MRI is performed first and then followed by a specialist referral may mean that some of those initial MRIs were performed unnecessarily.

Even though the study attempted to control for wait for orthopaedic appointment, the time from randomisation to first appointment (median 93 days, range 60–183 days vs. median 79 days, range 54–168 days) and subsequent treatment (median 337, range 219–458 days vs. median 294 days, range 174–468 days) was longer in the direct access to MRI group. Interestingly, there was a small but significant improvement in health-related quality of life (HRQoL) in the direct access to MRI group.

In this study, direct access to MRI seems to have had little impact on referral decision-making, as most patients also had a specialist appointment. Unlike in a rule-out test, the test had a high diagnostic yield, and although it may have distinguished possible diagnoses and subsequent treatment, in most cases, a positive result necessarily required specialist referral. Indeed, the authors note 'It is likely that MRI confirmed what GPs already knew, which was the need for a referral to an orthopaedic specialist.' However, an unintended consequence of direct access was to increase the usage of MRI scanning, as not all referrals direct to a specialist underwent a scan. In terms of the theory under test, it suggests the direct access to MRI scanning may increase patient waiting times and result in unnecessary tests, as GPs do not have the clinical skills to definitively diagnose joint problems. For the GP, a MRI scan may simply confirm their original diagnosis and that referral to a specialist is required. In contrast, specialists have the clinical skills to diagnose joint problems in some instances without the need for tests and direct referral to a specialist may, thus, prevent the use of unnecessary tests.

Watura and colleagues²²² (plus responses and rejoinders in the same journal)^{223–225}

This was a brief paper that reports the results of an audit of direct-access MRI imaging of the knee to one radiology department in a Welsh hospital. A direct-access service was introduced in January 1993 and the audit compared referrals via GP direct access up to May 1994 with referrals made via orthopaedic outpatients in the year preceding the introduction of the service.

The diagnostic yield from the scans did not differ significantly between the two groups, although the scan result alone was a poor discriminator of who had a subsequent referral to the orthopaedic clinic. The results table from the paper is reproduced here as *Table 7* for illustration.

The authors suggested that direct access to MRI reduced the number of referrals to specialists, as most with normal scans or minor abnormalities were not referred; 76 out of 165 (65%) were subsequently referred to the orthopaedic clinic. In addition, they suggest that direct access did not increase the workload of MRI as all patients, if referred to the orthopaedic surgeon, would have had a scan performed anyway. An additional bonus, from the authors' perspective, is that the availability of the scan results at consultation for those who were referred allows for more streamlined planning of care.

In terms of our programme theory, the suggestion is that direct access allows GPs to make decisions higher up the decision hierarchy. However, the scan result is a poor discriminator of who got a referral; not all abnormalities were referred and some normal scans received a referral. This suggests that referral decisions were not made solely on the basis of the scan result.

TABLE 7 General practitioner vs. orthopaedic referrals for MRI

Patients	General practitioner referral	Orthopaedic referral	p-value
Number referred	165	470	
Number (%) males	115 (70)	331 (70)	NS
Mean (range) age (years)	37 (9–83)	32 (2–84)	<0.001 ($\chi^2 + 2$ test)
Mean wait for scan (days)	19	14	<0.001 (Student's t-test)
Findings, n (%) of patients			
Cruciate/meniscal tears	73 (44)	213 (45)	NS
Followed up	65 (89)		
Referred to orthopaedic clinic	55 (75)		
Normal	43 (26)	134 (29)	NS
Followed up	37 (86)		
Referred to orthopaedic clinic	5 (12)		
'Other' abnormality	49 (30)	123 (26)	NS

NS, not significant.

Reproduced from Magnetic resonance imaging of the knee: direct access for general practitioners, Watura R, Lloyd DCF, Chawda S. *BMJ*, vol. 311, p. 1614, © 1995²²² with permission from BMJ Publishing Group Ltd.

The authors assumed that all referrals would have a MRI scan anyway. The DAMASK trial,^{219–221} discussed above, shows that this is not necessarily the case; MRI utilisation in those referred directly to a specialist was lower than that in the direct-access arm.

The role of direct-access MRI in this instance is shown to be the subject of disagreement when we investigate the responses to this paper on the letters pages of the journal. This debate is useful evidence to show that real divisions exist between generalists and specialists in their understanding of the value of direct-access MRI in the diagnosis of knee derangement. Responses from orthopaedic specialists suggest that they have access to knowledge and skill sets that allow them to use such tests more selectively than a generalist.

We believe that magnetic resonance imaging of the knee should be used selectively. Patients who have obvious clinical symptoms and signs of a meniscal tear do not need such imaging before having an arthroscopy. Imaging may, however, be useful to avoid a diagnostic arthroscopy . . . The selective use of magnetic resonance imaging helped us avoid unnecessary operations, plan surgical procedures and save money. These benefits would be lost if demand for inappropriate scans increased. We do not believe that magnetic resonance imaging should be used as a blanket test; it would be a shame if such imaging replaced the need for accurate clinical skills

*Reproduced with permission from Jeremy Southgate, from Southgate J, Thomas N. Direct access magnetic resonance imaging of the knee for GPs. Magnetic resonance imaging should be used selectively. *BMJ* 1996;**312**(7034):849; author reply 50²²⁵*

An orthopaedic consultant with an interest in the knee should . . . sort out most patients with knee pain on clinical grounds . . . for some clinical diagnoses . . ., magnetic resonance imaging has little proven clinical value. Not all general practitioners will understand the implications of the findings of magnetic resonance imaging.

*Mackenzie et al.*²²⁴

One consultant orthopaedic surgeon suggested that the assumption in the paper that all specialist consultations would necessitate a MRI was misplaced and, therefore, the cost savings of open access were overestimated. The author also pointed to the importance of clinical examination skills and, therefore, a need for more specialist orthopaedic skills.²²³

In terms of the theory under test, these letters point to ‘turf wars’ between GPs and specialists over who possesses the skills and knowledge necessary to make a diagnosis in this context. Indeed, the original authors’ reply to these letters highlighted the issue of professional boundaries: ‘It would be absurd to suggest that general practitioners are capable of prescribing expensive drugs with complex interactions and side effects but not capable of selecting patients for magnetic resonance imaging of the knee and using the results appropriately’.²²⁶

Chawda and colleagues²²⁷

This was a brief report of an observational study of the MRI service at Cardiff Royal Infirmary for patients referred for a lumbar MRI for sciatica or suspected spinal claudication. All hospital outpatient referrals (i.e. those initiated by consultants) and those directly referred by GPs were reviewed for a 2-year period (January 1993–December 1994). GPs were contacted in writing and by telephone to document subsequent management. This study design, using retrospective data collection, is open to the possibility of recall bias on the part of the GP and, as such, the findings should be treated with caution. Nonetheless, the papers provide further evidence with which to test the theories under investigation for a different condition.

Over the time period, GPs referred 459 patients to the direct-access service, compared with 435 patients who came for MRI via hospital outpatients. The differences in diagnosis and subsequent management between the two groups were shown in *Table 1* in the paper, which is reproduced here as *Table 8*.

TABLE 8 General practitioner vs. hospital outpatient referrals and GP patient management following MRI result

	GP referral for MRI, n (%)	Outpatient referral for MRI, n (%)	GP vs. outpatient referrals, p-value	GP management following scan	
				Hospital referral, n (%)	Conservatively managed, n (%)
Sex: male	244 (53)	234 (54)	NS (a)		
Female	215 (47)	201 (46)	NS (a)		
Mean age (years)	42	44	p < 0.025 ^b		
Mean waiting (days)	19	13	p < 0.001 ^b		
1a. Disc herniation corresponding with symptoms	182 (39.8)	165 (37.9)	NS (a)	113 (72.9)	42 (27.1)
1b. Disc herniation not corresponding with symptoms	46 (10.1)	55 (12.6)	NS (a)	20 (47.6)	22 (52.4)
2. Spinal stenosis	20 (4.4)	20 (4.6)	NS (a)	11 (57.9)	8 (42.1)
3. Spondylosis/degeneration	114 (24.9)	77 (17.7)	p < 0.01 (a)	27 (27)	73 (73)
4. Normal	57 (12.5)	68 (15.6)	NS (a)	17 (34)	33 (66)
5. ‘Other abnormalities’	38 (8.3)	50 (11.5)	NS (a)	12 (38.7)	19 (61.3)
Total	457	435		200	197

NS, no significant difference.

a t-test.

b Chi-squared test.

Reproduced with permission from the RCGP, from Chawda and colleagues.²²⁷

There was no significant difference between the groups in the percentage of scans reported as normal: 12.5% from GP referral and 15.6% from outpatient referral. There were no significant differences in the percentage of specific classifications of abnormalities reported except an increased incidence of spondylosis/degeneration in the GP referrals. Overall, this suggests that the case mix of patients referred for MRI by GPs and specialists in this context was broadly similar.

The results of the GP follow-up (86.7% response rate) suggested that 49% of those referred for MRI were conservatively managed subsequently in primary care. However, the test result was not a good discriminator of who was subsequently referred for specialist care, which suggests that the test alone was not driving the decision regarding subsequent management.

The authors concluded that 'prompt investigation allows GPs to make appropriate and informed management decisions. For those patients who are referred, the availability of the MRI scan result at consultation permits rapid assessment and avoids a follow-up appointment to discuss results, as would be necessary if the MRI were requested at the first outpatient attendance. The availability of a scan report could also assist hospital specialists in prioritizing outpatient appointments and reviewing the urgent cases more quickly' [reproduced with permission from the RCGP, from Chawda, S. J., Watura, R. and Lloyd, D. C. (1997) Magnetic resonance imaging of the lumbar spine: direct access for general practitioners, *British Journal of General Practice*, 47(422), 575–6.]

In terms of the theory under test, direct access to MRI of the lumbar spine appears to have some impact on subsequent referrals, with the GP using the results to direct subsequent management, which suggests that they are making decisions which may have previously been deferred to specialist via a referral. This may be because there is genuine uncertainty on the part of GPs and specialists prior to use of the test and the test clearly differentiates between diagnoses that can be managed in primary care and those that need specialist management. The authors also commented on the clear guidelines for the use of MRI in patients with acute back pain and concluded that 'direct access for MRI of the lumbar spine is a viable proposition if driven by protocols and used appropriately'.

Anderson and colleagues²²⁸

A study of diagnostic accuracy in 2-week-wait colorectal cancer referrals also suggested that specialists may have access to more sophisticated diagnostic tools which can aid in establishing diagnoses. In this study, the diagnostic accuracy of GP referral letters was compared with the accuracy of the patient's first surgical assessment. Surgical assessment showed significantly greater diagnostic accuracy and this was explained almost entirely by the use of a rigid sigmoidoscopy during rectal examination; rigid sigmoidoscopy made a significant contribution to managing 12% of referrals. The authors reported that rigid sigmoidoscopy is standard practice for outpatient assessment of suspected colorectal cancer but is rarely performed by GPs. They suggested that a referral for surgical assessment was, therefore, preferable to a straight-to-endoscopic test pathway, as some patients may be denied the opportunity for immediate diagnosis. *Box 41* provides an interim summary.

Clinical indicator tests

A clinical indicator test encompasses screening for disease, monitoring the progress of disease and/or monitoring an individual's response to treatment. Although these tests are used at different points in the patient's disease trajectory, the common theory underlying their use is to identify where the patient is on this trajectory in order to guide subsequent management and treatment. As screening tools, these tests aim to inform a decision about whether to intervene now or whether to continue to monitor the patient, and who is best placed (the GP or the consultant) to undertake this intervention. They are used when there is a high suspicion of a particular diagnosis from the patients' signs and symptoms and the test is used to confirm this and also to quantify the stage or severity of the condition in order to identify the patient's location on a disease trajectory.

BOX 41 Interim summary

The theory underlying direct access to differential diagnostic tests is that through bypassing the need to see a consultant, patients can have speedier access to the test. It is expected that the results of the test enable GPs to distinguish between those who can be managed in primary care and those who require an onward referral to secondary care, thus reducing the number of referrals to secondary care. However, our evidence synthesis suggests that direct access to tests such as MRI or endoscopy does not reduce referrals to secondary care and may result in unnecessary tests being performed and increased patient waits for treatment. This is because specialists have the necessary clinical skills and/or access to alternative and preferred diagnostic techniques that enable them to reach a differential diagnosis without the need for the direct-access test. Direct access to MRI appears to confirm what GPs already know: that the patient requires referral to a specialist.

In contrast, there is some evidence to suggest that direct access to MRI for suspected lumbar sciatic or claudication does enable GPs to distinguish between patients who can be managed in primary care and those who require a specialist referral, and did reduce referrals to secondary care. The difference in the outcome patterns between the two tests may be explained by the extent of agreement between specialists and GPs on the role and value of the test in the patient pathway and whether or not the test is designed to provide a differential diagnosis between conditions that require management in primary or secondary care. Although MRI scans and endoscopies can be interpreted by GPs, with guidance, they are not the initial tests of choice for specialists in diagnosing internal derangement of the knee or the presence of bowel cancer. In contrast, MRI for back problems appeared to be used with similar frequency by specialists and GPs in reaching a differential diagnosis in patients with back problems.

Furthermore, while MRI for knee problems may enable GPs to reach a differential diagnosis in these cases, the vast majority of these diagnoses are still perceived by GPs to require management by a specialist. In contrast, MRI for back problems does appear to distinguish between diagnoses that can be managed in primary care and those that require specialist referral. This evidence synthesis clearly shows that not all tests are created equal and that the function of the test in the patient pathway and differences in the practices, skills and experience of GPs and specialists influences whether the test produces intended or unintended consequences.

By providing GPs with direct access to these tests, it is assumed that GPs have the necessary skills to decide when and for whom such a test is warranted and then have the skills and confidence to interpret the results and decide which patients can continue to be managed in primary care and which patients require referral. As a result, more patients will be managed in primary care without the need for a specialist opinion and only those requiring specialist management are referred to secondary care.

As ever these programme theories have been the subject of debate on now-familiar grounds. For instance, a potential unintended consequence of providing GPs with direct access to clinical indicator tests is that GPs misinterpret the test results and refer patients inappropriately or fail to refer patients when a referral is warranted by the test results.

The literature gives us a number of examples in this typology: bone mineral density scanning for suspected osteoporosis, cardiac arrhythmia monitoring and echocardiography for suspected heart failure.

Dhillon and colleagues²²⁹

This paper was a RCT of GP direct access to dual X-ray absorptiometry (DXA) scanning versus hospital referral for specialist consultation in women with suspected osteoporosis and its impact on clinical decision-making. DXA scanning is used to diagnose osteoporosis or to assess the risk of osteoporosis developing. The results are used to guide the initiation of therapy. The authors stated that, at the time of writing, access to NHS-funded DXA scanning and osteoporosis services in the UK was variable. At the time of the study, access to DXA scanning was only via referral to hospital specialist. Direct access to DXA scanning would allow GPs to use the results to manage subsequent treatment and care without the need for specialist referral.

No referral guidelines were given to participating GPs; the authors assumed that they would follow one of the established guidelines, such as those published by the National Osteoporosis Society. GPs in participating practices selected patients who they felt warranted DXA scanning, and completed a structured referral form. On this form they were asked to state their management plan if access to a DXA scan was not available. Participants were then randomised to the study arm – direct access to DXA scan with a brief scan report and one of four management recommendations – or the control arm of DXA scan and routine rheumatology clinic consultation.

To note here is that this is not a comparison of usual care versus direct access; it is not clear if hospital referral would be sought in the absence of direct access. Many of the patients in the study had commenced treatment for osteoporosis on the basis of clinical history alone and it is unclear if specialist referral would be sought for all patients even without access to DXA scanning. Furthermore, the study design does not compare direct access to the scan with specialist referral. The aim of the study was to test the effect of two different routes to DXA scanning; all patients in the study received a DXA scan so the comparison is between whether this is accompanied by a scan report with management recommendations to the GP or a specialist consultation.

A total of 330 participants were randomised: 165 in each arm. The groups were similar at baseline in terms of demographics, reasons for referral, pre-referral treatment, health state and GPs' anticipated post-scan management. Significantly more participants in the control arm were recorded as having clinical risk factors and/or prior low-trauma fracture. On scan, there was no significant difference in bone mineral density (BMD) between the groups.

In the control arm, the GP received a detailed clinic letter giving the DXA result and a written opinion on management given by the specialist who saw the patient in the clinic. In the study arm (direct-access route), the GP received the DXA result and one of four management recommendations:

- lifestyle advice
- consider measures to prevent osteoporosis
- measures to treat established osteoporosis
- specialist referral.

The primary outcome measure was change of management following the DXA scan. Sixty per cent of all study participants had their management changed after the scan report; the proportion was similar across the arms. At 1-year follow-up there was no significant difference between groups in outcomes. The main difference was in waiting times, as those in the control arm had a longer time from referral to GP receipt of the report.

Resource use was assessed at 1-year follow-up and there was a significant difference in subsequent referrals to a specialist, with 17% in the direct-access arm and 9% in the control arm receiving a referral to a specialist rheumatology service. It is not clear if the referrals in the direct-access arm followed the advice with the scan results and this was not reported in the study.

The authors concluded that 'DXA scanning influences GPs' management of osteoporosis in two-thirds of patients'. However, there was no difference in GP management depending on whether the GP received the scan results and management plan or whether the GP received the scan result and the patient also saw a consultant. They also noted that 'The clinical outcome is similar following direct access or the current model of hospital referral, but direct access is more economically efficient'.

In terms of the theory under test, in this study, the GP management of patients with suspected osteoporosis following direct access to DXA scanning accompanied by a management plan from a specialist was similar to that when patients were referred to a specialist and scan. However, any necessary changes to treatment could be enacted earlier by the GP following direct access, as the results were available more promptly

than from the specialist referrals. In addition, direct access had an impact on specialist resource use, as only 17% of patients referred to direct-access DXA scanning had a subsequent referral to a specialist. This is in comparison with the control arm (referral to a specialist), where all patients had an initial consultation and a further 9% had subsequent appointments with a specialist.

This suggests that direct access can both speed up patient access to appropriate treatment plans and reduce the resource use of specialist clinics. It is difficult to quantify any potential savings in resource use of specialist clinics as it is not clear that all patients would have been referred to a specialist had the study not been recruiting. In this study, the referring GP did not have to interpret the results: the scan report was accompanied by a suggested management plan that stipulated if treatment should be commenced and identified which patients required specialist referral.

*Sim and colleagues*²³⁰

This paper was an observational study of the utilisation of an open-access service to bone density scanning (DXA) over a 12-month period. All referrals during the study period were assessed against agreed clinical referral criteria, supplied on a referral form completed by the GP at the point of referral, although even those who did not fit the agreed criteria still underwent the DXA scan. Following the direct-access scan a report was sent to the referring GP including comments from a consultant specialising in osteoporosis – this report gave a clinical judgement based on the DXA results and other clinical information available on the direct-access request form. The report gave management recommendations in one of four categories:

1. normal
2. borderline BMD with no treatment required
3. osteoporosis and treatment required
4. significant abnormal DXA result; refer to bone clinic.

All GPs who used the open-access service were sent a questionnaire regarding the first patient they referred during the study period.

During the study period, 586 patients were referred to the open-access DXA service. Of these, 26 were excluded from the study as the referral was for monitoring of treatment. Of the 560 included patients, 473 (84%) of referrals were judged as appropriate against the clinic referral criteria, suggesting that the GPs were reasonably good at following the referral guidelines. However, a number of referrals that did not match the criteria were also diagnosed positively from their scan, indicating that where GPs did not follow the guidelines this may well have been for good clinical reasons. In total, there were 229 (41%) of patients who had osteoporosis diagnosed by DXA. *Table 2* from the paper is reproduced here as *Table 9* to illustrate the results.

TABLE 9 Diagnosis of osteoporosis in referrals that met the referral criteria and those that did not meet the referral criteria

	Osteoporosis cases diagnosed and treatment required	No osteoporosis
Referral according to CRC	212	261
Referral which did not fulfil the CRC	17	70

p -value by χ^2 test is < 0.001

Sensitivity of CRC in predicting osteoporosis = $212/(212 + 17) = 93\%$

Specificity of CRC in predicting osteoporosis = $70/(70 + 261) = 21\%$

Reproduced with permission from Sim MFV, Stone M, Johansen A, Ho P, Pettit RJ, Evans WD. An analysis of an open access general practitioner bone densitometry service, *International Journal of Clinical Practice* 2004;**58**, 300–5.²³⁰
© 2004 Blackwell Publishing Ltd.

The consultant recommendations with the scan report identified 19 (3%) patients who required referral to the outpatient clinic, with the remainder suitable for management in primary care:

- 324 (58%) normal
- 7 (1%) borderline BMD with no treatment required
- 210 (38%) osteoporosis and treatment required
- 19 (3%) significant abnormal DXA result; refer to bone clinic.

Questionnaires were returned by 119 (77%) of the referring GPs, who reported that in 101 (85%) of the cases they would have referred to hospital outpatients if direct access had not been available. The scan and report identified 76 (64%) patients for whom treatment was not indicated. The actual GP management of these patients corresponded to the recommendations. In the case of confirmed diagnosis of osteoporosis, 37 of the 43 diagnosed patients (86%) were identified as suitable for management in primary care. The reports from the direct-access scan recommended referral to a specialist in only six patients (5%). However, the results show that 10% of patients had a referral, with six patients for whom management in primary care was indicated being referred for a specialist appointment. *Table 3* from the paper is reproduced here as *Table 10* for information.

In terms of the theory under test, this adds some explanatory value to the RCT cited above,²³¹ as the authors of that RCT suggested that many of the patients would have been referred to a specialist in the absence of direct access and, therefore, direct access can reduce the referrals to secondary care. In this study, direct access to DXA scanning enabled GPs to distinguish between patients who did not have disease and patients who had disease but could be managed in primary care, and thus reduce referrals to secondary care. Perhaps the reason for this was that the scan result was accompanied by guidance on the meaning of the scan result and guidance on management from the specialist.

In terms of professional boundaries, there was minimal transfer of decision-making responsibilities between the specialist and the GP beyond the request for a test itself. In this example the decision-making remained with the consultant specialist, albeit without him or her seeing the patient directly:

A DXA report with comments by a consultant specialising in osteoporosis was sent to the referring GP. The diagnosis on the report is a clinical one made by the specialist based on the DXA result [and] age, sex and any other relevant clinical information available on the request forms.

This suggests that the clinical information provided by the GP on the referral form was also important in enabling the consultant to formulate a diagnosis and management plan along with the scan result without seeing the patient. There is also some evidence of short-circuiting, with GPs still making referrals for a

TABLE 10 Recommendations of DXA reports and GP action

DXA reports (n = 119)	Specialists' recommendations (%)	GPs' action upon the DXA report (%)
Normal DXA report, patients reassured	75 (63)	75 (63) ^a
Borderline osteoporosis, monitor patient with repeat DXA later	1 (1)	1 (1)
Osteoporosis confirmed, treatments initiated as recommended	37 (31)	31 (26)
Referral to metabolic bone clinic	6 (5)	12 (10)

^a HRT, GPs were able to stop HRT in two patients because of their normal DXA. Reproduced with permission from Sim MFV, Stone M, Johansen A, Ho P, Pettit RJ, Evans WD. An analysis of an open access general practitioner bone densitometry service, *International Journal of Clinical Practice* 2004;**58**, 300–5.²³⁰ © 2004 Blackwell Publishing Ltd.

minority of patients with a positive diagnosis, even though the consultant has indicated that this could be managed in primary care. The study made no attempt to explain the source of this unintended effect.

Stock and colleagues²³²

This paper reported on a study that did not examine the effect of direct access to DXA scanning on referral behaviour, but is useful for exploring the impact of the format in which GPs receive test results. The authors hypothesised that a low use of DXA scanning among American primary care physicians may be because they had difficulty comprehending the often brief technical reports.

All primary care physicians ordering DXA scanning over a 7-month period were enrolled, and when one of their patients underwent the procedure, the physician was randomly assigned to receive either the usual short technical report or a long narrative clinical report for this and all subsequent requests. The clinical reports were written by a specialist endocrinologist. Of note here is that primary care physicians in this study included general internists, family practitioners and gynaecologists.

Physicians were interviewed by telephone after receiving at least two reports in their assigned format. Once physicians in the short-report arm had been interviewed, they received long reports for any subsequent scans and were interviewed again after having received at least two long reports. The interviews explored the physicians' understanding of bone densitometry and asked specific questions about the individual patients. The questions asked were not specified in the paper, although a standard pro forma was used.

A total of 68 physicians were enrolled, although 11 of these were excluded from the analysis as either they had moved out of area before follow-up or they did not meet the study protocol (the paper does not state what this protocol was). The final sample included 57 physicians (35 in the short-report arm and 22 in the long-report arm) who between them ordered 894 tests over the study period.

On interview, physicians who received the long report were more likely to understand the role of bone mineral density in the diagnosis of osteoporosis: 86% of physicians in long-report arm compared with 30% in the short-report arm. The authors report trends towards fewer formal referrals to specialists in the long-form arm (3% of patients vs. 11% of patients), although this did not reach statistical significance. All of the 20 physicians who received the long reports after initially receiving the short report preferred the long-report format.

In terms of the theory under test, this study suggested that physicians may find technical test results difficult to interpret and that this may impact on subsequent patient management. Physicians preferred clinical reports and this seemed to have had some impact on referral behaviour.

Skipsey and colleagues²³³

This paper reported the outcomes from a pilot study of an open-access service for cardiac arrhythmia monitoring from a regional centre in Scotland. The aim was to improve access for patients to monitoring and reduce unnecessary referrals to the cardiology clinic in low-risk arrhythmia patients. 'The definition of "high risk" was left to the discretion of the primary care physician or the consultant cardiologist', so there did not appear to be any guidelines to facilitate the referrals.

Data were collected over a 29-month period (2008–10) on all consecutive patients referred for cardiac arrhythmia monitoring. Patients came from two sources: GP direct access and 'redirected' consultant referrals where the GP had made a cardiology referral but the consultant identified the patient as eligible for the monitoring without clinical review.

A total of 280 patients were referred for monitoring, although 41 were excluded from the analysis (39 cancelled/did not attend and the equipment failed in two cases). Of the 239 patients evaluated, 165 (69%) came through direct access and 72 (31%) came via the 'redirected' route. However, it is not known how many GP referrals were made directly to specialists during this period and, of these, what percentage of patients were then referred for arrhythmia monitoring.

Of the 239 patients, 230 (96.2%) were returned directly to be managed by the GP. Nine patients were reviewed by the consultant cardiologist and three of these were returned to GP care without clinical review. However, follow-up data showed that 50 (21%) of the patients had a subsequent cardiology outpatient appointment. It would seem from the paper that the length of follow-up was not consistent for every patient (days to follow-up ranged from 216 to 1119) which would suggest that the data on subsequent outpatient usage should be treated with a little caution, although it is perhaps likely to be an underestimate. Of these 50 patients, 27 of them were returned to GP care after one clinic appointment. *Figure 2* from the paper, reproduced here as *Figure 32*, illustrates the patient flow in the study.

The authors conclude that:

there were no adverse events and thus the direct access service appeared to achieve its aim of safely reducing unnecessary referral of low risk patients with palpitations to the cardiology clinic. However, despite the availability of the service, many GPs still referred low risk patients directly to consultants.

The authors advocated both greater promotion and more effective 'policing' of this service in order to ensure that the benefits of the service were fully realised.

The authors note that 19 patients were referred to clinic after initially being cleared for GP management. They suggest that, as all these patients were discharged after a single clinic visit, no serious conditions were missed through the provision of open access.

In terms of the theory under test, this study suggests that despite the provision of direct access, there was evidence of GPs bypassing the service by referring directly to clinic. The authors suggested that better promotion of the service may increase the usage. However, the paper noted that no guidelines were given to aid the GP in deciding which patients were 'low risk' and, therefore, eligible for the service. It could be that without clear guidelines some GPs did not feel confident to make that decision and preferred to defer to the clinic. The results also suggested that many of the GPs made subsequent referrals even though the test had indicated the patient could be managed in primary care. The study made no attempt to explore why this might be the case.

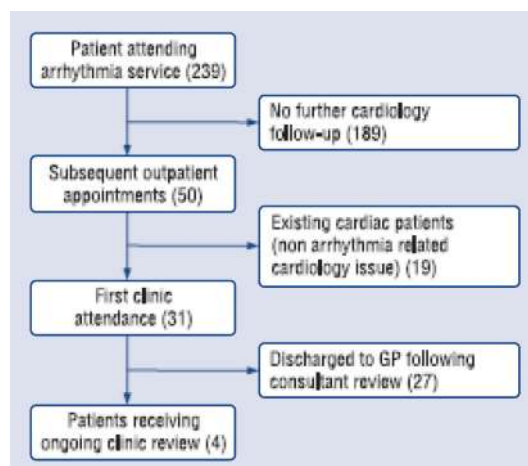


FIGURE 32 Patient outcomes following arrhythmia monitoring. Reproduced with permission from via Medica from Skipsey DA, Dawson FM, Breen C, Leslie SJ. Evaluation of a direct access cardiac arrhythmia monitoring service. *Cardiology Journal* 2012;19:70–5.²³³

van Gorp and colleagues²³⁴

This paper reported the results from a prospective cohort study in the Netherlands. GPs were given direct access to echocardiography and asked to complete a standardised request form for all patients, including an indication of management if direct access was not available. A telephone call to the GP following receipt of the results collected information on subsequent management. As no control group was utilised, the findings of the study need to be interpreted with caution, as it is not clear how patients would have been managed in the absence of direct access. Estimates of how many referrals to secondary care were avoided were based on GP response to a hypothetical question of how they would have treated patients without direct access. Despite this, the study provides a further layer of evidence to test the theories under review.

A total of 164 patients were referred to the direct-access service for echocardiography for suspected cardiac conditions. Of these, 156 patients consented to be included in the study, although one of these was ineligible as he was under age. The results are reported for only the 105 of these patients who had a complete referral form and interview data.

If the GP had not had access to the service then the referral request indicated that 97/105 (92%) would have been referred to a specialist. As a consequence of direct access, only 36 (34%) were actually referred.

Our study shows that open access echocardiography may lead to significantly less referrals to the cardiologist (34% vs. 92%, $p < 0.001$). GPs were able to manage more patients in primary care ($p < 0.001$) and in most patients (82%) GPs found the echocardiogram of benefit for decision making.

van Gorp and colleagues²³⁴

In relation to the theory under test, this study suggested that direct access to echocardiography reduced the number of referrals to secondary care. In terms of professional boundaries, consultants retained a significant amount of control over the decision-making process. Although GPs maintained responsibility for the management of patients, all tests were evaluated and reported on by a cardiologist, with those assessed as needing referral given specific advice to do so by the evaluating cardiologist. Indeed, in the discussion section, the authors noted:

At first the evaluating cardiologists did not . . . add any advice to the echocardiography results . . . Because other studies have suggested that GPs strongly appreciate the specialist's advice, the cardiologist was asked to provide advice with every echocardiogram.

van Gorp and colleagues²³⁴

This suggests that direct access to echocardiography may reduce referrals to secondary care and perhaps works best when the test result is accompanied by management guidance from the specialist to support GP decision-making. The paper did not explore the impact of increasing the amount of information given to GPs with the test results and so we are unable to comment on how much advice might be optimal.

van Heur and colleagues²³⁵

This paper reported on a retrospective review of management of patients following direct-access echocardiography for suspected heart failure; for these patients, echocardiogram was the gold-standard diagnostic test.

Participating GPs received short training on indications and restrictions for echocardiograms and on interpreting the results. The reports contained a summary of the results from a cardiologist and, if abnormalities were noted, a suggested management plan of four options:

- refer the patient to the cardiologist
- start or change medication
- repeat the echocardiogram within a few years
- consider endocarditis prophylaxis.

All echocardiogram requests and reports during a 4.5-year period (2002–7) were reviewed. In addition, a questionnaire was sent to all GPs who referred to the service to establish subsequent management.

During the study period, 625 patients were referred and GP questionnaires were returned for 517 (82.7%) patients. GPs varied in their adherence to the cardiologist's advice; *Table 11*, which is a reproduction of *Table 4* in the original paper, gives details.

The authors stated that the purpose of open-access echocardiogram is to lower the threshold for diagnostic testing in patients with suspected heart failure. They suggested that the referral rate in this study was low, compared with expected numbers based on population data. The authors suggested that this may be because GPs are too strict in their test ordering behaviour, or alternatively it may be that patients in whom heart failure is suspected are being referred to the specialist when they would be suitable for open access.

The authors concluded that 'open access echocardiography is a popular service to detect patients with heart failure and patients at risk for developing heart failure. Overall, GPs used the open access echocardiography service efficiently (i.e. with a high chance of relevant pathology), but efficiency decreased slightly over the years . . . Further specification of the indications for open access . . . might improve the service and make it clearer for GPs when to use it' (reproduced from van Heur and colleagues²³⁵ under the terms of the Creative Commons Attribution Licence 2.0).

TABLE 11 Management by GPs after open-access echocardiography

Indication	A	B	Management by GP	
	Patients with a relevant echocardiographic diagnosis, ^a of whom the GP returned the questionnaire on management	Advice of cardiologist to refer patient (% of A)	GP followed cardiologist's advice to refer patient (% of B)	GP referred patient without advice of cardiologist [% of (A – B)]
Whole group, <i>n</i> = 625	397	114 (28.7)	81 (71.1)	47 (16.6)
Cardiac murmur, <i>n</i> = 368	237	73 (30.8)	56 (76.7)	26 (15.9)
Dyspnoea, <i>n</i> = 198	128	43 (33.6)	30 (69.8)	16 (18.8)
Peripheral oedema, <i>n</i> = 105	75	35 (46.7)	27 (77.1)	12 (30.0)
Other, <i>n</i> = 137	84	19 (22.6)	8 (42.1)	12 (18.5)

GP, general practitioner; LVD, left ventricular dysfunction.

^a Relevant echocardiographic diagnosis: at least one relevant abnormality, i.e. valve disease or left ventricular dysfunction or left ventricular hypertrophy or pulmonary hypertension or septum defects.

Reproduced from van Heur and colleagues²³⁵ under the terms of the Creative Commons Attribution Licence 2.0.

In terms of the theory under consideration, this study suggested that the specialists retained much of the control over the decision-making and management of patients referred for echocardiography. Although the GPs received training in interpreting the results, the results were accompanied by a management plan, which suggests that decision-making following the test remains under specialist control. However, this advice was not always adhered to by the referring GP. This may point to evidence of a lack of consensus between GPs and consultants as to which patients can be managed in primary care. Although it is unclear from this study if all patients would have received a specialist referral should open access not have been available, the results do suggest that the service was useful in identifying those patients who could be managed in primary care and it is likely that this reduces the demand for specialist referral. However, evidence remains of short-circuit, with a number of patients receiving a subsequent specialist referral despite the advice with the results that this was not necessary.

*Fuat and colleagues*²³⁶

This paper reported the results of a qualitative study using focus groups to explore the beliefs, current practice and decision-making of GPs in the diagnosis and management of suspected heart failure. Participating GPs were purposively sampled to recruit a range of participants over a number of variables including age, sex, ethnicity, practice size and geography. Data were collected from 30 GPs over four focus groups and analysed using an adapted grounded theory approach.

The analysis identified three themes in the data: (1) uncertainty about clinical practice; (2) lack of awareness of relevant research evidence; and (3) influences of personal preference and local organisational factors. Although this study was not about direct access to diagnostic testing, the results do provide some explanatory value to our programme theories. The GPs discussed the difficulty in making a diagnosis in primary care, particularly as the patients tended to be elderly with multiple comorbid conditions. At the time of publication, access to diagnostic tests such as echocardiography was variable, although two-thirds of the GPs in the study had access to this test. The quotes from the GPs expressed ambivalence towards the value of echocardiography in making a diagnosis in primary care, as they did not feel confident in interpreting the results. Some GPs would, therefore, not use direct access even when this was available.

The problem with echocardiograms is that I really just don't understand them. I don't think of myself as being really that old, I mean I'm 43 . . . and when I went through my post-registration years echocardiograms just weren't around . . . I just don't know where I am with them. When does an ejection fraction of such and such per cent stop being reasonable and start being a problem?

I would rather refer than do an echocardiogram, the interpretation of which I am not confident with.

Reproduced from Barriers to accurate diagnosis and effective management of heart failure in primary care: qualitative study, Fuat A, Hungin APS, Murphy JJ, BMJ, vol. 326, p. 196, © 2003, with permission from BMJ Publishing Group Ltd²³⁶

The authors reported that GPs were less likely to use direct-access echocardiography when the reports they received were technical and lacked a clinical opinion.

In terms of the theory under test, GPs expressed difficulty in interpreting the results of the test and, therefore, direct access may not enable them to make appropriate decisions as to which patients can be managed in primary care and which require specialist referral. This can result in short-circuit of direct access, with some GPs preferring to make a referral to a specialist to make a diagnosis. The results resonate with those from Stock and colleagues²³² on DXA scanning in that GPs state a preference for clinical reports that give guidance on management decisions. If the results of the test are accompanied by guidance on management in the light of the results then short-circuit may be minimised and referrals for specialist opinion may be reduced. *Box 42* provides an interim summary.

BOX 42 Interim summary

Clinical indicator tests are used to identify the location of a patient on a particular disease trajectory in order to inform a decision about whether to intervene now or whether to continue to monitor the patient. This information is also used to decide who is best placed (the GP or the specialist) to undertake this intervention. By providing GPs with direct access to these tests, it is assumed that GPs have the necessary skills to decide when and for whom such a test is warranted and then have the skills and confidence to interpret the results and decide which patients can continue to be managed in primary care and which patients require referral. As a result, more patients will be managed in primary care without the need for a specialist opinion and only those requiring specialist management are referred to secondary care.

Our evidence synthesis suggests that providing GPs with direct access to clinical indicator scans can enable GPs to distinguish between patients requiring specialist referral and those who can be managed in primary care if specialists provide GPs with referral guidelines and guidance on subsequent management. As such, specialists retain a significant amount of control over the decision-making in the patient pathway by setting guidelines and management plans for GPs to follow. Without these guidelines, GPs may bypass direct access to the test and refer directly to the specialist because they do not know which patients can be referred to direct access or they feel unconfident in the interpretation of the test results.

However, even in the presence of guidelines, some patients may still receive a referral when the results suggest this is not necessary. This may be because of a lack of consensus between GPs and specialists regarding which patients are suitable for referral to secondary care.

Conclusion

The evidence clearly shows that 'not all tests are created equal'. Organising the synthesis around the typology of the functions of tests has allowed us to draw out the differences between tests and, consequently, their differential impact as a demand management intervention.

Direct access to differential diagnosis tests seemed to have limited impact on demand for specialist care. The majority of patients referred to direct access in this modality also received a referral to a specialist with the suggestion that the test result simply acted as confirmation for the GP opinion that referral was necessary. Even where there was some reduction in referrals as a consequence of direct access to a differential diagnosis test, there was disagreement over the value of the test in the patient's diagnostic pathway. It would seem, therefore, that in this instance the small gains in reducing demand for specialist care may be offset by an overuse of unnecessary diagnostic testing.

In contrast, the results showed that direct access to rule out and clinical indicator tests had the potential to substantially reduce the demand for specialist care. This is because these tests were able to clearly delineate between patients requiring specialist care and those that were suitable to be managed in primary care. The utility of these tests could be maximised by providing GPs with guidance on the interpretation of the results and specialist opinion on subsequent patient management.

However, even where the test results and the guidance were clear in identifying those patients who did not require specialist referral, some still received a referral. It would seem that, in these cases, either patients or referring GPs were not reassured sufficiently by the test results that their symptoms could be managed in primary care. Residual folk wisdom about high clinical expertise may still impede the introduction of rational decision systems.

Chapter 7 Learned counsel: can guidelines control and shape demand?

Introduction: the guidelines industry

This is the fourth of our 'empirical' chapters. We move from those demand management programmes utilising role or mandate change and enter the procedural domain. Our focus here is on the attempt to use formal guidelines in the attempt to curb demand and reduce inappropriate referrals. This topic has a rather different research history and we have amended our reviewing strategy accordingly.

Our first 'historical' observation is to note that, unusually, there is some consensus about the utility (or perhaps futility) of using guidelines in the role of demand dampener. Imison and Naylor's report comes to the unequivocal conclusion, from which few researchers have demurred, that the 'passive use of guidelines' has failed as a tool for reducing demand.⁹³ More anecdotal evidence on the matter was voiced in *Chapter 3* by the practitioner who asserted that they 'usually end up either in the bin or are added to the guideline mountain in the corner of the clinic'.¹³⁶ This provides a helpful push for our review towards the question of what constitutes the 'active' use of guidelines and how they can be incorporated into wider systems of demand management.

The second matter of record brings us to our reference to the guidelines industry. There are guidelines for every condition from abdominal pain to the zoster virus. There are guidelines for all sectors: clinical practice guidelines, public health guidelines, technology appraisal guidelines, self-care guidelines. There are guidelines for every point in the patient pathway: diagnosis, screening, referral, treatment, withdrawal from treatment. Production takes place at every level from cottage industry (e.g. guidelines for local practices and emergency teams) to national function (e.g. guidelines issued by Royal Colleges and National Institutes) and to global enterprise (e.g. guidelines from World Health Organization taskforces). The delivery formats are diverse, covering everything from the one-page poster to hundred-page manuals and, latterly, to e-guidelines. And yet there is a gap in the glut. Relatively few of the guidelines restrict themselves to demand management per se. As noted, they tend to the comprehensive, offering advice on all aspects of screening, diagnostics, testing, treatment and then referrals. The significant corollary is that the primary research on their utility has not tended to confine itself to demand issues and also has leaned towards consideration of their uptake in general terms.

Our third point of ground clearing is to note that the topic, as just conceived, has gathered its own research industry. Many hundreds of studies have amassed assessing guideline 'X' in order to discover the extent of compliance and the reasons why the advice is and is not followed. Investigation on this scale results inevitably in the need for synthesis and so the primary research is rapidly followed by many systematic reviews of guidelines from 'A' to 'Z' – trying to understand, in broader terms, what works in guideline production and to unearth the all-pervasive facilitators and impediments to their implementation.

The methodological point of this brief history is that our review is preceded by several significant reviews and that much is to be gained by conducting a review of reviews. To this end, the chapter has a different format. We offer (1) a very brief look at a review in the Cochrane tradition; (2) a closer examination of the mainstream method in this area – primary studies and their reviews in the 'barriers and facilitators' tradition; and (3) a rapid realist review in which we concentrate on how guidelines may take on a more active role within a whole-system approach to demand management.

Approach 1: Cochrane orthodoxy – guidelines as interventions

With the reminder that clinical guidelines have been in use and under research for decades, we begin by reproducing the conclusion of a recent study following ‘strict Cochrane methodology’:

Our results reveal that there are only a few rigorous studies which assess the effectiveness of a strategy to implement clinical guidelines in Europe. Moreover, the results are not consistent in showing which strategy is the most appropriate to facilitate their implementation. Therefore, further research is needed to develop more rigorous studies to evaluate health outcomes associated with the implementation of clinical guidelines; to assess the cost-effectiveness of implementing clinical guidelines; and to investigate the perspective of service users and health service staff.

Reprinted from Health Policy, vol. 107, Brusamento S, Legido-Quigley H, Panteli D, Turk E, Knai C, Saliba V, et al. Assessing the effectiveness of strategies to implement clinical guidelines for the management of chronic diseases at primary care level in EU member states: a systematic review, pp. 168–83, 2012, with permission from Elsevier²³⁷

The review unearths 2562 potential ‘records’, of which 2535 are rejected (on the usual grounds of failure to meet the methodological gold standard established for clinical trials).²³⁷ Although the remnants do not carry all the complexity described above, it turns out that the chosen 27 are still massively heterogeneous (local vs. national guidelines; single vs. multifaceted interventions; delivery variation – feedback, workshops, outreach, computerised decision systems, etc.). Unable to perform meta-analysis with such raw materials, the study settles for six dense pages that merely describe and summarise the ‘eligible’ studies. The methodological error, of course, is to assume that guidelines, which carry the imprint of every part of a health-care system, can be regarded as a form of treatment applied to institutions, which can be turned on and off and outcomes compared. Guidelines do not impact on health providers like a pill on patient. As we shall see, guidelines have meaning and impact only in the varied, shifting and uncontrollable reactions of their intended stakeholders. It may be concluded that approach 1, as is often the case, tells us much more about the dictates binding the particular review methodology than it does about the effectiveness of complex programmes.

It also goes without saying that the more-research-is-needed conclusion is of no use whatsoever to policy-makers. Just as misguided is the idea that, had enough gold-standard RCTs been unearthed, a mean effect calculation on their effectiveness would make a difference to the DMs. Health-care systems are awash with guidelines and the time has long passed for some aggregative verdict on whether or not they work.

Approach 2: thematic reviews – facilitators and barriers to guidelines

The second approach to research review considered here, which is sometimes characterised as a ‘thematic’ or ‘narrative’ perspective, can be considered the orthodox approach to research synthesis in the particular domain of guideline investigation. It finds widespread application because it responds directly to the available raw materials. Rather than finding them wanting, it builds foursquare on the research designs used in the primary studies. That mainstream research strategy is the self-styled ‘barriers and facilitators’ approach. The basic design may be summarised as follows. A particular guideline is chosen for investigation, as is a selected set of its intended users. They are then faced with a broad set of questions on their familiarity with, experience of, attitude towards and confidence in the said guideline. The questions are delivered in interviews, focus groups and surveys, though it is probably true to say that the majority are derived from small-scale case studies – so the questioning is predominantly non-directive and the responses are, in the main, discursive. Responses, usually via transcripts, are then subjected to thematic analysis. Initial codes are generated, candidate themes are extracted, data reduction and simplification is considered, overarching frameworks are compiled and member checking is conducted to support the veracity of the results.

The analysis is usually presented in the form a framework or typology. If surmounted, the extracted themes become 'facilitators'; if they present a stumbling block to implementation, they become 'barriers'. Although they have this common origin the ensuing frameworks vary greatly in the pattern of types and subtypes derived in the analysis. We present a brief synopsis from four characteristic studies in *Table 12*, noting a significant proviso that these represent drops in the primary research ocean. We have no space here to describe the eventuating themes in any detail – hopefully, their meaning will be reasonably self-evident. In the original studies the meaning, distinguishing features and rationale for each type is elucidated in greater depth. Illustrative, from-the-horse's-mouth quotations are provided in the exposition of each theme. For example, consider the two following miniature testimonies. In respect of a widely cited theme on the need for 'clarity' in guideline presentation a GP complains: 'a complicated piece of paper, it's no use to me. I'm a simple man and I need to have simple ideas'.²³⁹ On the idea that 'patient pressure' may encourage the physician to ignore guidelines, another GP explains that radiography referrals for back pain are in high demand as a source of 'illness legitimisation'. 'The patient can come home and say, "I had an X-ray" and then everybody will realise I have a pain in my back.'²³⁸

Here, then, are the bountiful raw materials for another type of review, one that will go on to compile, translate between and synthesise the extracted types and subtypes. Because they represent the daily struggles of actual users, these primary studies present rich practical insight of the fortunes and foibles of guideline implementation. But does this approach constitute valid and reliable evidence? We have no space to conduct a methodological evaluation of even the handful of inquiries above. By way of a 'composite quality appraisal' we offer a reminder of the classic challenges of such qualitative, thematic analysis. They tend to be localised case studies and carry problems of generalisability. Especially if conducted in focus groups, they face problems of panel selectivity and 'chatty bias'. The precise analytic details of code extraction and theme development are rarely reported and so there is the possibility that the chosen ensemble corresponds to some pre-existing conceptual framework preferred by the researchers. There is also the possibility of the imposition of a broader ideological or professional opposition to guidelines – perhaps noticeable in the main title of the Rashidian and colleagues' study, 'Falling on stony ground?'.²³⁹ We do not suppose that these challenges disqualify such primary research from secondary review, merely that they should be reflected in any subsequent analysis.

We arrive at the main business of this section – to provide an account of two recent reviews attempting to synthesise the entire barriers and facilitators literature.^{241,242} It should be noted that there are other reviews of this ilk, with a little history going back to 1999.²⁴³ The research strategy is complex. Basically, it is a cross between a systematic review (scoping and definitional work, searching for studies, quality appraising, extracting findings) and a thematic analysis (to be precise, a secondary thematic analysis of the various primary thematic analyses). The essential aim is to provide a master or meta-framework that captures and combines all existing frameworks. Cochrane and colleagues call them 'barrier groupings' (see *Box 43*).²⁴¹ Gagliardi and colleagues call it a 'final conceptual framework' (see *Table 13*).²⁴² *Box 43* contains some additional information on 'frequency', that is the number of times a particular theme has been discovered in the primary literature. *Table 13* provides a different set of themes and subthemes but also includes a useful third column providing brief examples explaining the coverage of each theme.

A conclusion that hits the eyeball is that there are encouraging similarities but also significant differences between the meta-frameworks. Familiarity with other reviews in the barriers tradition, as above, confirms this volatile picture. Our focus here is to examine the practical purchase of such reviews and this quest now seems to turn on a prior question of deciding between frameworks. Why might the DM follow one typology rather than another? Is it feasible to seek the optimal framework?

There is, of course, a large methodological literature on frameworks, classification systems, typologies, taxonomies and so forth.²⁴⁴ And within this there is the classic discussion on the different types of 'validity' – 'face', 'content', 'criterion', 'construct' and so forth. All of the frameworks under discussion (primary and secondary) have evident face validity. They originate in the lived experience of guideline users – so that notion, alas, is no help in deciding between them. It transpires that most methodological authorities place

TABLE 12 Typological reviews: four accounts of facilitators and barriers to guideline implementation

Espeland and Baerheim ²³⁸	Sheldon and colleagues ²³¹	Rashidian and colleagues ²³⁹	Lugtenberg and colleagues ²⁴⁰
Knowledge related	Characteristics of guidelines:	Theme 1: credibility of content of clinical guideline	Knowledge:
Lack of knowledge of the guideline	<ul style="list-style-type: none"> strong professional support stable and convincing evidence base no increased or unfunded costs good systems for tracking guidance professionals involved are not isolated guidance clear and reflects clinical context 	<ul style="list-style-type: none"> evidence based flexible 	<ul style="list-style-type: none"> lack of knowledge lack of awareness familiarity
Attitude/feeling related		Theme 2: credibility of source of clinical guideline	Attitude:
Lack of agreement with the guideline:	<ul style="list-style-type: none"> commitment to guidance lead clinician identified to implement proactive audit of costs of implementation responsibility for implementation vested locally strong clinical governance culture of consensus recognise legitimacy of provider involvement of clinicians in guidelines financial stability expectation that compliance in mandatory targeted audit of non-compliance 	<ul style="list-style-type: none"> national professional bodies national governmental bodies published in respected sources 	<ul style="list-style-type: none"> lack of agreement with recommendation lack of applicability lack of outcome expectancy lack of motivation (practice inertia) behaviour patient factors patients' preferences/demands patients' ability/behaviour guideline factors unclear/ambiguous incomplete/not up to date not easy to use/too complex environmental factors lack of time/time pressure lack of resources/materials organisational constraints lack of reimbursement
<ul style="list-style-type: none"> lack of agreement with its decision criteria lack of outcome expectancy lack of process expectancy lack of feelings expectancy lack of self-efficacy lack of motivation/inertia of previous practice 	Characteristics of recipients:	Theme 3: presentation of clinical guidelines	
External		<ul style="list-style-type: none"> simple systematic presentation 	
Guideline related: unclear or impractical to use	<ul style="list-style-type: none"> responsibility for implementation strong clinical governance culture of consensus recognise legitimacy of provider involvement of clinicians in guidelines financial stability expectation that compliance in mandatory targeted audit of non-compliance 	Theme 4: influential people in implementation	
Patient-related pressure		<ul style="list-style-type: none"> presence of practice nurses/primary care team/primary care organisations/pharmacists/prescribing advisers 	
Setting related:		Theme 5: organisational factors	
<ul style="list-style-type: none"> lack of time lack of other practice resources increased costs increased malpractice liability external pressures in the health-care system improper access to health-care services 	<ul style="list-style-type: none"> practice characteristics information technology 	Theme 6: disease characteristics	
		<ul style="list-style-type: none"> rare or 'simple' disease 	
		Theme 7: dissemination strategy	
		<ul style="list-style-type: none"> ownership – local vs. national guideline perceived need of first contact enforced or supporting implementation 	

BOX 43 Cochrane and colleagues²⁴¹ 'barrier groupings': categories (frequency)**Cognitive/behavioural barriers (65)**

Knowledge (38).

Awareness (13).

Skill/expertise (9).

Critical appraisal skills (5).

Attitudinal/rational-emotive barriers (58)

Efficacy/perceived competence (16).

Perceived/outcome expectancy (16).

Confidence in abilities (15).

Authority (9).

Accurate self-assessment (2).

Health-care professional/physician barriers (62)

Characteristics (29).

Age/maturity of practice (11).

Professional boundaries (7).

Legal issues (5).

Peer influence, models (5).

Gender (3).

Inertia (2).

Clinical practice guidelines/evidence barriers (41)

Utility (11).

Evidence/disagree content (11).

Access (10).

Structure (5).

Local applicability (4).

BOX 43 Cochrane and colleagues²⁴¹ 'barrier groupings': categories (frequency) (*continued*)**Patient barriers (30)**

Patient characteristics/factors (20).

Patient adherence (10).

Support/resource barriers (69)

Time (31).

Support (15).

Costs/funding issues (12).

Resources (11).

System/process barriers (62)

Organisational (20).

System (17).

HR/workload/overload (10).

Team structure/work (9).

Referral process (6)

Reproduced with permission from Cochrane LJ, Olson CA, Murray S, Dupuis M, Tooman T, Hayes S. Gaps between knowing and doing: understanding and assessing the barriers to optimal health care. *The Journal of Continuing Education in the Health Professions* 2007;**27**:94–102. © 2007 Wiley Periodicals, Inc.²⁴¹

TABLE 13 Gagliardi and colleagues²⁴² final framework of guideline implementability

Domain	Element	Examples
Usability	Navigation	Table of contents
	Evidence format	Narrative, tabulated or both
	Recommendation format	Narrative, graphic (algorithms) or both; recommendation summary (single list in full or summary version)
Adaptability	Alternate versions	Summary (print, electronic for PDA); patient (tailored for patients/caregivers); published (journal)
Validity	Number of references	Total number of distinct references to evidence on which recommendations are based
	Evidence graded	A system is used to categorise quality of evidence supporting each recommendation
	Number of recommendations	Total number of distinct recommendations (subrecommendations considered same)
Applicability	Individualisation	Clinical information (indications, criteria, risk factors, drug dosing) that facilitates application of the recommendations explicitly highlighted as tips or practical issues using subtitles or text boxes, or summarised in tables and referred to in recommendations or narrative contextualising recommendations
Communicability	Patient education or involvement	Informational or educational resources for patients/caregivers, questions for clinicians to facilitate discussion, or contact information (telephone, fax, e-mail or URL) to acquire informational or educational resources
Accommodation	Objective	Explicitly stated purpose of guideline (clinical decision-making, education, policy, quality improvement)
	Users	Who would deliver/enable delivery of recommendations (individuals, teams, departments, institutions, managers, policy-makers, internal/external agents), who would receive the services (patients/caregivers)
	User needs/values	Identification of stakeholder needs, perspectives, interests or values
	Technical	Equipment or technology needed, or the way services should be organised to deliver recommendations
	Regulatory	Industrial standards for equipment or technology, or policy regarding their use
	Human resources	Type and number of health professionals needed to deliver recommended services
	Professional	Education, training or competencies needed by clinicians/staff to deliver recommendations
	Impact	Anticipated changes in workflow or processes during/after adoption of recommendations
	Costs	Direct or productivity costs incurred as a result of acquiring resources or training needed to accommodate recommendations, or as a result of service reductions during transition from old to new processes
	Implementation	Barriers/facilitators
Tools		Instructions, tools or templates to tailor guideline/recommendations for local context; point-of-care templates/forms (clinical assessment, standard orders)
Strategies		Possible mechanisms by which to implement guideline/recommendations
Evaluation	Monitoring	Suggestions for evaluating compliance with organisation, delivery and outcomes of recommendations, including programme evaluation, audit tools and performance measures/quality indicators

Reproduced from Gagliardi and colleagues²⁴² under the terms of the Creative Commons Attribution Licence (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

faith in 'construct validity' (there are various antonyms). Simply stated, this postulates that the worth of a measure is extrinsic to the measure and belongs to the inferences drawn using the measure. If one tests a theory using a particular measure and that theory is corroborated then this also adds validity to the measure. A measure becomes validated over time – the more theories it supports and the more those theories are successful, the stronger the faith in the measure: 'The best construct is the one around which we can build the greatest number of inferences, in the most direct fashion'.²⁴⁵

If one follows this advice (and we do) this allocates the assessment of the validity of any guidance framework to the utility of the model rather than the construction of the model. And this returns us to the central question of the chapter: how are users supposed to make use of all of this evidence on guideline effectiveness? We begin by considering the expectations on this score of one our review teams. In the second half of Gagliardi and colleagues' paper there is a 'test' of 20 existing specialist guidelines (on the management of diabetes, hypertension, leg ulcer and heart failure) to see if they fit with the 'extended model' emerging from the typological review.²⁴² Results are disappointing. The various guidelines under scrutiny are largely adequate in terms of presenting 'graded evidence' but few examples contained 'additional features that could improve guideline usage'. In other words existing guidelines fare much better in their coverage of items in the upper sections of *Table 13*.

A curiously limited notion of guideline validity is implied in this test. What the Gagliardi review seeks to inform is the business of guideline construction. The evidence on why potential users of guidelines follow or fail to follow them is returned on itself in order to improve the presentation and content of guidelines. A comprehensive, master framework is devised in the expectation that future guidelines should conform to that rubric. A grammatical shift from noun to adjective is introduced to establish this goal – guidelines vary in their 'implementability' and this evidence-endorsed template will help get their content right. The guideline industry is certainly promulgated and perhaps regulated as a result of such frameworks. Powerful collaborations have gathered promoting these 'international tools for the rating and assessment of practice guidelines' such as the AGREE template (www.agreetrust.org/).

But all of this is indeed a curious interpretation of construct validity and guideline utility because the main empirical lesson from the 'barriers and facilitators' investigations is that guidelines only have paper authority. The reason why guidelines fail is little to do with their content and format (their implementability) but mostly due to complex decision structures in which they are embedded (their implementation). However perfect their presentation, however comprehensive their coverage, however true to template, there is no reason to suppose that guidelines will be followed. This is the lesson that emerges from all empirical studies. What matters is the reception that awaits guidance when it has left the page and enters the clinic.

This brings us to a second, and perhaps more commonplace, expectation about how the thematic analysis on guideline effectiveness might be put to use. In this version, the review is said to provide the most authoritative checklist of barriers/facilitators to guideline implementation in expectation that it presents a 'to do' list for policy-makers, managers and practitioners. The role of barrier and facilitator reviews, on this view, is to provide strategic overviews; they are the design tools itemising what is required in a comprehensive planning process. In the context of our review on the utility of guidelines for demand management it would be possible to present a similar framework of 'dos and don'ts' as advice to the policy-maker about to set in train some new guidance. The obvious question is raised: are these 'barrier groupings' an effective planning tool?

In any review of the literature one encounters critical appraisals. We thus noted that the checklist perspective has come in for its fair share of criticism, most notably in an article with a telling title: *Is the metaphor of 'barriers to change' useful in understanding implementation?*²⁴⁶ This paper takes us back to the raw materials similar to those noted in many of the primary studies on guidelines. Prompted by questions about why guidance might be ignored, a GP in the Checkland study replies: 'But if anybody thinks that things arrive here, somebody has the time to look at it and then spread it as useful information

that everybody else thinks sensibly about, they've got another thing coming'.²⁴⁶ Now 'removing' this barrier to change is usually discussed in terms of simplifying guideline presentation, providing guideline summaries, improving the channelling of recommendations and so on. The authors, however, submit the rival interpretation that the 'time pressure barrier' is simply an underlying organisational reality and that the source of the problem and its solution may lie there.

The moral of the tale is that barriers are not something to be ticked off and torn down one at a time. What the 'barriers' primary studies are actually describing are personal, social and institutional *interrelationships*. Barriers interlock because a change in one part of a complex system will always trigger change in another and then another. This suggests a rather different role for reviews of guideline effectiveness. What the synthesis should be studying and explaining is why some barriers are more intractable than others and why solutions always have emergent effects. Solving barrier A might exacerbate barrier B, solving barrier C might create unintended consequence D, introducing facilitator E might be crushed by impediment F. We rather suppose that real DMs inhabit this alphabet soup of system strains and we turn next to our review that attempts to decipher the interconnections.

Approach 3: a rapid realist review

We devote the remainder of the chapter to exploring the potential of a realist synthesis of the primary research on guideline effectiveness. The first question, as ever, is to discover the programme theories that underpin the development of guidelines and what is an appropriate way to test them. We start with an enormous advantage. The vast literature on barriers and facilitators provides us with a wealth of provisional programme hypotheses, starting at square one with the elementary idea that a guideline that is simply stated is more likely to be understood and to find use. Our model thus starts with familiar ingredients (thanks be to all that typological work). But we do not seek to repeat reviewing history by working sequentially through all of the themes and subthemes of guideline implementability already identified. The vital empirical test, therefore, is not about whether or not a particular application of a guideline fulfils the checklist of facilitators and avoids all barriers. The checklist approach, as we have just seen, atomises guidelines into separate themes without appreciating that they are inextricably linked. Our review treats guidelines as a complex, multicomponent system thrust into a pre-existing, complex health-care system. The goal, as in the previous chapters, is to examine the tensions and harmonies with the emerging system (*Figure 33*).

Our model thus begins by identifying and compressing 'guideline theory' into four broad 'if-then' phases, which must be met for successful outcomes. The first concentrates on guideline production and considers the qualities of its production and of its producers. Only if these are sufficiently authoritative will guidelines reach square one – they get 'on the table'. The recommendations then pass into potential usage and to the next set of theories. Only if there is a close fit between the paper advice and the circumstances of the presenting patient will the guidelines reach phase 2 – the point of applicability. However, even if there is good initial 'fit' with the patient's condition the guideline may or may not be followed. All depends on the intricacies of phase 3 and whether or not there is consonance between its recommendations and the experience, custom and practice of the clinical user. Even if they survive this internal dialogue, the guidelines still may or may not be followed at phase 4. Recommended actions may have to overcome significant external resistance from further stakeholders, most notably from the patient and then from other resource-constrained providers in the system.

To repeat, these 'phase theories' are simple distillations of material from the barriers and facilitators literature. The difference is that we treat them as part of an interlocking system (represented by the linkages in *Figure 33*). In classic system terminology – the whole is more than the sum of the parts. Each phase represents a necessary condition to reach the next, but, as every manager knows, there are harmonies and tensions throughout the system. Success in one phase can actually make it more difficult to

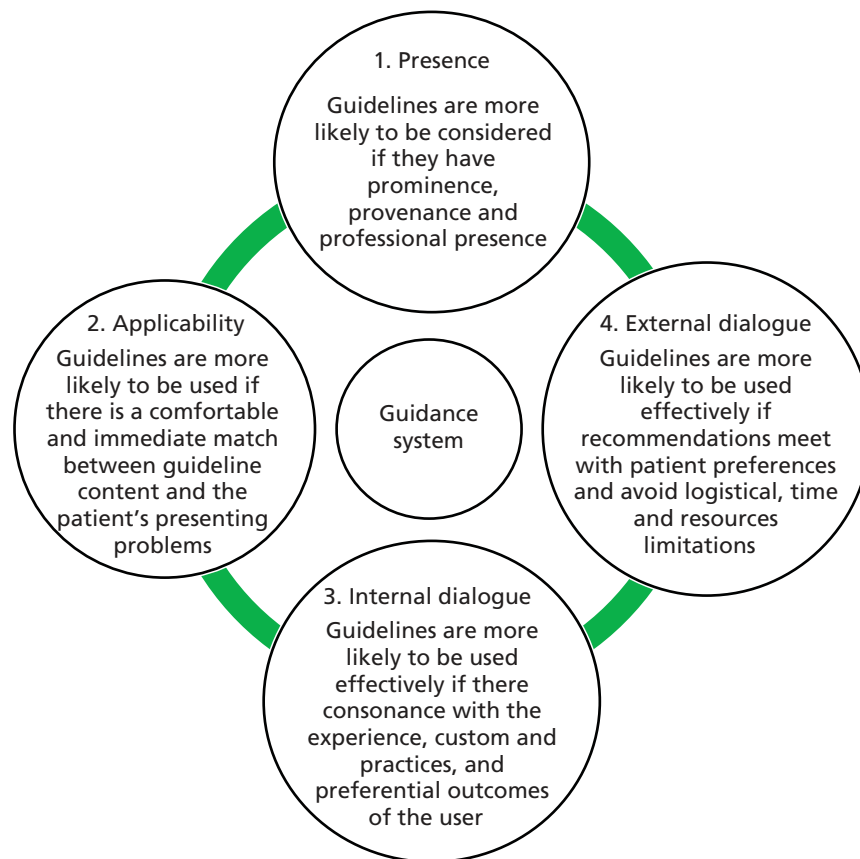


FIGURE 33 A phases and interdependence model.

achieve another. Success in one phase can drive unforeseen consequences into another. Success in one phase may be squashed in another.

The focus of our rapid review thus turns to system functions and dysfunctions (or 'system strains' for short). Our mental image of the policy-maker is not one of the master planner, designing a guideline blueprint from scratch; rather, it is of the manager coping with many existing guidance systems and attempting to integrate them into wider care regimes. They will face a daily diet of system strains – is there an evidence base that might provide decision support in these circumstances? We provide examples of four system strains:

1. the tension in using simple guidelines for complex comorbidity
2. the tension between (inter)national credibility of and local control over guidelines
3. the tension between patient choice and top-down guidelines
4. the tensions involved when there are competing guidelines and contending targets.

What follows are some indicative data from a realist review; we extract only a few key findings and discuss their implications.

System strain 1: simple guidelines versus comorbid patients

A recurrent refrain in all the aforementioned research is that guidelines are more likely to be implemented if the presentational format is straightforward, intelligible, comprehensible, uncomplicated and so on. The beginnings of a system strain lurk here when through the surgery door walks patient George whose ills are far from uncomplicated. Several studies have explored the problems of fitting a patient to guidelines when that patient has comorbidities and, more especially, if that patient happens to be frail or elderly.

Primary studies group, first, into those examining the texts of existing guidelines within a clinical domain and assessing the extent to which they deal with comorbidity. Typical results in the study of guidance on chronic conditions are as follows:

Half the guidelines addressed treatment for older patients or for patients with one comorbid condition. But only one addressed treatment for older patients with comorbid conditions.

Vitry and Zhang²⁴⁷

Of the 20 guidelines, 17 (85%) addressed the issue of comorbidity and 14 (70%) provided specific recommendations on comorbidity. In general, the guidelines included few recommendations on patients with comorbidity (mean 3 recommendations per guideline). Of the 59 comorbidity-related recommendations provided, 46 (78%) addressed concordant comorbidities, 8 (14%) discordant comorbidities, and for 5 (8%) the type of comorbidity was not specified.

Reproduced from Lugtenberg and colleagues.²⁴⁰ © 2011 Lugtenberg and colleagues. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 Licence, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited

A basic tension is thus recognised in the literature, which can then be explored for potential resolutions. The simplest adaptation is to increase the elderly, comorbid patient's exposure to multiple guidelines and several studies cover the potential costs and unintended conflicts of following more than one guideline simultaneously. Caughey and colleagues used treatment-dispensing data showing that 16% of people with diabetes being treated for other conditions received medicine with adverse effects on diabetes.²⁴⁸ Hughes and colleagues²⁴⁹ mounted a simulation exercise on the consequences of following the explicit recommendations of two or more guidelines demonstrating a significant hike in the treatment burden, especially on self-care regimes in an elderly comorbid population ill equipped to meet such demands. Boyd and colleagues²⁵⁰ carried out an early exercise following a hypothetical comorbid patient who, following all relevant guidelines, would be prescribed 12 medications at a cost of US\$406 per month, some with possible adverse effects.

Neither is this system strain resolved by simply increasing attention to comorbidity within condition-specific guidelines, as recommended in both the primary and the secondary research.^{240,251} This suggestion would create an elemental strain on the evidence base, unlikely to be resolved, which reserves a place of honour for clinical trials in which patients with multiple morbidities are generally excluded. It also sends us full circle to the original conundrum – addressing comorbidity within a guideline would inevitably increase the complexity of guidelines, which is already an established standard deterrent to their usage.²⁵²

Even these simple nuggets of evidence show that guidelines are constantly being made and remade under system strain. A further recent refinement to guideline logic suggests a shift from 'disease specific recommendations' to 'patient centred guidelines'.^{253–255} To be more precise, such guidance places attention on carefully selected subgroups, for example 'decision making on care of the elderly with condition X'. Guidelines here focus much more on choosing and prioritising treatment and so make useful ground in reducing the tension inherent in following multiple combinations of condition specific guidance. However, end-of-life care has complexities of its own which, if guidelines are to be comprehensive, will then be drawn into the equation. Consider the key issue of life expectancy and the following recommendation

from Durso: 'Estimate the patient's approximate life expectancy compared to the median for individuals of that age-sex cohort by considering the presence or absence of unusually good or poor health and function'.²⁵⁵ We pass no comment on whether or not this meets the simplicity rule, noting only the existence of further such algorithms in similar guidelines.²⁵⁶

At this point the evidence trail grows thinner and much of the literature on these niche ideas may best be described as proposals, prototypes and product descriptions. But we can summon other aspects of the evidence base already amassed to note basic tensions that are likely to surface in relation to these more specialised guideline configurations. There is already a problem with physicians being bombarded with guidelines on X, Y, Z. If, thanks to patient complexity, these become subdivided into guideline X₁, X₂, Y₁, Y₂, Y₃, Z₁, Z₂, with supplementary decision rules on navigating to the appropriate pathway, there is an obvious emergent further problem afoot in the realm of guidance fatigue.¹⁷¹

In conclusion, we observe that both practitioners and researchers are alert to this particular system strain. The chain of adaptation and readaptation teaches us that comorbidity represents an ineluctable limitation on the implementation of guidelines. However, with the move from 'disease specific' to 'patient specific' and 'decision specific' guidelines it is also possible to discern more positive and more subtle efforts to strike a balance between the rule and the recipient. Work is only just beginning to make progress on which decisions cause the most consternation and how to direct these nuggets of information on treatment alternatives to the deluged physician.

System strain 2: the tension between (inter)national credibility and local control over guidelines

As noted in *Box 43* and *Table 13*, the 'credibility' of guidelines is often cited as a key facilitator in influencing their uptake. Credibility, however, may be nurtured in quite different quarters. One programme theory here begins with the notion that guideline development requires major methodological resources. Accordingly, guidelines endorsed by national or global professional organisations are seen as more trustworthy by the clinical community and are, in turn, more likely to be implemented (e.g. The Royal College of Physicians' national guideline for stroke²⁵⁷). However, nationally developed guidelines may lack applicability and relevance to local contextual factors. Programme theory 2 thus posits that involving local practitioners is the key source of credibility because the guidance will include intelligence on the prevalence of the condition in the local community, on the local availability of services and resources, and on current inter- and intraorganisational relationships (e.g. the Oxfordshire Regional Genetics Service referral guidelines²⁵⁸).

Here we have the beginnings of a classic system strain, sometimes termed the 'glocalism' paradox. To address this tension, the local adaptation of nationally developed guidelines has been widely advocated as a potential solution. Programme theory 3 thus seeks a 'best of both worlds' solution and considerable resources have been expended in pursuing it. The best known of these initiatives is the international ADAPTE collaboration (www.g-i-n.net/working-groups/adaptation/history). The task of 'customising' a guideline so that the global becomes the local is itself a feat of organisation. Decisions need to be made about topics, organising committees, source documentation, consultation rules, format and promotion. The ADAPTE process has a whole series of phases and modules, numbering 24 steps in all.²⁵⁹ At the other end of the scale, adaptation can be ad hoc, locally initiated and focused on particular units with identifiable users.²⁶⁰

While all of this is proof positive for our thesis that reviewers of the guidelines literature must anticipate a moving target, it opens up another system tension in need of explanatory synthesis. There is copious material comparing guideline content before and after implementation. Some of this considers whether or not adapted guidelines are of the same standard as those with a national pedigree. Rowe, for instance, suggests that locally developed guidelines are poor in coverage according to AGREE standards.²⁶¹ By contrast, Shye and colleagues suggest that the process of local adaptation does not seriously distort the clinical validity of the original guideline.²⁶² Other studies seek to unearth subtle, qualitative differences between the national and local

instruments. Sometimes, the adapted guidelines have almost identical coverage and content to the originals.²⁶³ Other research notes the adaptations are real: specialising in topics with local priorities,²⁶² having shorter and more accessible formats^{262,264,265} and, perhaps most significantly, providing additional information on the availability of local services and thresholds for referral.²⁶⁴

There are also many process evaluations, which trace the practical steps and resources involved in guideline adaptation. Sometimes guidelines are adapted through a series of short workshops or meetings with local primary and secondary care clinicians.^{262–265} For example, Capdenat St Martin and colleagues report on a process that took 2 months from initiation to completion.²⁶⁰ Liaw and colleagues report a case where the adaptation was undertaken through 12 hours of group discussion.²⁶⁵ By contrast, other studies have followed the processes involved using formal templates, such as ADAPTE, described above, which attempt to standardise the process of guideline adaptation.^{266,267} Key evidence emerging here suggests that full-blown formal adaptation is a resource-intensive and costly business – involving a need to revisit the original evidence, identify additional research evidence to support local content and provide methodological support to critically appraise the final product.²⁵⁹

The major issue lurking within the ‘glocalism’ issue is, of course, whether locally adapted guidelines improve outcomes. Do users pay more attention to and act on guidelines that have been locally adapted? Yet another outcrop of studies tackles this problem and we begin a mini-review by noting a considerable methodological difficulty. The basic design involves manipulating a situation whereby identical groups of practitioners are exposed to existing and to adapted versions of the guidelines and then observing differences in their understanding and action. The whole of the facilitators and barriers literature teaches us that achieving this clean difference will be demanding because the actuality of ‘exposure’ to a guideline is so diverse. For illustration, we examine findings from two such studies.

Silagy and colleagues²⁶³ conducted a cluster RCT on guidelines for stroke prevention and on urinary tract symptoms. Nationally and locally designed guidelines were allocated randomly to two divisions of general practice in Adelaide, South Australia. The local guidance included additional information regarding the availability of divisional resources and was presented with more ‘user-friendly’ design and formatting. Pre-trial measures were collected from a postal survey of 200 GPs from each division, covering knowledge, action and reported practices in respect of the two conditions. This was repeated 3 months after the dissemination phase. Dissemination, identical for both versions, included mailing shots, newsletter articles, prompt sheets, educational workshops and web-links. Considerable change in GPs’ usage of guidance was noted across the before-and-after measures – observable, disappointingly, in both arms of the trial. The authors concluded that although they found significant changes in knowledge, attitudes and reported practice as a result of guidelines dissemination, the local adaptation of the guidelines did not produce any additional effect. They argue that efforts to promote guideline implementation should focus on multifaceted dissemination strategies, rather than local adaptation per se.

A pioneering French study by Capdenat Saint-Martin and colleagues also throws light on mechanisms through which guideline adaptation may work.²⁶⁰ The background here was the perceived overuse of preoperative tests for anaesthetic risk. National guidelines had shown that they were costly, had highly restrictive diagnostic use and did not add significantly to the safety of operations. Capdenat Saint-Martin’s team thus undertook a process of local adaptation of these guidelines at the hospital level (15 surgical wards). The organisational structures (decision trees) responsible for ordering the tests were mapped and team discussions of the new recommendations were set up in the appropriate branches. Before-and-after measures were taken of referrals for such preoperative tests and in the targeted low-risk groups requests fell from 80% to 48%. The question of attribution raises its troublesome head. For the trialists, we lack a control group here and simply cannot say that the adapted guideline bore responsibility for the change. The French team drew a different, system-based lesson. The people responsible for adapting the guidelines were the same people who organised the implementation of the new guidelines: ‘We think that the main contribution of this work is the linking of the process of local adaptation to an analysis of the organisational aspects of the practice and the emphasis we placed on the organisational aspects of change’.²⁶⁰

Many other studies of the impact of local adaptation have followed but these two studies above provide indication of the direction of travel. Whether guideline content is de novo or de integro seems to be far less crucial than its passage from the text to consulting room, through the conversations between its users and onto the external constraints on its usage. The take-home messages seem to be (1) adapt guidelines to kindle interest rather than to impart new knowledge and (2) the more local the adaptation, the greater the number of interested parties who will be drawn directly into the scheme.

System strain 3: the tension between authoritative, population-based guidelines and discretionary, individual-based patient choice

Clinical guidelines are derived from the assimilation of evidence at the level of the population. The potential risks and benefits of referral or treatment are expressed in terms of population estimates; recommendations are general and focused on the average patient. By contrast, GPs and patients make decisions on an individual basis and patients vary in terms of their preferences for referral and treatment. There is a considerable anecdotal literature in the practitioner journals about ‘turning a blind eye to guidelines’ by referring the worried well patient onwards for further advice or by complying with their world-weary wishes to forego any recommended interventions.²⁶⁸ This system strain also occurs in macro policy. Recent years have seen an introduction of procedures giving patients more choice on the mode, location and timing of treatment.²⁶⁹ Clinical guidelines, paradigmatically, seek to impose order on treatment modalities.

The first and most widely evaluated solution to this tension is to involve patients in the process of developing the guidelines. It is hypothesised that incorporating patient knowledge into the guideline development process will make the guidelines more applicable to everyday practice and increase their legitimacy and, thus, improve their uptake. However, the process through which patients are involved is a matter of debate. Van Wersch and colleagues²⁷⁰ compared a number of different methods (guideline group membership, one-off meetings, workshop series and consumer advocates) and concluded that no single method alone would be sufficient to optimise patient involvement. Providing intensive support to patients or having a separate working group feeding into the main guideline group may succeed in ensuring that key patient concerns are addressed within the published guideline. But this approach is resource intensive and, accordingly, some patients report that their input feels marginal and tokenistic.²⁷¹

Van de Bovenkamp and Trappenburg’s review²⁷² and also a later empirical study²⁷³ highlighted crucial practical impediments to including the patients’ voice in guideline development. Patients may have difficulty understanding medical jargon and, consequently, provide little input into the process. Restricting involvement only to patients who are able to understand the scientific literature may call into question the representativeness of this input. Training patients to understand the medical literature enlists them as fellow ‘academics’ and may distance them from their experiential knowledge that is the basis for their involvement. They may become professionalised and fail to represent true patients ‘in the wild’.

The degree of fit between the patient perspective and those of other stakeholders involved in guideline development has also come under empirical scrutiny.²⁷⁴ Case scenarios of patients with knee and urinary tract symptoms of different severity were presented to a guideline development group consisting of surgeons, GPs, health-care professionals *and* patients. The research examined which patients were deemed appropriate for referral with the idea of testing whether or not the ensuing guidelines would be consensual. Interestingly, patient and surgeon viewpoints came significantly closer to those of the GPs in some instances. Overall, however, the results are described as a ‘juggling act’, with different stakeholders providing discordant ratings across the differing scenarios.

The fundamental and perhaps irresolvable problem with involving patients in guideline production is that patient’s experiential knowledge, which is inherently individual, subjective and fluid is incommensurate

with guidelines based on knowledge considered to be objective, valid and scientific. Accordingly, policy and the research literature has gone in search of ways of 'reconfiguring the tension' by locating the patient perspective at other points in the guidance machinery.

An alternative strategy is to have explicit statements within guidelines that advise on making room for individual patient preferences. Van der Bovenkamp and Trappenburg argued that this can be accomplished by incorporating sections on doctor–patient communication within the guideline or by having a patient 'version' of the guideline to support their active involvement in treatment decisions.²⁷² Clarke's research pinpointed a specific niche in which the two activities may be combined.²⁷⁴ They argued that guidelines could identify and differentiate levels of symptom severity, as patients' preferences for referral depended on the severity of their symptoms.²⁷⁴ We were unable to discover any empirical evaluations of these specific manoeuvres but they are likely to reintroduce a strain already discussed. Guidelines that come in competing versions or are laced with too many contingencies and caveats are less likely to be followed.

Van der Weijden and colleagues' solution was the collaborative development of clinical guidelines *with* patient decision aids.²⁷⁵ The latter already have considerable authority and autonomy and are aimed directly at patient concerns. Their development, moreover, has been every bit as rapid as that of clinical guidelines.²⁷⁶ Van der Weijden and colleagues set out a framework to specify another niche condition in which a marriage of these distant cousins might be fruitful.²⁷⁵ Where there is low uncertainty about the factors that influence decision-making, such as agreement about treatment risks and benefits, it is possible to make strong recommendations on the ideal course of action in the guidelines. In these circumstances, decision aids can be used in parallel to support adherence to the guideline and behaviour change. However, some decisions entail a high level of uncertainty, for example because there are two or more treatment options of equal attractiveness regarding the outcome. For these 'utility sensitive' decisions, guidelines should make only conditional recommendations, which should be mirrored in decision aids that recommend 'careful deliberation'.²⁷⁵

Again we are presented with an attractive, but largely untried, programme theory. Several stringent tests, however, have already been suggested in materials reviewed to this point. The patients' use of decision aids and the physicians' use of guidelines would require considerable co-ordination. Shared decision-making in routine clinical practice is in its infancy and usage is deeply stratified. Han and colleagues, for instance, reported low levels of shared decision-making in prostate-specific antigen screening with 'significant variations by black race, Hispanic race, higher education and health insurance'.²⁷⁷ Although co-ordination may be relatively straightforward for individuals faced with a single condition, it would fall under the standard problems in dealing with comorbidities or multimorbidities. The 'simplicity rule' is once again under threat. Guidelines, like hymn sheets, work better in the singular. We are driven to a similar conclusion – solving a tension in one part of the system may overlook or even give rise to different system strains. We should anticipate that the patient's consent to guidelines will remain a matter for piecemeal solution in another niche condition, namely in the doctor's surgery.

System strain 4: tensions resulting from guideline oversupply – a new guideline can swamp a routine system

It is well established in the barriers-and-facilitators literature that the 'prominence' of a guideline can influence its uptake. Guidelines with the highest profile are much more likely to be implemented, with the accompanying caveats already noted about their requisite applicability and synchronicity. Coiera suggested that we now have a 'clinical information marketplace', where evidence-based recommendations are in competition for the clinician's attention and, with the proliferation of guideline production, we are in a situation of information oversupply; 'the amount of information that can be accessed or "consumed" is fundamentally limited by human attention'.²⁷⁸ Another paper in the barriers tradition focuses on a significant aspect of the way in which GPs consume guidelines, namely that they often do so 'haphazardly' or 'opportunistically'.²⁷⁹ Especially in preventative care, where there are no clear daily routines into which

guidelines on health promotion and disease prevention can be inserted, the response to the marketplace may be capricious and piecemeal.

One gathering and somewhat paradoxical response to information overload is that the 'guidelines on guidelines' also come replete with advice on how to position clinical guidance institutionally and how to increase its visibility in the daily business of the clinic and the surgery. We address this issue of guideline rivalry with a brief review of the evidence on a highly visible episode. A guideline will have obvious prominence if it is introduced as a national priority and accompanied by strong ideological wind. From a systems perspective, however, the key issue is not about tactics for assembling the requisite political leverage. The new guidance will always sit alongside a range of other clinical decision supports and so the key question is about displacement effects of new mandates.

In 2000, in response to comparatively unfavourable mortality rates from cancer, the UK government introduced the '2-week wait' rule for suspected-cancer referrals, accompanied by widely distributed guidelines for urgent referral across a number of specialities. Several studies went on to identify problems with the 2-week mandate. Secondary care services, working under a pre-existing guidance regime, were often ill prepared for the change in referrals and suffered from 'system swamping'. Malik and colleagues tracked all referrals to a bone and tissue soft tumour service over a 2-year period.²⁸⁰ They also reviewed the case notes of those referred under the guidelines for urgent review by the GP:

40 patients were referred under the guideline between January 2004 and December 2005. Ten of these patients (25%) had malignant tumours, compared with 243 of 507 (48%) of those referred from other sources. In 9 of the 40 cases, the patient did not meet the criteria for urgent referral. Although this target has focused attention on shortening the time to diagnosis and treatment, prioritising patients referred from general practitioners has the potential to disadvantage those with malignant tumours referred from other sources

Reproduced under the terms of the Creative Commons Attribution Licence from Malik and colleagues²⁸⁰

This pattern recurs in other cancer services. Jones and colleagues reported the results of two separate audits of referrals for patients with suspected colorectal cancer, which suggested that the 2-week waiting standard was being met at the expense of a substantial increase in waiting times for routine referrals, even though fewer than one-third of patients were referred through the 2-week wait.²⁸¹ Potter and colleagues reported on a prospective cohort study of referrals for suspected breast cancer in Bristol.²⁸² Over a 6-year period (1999–2005), the referral patterns had changed, with a higher proportion referred as urgent using the 2-week-wait system. Over the same period the odds of receiving a positive diagnosis in the urgent referral group fell and the odds increased in the routine referral cases. Moreover, although the target wait was achieved for the urgent referrals, the waiting times for routine referrals were extended.

How is all this to be explained? Cornford and colleagues provided an explanation with masterful brevity: 'Patients want early referral to assure them that they do not have cancer, specialists want early referral for patients with cancer, and GPs want both'.²⁸³ To address this rather more systematically, it is useful to return to *Figure 33*. We discover that concomitant guidelines can carry subtly different agendas. The impact of any guideline depends on whether or not it squares with GP presumptions, with patient proclivities, with specialist preferences and with further resource constraints. Different mandates may thus be seized on, sometimes opportunistically, sometimes ruthlessly and sometimes unintentionally by different stakeholders, and the result may be broader system strain. Such system dysfunction may result even if the original objectives are quite benign. In a further paper, Cornford and Oswald suggested that the 2-week rule was flawed from the start, being based on two conflicting agendas: the need to limit referrals to high-risk patients but also the need to increase referrals to achieve early detection.²⁸⁴

It is not all bad news. Barwick and colleagues²⁸⁵ provided an update to the study by Malik and colleagues.²⁸⁰ Over the years there had been a gradual accommodation to and a subsequent increase in admissions under the 2-week rule. A higher proportion of them were also considered appropriate, given the resulting detection of a malignant tumour. The folk explanation here is that change takes time to bed down.

The systems explanation is that the change in guidance reverberated across the entire system and thoroughgoing reorganisation was required to bring compatibility to the different referral routes. The referral centre in question went on to introduce a multidisciplinary team co-ordinator, who monitored care across intakes, who established a one-stop shop for access to ultrasound and biopsies and who was empowered to 'stop the clock' on patients who 'clearly do not have a malignancy'.

The insertion of the 2-week rule is perhaps an acute example of how guideline overload may lead to competition for resources among stakeholders. Cookson and colleagues, however, suggested that it may be the norm: 'seemingly robust professional advice is distributed to clinicians, which, if adopted, will distort resource allocation and waste scarce resources'.²⁸⁶ A high-quality guideline may attempt to make suggestions on workflow changes and resources allocations but, in zero-sum games, the costs involved in implementing a new process will always generate costs elsewhere in the system. Competition and displacement effects are the norm and the guideline implementation challenge is to anticipate them.

Conclusion

Here we attempt to draw together the key implications of the above analysis, which we separate into three paragraphs: substantive, methodological and procedural.

Given the tumultuous increase in their usage, guidelines should be regarded as part of the fabric of service delivery rather than as separate interventions with unique objectives. When a new guideline is introduced, or when an old one is updated, it sits alongside a range of other organisational controls, rules, norms, customs, practices, targets *and* guidelines. Guidelines will always be a constituent part of a system of governance and their destiny rests on how well they are absorbed into that system.

There has been an equally tumultuous increase in the research on guidelines and it is important to find the appropriate means of synthesising the burgeoning evidence. We have argued that it is futile to try to answer questions about whether or not guidelines work. The only reasonable answer is that 'it all depends'. We have also pointed to the limitations of trying to seek a master framework, listing all of the factors, on which success depends. There are scores and scores and scores of barriers and facilitators that help and hinder guideline implementation, but these do not resolve into some sort of winning formula because the factors identified are always interdependent. Dealing with barrier A will always have effects, anticipated and unanticipated, on enabler B, and so on. This interconnectivity does not leave systematic review with the task of describing unending, unforeseeable change. History does repeat itself and a raft of discernible system strains can be detected as guidelines are introduced. There are oft-repeated tensions between semiautonomous stakeholders. There are inevitable strains between blanket recommendations and limited resources. There are persistent frictions between national ambitions and local experiences. System change does have a pattern and in the paper we have described a method for analysing how some of the more familiar strains evolve and resolve.

This brings us finally to the business of research utilisation. Evidence is supposed to inform policy and practice and this ambition necessitates a realistic understanding of the roles of the policy-maker, the manager and the practitioner. Our understanding, in this domain, is that the time has long passed for high arbitration about whether or not to have guidelines. DMs are already awash with guidelines and so they are not sitting, Pilate-like, awaiting the definitive verdict about their effectiveness. We also presume that very few DMs work in splendid isolation and have the task of implementing an entire guideline system from scratch. They are already inundated with guidelines and they do not operate by ticking off ratified checklists about how they should be managed. We do suppose that the key business is system improvement. Policy-makers, managers and practitioners are awash with guidelines. Accordingly, the key task is to dovetail the latest manifestation into an existing organisational structure. The really difficult activity is to engineer and re-engineer guidelines so that they work smoothly with the pre-existing system. System technicians want to know how to get the grease to the squeak – they want to know about system strains and how to resolve them.

Chapter 8 Conclusions: facing the challenge of complexity

This chapter has two main sections: a summary of key learning points in the form of ‘design principles’ for demand management schemes, followed by a ‘modest recommendation’ on a promising technique that is finding increasing use in the development of complex service delivery interventions.

Lessons learned

We commence with some brief remarks on our topic, our study and the nature of the lessons learnt. The evidence amassed in this report confirms our preliminary hypothesis that demand management constitutes a wicked problem, which has defied clear and easily reproducible solutions. In scouring the evidence base we have unearthed an ever-growing list of contingencies that have to be overcome if such schemes are to succeed. We consider that the task of the impartial reviewer is to be the bearer of reasoned judgements, regardless of whether they constitute good or bad news. And, in this respect, the negative inference to be drawn from our review is that there are no bullets (magic or silver) to be fired that have the omnipresent capacity to slay rising demand. There are no preferred interventions that have the capacity to outperform all others. There is no best practice manual to be found out there or in this report to guide health-care personnel on when, where and how to make referrals. With equal emphasis it is worth recalling a general principle mentioned earlier, namely that the designation ‘wicked problem’ does not signify a hopeless, unfathomable, insoluble task. What this means is that there are many, diverse, hard-won, local and adaptive solutions. The key evidence in this review comes in the form of detailed expositions of the immense difficulties and occasional, bespoke successes, in bringing into equilibrium the interlocking systems on which sustainable change depends.

How can this highly nuanced evidence be marshalled for benefit of future policy-makers and practitioners? Hopefully, we have presented evidence enough to show that our recommendation cannot be neatly parcelled up into a league table that differentiates intervention winners and losers. Some previous reviews of demand management interventions have attempted to work to this formula, seeking to discover whether or not programmes of type A work better than type B or type C, and so on. We present evidence in *Appendix 1* to show that these reviews always end in indeterminacy rather than an inventory of programmes to be cherished and a hit-list of those to be chastised. In exactly the same way, it would have been nonsensical for us to conjure up claims that ‘referral management centres’ work better than ‘GPs with special interests’, or that ‘guidelines’ are inferior to ‘direct access to tests’ and so on. The proper answer to the ‘what works’ question is always ‘it depends’, and the job of research synthesis, we submit, is to deepen our understanding of the contingencies. And this is what we have attempted in the main body of the review.

So, how does one convey the lessons learnt from a patchwork quilt of success and failure? Another approach, much more consonant with a complex and mixed picture of outcomes, is to present a listing of the ‘common denominators’ of success, those features of implementation that make a difference. The logic here is captured by adapting the old adage, ‘It ain’t what you do, it’s the way that you do it’. Although this proposition underscores perfectly the predicament of demand management, these laundry lists of implementation ‘must dos’ have a tendency to the tautological and the banal. It is indeed the case that we have shown that demand management requires ‘good communication’, ‘strong leadership’, ‘careful forward-planning’, and so on. We also know that there is evidence aplenty to show that the way forward lies in ‘local’, ‘adaptive’ solutions. Alas, this advice applies to almost everything under the sun and supplies little more than a moral message to the practitioner.

Another familiar approach to research recommendations is to present vignettes of successful 'case studies'. These feature, for instance, in the NHS Institute for Innovation and Improvement's document, *Improving Patient Flows in the NHS: Case Studies in Reducing Delay*.²⁸⁷ What the case studies are able to show is the complexity of problem and response. There are many similar instances, recounted in the main body of our review, in which groups of practitioners have fashioned and refashioned an intervention to provide important local, adaptive solutions. Again, there is a drawback. The highly specific nature of the 'exemplary cases' may have a tendency to draw a different, localised retort – namely, that it is not really like that in *my* bailiwick. My clinical domain, my patient pathway, my management structure, my funding streams are all different.

Accordingly, what we have tried to present in these conclusions is a mixture of these two approaches, grasping their advantages and skirting some of the difficulties. Rather than present self-evident lists of common factors that lie behind intervention success, we have generated a series of propositions, drawn from our analysis, that explain why particular processes or features contribute to successful outcomes. In other words, we attempt to convey the core explanations embodied in the solutions and it is these design principles that may be transferable. In order to make this exercise reasonably economical we have restricted ourselves to 12 design principles, two drawn from each analytic chapter in the report. We also add vignettes, miniature case studies from the previous reviews showing how each principle may be realised. What this mix of design principle and local practice attempts to portray are some common predicaments that apply in common demand management interventions, pointing to the strains that have to be overcome and the interdependencies that have to be forged. We hope that these learning points echo with the struggles of practitioners as they design their own referral systems.

Design principles in demand management

In order to preserve the evolution of our own argument, we present our key lessons in the same chapter running order as in the review. It will be recalled that *Chapter 2* presented the background to the demand management problem and *Chapter 3* presented an overview of potential solutions. The remaining chapters, the evidence chapters, reviewed the primary studies on a selection of key interventions. Each chapter posed the same question – can the particular intervention 'control and shape demand?' – and each chapter examined the many contingencies that have to be battled through to achieve this desired end. Readers are referred back to the conclusions of each chapter for our detailed results. Here, we reduce our findings to a brief set of summary propositions, which explain the design principle that needs to be in place in order to meet the policy aspiration. These key design ideas are summarised in *Box 44* at a level of abstraction that allows for transfer and application across wide range of service remodelling. They are followed by key vignettes, which are intended to exemplify practical applications of the each principle. They are also presented on a chapter-by-chapter basis.

Chapter 2 gathered evidence on the manifold interconnected processes that generate demand for health care. It presented us with two daunting hypotheses. The first is that multiple, intertwined problems are unlikely to yield to singular solutions, however well aimed. The demand for health care may be regarded as a swelling punch bag. Landing a blow in respect of one problem may simply be absorbed as other vicissitudes gather. Our second observation is that health service provision is itself a magnet for demand. Many of the solutions to demand inflation involve the creation of new roles and services. The solution always has the potential to attract further demand (*Box 45*).

Chapter 3 examined the programme theories that underpin the small army of interventions that have attempted to quell the inflation in demand for services. Interventions are theory incarnate. Practitioners do not simply implement programmes. They think them through. Much is to be gained by tracing this perpetual quest to improve programme effectiveness. When the sprint of implementation turns to the middle distance and then to the marathon we noted the tendency of programme theories to become whole-system theories. The programmes that appear to survive are the ones which adopt a solution, and then adapt that solution in the face of its unintended consequences, and then adapt the adaptation, and so on (*Box 46*).

BOX 44 Key design principles in demand and referral management**Chapter 2 The runaway train: the multiple, intertwined causes of growth in demand for health care**

- Solutions to health service delivery problems can be only as good as the diagnosis of the underlying causes. As always, basic science needs to precede the engineering applications.
- The reliance on common sense diagnosis of demand and capacity imbalance as well as the lack of understanding of their complex interconnected roots has led to partial solutions, which often displace rather than solve the underlying problem.

Chapter 3 The policy response: charting the family of purported solutions

- Demand management should be understood as an evolving and highly variegated family of policy conjectures. They are usually mounted as one-off, single-lever interventions aimed at change in staff motivations or routine procedures or professional roles or administrative arrangement or strategic ambitions.
- Regardless of this primary aim all demand management interventions reverberate around the entire service apparatus and require system-level planning and implementation.

Chapter 4 Organisational change: referral management centres – can they control and shape demand?

- Significant organisational change requires a smooth transition of governance arrangements and this can be achieved only with endorsement from the key incumbents in the previous regime.
- Detailed and prompt attention must be paid to every ensuing modification in staff functions and responsibilities. Such changes need to be acknowledged, justified and then organised with bespoke and agreed guidance.

Chapter 5 Role change: general practitioners with special interests – can they control and shape demand?

- Design for equidistance between new and old roles and a closely defined case load so that special-interest roles avoid passive support or interprofessional rivalry and involve 'selective substitution' at key decision points.
- The new roles should be normalised and positioned with careful negotiation, training and recruitment. Even a well-designed GPwSI service can increase handovers and become a holding station, furthering supply-induced demand.

Chapter 6 Procedural change: direct access to the results of clinical tests – can it control and shape demand?

- General Practitioner direct access to rule-out tests and clinical indicators tests leads to greater efficiencies in demand management than GP direct access to differential diagnostic tests.
- Clear guidance from specialists indicating which patients should be referred for the tests and how they should then be managed are required to ensure patients that can be managed in primary care are diverted away from secondary care.

Chapter 7 Learned counsel: can guidelines control and shape demand?

- The passive dissemination of guidelines is ineffective. Instead, ownership of guidelines should be instilled and installed within the division of labour responsible for their implementation.
- Rather than adapting service delivery arrangement to generic guidelines, guidelines should be adapted to meet evolving local delivery requirements.

BOX 45 Vignette: supply-induced demand

Additional services, even if they are delivered in the name of prevention or screening may inflate demand. One example, reported in *Chapter 2*, is illustrated in the result of a two-stage randomised trial.⁴⁷ Doctors working in a university hospital clinic were randomised to receive income in two modes: by salary or by fee per service. Patients attending the clinic were also randomly assigned, in this case to the different payment regimes. The key result was that for the same population of patients the fee-per-service doctors scheduled almost 30% more return visits than did those receiving a salary. Most of the discrepancy was attributable to a 50% increase in the scheduling of 'well child' visits, which the authors describe as having 'doubtful' medical worth.

Increasing the number of patient pathways, as planned in many triage processes, may reduce waiting times and increase the appropriateness of the service received, but these gains are often won at the expense of an increase in the total volume of referrals. A 'speedy evaluation cardiology service', reported in *Chapter 4*, triaged patients along four separate routes, including a new nurse practitioner-led scheme.¹⁶⁶ Following the introduction of the service, the waiting time to first appointment reduced from 71 to 33 days. The total number of eventual referrals to cardiology services increased from 1512 to 2574. The volume of case seen by the new triage service also doubled in the first 2 years following its introduction.

BOX 46 Vignette: whole-system adaptation

A *Pulse* article by Wootton and Whiting from September 2012 argued that truculent resistance to organisational restructuring in the form of a RMC (Gateway) was gradually overcome through role and motivational change:

*... the vast majority of GPs in Manchester now use the gateway ... a number of factors we believe have led to this almost universal acceptance. The first is that this was not something introduced suddenly and on a whim. It was developed over a long period of time and after plenty of discussion and collaboration with primary care through local meetings, the CCGs and the LMC. We showed GPs it would be effective through the smaller pilot at the beginning. One very important factor in setting up the gateway was the use of local GPwSIs and senior, well-known GPs. It meant the other GPs respected the decisions made and the feedback they were receiving. Because the advice is provided in a very professional and supported way, the GPs saw it as more of a CPD tool than a criticism of their decisions.*¹¹⁵

Another *Pulse* article, reported in *Chapter 3*, made the case that the real effectiveness of the introduction of a system to peer review the accuracy and pertinence of GP referrals lay not in its manifest function but in the latent activities that surrounded it.¹²⁴ The author argued that the educational aspect of the review service is the key change. Rather than passively accepting or stubbornly rejecting the peer reviewer's opinion, there was a weekly meeting to discuss those referrals where an alternative option was recommended, with registrars presenting the case. 'We issue CPD certificates with these meetings, so instead of becoming a chore it is part of GPs' career development'.

CPD, continuing professional development. Quotes reproduced from Pulse with permission

Chapter 4 examined the theory that RMCs and other centralised triaging services could bring efficiencies to demand management by rationalising decision-making along the habitual, heterogeneous referral pathways that exist presently between primary and secondary care. RMCs only function to this end in the presence of a cluster of conducive conditions. Major administrative reforms generate power struggles and there is need for extensive collaboration between NHS managers and clinicians to agree on the governance structures. Continuity in the form of the recruitment of experienced local GPs (rather than adjunct professionals) to help manage the triage service is also found to be important. All of the protocols and guidelines that define the RMC's remit and functioning should be codeveloped rather than imposed (*Box 47*).

Chapter 5 examined the theory that an intermediate professional role, the GPwSI, blending the expertise of the GP and the consultant, could manage demand in the system by siphoning off and dealing with an intermediate case mix appropriate to their medical experience and proficiency. This ambition failed to realise and the new role descended into an administrative support function in situations where the consultant remained in relative control of referral decisions and the protocols that govern them. The intermediary function also failed when GPs retained referral habits either by maintaining direct referrals to secondary care or by using the new GPwSI pathway to offload many cases for purposes of patient reassurance. The new function comes to fruition only when there is a steadily negotiated division of labour, recruitment strategy, case mix and physical spacing (and political truce) (*Box 48*).

BOX 47 Vignette: reconfiguring the administrative apparatus

The RMC set up in Kernow, Cornwall, provides one such example of collaboration between NHS managers and clinicians in agreeing on the remit and function of the service. Although the service was owned and funded by the local PCT, the clinical direction of the service was set by a CIC board consisting of five locality GP leads, a LMC representative and the PCT's director of commissioning. The guidelines determining how triage decisions were undertaken were drawn up by the CIC board and secondary care consultants. The idea behind these guidelines was to improve GPs' confidence in managing patients themselves and carrying out investigations before referring patients. The perception was that RMCs were successful because GPs were considering more carefully whether or not patients should be referred, rather than the RMS rejecting large numbers of referrals.¹⁵⁸

Case study A in Imison and Naylor's King's Fund report on referral management also shows how important these governance structures are in setting the right conditions for one of the key mechanisms through which RMCs are considered to work – through peer scrutiny and feedback on referrals.⁹³ Here, again, the clinical direction of the RMC was set by local clinicians and local GPs expressed few concerns about the triage process. Imison and Naylor report that those working in the RMC perceived that this was because clinical risk was managed through 'the use of clinical pathways, knowledge gained through experience of being a triager, the availability of peer support and the availability of a robust tracking system for referrals'. GP triagers perceived that the RMC worked because local GPs 'really have the feeling that one of your peers is going to be reading the letter you write' (triage GP, site A), while local GPs felt 'I think very clearly now about what I am referring for ... what background information does the triager need to know?' (Reproduced with permission from The King's Fund from Imison C, Naylor C. *Referral Management: Lessons For Success*. London: The King's Fund; 2010.⁹³)

BOX 48 Vignette: establishing the intermediate function

All studies show that striking a fine balance between roles is the key to the acceptance of GPwSI services. Martin's inquiry displayed the gulf between the professional identities in primary and secondary care ('we speak a different language') and shows that consultant acceptance of the GPwSI services was the key determinant of the level of access for support, supervision, ongoing training, etc. Without consultant acceptance, GPwSIs end up working in isolation and with insufficient clinical back-up.¹⁹³

Rosen and colleagues' study, reported in *Chapter 5*, reported on four highly contrasting GPwSI services.¹⁷⁷ Characteristics of the cases in which the intermediate role was more securely established included (1) progressive change in the patients seen by GPwSIs (e.g. a musculoskeletal service 'graduating' from hand problems to collaborative work on knees, ankles and shoulders); (2) services changes labelled as pilots – and were, therefore, potentially reversible – were easier to accept than top-down 'diktats'; (3) recruitment of GPs with existing working relations with specialists, who 'know how we work, what treatments we use, what guidelines we have for different conditions'; (4) clinical locations which privilege the opinions and uphold the status of GPwSIs in the eyes of the patient; and (5) co-operative working, avoiding major territorial claims and power struggles.

Chapter 6 examined the theory that providing GPs with direct access to tests can enable GPs to distinguish between those patients who can be managed in primary care and those who require referral to secondary care. GP direct access to tests, designed to 'rule out' serious pathology, or 'clinical indicator tests', designed to identify where patients were in a disease trajectory, led to greater efficiencies in demand management than did GP direct access to tests designed to provide a differential diagnosis. This is because 'rule-out' tests and 'clinical indicator tests' provide a clear distinction between patients requiring specialist care and those who can be managed in primary care. In contrast, differential diagnostic tests often differentiate between different diagnoses that all require referral to secondary care. Furthermore, there is greater consensus between consultants and GPs regarding the timing and function of rule-out tests and clinical indicator tests in the patient pathway. The differential diagnostic tests that GPs are provided with direct access to are not the diagnostic tests of choice for specialists, who may utilise their clinical skills or other tests to reach a diagnosis. However, the improvements in patient flow resulting from GP direct access to rule-out or clinical indicator tests are realised only when there is clear guidance from specialists indicating which patients should be referred and how they should be subsequently managed (*Box 49*).

Chapter 7 examined the role of guidelines in moderating demand, the evidence revealing, repeatedly, an elementary design principle that they work more effectively if they are adapted to local circumstances. The process of adaptation, however, requires far more than the rewriting of guidance to a local rubric. The guidelines need to be mapped onto local organisational structures. A division of labour should be agreed on which team members are responsible for which aspect of the guidance. The people responsible for adapting the guidelines should be the same people who organise the implementation of the new guidelines. Without this level of accommodation the documentation is likely to be added to 'the guideline mountain in the corner of the clinic' (*Box 50*).¹³⁴

BOX 49 Vignette: differentiating patients according to test results

For direct access to testing to be effective in reducing demand for specialist care, the test must be able to differentiate between patients who need specialist input and those who do not. The DAMASK trial showed that, where tests acted simply as confirmation that a referral was necessary, they were unlikely to be effective in reducing demand. Indeed, in this instance direct access had no impact on demand for specialist care but did increase the demand for the test: many of the patients referred directly to the specialist were diagnosed through clinical examination without the need for a MRI.²²⁰

Even where the test itself can differentiate between patients who need specialist input and those who do not, this is dependent on the recipient of the results having the skills to interpret the results appropriately. Direct access to testing works best when the specialist maintains some input in patient care via the setting of clear guidance on the management of patients in the light of the test result. The paper by van Gurp and colleagues shows the value of receiving specialist advice with the test results.²³³ At the outset of the evaluation of a direct-access echocardiography service the results were returned to the GPs without advice attached unless the results showed pathology which required referral. During the cohort study the cardiologists were asked to provide advice with every echocardiogram and the reporting evolved to include an online report sheet which the cardiologist was required to complete before the results could be released to the GP. This report thus supported the GP in managing those patients in primary care rather than simply identifying which patients should be referred to a specialist.

BOX 50 Vignette: the local adaptation and actualisation of guidelines into individual roles

Capdenat Saint-Martin and colleagues' case study, reported in *Chapter 7*, embodied this design principle.²⁶⁰ The background here was the perceived overuse of preoperative tests for anaesthetic risk. National guidelines had shown that they were costly and had highly restrictive diagnostic use and yet referrals for such tests remained stubbornly high. The French team undertook a process of local adaptation of these underused guidelines. The organisational structures (decision trees) responsible for ordering the tests were mapped and team discussions were undertaken of how to refashion guidance to make them appropriate to the specific needs of each local stakeholder. Measures were taken of referrals for such preoperative tests before and after the introduction of the local negotiations and in the targeted low-risk groups requests fell from 80% to 48%. Capdenat Saint-Martin's team draw a system-based lesson. 'We think that the main contribution of this work is the linking of the process of local adaptation to an analysis of the organisational aspects of the practice and the emphasis we placed on the organisational aspects of change.'

Group model building

This coda provides a modest recommendation for a tool that may be of assistance in future demand management interventions. The idea follows logically from all of the previous analyses, though it was not formally part of our investigations.

There is a remorseless order to all the ventures we have reviewed attempting to manage demand and rationalise referrals in health-care services. Regardless of the point of origin – be it a new guideline or a fresh post or a novel procedure or a management remodelling or a government imperative – the change reverberates around the whole service. The system is thrown out of equilibrium, with both intended and unintended outcomes. Often the positive and the negative occur simultaneously, with the gains in speed and accuracy of referrals accruing in one patient pathway being matched by overload and imprecision along another. What then happens, post intervention, is that key personnel respond to the resultant of the various intended and unintended transitions. In other words there is always an adaptive and gradual response: the new guideline is ignored until it is embedded in job descriptions; the new role becomes low-grade support until local political conciliation is achieved; the new centre is bypassed until its providers and recipients become involved in its management. These adaptations are endless. Sometimes they occur so distantly from the original initiative that it is hard to envisage them as part of programme implementation. Nevertheless, we have evidence aplenty to show that improvement only occurs thanks to this perpetual process of ‘snagging’ and ‘desnagging’ across the entire referral system.

How can this insight be translated into a way of building programmes whose fate is always determined in the wider health-care system? One starting point is to revert to a ‘system change’ perspective in evaluating any intervention. The idea is neatly encapsulated in a quotation from a recent paper on ‘complex responsive processes’.²⁸⁸ This is defined as ‘a perspective in which organisations are regarded as patterns of conversations between independent individuals . . . Learning can be understood as the qualitative change in conversational patterns’. Although we might prefer to label the active ingredients here as ‘groups’ and ‘negotiations’, the point holds that it is the interaction between different sectors and segments of health service delivery that seals the fate of a particular change initiative.

This insight has been much seized on. Occasionally it has involved throwing out the baby with the bathwater. The recent NHS Energy for Change initiative recognises, quite correctly, that many change initiatives simply fade away and that they work in solo or solitary settings, and they meet with determined defiance attempts to spread and sustain improvement in quality. The system-level solution in this consultancy report is to create a widespread change culture, which is to be fed by the development of ‘organisational energy’. This goal, in turn, is met by applying a fivefold framework that will nurture such energy for change across the entire system. The development plan is to measure and then generate more energy across its five domains of social, spiritual, psychological, physical and intellectual activity.²⁸⁹

Although we have encountered much fading away and plenty of stubborn resistance in our review, we have no evidence pointing to the fact that energy depletion is the core of the problem. The demand management problem in every case is a practical issue – disagreement over the implementation of an intervention that is supposed to change the way a service is run. In all the would-be examples of demand control it is the fate of specific ideas (programme theories) that enter the boiling pot of negotiation. And insofar as they emerge constructively from those negotiations it is because the underlying idea has been transformed progressively into something that is more workable. System-wide solutions are about concrete ideas and practices and not abstractions such as energy diffusion.

The real system problem that we have encountered, over and again, is well summarised in a paper recognising that large institutions contain networks of semi-independent organisations.²⁹⁰ The paper argues that the interactions and interdependencies within these networks are so complex that it is virtually impossible to control and design them centrally. There is no single command and control structure, with the result that management imperatives always generate intended and unintended effects as they are negotiated and renegotiated. Top-down change most often resolves itself into bottom-up, emergent solutions. This pen-picture is a perfect summary of the predicament of all the demand and referral initiatives we have reviewed.

It should be emphasised that such analysis is not a defeatist cry. It acknowledges that solutions, gains and positive outcomes can follow major interventions but recognises that those solutions are adaptive and local. Our evidence is particularly pertinent to this final adjective. Unintended consequences may well flow because of local funding problems, condition-specific patient pathways and context-specific peculiarities in line management and, indeed, due to that particularly awkward individual who finds himself landed with unwelcome responsibilities. Negotiated solutions follow a similarly tortuous and provincial path.

Accurate as is the picture painted by this scenario, it has a wait-and-see quality that exponents of evidence-based policy may find disappointing. Is there any way of speeding and strengthening the emerging solutions? Is there any way of avoiding false trails and cul-de-sacs? One emerging and increasingly celebrated strategy for harnessing and quickening these emerging, bottom-up solutions goes by the name of 'group model building',²⁰⁰ and we close the report with a brief example and a preliminary assessment.

Group model building, having found much use in the business sector, has more recently been applied in health-care settings.²⁹¹ Reminiscent of the initiatives reviewed in this study, Pieters and colleagues²⁹¹ study the redesign of the care process in the attempt to produce 'more efficient and patient-friendly flows'. The system in question, the Dutch obstetric service, is characterised as having a large number of semiautonomous professional stakeholders, harnessed together in diverse and ever-changing constellations. The system operates to a basic model, which differentiates pregnant women into three risk groups – each one cared for in different locations and by different groups of professionals. As in the systems we have studied, the responsible networks are numerous (radiographers, laboratory technicians, general nurses, gynaecologists, assistant gynaecologists, midwives, obstetric nurses, maternity nurses and paediatricians) and their functions are spread over a considerable period (antenatal, intrapartum and postpartum care).

Relatively poor perinatal and maternal mortality rates, together with a perceived lack of co-ordination, co-operation and exchange across these stakeholders, drove the need for reform. This led to a voluntary, exploratory exercise, conducted in the action research tradition. The group model building element has the task of creating the 'bottom-up emergence of successful network relations' to carry forward the reforms. Simply described, it consists of an iterative succession of workshops, conducted within and between the various professional groups listed above. Over time the focus of discussion ranges from problem analysis, to defining potential improvements, to trouble-shooting those improvements, and then to the first stages of implementing the emerging schema. Discussion centres around the production of 'process maps' incorporating the roles and responsibilities of the various stakeholders (the complexity of such maps has been illustrated via several examples in earlier chapters). The process is orchestrated by a facilitator who leads, records, draws out models (literally) and acts as intermediary between the various groups.

The idea of the exercise, in realist terms, is to gain a sound understanding of the programme theory underlying a potential improvement – as seen through the ideas of each grouping. They are thus urged to 'map' interventions according to the following typical format (*Box 51*).

The succession of meetings begins to explore the consensus (or lack of it) between respective maps. Pieters and colleagues²⁹¹ describe how a particular map was perceived as privileging the perspective of the gynaecologist and how it generated suspicion from other groups and how, in turn, the comparison with other maps led to potential compromises that could be further articulated and tested. The humble facilitator is charged with mapping the emerging maps (hence the term 'group model building'). The process ends with a range of 'improvement proposals', the expectation being that, as they had been through this process of road testing (on the local byways to boot), the chances of successful implementation are enhanced.

BOX 51 Format description of improvement

- Description of the improvement.
- Problems dealt with.
- Unwanted consequences.
- Hypothesis why the improvement should work.
- Relations with other improvements.
- Necessary conditions.
- Advantages for gynaecologists, midwives, assistants, pregnant women, care process and final outcomes of care.
- Needed effort from gynaecologists, midwives, assistants, pregnant women, care process and final outcomes of care.

Clearly such a process chimes with our findings about the emergent, local adaptation of demand management interventions. It is, of course, a paper-and-pencil simulation of a genuine implementation process. The rotation of workshops is a mere simulacrum of the real hide-and-seek of interdisciplinary working. Nevertheless, group model building has the potential to shift both from an ex-post to an ex-ante as well as from a top-down to bottom-up appraisal of service reorganisation. At its best it can anticipate some of the problems that reveal themselves only after ambiguous interventions are launched on unsuspecting audiences (c.f. the ambivalent beginning of the GPwSI interventions).

And at its worst? There are caveats, of course. Group model building has its roots in management consultancy and the intricate, participatory research processes described above can be reduced to standardised, off-the-peg events. Some consultants use patented and expensive software tools and handbooks to facilitate workshops and map making (others such as Scriptapedia are available under a Creative Commons Licence). The other obvious consideration is that group model building, itself a form of intervention, will work only in certain hands and in certain circumstances and in certain respects. Much depends, for instance, on the skill and experience of the facilitator. Above all, the success of the entire exercise depends on the willingness of all stakeholders to collaborate. In this respect, group model building might be said to be a perfect imitation of the real thing. As we have seen in the body of the review, turf wars are a perpetual problem in service reorganisation and may end up as the block over which everything stumbles.

Given the need for delicate relationship building, there is a danger then that group model building becomes regarded as 'more art than science' and in this respect we note welcome efforts to evaluate and review the effectiveness of the approach.^{292,293} It is of interest to note that they use a realist evaluation and a prototypical realist syntheses approach in making these assessments.

Our conclusions thus end with a 'modest recommendation'. The main body of our review has explored many interventions mounted in the name of demand and referral management. The evidence has pointed to the many, many different nuances required in their implementation and the need for careful targeting to the appropriate stakeholders and institutional spaces. One of the elemental lessons learnt has been about the importance of local collaboration. Group model building offers considerable potential in realising this goal.

Recommendations for future research

Our synthesis has identified some of the contingencies that influence the success of demand management interventions. We concluded that demand management interventions require local, adaptive solutions in order to make them work. Implementation requires stakeholders to take account of the whole system within which any collection of interventions is placed. With this in mind, we caution against future research endeavours that attempt to identify preferred interventions which have the capacity to outperform all others. However, here we make a small number of recommendations about the possible approaches future research may take.

As we anticipated in our original proposal, the literature on demand management interventions is uneven; the guidelines literature is expansive and spans the full range of research methodologies, while the literature on RMCs and the peer review of referrals is much smaller and is often limited to case studies or observational studies. Although our synthesis has demonstrated that important lessons can be learned from this literature, more detailed mixed-methods studies comparing the implementation of different configurations of demand management interventions in different local contexts would be valuable. Such studies should also take a systems approach and consider how such interventions influence and are influenced by the entire system within which they are placed. Such studies should also recognise that, in the real world, demand management interventions are rarely implemented in the singular and consist of a configuration of different interventions (e.g. peer review, guidelines and triage). What is needed is a closer examination of the processes through which these configurations are adapted to local circumstances. This would enable a deeper exploration of the contexts within and mechanisms through which such interventions become integrated into local practice.

There is also scope for further evidence synthesis. We examined the fortunes of GPwSI services and whether or not they provided support of substitution for the work of consultants. The substitution theory is ubiquitous in health and social care and a realist synthesis to compare the fortunes of different practitioners placed at different professional boundaries (e.g. nurses/doctors, dentists/dental care practitioners, radiologists/radiographers and so on) would be valuable to identify the contexts and mechanisms through which substitution, support or short-circuit occurs. While developing our programme theories for RMCs, we also touched on the literature exploring whether rule-based systems embodied in protocols offer a support or substitution to clinical expertise, for example, such as those protocols followed by nurses working in NHS Direct or NHS 111 services. A realist synthesis of this literature would also be useful to explain the different outcome patterns observed in the extent to which these services are able to curb or redirect demand for primary and emergency care. We would argue that the systems theory developed in our review (see *Chapter 3*) offers a useful 'reusable conceptual platform'¹¹ to explore these questions.

Chapter 9 'Thinking it through': prompts for practitioners

Our concluding chapter is aimed at policy-makers, planners, managers and practitioners with responsibility for the throughput and planned care of patients. Matching demand with capacity is a challenge throughout the entirety of health-care provision and your duties may be located at any point in the interface between primary and secondary care.

The recommendations that follow are 'evidence based' and result from a review of the vast array of research on schemes attempting to tame the seemingly inevitable rise in demand and in referrals for all health-care services. None of that evidence is presented here and we rest content to unveil the following conclusions using those weasel words, 'research has shown'. Interested (or disbelieving) readers should consult the main report for the empirical details.

What are the major implications of our investigation? A constant finding of our inquiry is that demand cannot be managed by rote. There are no silver bullets. There is no manual of best practices to be blindly imitated. Every pressure point is to some extent unique. There are, however, many instances in which available capacity has been well managed and this has usually followed through trial and error as practitioners think through a succession of challenges that apply in their corner of the patient pathway. Different practitioners occupy quite different intersections in the system but what is transferable is the very process of 'thinking through' all of the various conditions and caveats. Our recommendations are, thus, couched as series of prompts designed to get you pondering. They are decision aids intended to help *you* to solve the problem.

Prompt 1: what (exactly) is the problem?

A problem well put is half solved.

John Dewey²⁹⁴

Faced with soaring demand and rising waiting lists, as well as all of the accompanying angst, you may consider that 'the problem' is obvious. It lies with nagging patients or officious clinicians or obdurate bureaucrats. Or, given the elephantine nature of modern health services, the underlying cause might seem more amorphous: the glitches are inherent in 'the system'. Given, however, that a problem well understood is a problem well on the way to resolution, a close consideration of manifold causes of demand inflation is a key first step. The literature is replete with explanations. Which of the following apply in your bailiwick?

1. Irregularity of demand? Are queues unpredictable, with services sometimes over-run and at other times having spare capacity?
2. Professional closure? Is 'turf protection' a problem with opposition about referral and treatment priorities across different stakeholders?
3. Personalisation of referral decisions? Are clinical referral decisions sometimes clouded by the inevitable intrusion of social and emotional judgements?
4. Self-propelling diagnostic cascades? The first step in referral decisions is often the diagnostic test. These have multiplied in number and sophistication. They are often 'nested' – thus automatically propelling patients along the referral chain.
5. Supply-induced demand? New treatments and services roll constantly off the production line. It is generally in the physician's interest to promote them and the manager's job to provide access, thus creating additional demand that then may need to be checked.

6. Changing demographics? Populations evolve and their changing composition has a direct effect on demand. The so-called 'demographic time-bomb' of ageing is the best known example but there are numerous other bulges as populations and provision shift.
7. The internet and the informed patient. Patients' expectations and medical knowledge have grown rapidly in the 'information age'. Concomitant innovations in policy such as 'shared decision-making' and the 'choice framework' may also fan demand.

Some of these causes may simply not apply within medicine's many specialities, but our experience is that most practitioners display an instant sense of recognition when presented with this list. Research shows, moreover, that these multiple drivers of demand often interlock and reinforce each other. In the design of any new demand and referral management schemes it is important to:



Think through the full array of underlying causes. Use the sevenfold typology above to prompt analysis of the particular bulge in demand that you may be confronting. The immediate challenge is likely to have deeper roots.

Prompt 2: what are the options?

My washing machine overwhelms me with its options and its sophistication.

Uma Thurman²⁹⁵

Health services are a mite more complex than washing machines, with the options for reform multiplying and evolving rather than being fixed and pre-programmed. It is, nevertheless, important to think through all of the purported and existing options mounted in the name of demand management as the basis for planning and preparing for any fresh challenge.

Research reveals the remarkable array of schemes, interventions and strategies mounted in the name of demand management. *Box 52* gives a brief list of options:

Is it possible to digest the lessons from this little lot without, Thurman-like, being completely overwhelmed? It is useful first of all to think through the basic mechanics of the proposed reforms: what exactly is it that comes under change? Beginning with this simple question reveals quite different starting points. Ambitions vary from the macro to the micro: some interventions seek to introduce new centres to administer the entire referral apparatus; some introduce new roles and key personal with specific responsibility for managing demand; some set in place new procedures and guidelines in the hope of standardising referrals; and some offer incentives to improve to motivation of staff towards specific referral targets. There are powerful advocates of all of these measures.

Research reveals a vital lesson in relation to these varied ambitions, captured rather well in the aphorism 'you can't have one without the other'. Sustainable change requires co-ordinated action at the macro, meso and micro levels. Starting at the top, the evidence shows that organisational change is resisted unless new roles are properly negotiated, procedural modifications are agreed, and staffs are motivated to work under the new arrangements. Starting at the bottom, newly incentivised staff will be effective only if they operate with procedural elbow-room to accommodate their new resolve and targets; with the co-operation of colleagues elsewhere in the referral chain; and if there is harmony with co-existing organisational structures and strategies.

BOX 52 Demand management options

Referral management centres.

Triage services and gateways.

Revising fundholding.

Increasing practice-based services and diagnostic testing.

Joint-working, shared care and consultation liaison.

The creation of intermediary roles between primary and secondary care.

Attachment of specialists to primary care teams.

Peer review and audit of referrals.

Educational interventions.

Guidelines for practitioners.

Information for patients.

Shifting venues of specialist care into the community.

Financial incentives to meet targets.

Increasing the use of telemedicine.

Such co-ordination is axiomatic in change management. Policy architects and managers are often driven to eliminate specific inefficiencies within a service. But, whatever the starting point, it is important to:



Think through the system-wide implications of any innovation. Change has no resting point. Change will always be resisted at some point. It is impossible to anticipate all of the consequence of any reform as it travels through the system but this is no reason to ignore the challenge. This starting point will instil the idea that successful interventions rely on distributed leadership and shared responsibilities.

Prompt 3: will reorganisation work?

New leaders who are intoxicated with the prospect of change further fuel perpetual cycles of reorganization.

Andrew Oxman and colleagues²⁹⁶

Perhaps the most characteristic and most ambitious intervention in the demand management repertoire is the introduction of RMCs. These centralised triaging services or gateways often experienced a troubled introduction, and research shows that the ones that have stayed the course have undergone considerable adaptation from the prototypes.

The core idea behind the introduction of these centralised services was to rationalise the hitherto highly variable GP practices in making referrals. Having a professional service to review referral letters would prompt local GPs to think more carefully about whether or not the referral was necessary, ensure that they had undertaken all the relevant clinical tests prior to referral and consider whether or not the content of their referral letters contained all the necessary information. Without such systematic feedback, it was supposed that GPs would gain little insight into why referrals were redirected or deemed inappropriate and would not change their referral routines.

This menu of aspirations depended crucially on clinical engagement. In order to work successfully, the remit and operation of the RMC had to meet clinical as well as managerial needs. Research shows that this core requirement proved elusive. Several RMCs fell by the wayside. In the ones that prospered, sustainability was fostered by a prolonged and collaborative series of readjustments. In practical terms, clinical engagement was secured by thinking through the following:

- balancing financial ownership of RMCs by PCTs and clinicians, for example by setting up a CIC
- negotiating the precise percentages and type of referrals which should be directed to the RMC
- collaborating in the authorship of referral and triage guidelines and in the redesign of care pathways
- identifying those clinical leads who were able to secure the buy-in of local consultants *and* GPs
- cultivating and publicising the RMC via face-to-face meetings with local GPs
- negotiating clear agreements about the time scale under which triage would be undertaken
- installing robust tracking systems to ensure that referrals were not lost in the system.

The success of the intervention, in short, rests on a whole series of mini-interventions. Neither is it possible to treat the subsidiary interventions in a tick-box manner. Each one needs to be thought through. Research revealed another set of requirements to deal, for instance, with immediate capacity issues. For example:

- recruiting sufficient, locally respected GPs who were willing to take on the role of RMC triagers
- providing local GPs with direct access to some forms of clinical testing to ensure that all the relevant checks had been undertaken prior to referral
- extending local expertise and auxiliary services – such as GPwSIs or Allied Health Professionals – to cover a more sophisticated set of referral routes.



Organisational change has relentless capacity to end as reorganisation unless a cascade of consequences is thought through. If alternative management systems are provided without any positive attempts engage physicians in the making of the system, and without providing good reasons to change, and without evidence of demonstrable gains, the new layer of service risks creating supply-induced demand and increasing overall referral volumes.

Prompt 4: can intermediaries do the trick?

The go-between wears out a thousand sandals.

Japanese proverb²⁹⁷

Intermediate professional roles have multiplied throughout all health services. The impetus for such change lies in the bright idea of adjusting the workload between senior and junior DMs. Historic work routines often involve a rigid division of labour in which the 'superior' carries a number of tasks which do not really require specialist skills and expertise but are not passed on to an unqualified 'subordinate'. A newly trained and accredited intermediary, the idea goes, can take over a proportion of the decision-making with benefits accruing on workload, value for money and consistency of judgement. One of the major attempts to implement this policy was the introduction of the role of GPwSIs who, in blending the expertise of the GP and the consultant, would improve the management of demand in the system by siphoning off and dealing with an intermediate case mix appropriate to their experience and proficiency.

Research shows that the realisation of this bright idea has met with mixed success, with some GPwSIs wandering awkwardly between pillar and post. Sometimes the new role descended into administrative support if consultants remained in overall control. Sometimes GPs retained referral habits and bypassed or misused the intermediary. Once again, the key to effectiveness is clinical engagement and the development of co-operative working, thus avoiding major territorial claims and power struggles. This, of course, is easier said than done. Research reveals yet again the slow, adaptive, sandal-destroying processes involved as partners have thought through barriers to integrated working. There is no one formula for perfecting the division of labour; case histories reveal many different routes, four of which are summarised here:

- Working and reworking the 'case mix'. The precise area of expertise that comes with the 'special interest' has to be carefully negotiated. One method is for GPwSIs to graduate their responsibilities, starting with tightly defined ailments and symptoms and adding others incrementally.
- General practitioner with special interest posts were funded from quite different NHS posts. Recruitment 'from within' had the advantage of building on existing networks and of prior understanding of existing custom and practice (but also of its weaknesses).
- The new role depends on maintaining some equidistance from consultant and GPs, while at the same time being recognised and valued by the patient. Choosing the right location for GPwSI operations has proved significant.
- The GPwSIs' clinical obligations become clarified via their training and management duties. Having a parallel educative function allows for the coproduction of protocols and guidance on how GPs can improve their referrals.



The introduction of intermediate roles depends for its success on thinking through every detail of the 'job description'. GPwSIs were introduced enthusiastically on the back of a dozen vague ambitions and often ended up working in isolation, with insufficient clinical back-up. To meet a specific challenge like referral management their roles and responsibilities need to be negotiated, torn up and renegotiated to accommodate changing local contingencies and sensibilities.

Prompt 5: will direct access to tests results reduce excess referrals?

Don't cut out the middleman – become one.

*Entrepreneur Magazine*²⁹⁸

The aim of providing GPs with direct access to diagnostic tests that were previously under the remit of consultants is to enable them to distinguish between patients who require management by a secondary care specialist and those who can be managed in primary care. It is a time-honoured idea that has met with mixed success in reducing referrals. Alas, not all tests are created equally and research shows that the precise function and timing of the test is crucial in determining whether or not GP direct access makes a difference.

Researchers have devised many classifications to differentiate the purposes of diagnostic testing. A simple distinction may be made between (1) rule-out tests that screen for serious disorders in the expectation that most patients receive a negative result, (2) differential diagnosis tests to distinguish between several potential conditions in order to initiate the appropriate management regime and (3) clinical indicator tests to measure the progress of known disease in order to refine its treatment or management. Research has shown, as a broad rule of thumb, that direct access to 'differential diagnostic' tests had limited impact on demand for specialist care, while direct access to 'rule-out' and 'clinical indicator' tests had greater capacity to reduce such demand. The immediate practical implication of this finding is that before negotiating direct access to a test it is important to think through carefully its precise function. Although useful, this is too crude a rule of thumb to operate unthinkingly, for there are winners and losers within each type.

Further considerations and expectations need to be managed, such as the following unanticipated consequences. Regardless of the type of tests, research revealed that a small proportion of patients end up with a secondary care consultation even if the results indicate its redundancy. Even in seemingly decisive rule-out tests, some patients pressed for, and some GPs acceded to, requests for further 'expert' reassurance. A further unintended outcome is supply-induced demand: widening access to tests may prompt unnecessary, 'shotgun' testing.

As with all of the other demand management schemes, a range of supporting practices enhances the value of direct access. Especially when GPs may lack the skills and experience to reach diagnostic conclusions (as in the differential diagnostic tests) and, more especially, when there is disagreement about the value and role of a test at a particular point in the patient pathway, managers should ensure such schemes are also accompanied by:

- clear referral guidelines to advise GPs which patients should be referred for the test
- guidance on how GPs should interpret the test results
- guidance on how patients should be subsequently managed, including patients who may not be reassured by the test results.



Test results, however clinically and statistically precise, do not speak for themselves. It is important to think through how they will be interpreted under direct-access schemes, given the multitude of further pressures on the GP. Supplementary, actionable, bipartisan guidance need to be provided on who to test, what the test does and what follows from the results.

Prompt 6: will guidelines be followed?

We have all a better guide in ourselves, if we would attend to it, than any other person can be.

Jane Austin²⁹⁹

There are clinical guidelines for every medical condition and for every point in the treatment pathway and, thus, no shortage of formal, professionally endorsed advice on when and how to make referrals. Research on the impact of guidelines on practice comes to a crushing and consensual conclusion, namely that the 'passive' dissemination of guidelines has failed as a tool for reducing demand. However thoroughly researched and immaculately presented, guidelines will be left on the shelf unless there are further 'active' steps to ensure that the advice is followed.

Research has also made useful strides in determining what does and what does not contribute to active utilisation. Guideline uptake is not enhanced if the documentation is enlarged, as it very often is these days, to offer suggestions on overcoming the potential 'barriers' to guideline implementation. Paper advice remains paper advice, however thorough and exhaustive. Also ineffective is the much-used process of 'adapting' guidelines – by having local expert panels rewrite the advice of national expert panels to local conditions. Paper advice remains paper advice.

Effective guidelines obey Austin's maxim and depend on practitioners locating useful rules within their own practice. Broadly speaking, guidelines are followed if they are constructed by the same personnel responsible for implementing them. The micro-logistics required to enact a guideline need to be carefully specified and mapped step by step into local organisational structures. Managers need to ask who precisely is responsible for which referral decisions, who is responsible for responding to that decision, and so on. The guidelines then become tailored to the needs and responsibilities of each post holder in the decision-making tree. Even then, the guidance is not issued by decree. Research shows that the real implications of each element of a guideline becomes fully intelligible and actionable only if they are negotiated with each post holder.



Think through the process of getting paper advice into practical decisions. Guidelines, by their very nature, are couched as generalities on the extent of the medical condition, on test sensitivity and so on. Practice responds to the workaday specifics of patient characteristics, available resources and so on. Guidance becomes effective only when translated from the page into an agreed action agenda across each member of the team responsible for a service.

Prompt 7: can productive change be accelerated?

All the real work is done in the rehearsal phase.

Donald Pleasence³⁰⁰

Research has shown, as recounted repeatedly above, that demand management schemes become effective in the longer term only after a period of reconsideration and adjustment. Initial plans often require auxiliary measures to bring them to fruition. A change at one point in a complex system always has implications at dozens of other junctures and these are hard to predict. These findings represent more than a plea for patience. Effective change management is not just a case of waiting for interventions to 'bed down'. Some programmes will have fatal unintended consequences; others will simply be overwhelmed by adverse conditions elsewhere in the system. Is it possible to distinguish those ideas that have staying power?

There is hope. It resides, inevitably, within the theme of 'thinking through'. One can think retrospectively (if only?) or concurrently (let's try?) or prospectively (what if?). The final mode, under a notion sometimes called group model building, may help to identify and build interventions with stamina. The idea, simply said, is to put policy proposals through a series of rehearsals. Any demand management scheme is dependent for its success on the reactions of a whole range of different professional groups. Under group model building, the plan is to bring together representatives to unpick and challenge the underlying ideas behind a new proposal. The group, by turns, questions whether or not the assumed diagnosis of the problem is valid, it evaluates alternative solutions, it maps the roles and responsibilities of the various stakeholders in relation to emerging proposals and it trouble-shoots the surviving ideas to root out potential antagonisms and power plays. All ends, if things go well, in an adapted, bottom-up version of the original plans.

Things may not go well, of course. Group model building, like any other intervention, will only work in the right circumstances and will fail with the inadequate selection of stakeholders, with domineering panellists and if *it* is managed without foresight.



Think ahead in the planning of any demand management scheme by simulating its system-wide implications. This involves considerably more than putting out policy proposal to 'consultation'. It involves a systematic series of 'rehearsals' – repeated, representative forums attempting to iron out the deep creases in the initial plans. Group model building can never imitate the rough and tumble of real interdisciplinary working but it can initiate and inspire an adaptive process that carries onwards into actual implementation.

Acknowledgements

Contributions of authors

Ray Pawson, Joanne Greenhalgh and **Cathy Brennan** conducted the searches, selected and reviewed the papers, conducted the synthesis and wrote the report.

Other contributions from the project team

Judy Wright advised on search strategies. Roberta Longo advised on economic theories underlying demand and capacity problems explored in the review. Laurence Wood chaired the patient group who reviewed and commented on our theories and synthesis. He also advised on stakeholder involvement and dissemination strategies. Lisa Kundi and Hugh Sturgess advised on our theories and reviewed and commented on our synthesis; they also advised on dissemination strategies. Liz Glidewell helped to conduct initial searches and contributed to identification of initial theories. Allan Best reviewed and commented on our report and contributed to discussion on systems thinking throughout the report and on group model building in the conclusions chapter.

Project team

Professor Ray Pawson, Professor of Social Research Methodology, School of Sociology and Social Policy, University of Leeds, Leeds, UK.

Dr Joanne Greenhalgh, Principal Research Fellow in Healthcare Evaluation, School of Sociology and Social Policy, University of Leeds, Leeds, UK.

Dr Cathy Brennan, Lecturer in Public Health, Leeds Institute of Health Sciences, University of Leeds, Leeds, UK.

Dr Roberta Longo, Lecturer in Health Economics, Leeds Institute of Health Sciences, University of Leeds, Leeds, UK.

Judy Wright, Senior Information Specialist, Leeds Institute of Health Sciences, University of Leeds, Leeds, UK.

Dr Elizabeth Glidewell, Lecturer in Primary Care, Leeds Institute of Health Sciences, University of Leeds, Leeds, UK.

Lisa Kundi, Locality Manager, NHS Leeds West Clinical Commissioning Group, Leeds, UK.

Laurence Wood, patient and public representative.

Dr Hugh Sturgess, Director, Pennine MSK Partnership Ltd, Oldham, UK.

Dr Allan Best, Managing Director, InSource, Centre for Clinical Epidemiology and Evaluation, University of British Columbia, Vancouver, BC, Canada.

We would also like to acknowledge contributions from Roshanak Mehdipanah, visiting PhD student from the Department of Experimental and Health Sciences, Universitat Pompeu Fabra, Barcelona, Spain, and Lucy Bailey, undergraduate research scholar, School of Sociology and Social Policy, University of Leeds, Leeds, UK.

Patient group

Mike Edwards.

Elieen Exeter.

Rosemary Hassaman.

Gill Riley.

Advisory group

Professor Jenny Hewison, Professor of the Psychology of Health Care, Leeds Institute of Health Sciences, University of Leeds, Leeds, UK.

Dr Geoff Wong, Clinical Senior Lecturer in Primary Healthcare, Queen Mary University London, London, UK.

Gordon Tollefson, patient and public representative.

Leaf Mobbs, (then) Head of Commissioning, NHS Leeds West Clinical Commissioning Group, Leeds, UK.

Publication

Pawson R, Greenhalgh J, Brennan C, Gildewell E. Do reviews of healthcare interventions teach us how to improve healthcare systems? *Soc Sci Med* 2014;**114**:129–37.

Data sharing statement

All of the data is in the original documents on which this report is based.

References

1. Suri H, Clarke D. Advancements in research synthesis methods: from a methodologically inclusive perspective. *Rev Educ Res* 2009;**79**:395–430. <http://dx.doi.org/10.3102/0034654308326349>
2. Hilsdon M, Foster C, Cavill N, Crombie H, Naidoo B. *The Effectiveness of Public Health Interventions for Increasing Physical Activity Among Adults: A Review of Reviews*. 2nd edn. Health Development Agency; 2005.
3. Pawson R. *Evidence Based Policy: A Realist Perspective*. London: Sage; 2006.
4. Wong G, Greenhalgh T, Westhorp G, Pawson R. Development of methodological guidance, publication standards and training materials for realist and meta-narrative reviews: the RAMESES (Realist And Meta-narrative Evidence Syntheses – Evolving Standards) project. *Health Serv Deliv Res* 2014;**2**(30).
5. Wong G, Greenhalgh T, Westhorp G, Buckingham J, Pawson R. RAMESES publication standards: realist synthesis. *BMC Med* 2013;**11**:21. <http://dx.doi.org/10.1186/1741-7015-11-21>
6. Greenhalgh T, Robert G, Macfarlane F, Bate P, Kyriakidou O. Diffusion of innovations in service organizations: systematic review and recommendations. *Milbank Q* 2004;**82**:581–629. <http://dx.doi.org/10.1111/j.0887-378X.2004.00325.x>
7. Rittel H, Webber M. Dilemmas in a general theory of planning. *Policy Sci* 1973;**4**:155–69. <http://dx.doi.org/10.1007/BF01405730>
8. Byrne D, Callaghan G. *Complexity Theory and the Social Sciences: The State of the Art*. London: Routledge; 2013.
9. Centre for Reviews and Dissemination. *CRD's Guidance for Undertaking Reviews in Health Care 2009*. URL: www.york.ac.uk/inst/crd/pdf/Systematic_Reviews.pdf (accessed 13 August 2015).
10. Craig P, Dieppe P, Macintyre S, Mitchie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ* 2008;**337**:a1655. <http://dx.doi.org/10.1136/bmj.a1655>
11. Pawson R. *The Science of Evaluation: A Realist Manifesto*. London: Sage; 2013.
12. Kannampallil T, Schauer G, Cohen T, Patel V. Considering complexity in healthcare systems. *J Biomed Inform* 2011;**44**:943–7. <http://dx.doi.org/10.1016/j.jbi.2011.06.006>
13. Willis C, Best A, Riley B, Herbert A, Miller J, Howland A. Systems thinking for transformational healthcare. *Evid Policy* 2014;**10**:113–26. <http://dx.doi.org/10.1332/174426413X662815>
14. Department of Health. *The NHS Plan: A Plan for Investment, A Plan for Reform*. Department of Health; 2000.
15. Sun Tzu, Griffiths SB. *The Art of War*. Oxford: Clarendon Press; 1964.
16. Burton R. *The Anatomy of Melancholy*. London: Dent; 1932.
17. Pandit JJ, Pandit M, Reynard JM. Understanding waiting lists as the matching of surgical capacity to demand: are we wasting enough surgical time? *Anaesthesia* 2010;**65**:625–40. <http://dx.doi.org/10.1111/j.1365-2044.2010.06278.x>
18. Bain CA, Taylor PG, McDonnell G, Georgiou A. Myths of ideal hospital occupancy. *Med J Aust* 2010;**193**:311.
19. Erlang AK. The theory of probabilities and telephone conversations. *Nyt Tidsskrift for Matematik B* 1909;**20**:33–9.

20. NHS Institution for Innovation and Improvement. *Improving Patient Flow in the NHS Case Studies on Reducing Delays*. URL: www.carebydesign.org/files/no_delays_achiever_case_studies.pdf (accessed 13 August 2015).
21. Johnson T. *Professions and Power*. London: Heinemann; 1972.
22. Turner S. *Medical Power and Social Knowledge*. London: Sage; 1987.
23. Timmermans S, Oh H. The continued social transformation of the medical profession. *J Health Soc Behav* 2010;**51**:S94–106. <http://dx.doi.org/10.1177/0022146510383500>
24. Numerto D, Salvertori D, Fattore G. The impact of management on medical professionalism: a review. *Sociol Health Illn* 2011;**34**:626–44. <http://dx.doi.org/10.1111/j.1467-9566.2011.01393.x>
25. Foote JL, Houston DJ, North NH. Betwixt and between: ritual and the management of an ultrasound waiting list. *Health Care Anal* 2002;**10**:357–77. <http://dx.doi.org/10.1023/A:1023427210009>
26. Foote JL, Houston DJ, North NH. Towards a systemic understanding of a hospital waiting list. *J Health Organ Manag* 2004;**18**:140–54. <http://dx.doi.org/10.1108/14777260410548400>
27. Abel GM, Thompson LE. General Practitioners, specialists and surveillance guidelines: interpreting the socio-clinical context of decision-making. *Health Risk Soc* 2011;**13**:547–59. <http://dx.doi.org/10.1080/13698575.2011.615826>
28. Currie G, Finn R, Martin G. Professional competition and modernizing the clinical workforce in the NHS. *Work Employment Soc* 2009;**23**:267–84. <http://dx.doi.org/10.1177/0950017009102858>
29. Cummins RO, Jarman B, White PM. Do general practitioners have different 'referral thresholds'? *Br Med J* 1981;**282**:1037–9. <http://dx.doi.org/10.1136/bmj.282.6269.1037>
30. King N, Bailey J, Newton P. Analysing general practitioners' referral decisions. I. Developing an analytical framework. *Fam Pract Manag* 1994;**11**:3–8. <http://dx.doi.org/10.1093/fampra/11.1.3>
31. Morgan M, Jenkins L, Ridsdale L. Patient pressure for referral for headache: a qualitative study of GPs' referral behaviour. *Br J Gen Pract* 2007;**57**:29–35.
32. O'Donnell Ca. Variation in GP referral rates: what can we learn from the literature? *Fam Pract* 2000;**17**:462–71. <http://dx.doi.org/10.1093/fampra/17.6.462>
33. Sigel P, Leiper R. GP views of their management and referral of psychological problems: a qualitative study. *Psychol Psychother* 2004;**77**:279–95. <http://dx.doi.org/10.1348/1476083041839394>
34. Morgan D. Psychiatric cases: an ethnography of the referral process. *Psychol Med* 1989;**19**:743–53. <http://dx.doi.org/10.1017/S003329170002434X>
35. Nandy S, Chalmers-Watson C, Gantley M, Underwood M. Referral for minor mental illness: a qualitative study. *Br J Gen Pract* 2001;**51**:461–5.
36. Mold J, Stein H. The cascade effect in the clinical care of patients. *N Engl J Med* 1986;**314**:512–14. <http://dx.doi.org/10.1056/NEJM198602203140809>
37. Harris JC. *Uncle Remus and Br'er Rabbit*. New York, NY: Frederick A Stokes Company; 1907.
38. Deyo Ra. Cascade effects of medical technology. *Ann Rev Public Health* 2002;**23**:23–44. <http://dx.doi.org/10.1146/annurev.publhealth.23.092101.134534>
39. Auditor General for Scotland. *Key Messages: Review of NHS Diagnostic Services, Prepared For The Auditor General For Scotland*. Edinburgh: Audit Scotland; 2008.

40. National Audit Office. *Managing High Value Capital Equipment in the NHS in England*. London: The Stationery Office; 2011. URL: www.nao.org.uk/wp-content/uploads/2011/03/1011822.pdf (accessed 13 August 2015).
41. Lundberg G. Diminishing returns on the road to diagnostic certainty. *JAMA* 1991;**265**:2229–31. <http://dx.doi.org/10.1001/jama.1991.03460170083038>
42. Venalovich V. How much routine preoperative laboratory testing is enough? *Am Coll Med Res Q* 1993;**8**:144–51.
43. Newman T, Browner A, Hulley S. The case against childhood cholesterol screening. *JAMA* 1990;**264**:3039–43. <http://dx.doi.org/10.1001/jama.1990.03450230075032>
44. Torgerson D, Spencer A. Marginal costs and benefits. *BMJ* 1996;**312**:35. <http://dx.doi.org/10.1136/bmj.312.7022.35>
45. Ganott MA, Sumkin JH, King JL, Klym AH, Catullo VJ, Cohen CS, et al. Screening mammography: do women prefer a higher recall rate given the possibility of earlier detection of cancer? *Radiology* 2006;**238**:793–800. <http://dx.doi.org/10.1148/radiol.2383050852>
46. Williams A. Priority setting in public and private health care. *J Health Econ* 1988;**7**:173–83. [http://dx.doi.org/10.1016/0167-6296\(88\)90015-X](http://dx.doi.org/10.1016/0167-6296(88)90015-X)
47. Hickson GB, Altemeier WA, Perrin JM. Physician reimbursement by salary or fee-for-service: effect on physician practice behavior in a randomized prospective study. *Pediatrics* 1987;**80**:344–50.
48. Robertson IK, Richardson JR. Coronary angiography and coronary artery revascularisation rates in public and private hospital patients after acute myocardial infarction. *Med J Aust* 2000;**173**:291–5.
49. Richardson J. Supply and demand for medical care: or, is the health care market perverse? *Aust Econ Rev* 2001;**34**:336–52. <http://dx.doi.org/10.1111/1467-8462.00203>
50. Glover A. The incidence of tonsillectomy in school children. *Int J Epidemiol* 1938;**37**:9–19. <http://dx.doi.org/10.1093/ije/dym258>
51. Glover JA. Monthly Bulletin of the Ministry of Health 9: 62. In Morris J. *Uses of Epidemiology*. Baltimore, MD: Williams & Wilkins; 1964.
52. Mulley A. The inconvenient truths about supplier induced demand and unwarranted variation in medical practice. *BMJ* 2009;**339**:b4073. <http://dx.doi.org/10.1136/bmj.b4073>
53. Cain M, Mittman R. *Diffusion of Innovation in Health Care*. Oakland, CA: California Healthcare Foundation; 2002.
54. Rogers E. *Diffusion of Innovations*. New York, NY: New York Free Press; 1995.
55. NHS London Health Programmes. *Polyclinics Learning and Development*. 2008. URL: www.londonhp.nhs.uk/wp-content/uploads/2011/03/Polyclinics-Learning-and-Development-Workshop-2.pdf (accessed 9 July 2015).
56. Gruber J, Owings M. Physician financial incentives and cesarean section delivery. *Rand J Econ* 1996;**27**:99–123. <http://dx.doi.org/10.2307/2555794>
57. Hughes D, Yule B. The effect of per item fees on the behaviour of general practitioners. *J Health Econ* 1992;**11**:413–37. [http://dx.doi.org/10.1016/0167-6296\(92\)90014-R](http://dx.doi.org/10.1016/0167-6296(92)90014-R)
58. Sirovich B, Gallagher PM, Wennberg DE, Fisher ES. Discretionary decision making by primary care physicians and the cost of U.S. health care. *Health Affairs* 2008;**27**:813–23. <http://dx.doi.org/10.1377/hlthaff.27.3.813>

59. Wennberg JE, Barnes BA, Zubkoff M. Professional uncertainty and the problem of supplier induced demand. *Soc Sci Med* 1982;**16**:811–24. [http://dx.doi.org/10.1016/0277-9536\(82\)90234-9](http://dx.doi.org/10.1016/0277-9536(82)90234-9)
60. Dill MJ, Salsberg ES. *The Complexities of Physician Supply and Demand: Projections Through 2025*. Washington, DC: Association of American Medical Colleges; 2008.
61. Abel-Smith B, Titmuss R. *The Cost of the National Health Service in England and Wales*. Cambridge: Cambridge University Press; 1956.
62. Marzouk M. Aging, age specific healthcare costs and the future care burden in Canada. *Can Pub Policy* 1991;**17**:490–506. <http://dx.doi.org/10.2307/3551709>
63. Strunk BC, Ginsburg PB, Banker MI. The effect of population aging on future hospital demand. *Health Affairs* 2006;**25**:w141–9. <http://dx.doi.org/10.1377/hlthaff.25.w141>
64. Evans RG. Illusions of necessity: evading responsibility for choice in health care. *J Health Politics Policy Law* 1985;**10**:439–67. <http://dx.doi.org/10.1215/03616878-10-3-439>
65. Gray A. Population ageing and health care expenditure. *Ageing Horizons* 2005;**2**:15–20.
66. Zweifel P, Felder S, Meiers M. Ageing of population and health care expenditure: a red herring? *Health Econ* 1999;**8**:485–96. [http://dx.doi.org/10.1002/\(SICI\)1099-1050\(199909\)8:6<485::AID-HEC461>3.0.CO;2-4](http://dx.doi.org/10.1002/(SICI)1099-1050(199909)8:6<485::AID-HEC461>3.0.CO;2-4)
67. Lubitz J, Prihoda R. The use and costs of Medicare services in the last 2 years of life. *Health Care Financ Rev* 1984;**5**:117–31.
68. Stooker T, van Acht JW, van Barneveld EM, van Vliet RC, van Hout BA, Hessing DJ, et al. Costs in the last year of life in The Netherlands. *Inquiry* 2001;**38**:73–80. http://dx.doi.org/10.5034/inquiryjnl_38.1.73
69. Levinsky NG, Yu W, Ash a, Moskowitz M, Gazelle G, Saynina O, et al. Influence of age on Medicare expenditures and medical care in the last year of life. *JAMA* 2001;**286**:1349–55. <http://dx.doi.org/10.1001/jama.286.11.1349>
70. Perls T, Wood E. Acute care of the oldest olds. *Arch Intern Med* 1996;**156**:754–60. <http://dx.doi.org/10.1001/archinte.1996.00440070078009>
71. Werblow A, Felder S, Zweifel P. Population ageing and health care expenditure: a school of 'red herrings'? *Health Econ* 2007;**16**:1109–26. <http://dx.doi.org/10.1002/hec.1213>
72. Caughey GE, Vitry AI, Gilbert AL, Roughead EE. Prevalence of comorbidity of chronic diseases in Australia. *BMC Public Health* 2008;**8**:221. <http://dx.doi.org/10.1186/1471-2458-8-221>
73. Reinhardt UE. Does the aging of the population really drive the demand for health care? *Health Aff* 2003;**22**:27–39. <http://dx.doi.org/10.1377/hlthaff.22.6.27>
74. Rice RE. Influences, usage, and outcomes of Internet health information searching: multivariate results from the Pew surveys. *Int J Med Inf* 2006;**75**:8–28. <http://dx.doi.org/10.1016/j.ijmedinf.2005.07.032>
75. Miller EA, West DM. Where's the revolution? Digital technology and health care in the internet age. *J Health Politics Policy Law* 2009;**34**:261–84. <http://dx.doi.org/10.1215/03616878-2008-046>
76. Atkinson NL, Saperstein SL, Pleis J. Using the internet for health-related activities: findings from a national probability sample. *J Med Internet Res* 2009;**11**:e4. <http://dx.doi.org/10.2196/jmir.1035>
77. Suziedelyte A. How does searching for health information on the Internet affect individuals' demand for health care services? *Soc Sci Med* 2012;**75**:1828–35. <http://dx.doi.org/10.1016/j.socscimed.2012.07.022>

78. Murray J, Majeed A, Khan MS, Lee JT, Nelson P. Use of the NHS Choices website for primary care consultations: results from online and general practice surveys. *JRSM Short Reports* 2011;**2**:56. <http://dx.doi.org/10.1258/shorts.2011.011078>
79. Stevenson FA, Kerr C, Murray E, Nazareth I. Information from the Internet and the doctor–patient relationship: the patient perspective – a qualitative study. *BMC Fam Pract* 2007;**8**:47. <http://dx.doi.org/10.1186/1471-2296-8-47>
80. Carrns A. On the Internet, Diseases are Rampant, Playing to Worries of Hypochondriacs. *The Wall Street Journal*, 5 October 1999. URL: <http://online.wsj.com/article/SB939076866193196830.html> (accessed 9 July 2015).
81. Shih YC, Tai-Seale M. Physicians' perception of demand-induced supply in the information age: a latent class model analysis. *Health Econ* 2012;**21**:252–69. <http://dx.doi.org/10.1002/hec.1710>
82. Sommerhalder K, Abraham A, Zufferey MC, Barth J, Abel T. Internet information and medical consultations: experiences from patients' and physicians' perspectives. *Patient Educ Couns* 2009;**77**:266–71. <http://dx.doi.org/10.1016/j.pec.2009.03.028>
83. Fang H, Miller NH, Rizzo J, Zeckhauser R. Demanding customers: consumerist patients and quality of care. *BE J Econ Anal Policy* 2011;**11**. <http://dx.doi.org/10.2202/1935-1682.2966>
84. Inter-university Consortium for Political and Social Research. *Community Tracking Study Physician Survey, 2000–2001*. URL: www.icpsr.umich.edu/icpsrweb/HMCA/studies/3820 (accessed 30 July 2013).
85. Victoor A, Delnoij DM, Friele RD, Rademakers JJ. Determinants of patient choice of healthcare providers: a scoping review. *BMC Health Serv Res* 2012;**12**:272. <http://dx.doi.org/10.1186/1472-6963-12-272>
86. Dixon A, Robertson R, Appleby J, Devlin N, Burge P, Magee H. *Patient Choice: How Patients Choose and How Providers Respond*. The King's Fund; 2010.
87. Peckham S, Sanderson M, Entwistle V, Thompson A, et al. *A Comparative Study of the Construction and Implementation of Patient Choice Policies in the UK*. Final report. National Institute for Health Research (NIHR) Service Delivery and Organisation programme. Southampton; NIHR: 2011.
88. Ford S, Schofield T, Hope T. What are the ingredients for a successful evidence-based patient choice consultation? A qualitative study. *Soc Sci Med* 2003;**56**:589–602. [http://dx.doi.org/10.1016/S0277-9536\(02\)00056-4](http://dx.doi.org/10.1016/S0277-9536(02)00056-4)
89. Elwyn G, Edwards A, Kinnersley P, Grol R. Shared decision making and the concept of equipoise: the competences of involving patients in healthcare choices. *Br J Gen Pract* 2000;**50**:892–9.
90. NHS England. *National Programmes of Care and Clinical Reference Groups*. URL: www.england.nhs.uk/npc-crg (accessed 12 July 2013).
91. Griffiths F, Green E, Tsouroufli M. The nature of medical evidence and its inherent uncertainty for the clinical consultation: qualitative study. *BMJ* 2005;**330**:511. <http://dx.doi.org/10.1136/bmj.38336.482720.8F>
92. Pawson R, Greenhalgh J, Brennan C, Glidewell E. Do reviews of healthcare interventions teach us how to improve healthcare systems? *Soc Sci Med* 2014;**114**:129–37. <http://dx.doi.org/10.1016/j.socscimed.2014.05.032>
93. Imison C, Naylor C. *Referral Management: Lessons For Success*. London: The King's Fund, 2010.
94. Coulter A. Managing demand at the interface between primary and secondary care. *BMJ* 1998;**316**:1974. <http://dx.doi.org/10.1136/bmj.316.7149.1974>

95. Roland M, McDonald R, Sibbald B. *Outpatient Services and Primary Care: A Scoping Review of Research into Strategies for Improving Outpatient Effectiveness and Efficiency*. National Institute for Health Research (NIHR) Service Delivery and Organisation Programme. Southampton; NIHR: 2006.
96. Pencheon D. Managing demand: matching demand and supply fairly and efficiently. *BMJ* 1998;**316**:1665–7. <http://dx.doi.org/10.1136/bmj.316.7145.1665>
97. Willmott R. The place of culture in organization theory: introducing the morphogenetic approach. *Organization* 2000;**7**:95–128. <http://dx.doi.org/10.1177/135050840071006>
98. Zaltman G, Duncan R, Holbeck J. *Innovations and Organisations*. New York, NY: Wiley; 1973.
99. Jones D, Mitchell A. *Lean Thinking for the NHS*. London: Lean Enterprise Academy UK, NHS Confederation; 2006.
100. NHS Scotland. *Patient Pathway Management: Referral Facilitation*. Edinburgh: The Scottish Government; 2007.
101. NHS Modernisation Agency. *Big Wizard*. 2002. URL: www.natpact.info/demand_management/wizards/big_wizard/index.php (accessed 30 July 2013).
102. NHS Modernisation Agency. *Little Wizard*. 2002. URL: www.natpact.info/uploads/PDF%20Little%20Wizard.pdf (accessed 30 July 2013).
103. NHS Institute for Innovation and Improvement. *Improving Patient Flows in the NHS: Case Studies in Reducing Delay*. URL: www.carebydesign.org/files/no_delays_achiever_case_studies.pdf (accessed 30 July 2013).
104. NHS Institute for Innovation and Improvement. *Reducing Delays in Patient Care: Seven Ways to No Delays*. URL: www.institute.nhs.uk/quality_and_service_improvement_tools/quality_and_service_improvement_tools/reducing_delays_in_patient_care.html (accessed 30 July 2013).
105. NHS Institute for Innovation and Improvement. *Productive General Practice: Releasing Time*. URL: www.institute.nhs.uk/productive_general_practice/general/productive_general_practice_homepage.html (accessed 30 July 2013).
106. NHS Scotland. *Patient Pathway Management: Referral Facilitation*. URL: www.scotland.gov.uk/Publications/2007/03/19092814/12 (accessed 30 July 2013).
107. Health Service Executive Ireland. *Report and Recommendation on Patient Referrals from General Practice to Outpatient and Radiology Services*. URL: www.cappagh.ie/sites/default/files/pdf/gp_referral_report.pdf (accessed 30 July 2013).
108. Hage J, Aiken M. *Social Change in Complex Organizations*. New York, NY: Random House; 1970.
109. NHS Modernisation Agency. *Improvement Leader's Guide to Involving Patients and Carers*. London: Ancient House Printing Group; 2002.
110. NHS Scotland. *Patient Flow in Planned Care: Admission, Discharge, Length of Stay and Follow-Up*. Edinburgh: The Scottish Government; 2007.
111. Department of Health. *Case Study: Fortis Healthcare*. London: Department of Health; 2011. URL: <http://healthandcare.dh.gov.uk/case-study-fortis/> (accessed 30 July 2013).
112. Somerset LMC. *LMC Position Paper on Referral Management Centres*. Somerset LMC; 2004.
113. Fashanu B. Managing the referral system is a good thing, despite the current shortfalls. *BMJ* 2006;**332**:844.
114. Nowotny S. Gatekeeper Job is One for Trained GP. *Pulse*, 10 August 2011.

115. Wootton S, Whiting M. Creating Ownership of a Referral Gateway. *Pulse*, 3 September 2012.
116. Williams S, Ryan D, Price D, Langley C, Fletcher M, Everden P. General practitioners with a special clinical interest: a model for improving respiratory disease management. *Br J Gen Pract* 2002;**52**:838–43.
117. Gerada C, Wright N, Keen J. The general practitioner with a special interest: new opportunities or the end of the generalist practitioner? *Br J Gen Pract* 2002;**52**:796–8.
118. Nocon A, Leese B. The role of UK general practitioners with special clinical interests: implications for policy and service delivery. *Br J Gen Pract* 2004;**54**:50–6.
119. Honey S, Small N, Studies H, Ali S. General practitioners with special interests: the potential benefits and possible risks for primary care. *Prim Health Care Res Dev* 2005;**6**:5–16.
<http://dx.doi.org/10.1191/1463423605pc229oa>
120. Carlisle D. Do GPs with Special Interests Have a Future in the NHS? *Health Service Journal*, 7 April 2008.
121. Graham-Brown R. GPwSIs: a useful addition, at a cost, but not a substitute (rapid response). *BMJ* 2005;**331**:1441.
122. Bamji AN. Evaluating GPSIs. www.bmj.com/rapid-response/2011/10/30/evaluating-gpsis (accessed 31 July 2013).
123. Rosen R, Stevens R, Jones RB. General practitioners with special interests: a potentially valuable asset, which requires evaluation. *BMJ* 2003;**327**:460–2. <http://dx.doi.org/10.1136/bmj.327.7413.460>
124. Wilczynski P. Taking the Isolation out of Referral Decisions. *Pulse*, 3 September 2012.
125. Everington S. Case Study: Referral Balanced Scorecards. *Pulse*, 9 August 2011.
126. de Marco P, Dain C, Lockwood T, Roland M. How valuable is feedback of information on hospital referral patterns? *BMJ* 1993;**307**:1465–6. <http://dx.doi.org/10.1136/bmj.307.6917.1465>
127. Evans E, Aiking H, Edwards A. Reducing variation in general practitioner referral rates through clinical engagement and peer review of referrals: a service improvement project. *Qual Prim Care* 2011;**19**:263–72.
128. Broughton R, Rathbone B. *What Makes A Good Clinical Guideline?* Hayward Medical Communications; 1999. URL: www.medicine.ox.ac.uk/bandolier/painres/download/whatis/whatareclinguide.pdf (accessed 13 August 2015).
129. Woolf SH, Grol R, Hutchinson A, Eccles M, Grimshaw J. Clinical guidelines: potential benefits, limitations, and harms of clinical guidelines. *BMJ* 1999;**318**:527–30. <http://dx.doi.org/10.1136/bmj.318.7182.527>
130. National Institute for Health and Care Excellence. *The Guideline Development Process: An Overview for Stakeholders, the Public and the NHS*. London: NICE; 2007.
131. Sherifi J. Protocols Reduce General Practice to a Maoist Collective. *Pulse*, 16 November 2011.
132. Carlsen B, Glenton C, Pope C. Thou shalt versus thou shalt not: a meta-synthesis of GPs' attitudes to clinical practice guidelines. *Br J Gen Pract* 2007;**57**:971–8. <http://dx.doi.org/10.3399/096016407782604820>
133. Rappolt SG. Clinical guidelines and the fate of medical autonomy in Ontario. *Soc Sci Med* 1997;**44**:977–87. [http://dx.doi.org/10.1016/S0277-9536\(96\)00223-7](http://dx.doi.org/10.1016/S0277-9536(96)00223-7)
134. McLellan A. National GP referral guidelines are needed to reduce inequalities. *Health Service Journal*, 29 April 2010.

135. Docherty MJ, Choudhury M, Littlejohns P. NICE 'referral advice' recommendations database. *BMJ* 2011;**342**:d129. <http://dx.doi.org/10.1136/bmj.d129>
136. Sibbald B. Direct access to diagnostic services. *Br J Gen Pract* 2009;**59**:562. <http://dx.doi.org/10.3399/bjgp09X420563>
137. Department of Health. *Direct Access Diagnostic Tests*. Department of Health; 2011. URL: <http://healthandcare.dh.gov.uk/direct-access-diagnostic-tests/> (accessed 13 August 2015).
138. Robinson S. Exclusive: PCTs Restrict Access to Diagnostic Scans. *GP Online*, 8 September 2011. URL: www.gponline.com/exclusive-pcts-restrict-access-diagnostic-scans/article/1089428 (accessed 13 August 2015).
139. Nowotny S. MRI Study Boosts Drive for Direct GP Access to Diagnostic Tests. *Pulse*, 10 November 2008.
140. Poole S. Five Questions You Should Ask before Offering Referral Incentives. *Pulse*, 4 October 2012.
141. British Medical Association. *Conflicts of Interest in the New Commissioning System: Doctors as Providers*. BMA; 2013.
142. British Medical Association. *GP Referral Incentive Schemes: GPC Guidance for GPs and LMCs*. London: BMA; 2009.
143. Roland M. linking physicians' pay to the quality of care – a major experiment in the United Kingdom. *N Engl J Med* 2004;**351**:1448–54. <http://dx.doi.org/10.1056/NEJMhpr041294>
144. Eaton L. BMA backs GPs in their objections to financial incentives to limit hospital referrals. *BMJ* 2008;**337**:a2306. <http://dx.doi.org/10.1136/bmj.a2306>
145. Archer M. *Realist Social Theory: The Morphogenetic Approach*. Cambridge: Cambridge University Press; 1995. <http://dx.doi.org/10.1017/CBO9780511557675>
146. Archer M. *Structure, Agency and the Internal Conversation*. Cambridge: Cambridge University Press; 2003. <http://dx.doi.org/10.1017/cbo9781139087315>
147. Drucker P, Hesselbein F, Snyder Kuhl J. *Peter Drucker's Five Most Important Questions: Enduring Wisdom for Today's Leaders*. Hoboken, NJ: John Wiley & Sons, Inc.; 2015.
148. Davies M, Elwyn G. Referral management centres: promising innovations or Trojan horses? *BMJ* 2006;**332**:844–6. <http://dx.doi.org/10.1136/bmj.332.7545.844>
149. Coleman A, Harrison S, Checkland K, Hiroeh U. *Practice Based Commissioning: Report of a Survey of Primary Care Trusts*. Manchester: National Primary Care Research and Development Centre; 2007.
150. Harrison S, Wood B. Scientific-bureaucratic medicine and U.K. health policy. *Policy Stud Rev* 2000;**17**:25–42. <http://dx.doi.org/10.1111/j.1541-1338.2000.tb00955.x>
151. Abbott A. *The System of Professions – A Study of the Division of Expert Labour*. London: University of Chicago Press; 1988.
152. Abbott A. Things of boundaries. *Sociol Res* 1995;**62**:857–82.
153. Hughes EC. *Men and their Work*. Toronto, ON: Free Press of Glencoe; 1958.
154. Pawson R. Digging for nuggets: how 'bad' research can yield 'good' evidence. *Int J Soc Sci Methodol* 2006;**9**:127–42. <http://dx.doi.org/10.1080/13645570600595314>
155. White R. Striving for 'Right First Time' Referrals. *Pulse*, 8 October 2012.

156. Offredy M, Cleary M, Bland A, Donovan A, Kelshiker A. Improving health and care for patients by redesigning services: the development and implementation of a clinical assessment service in Harrow Primary Care Trust. *Qual Prim Care* 2008;**16**:95–102.
157. Moss C, Evans M, Roan H. Implementing a referral management system. *Health Service Journal*, 8 October 2012.
158. CRG Research/Cardiff University. *Evaluation of Referral Management Pilots in Wales*. Cardiff: NLIAH; 2006.
159. Wootton S, Gomm N, Strobly J, Madhok R. *Individual Funding Requests (Non-commissioned Procedures) and Referral Gateway Initiative*. Paper L prepared for the NHS Manchester Board Meeting; NHS Manchester; 7 July 2010.
160. Whiting M, Wootton S. Introducing a GP-Led Referral Gateway. *Pulse*, 18 May 2011.
161. Anonymous. Gateways using Nurses to Screen Referrals. *Pulse*, 10 August 2011.
162. Anonymous. GPs Axe Remote Referral Gateway Run by Private Firm. *Pulse*, 2 March 2012.
163. Fink P, Whiting M, Eeckelaers M, Tamkin B. Referrals Gateway Worked Well. *Pulse*, 20 March 2012.
164. Harmoni. URL: www.harmoni.co.uk/nhs-manchester-referral-management-centre-rmc (accessed 6 January 2013).
165. Cox JM, Steel N, Clark AB, Kumaravel B, Bachmann MO. Do referral-management schemes reduce hospital outpatient attendances? Time-series evaluation of primary care referral management. *Br J Gen Pract* 2013;**63**:e386–92. <http://dx.doi.org/10.3399/bjgp13X668177>
166. Bungard TJ, Smigorowsky MJ, Lalonde LD, Hogan T, Doliszny KM, Gebreyesus G, *et al*. Cardiac EASE (Ensuring Access and Speedy Evaluation) – the impact of a single-point-of-entry multidisciplinary outpatient cardiology consultation program on wait times in Canada. *Can J Cardiol* 2009;**25**:697–702. [http://dx.doi.org/10.1016/S0828-282X\(09\)70530-6](http://dx.doi.org/10.1016/S0828-282X(09)70530-6)
167. Rymaszewski LA, Sharma S, McGill PE, Murdoch A, Freeman S, Loh T. A team approach to musculo-skeletal disorders. *Ann R Coll Surg Engl* 2005;**87**:174–80. <http://dx.doi.org/10.1308/1478708051793>
168. Maddison P, Jones J, Breslin A, Barton C, Fleur J, Lewis R, *et al*. Improved access and targeting of musculoskeletal services in northwest Wales: targeted early access to musculoskeletal services (TEAMS) programme. *BMJ* 2004;**329**:1325–7. <http://dx.doi.org/10.1136/bmj.329.7478.1325>
169. Raine R, Carter S, Sensky T, Black N. 'Referral into a void': opinions of general practitioners and others on single point of access to mental health care. *J R Soc Med* 2005;**98**:153–7. <http://dx.doi.org/10.1258/jrsm.98.4.153>
170. Xiang A, Smith H, Hine P, Mason K, Lanza S, Cave A, *et al*. Impact of a referral management 'gateway' on the quality of referral letters; a retrospective time series cross sectional review. *BMC Health Serv Res* 2013;**13**:310. <http://dx.doi.org/10.1186/1472-6963-13-310>
171. Grol R, Rooijackers-Lemmers N, van Kaathoven L, Wollersheim H, Mokka H. Communication at the interface: do better referral letters produce better consultant replies? *Br J Gen Pract* 2003;**53**:217–9.
172. Royal College of General Practitioners. *Profile of UK General Practitioners*. London: RCGP; 2006.
173. Department of Health. *Shifting the Balance of Power within the NHS*. London: Department of Health; 2001.
174. Department of Health. *Implementing a Scheme for General Practitioners with Special Interests*. Department of Health; 2002.

175. Department of Health. *Guidelines for the Appointments of General Practitioners with a Special Interest with a Role in Service Development*. London: Department of Health; 2002.
176. Pinnock H, Netuveli G, Price D, Sheikh A. General practitioners with a special interest in respiratory medicine: national survey of UK primary care organisations. *BMC Health Serv Res* 2005;**5**:40. <http://dx.doi.org/10.1186/1472-6963-5-40>
177. Rosen R, Jones R, Tomlin Z, Cavanagh M-R. *Evaluation of General Practitioners with Special Interests: Access, Cost Evaluation and Satisfaction with Services*. London: NCCSDO; 2005.
178. Pickard S. The professionalization of general practitioners with a special interest: rationalization, restratification and governmentality. *Sociology* 2009;**43**:250–67. <http://dx.doi.org/10.1177/0038038508101164>
179. Department of Health. *Our Health Our Say: A New Direction for Community Services*. London: The Stationery Office; 2006.
180. Zwijnenberg NC, Bours GJ. Nurse practitioners and physician assistants in Dutch hospitals: their role, extent of substitution and facilitators and barriers experienced in the reallocation of tasks. *J Adv Nurs* 2012;**68**:1235–46. <http://dx.doi.org/10.1111/j.1365-2648.2011.05823.x>
181. Thompson C, Dowding D, editors. *Essential Decision Making and Clinical Judgement for Nurses*. Edinburgh: Churchill Livingstone; 2009.
182. Hughes EC. Social role and the division of labor. *Midwest Sociol* 1956;**18**:3–7.
183. Gilbert R, Franks G, Watkin S. The proportion of general practitioner referrals to a hospital Respiratory Medicine clinic suitable to be seen in a GPwSI Respiratory Clinic. *Prim Care Respir J* 2005;**14**:314–19. <http://dx.doi.org/10.1016/j.pcrj.2005.04.006>
184. Sanderson D. *Evaluation of the GPs with Special Interests (GPwSIs) Pilot Projects within the Action on ENT Programme*. York: York Health Consortium; 2002.
185. Rogers BA, Kabir C, Bradley N. An audit of orthopaedic referrals via multi-professional triage teams. *Ann R Coll Surg Engl* 2008;**90**:671–4. <http://dx.doi.org/10.1308/003588408X318174>
186. Salisbury C, Noble A, Horrocks S, Crosby Z, Harrison V, Coast J, et al. Evaluation of a general practitioner with special interest service for dermatology: randomised controlled trial. *BMJ* 2005;**331**:1441–6. <http://dx.doi.org/10.1136/bmj.38670.494734.7C>
187. Coast J, Noble S, Noble A, Horrocks S, Asim O, Peters TJ, et al. Economic evaluation of a general practitioner with special interests led dermatology service in primary care. *BMJ* 2005;**331**:1444–9. <http://dx.doi.org/10.1136/bmj.38676.446910.7C>
188. Jones K. Impact of an interest in asthma on prescribing costs in general practice. *Qual Health Care* 1992;**1**:110–13. <http://dx.doi.org/10.1136/qshc.1.2.110>
189. Ridsdale L, Doherty J, McCrone P, Seed P. A new GP with special interest headache service: observational study. *Br J Gen Pract* 2008;**58**:478–83. <http://dx.doi.org/10.3399/bjgp08X319440>
190. Wilson E, Cox K. Community palliative care development: evaluating the role and impact of a general practitioner with a special interest in palliative medicine. *Palliat Med* 2007;**21**:527–35. <http://dx.doi.org/10.1177/0269216307082018>
191. Kernick DP. Developing intermediate care provided by general practitioners with a special interest: the economic perspective. *Br J Gen Pract* 2003;**53**:553–6.
192. Elliot S, Kernick D. Why do GPs with a special interest in headache investigate headache presentations with neuroradiology and what do they find? *J Headache Pain* 2011;**12**:625–8. <http://dx.doi.org/10.1007/s10194-011-0375-8>

193. Martin GP, Currie G, Finn R. Reconfiguring or reproducing intra-professional boundaries? Specialist expertise, generalist knowledge and the 'modernization' of the medical workforce. *Soc Sci Med* 2009;**68**:1191–8. <http://dx.doi.org/10.1016/j.socscimed.2009.01.006>
194. Gerada C. *Changing Partnerships*. London: RCGP; 2009.
195. Lester H, Campbell SM, McDonald R. The present state and future direction of primary care: a qualitative study of GPs' views. *Br J Gen Pract* 2009;**59**:908–15. <http://dx.doi.org/10.3399/bjgp09X473060>
196. Newman CE, Kidd MR, de Wit JB, Reynolds RH, Canavan PG, Kippax SC. What moves a family doctor to specialise in HIV? Interviews with Australian policy key informants. *Cult Health Sex* 2011;**13**:1151–64. <http://dx.doi.org/10.1080/13691058.2011.607904>
197. Jones R, Bartholomew J. General practitioners with special clinical interests: a cross-sectional survey. *Br J Gen Pract* 2002;**52**:833–4.
198. Lindblom CE. The science of 'muddling through'. *Public Admin Rev* 1959;**19**:79–88. <http://dx.doi.org/10.2307/973677>
199. Pieters A, Oirschot C, Akkermans H. No cure for all evils: Dutch obstetric care and limits to the applicability of the focused factory concept in health care. *Int J Oper Prod Manag* 2010;**30**:1112–39. <http://dx.doi.org/10.1108/01443571011087350>
200. Hovmand PS, Andersen DF, Rouwette E, Richardson GP, Rux K, Calhoun A. Group model-building 'scripts' as a collaborative planning tool. *Syst Res Behav Sci* 2012;**29**:179–93. <http://dx.doi.org/10.1002/sres.2105>
201. Clarke AH, Rickards DF. Economics of open access to diagnostic services. *Lancet* 1965;**286**:336–7. [http://dx.doi.org/10.1016/S0140-6736\(65\)90302-8](http://dx.doi.org/10.1016/S0140-6736(65)90302-8)
202. Grewal K, Leung YW, Safai P, Stewart DE, Anand S, Gupta M, et al. Access to cardiac rehabilitation among South-Asian patients by referral method: a qualitative study. *Rehabil Nurs* 2010;**35**:106–12. <http://dx.doi.org/10.1002/j.2048-7940.2010.tb00285.x>
203. Eimerl TS. Direct-access diagnostic facilities in general practice. *Lancet* 1962;**1**:851–2. [http://dx.doi.org/10.1016/S0140-6736\(62\)91861-5](http://dx.doi.org/10.1016/S0140-6736(62)91861-5)
204. Green RH. General practitioners and open-access pathology services. *J R Coll Gen Pract* 1973;**23**:316–25.
205. Department of Health. *Delivering Care Closer to Home: Meeting the Challenge*. Department of Health; 2008.
206. Department of Health. *Equity and Excellence: Liberating the NHS*. London: Department of Health; 2010.
207. David Laszlo Partnership. *Aligning Diagnostics: Strategic Guidance For Pathology Services Reconfiguration*. Department of Health; 2012.
208. Birmingham South Central Clinical Commissioning Group. *Clinical Pathways*. Birmingham: Birmingham South Central Clinical Commissioning Group; 2013.
209. Bramble MG. Open access endoscopy – a nationwide survey of current practice. *Gut* 1992;**33**:282–5. <http://dx.doi.org/10.1136/gut.33.2.282>
210. Agency for Healthcare Research and Quality (AHRQ). *Methods Guide for Medical Test Reviews*. Rockville, MD: AHRQ; 2012.
211. Simpson GC, Forbes K, Teasdale E, Tyagi A, Santosh C. Impact of GP direct-access computerised tomography for the investigation of chronic daily headache. *Br J Gen Pract* 2010;**60**:897–901. <http://dx.doi.org/10.3399/bjgp10X544069>

212. Taylor TR, Evangelou N, Porter H, Lenthall R. Primary care direct access MRI for the investigation of chronic headache. *Clin Radiol* 2012;**67**:24–7. <http://dx.doi.org/10.1016/j.crad.2011.02.006>
213. Thomas R, Cook A, Main G, Taylor T, Galizia Caruana E, Swingler R. Primary care access to computed tomography for chronic headache. *Br J Gen Pract* 2010;**60**:426–30. <http://dx.doi.org/10.3399/bjgp10X502146>
214. Benamore RE, Wright D, Britton I. Is primary care access to CT brain examinations effective? *Clin Radiol* 2005;**60**:1083–9. <http://dx.doi.org/10.1016/j.crad.2005.05.010>
215. Wong BC, Chan CK, Wong KW, Wong WM, Yuen MF, Lai KC, *et al.* Evaluation of a new referral system for the management of dyspepsia in Hong Kong: role of open-access upper endoscopy. *J Gastroenterol Hepatol* 2000;**15**:1251–6. <http://dx.doi.org/10.1046/j.1440-1746.2000.02353.x>
216. Todd JA, Zubir MA, Goudie BM, Johnston DA. Response to open access endoscopy findings by general practitioners guidelines need education for implementation. *Scott Med J* 2000;**45**:49–50.
217. Ridsdale L, Clark LV, Dowson AJ, Goldstein LH, Jenkins L, McCrone P, *et al.* How do patients referred to neurologists for headache differ from those managed in primary care? *Br J Gen Pract* 2007;**57**:388–95.
218. Brealey SD, Atwell C, Bryan S, Coulton S, Cox H, Cross B, *et al.* The DAMASK trial protocol: a pragmatic randomised trial to evaluate whether GPs should have direct access to MRI for patients with suspected internal derangement of the knee. *BMC Health Serv Res* 2006;**6**:133. <http://dx.doi.org/10.1186/1472-6963-6-133>
219. Brealey SD, DAMASK Trial Team. Influence of magnetic resonance imaging of the knee on GPs' decisions: a randomised trial. *Br J Gen Pract* 2007;**57**:622–9.
220. DAMASK Trial Team. Effectiveness of GP access to magnetic resonance imaging of the knee: a randomised trial. *Br J Gen Pract* 2008;**58**:e1–8; discussion 774. <http://dx.doi.org/10.3399/bjgp08X342651>
221. DAMASK Trial Team. Cost-effectiveness of magnetic resonance imaging of the knee for patients presenting in primary care. *Br J Gen Pract* 2008;**58**:e10–6. <http://dx.doi.org/10.3399/bjgp08X342660>
222. Watura R, Lloyd DCF, Chawda S. Magnetic resonance imaging of the knee: direct access for general practitioners. *BMJ* 1995;**311**:1614. <http://dx.doi.org/10.1136/bmj.311.7020.1614>
223. Ireland J. Direct access magnetic resonance imaging of the knee for GPs. Paper highlights shortage of orthopaedic skill. *BMJ* 1996;**312**:850. <http://dx.doi.org/10.1136/bmj.312.7034.850>
224. Mackenzie R, Edwards D, Dixon AK. Direct access magnetic resonance imaging of the knee for GPs. Patients should be seen by an experienced orthopaedic surgeon. *BMJ* 1996;**312**:849–50. <http://dx.doi.org/10.1136/bmj.312.7034.849b>
225. Southgate J, Thomas N. Direct access magnetic resonance imaging of the knee for GPs. Magnetic resonance imaging should be used selectively. *BMJ* 1996;**312**:849; author reply 50. <http://dx.doi.org/10.1136/bmj.312.7034.849a>
226. Watura R, Lloyd D, Chawda S. Author's reply. *BMJ* 1996;**312**:850. <http://dx.doi.org/10.1136/bmj.312.7034.850a>
227. Chawda SJ, Watura R, Lloyd DC. Magnetic resonance imaging of the lumbar spine: direct access for general practitioners. *Br J Gen Pract* 1997;**47**:575–6.
228. Anderson O, Afolayan JO, Ni Z, Bates T. Surgical vs general practitioner assessment: diagnostic accuracy in 2-week-wait colorectal cancer referrals. *Colorectal Dis* 2011;**13**:e212–5. <http://dx.doi.org/10.1111/j.1463-1318.2011.02617.x>

229. Dhillon V, Creiger J, Hannan J, Hurst N, Nuki G. The effect of DXA scanning on clinical decision making by general practitioners: a randomized, prospective trial of direct access versus referral to a hospital consultant. *Osteoporos Int* 2003;**14**:326–33. <http://dx.doi.org/10.1007/s00198-002-1371-2>
230. Sim MFV, Stone M, Johansen A, Ho P, Pettit RJ, Evans WD. An analysis of an open access general practitioner bone densitometry service. *Int J Clin Pract* 2004;**58**:300–5. <http://dx.doi.org/10.1111/j.1368-5031.2004.00151.x>
231. Sheldon TA, Cullum N, Dawson D, Lankshear A, Lowson K, Watt I, *et al*. What's the evidence that NICE guidance has been implemented? Results from a national evaluation using time series analysis, audit of patients' notes, and interviews. *BMJ* 2004;**329**:999. <http://dx.doi.org/10.1136/bmj.329.7473.999>
232. Stock JL, Waud CE, Coderre JA, Overdorf JH, Janikas JS, Heiniluoma KM, *et al*. Clinical reporting to primary care physicians leads to increased use and understanding of bone densitometry and affects the management of osteoporosis. A randomized trial. *Ann Intern Med* 1998;**128**:996–9. http://dx.doi.org/10.7326/0003-4819-128-12_Part_1-199806150-00006
233. Skipsey DA, Dawson FM, Breen C, Leslie SJ. Evaluation of a direct access cardiac arrhythmia monitoring service. *Cardiol J* 2012;**19**:70–5. <http://dx.doi.org/10.5603/CJ.2012.0010>
234. van Gorp N, Boonman-De Winter LJ, Meijer Timmerman Thijssen DW, Stoffers HE. Benefits of an open access echocardiography service: a Dutch prospective cohort study. *Netherlands Heart J* 2013;**21**:399–405. <http://dx.doi.org/10.1007/s12471-013-0416-9>
235. van Heur LM, Baur LH, Tent M, Lodewijks-van der Bolt CL, Streppel M, Winkens RA, *et al*. Evaluation of an open access echocardiography service in the Netherlands: a mixed methods study of indications, outcomes, patient management and trends. *BMC Health Serv Res* 2010;**10**:37. <http://dx.doi.org/10.1186/1472-6963-10-37>
236. Fuat A, Hungin APS, Murphy JJ. Barriers to accurate diagnosis and effective management of heart failure in primary care: qualitative study. *BMJ* 2003;**326**:196. <http://dx.doi.org/10.1136/bmj.326.7382.196>
237. Brusamento S, Legido-Quigley H, Panteli D, Turk E, Knai C, Saliba V, *et al*. Assessing the effectiveness of strategies to implement clinical guidelines for the management of chronic diseases at primary care level in EU Member States: a systematic review. *Health Policy* 2012;**107**:168–83. <http://dx.doi.org/10.1016/j.healthpol.2012.08.005>
238. Espeland A, Baerheim A. Factors affecting general practitioners' decisions about plain radiography for back pain: implications for classification of guideline barriers – a qualitative study. *BMC Health Serv Res* 2003;**3**:8. <http://dx.doi.org/10.1186/1472-6963-3-8>
239. Rashidian A, Eccles MP, Russell I. Falling on stony ground? A qualitative study of implementation of clinical guidelines' prescribing recommendations in primary care. *Health Policy* 2008;**85**:148–61. <http://dx.doi.org/10.1016/j.healthpol.2007.07.011>
240. Lugtenberg M, Burgers JS, Clancy C, Westert GP, Schneider EC. Current guidelines have limited applicability to patients with comorbid conditions: a systematic analysis of evidence-based guidelines. *PLOS ONE* 2011;**6**:e25987. <http://dx.doi.org/10.1371/journal.pone.0025987>
241. Cochrane LJ, Olson CA, Murray S, Dupuis M, Tooman T, Hayes S. Gaps between knowing and doing: understanding and assessing the barriers to optimal health care. *J Contin Educ Health Prof* 2007;**27**:94–102. <http://dx.doi.org/10.1002/chp.106>
242. Gagliardi AR, Brouwers MC, Palda VA, Lemieux-Charles L, Grimshaw JM. How can we improve guideline use? A conceptual framework of implementability. *Implement Sci* 2011;**6**:26. <http://dx.doi.org/10.1186/1748-5908-6-26>

243. Cabana MD, Rand CS, Powe NR, Wu AW, Wilson MH, Abboud PA, *et al.* Why don't physicians follow clinical practice guidelines? A framework for improvement. *JAMA* 1999;**282**:1458–65. <http://dx.doi.org/10.1001/jama.282.15.1458>
244. DeVellis RF. *Scale Development: Theory and Applications*. 2nd edn. London: Sage Publications; 2003.
245. Cronbach LJ, Meehl PE. Construct validity in psychological tests. *Psychol Bull* 1955;**52**:281–302. <http://dx.doi.org/10.1037/h0040957>
246. Checkland K, Harrison S, Marshall M. Is the metaphor of 'barriers to change' useful in understanding implementation? Evidence from general medical practice. *J Health Serv Res Policy* 2007;**12**:95–100. <http://dx.doi.org/10.1258/135581907780279657>
247. Vitry AI, Zhang Y. Quality of Australian clinical guidelines and relevance to the care of older people with multiple comorbid conditions. *Med J Aust* 2008;**189**:360–5.
248. Caughey GE, Ramsay EN, Vitry AI, Gilbert AL, Luszcz MA, Ryan P, *et al.* Comorbid chronic diseases, discordant impact on mortality in older people: a 14-year longitudinal population study. *J Epidemiol Commun Health* 2010;**64**:1036–42. <http://dx.doi.org/10.1136/jech.2009.088260>
249. Hughes LD, McMurdo ME, Guthrie B. Guidelines for people not for diseases: the challenges of applying UK clinical guidelines to people with multimorbidity. *Age Ageing* 2013;**42**:62–9. <http://dx.doi.org/10.1093/ageing/afs100>
250. Boyd CM, Darer J, Boult C, Fried LP, Boult L, Wu AW. Clinical practice guidelines and quality of care for older patients with multiple comorbid diseases: implications for pay for performance. *JAMA* 2005;**294**:716–24. <http://dx.doi.org/10.1001/jama.294.6.716>
251. Mutasingwa DR, Ge H, Upshur RE. How applicable are clinical practice guidelines to elderly patients with comorbidities? *Can Fam Phys* 2011;**57**:e253–62.
252. Ashwell M. Plea for simplicity: use of waist-to-height ratio as a primary screening tool to assess cardiometabolic risk. *Clin Obesity* 2012;**2**:3–5. <http://dx.doi.org/10.1111/j.1758-8111.2012.00037.x>
253. Balducci L. New paradigms for treating elderly patients with cancer: the comprehensive geriatric assessment and guidelines for supportive care. *J Support Oncol* 2003;**1**(Suppl. 2):30–7.
254. Dawes M. Co-morbidity: we need a guideline for each patient not a guideline for each disease. *Fam Pract* 2010;**27**:1–2. <http://dx.doi.org/10.1093/fampra/cmp106>
255. Durso SC. Using clinical guidelines designed for older adults with diabetes mellitus and complex health status. *JAMA* 2006;**295**:1935–40. <http://dx.doi.org/10.1001/jama.295.16.1935>
256. Braithwaite RS, Fiellin D, Justice AC. The payoff time: a flexible framework to help clinicians decide when patients with comorbid disease are not likely to benefit from practice guidelines. *Med Care* 2009;**47**:610–17. <http://dx.doi.org/10.1097/MLR.0b013e31819748d5>
257. Intercollegiate Stroke Working Party. *National Clinical Guideline for Stroke*. 4th edn. London: Royal College of Physicians; 2012.
258. Oxfordshire Regional Cancer Genetics Service. Criteria for Referral from Primary Care to the Oxford Regional Cancer Genetic Service. 2005. URL: www.ouh.nhs.uk/services/referrals/genetics/documents/orcgs-criteria.pdf (accessed 13 August 2015).
259. Harrison MB, Graham ID, van den Hoek J, Dogherty EJ, Carley ME, Angus V. Guideline adaptation and implementation planning: a prospective observational study. *Implement Sci* 2013;**8**:49. <http://dx.doi.org/10.1186/1748-5908-8-49>

260. Capdenat Saint-Martin E, Michel P, Raymond JM, Iskandar H, Chevalier C, Petitpierre MN, *et al.* Description of local adaptation of national guidelines and of active feedback for rationalising preoperative screening in patients at low risk from anaesthetics in a French university hospital. *Qual Health Care* 1998;**7**:5–11. <http://dx.doi.org/10.1136/qshc.7.1.5>
261. Rowe RE. Local guidelines for the transfer of women from midwifery unit to obstetric unit during labour in England: a systematic appraisal of their quality. *Qual Saf Health Care* 2010;**19**:90–4. <http://dx.doi.org/10.1136/qshc.2008.030239>
262. Shye D, Porath A, Brown JB. Adapting a national guideline for local use: a comparative case study in a US and an Israeli health maintenance organization. *J Health Serv Res Policy* 2000;**5**:148–55.
263. Silagy CA, Weller DP, Lapsley H, Middleton P, Shelby-James T, Fazekas B. The effectiveness of local adaptation of nationally produced clinical practice guidelines. *Fam Pract* 2002;**19**:223–30. <http://dx.doi.org/10.1093/fampra/19.3.223>
264. Croudace T, Evans J, Harrison G, Sharp DJ, Wilkinson E, McCann G, *et al.* Impact of the ICD-10 2 Health Care (PHC) diagnostic and management guidelines for mental disorders on detection and outcome in primary care. Cluster randomised controlled trial. *Br J Psychiatry* 2003;**182**:20–30. <http://dx.doi.org/10.1192/bjp.182.1.20>
265. Liaw S-T, Sulaiman ND, Barton CA, Chondros P, Harris CA, Sawyer S, *et al.* An interactive workshop plus locally adapted guidelines can improve general practitioners asthma management and knowledge: a cluster randomised trial in the Australian setting. *BMC Fam Pract* 2008;**9**:22. <http://dx.doi.org/10.1186/1471-2296-9-22>
266. Harstall C, Taenzer P, Angus DK, Moga C, Schuller T, Scott NA. Creating a multidisciplinary low back pain guideline: anatomy of a guideline adaptation process. *J Eval Clin Pract* 2011;**17**:693–704. <http://dx.doi.org/10.1111/j.1365-2753.2010.01420.x>
267. Graham ID, Harrison MB, Lorimer K, Piercianowski T, Friedberg E, Buchanan M, *et al.* Adapting national and international leg ulcer practice guidelines for local use: the Ontario Leg Ulcer Community Care Protocol. *Adv Skin Wound Care* 2005;**18**:307–18. <http://dx.doi.org/10.1097/00129334-200507000-00011>
268. Vause J. *Investigating the Worried Well and Missing the Blooming Obvious. Ideas, Observations and a Few Curly Questions from Invited Bloggers.* URL: www.nzdoctor.co.nz/login.aspx?ReturnURL=/blogs/2013/september-2013/23/investigating-the-worried-well-and-missing-the-blooming-obvious.aspx&returnid=46521 (accessed 13 August 2015).
269. Sanderson M, Allen P, Peckham S, Hughes D, Brown M, Kelly G, *et al.* Divergence of NHS choice policy in the UK: what difference has patient choice policy in England made? *J Health Serv Res Policy* 2013;**18**:202–8. <http://dx.doi.org/10.1177/1355819613492716>
270. van Wersch A, Eccles M. Involvement of consumers in the development of evidence based clinical guidelines: practical experiences from the North of England evidence based guideline development programme. *Qual Health Care* 2001;**10**:10–6. <http://dx.doi.org/10.1136/qhc.10.1.10>
271. Tong A, Lopez-Vargas P, Howell M, Phoon R, Johnson D, Campbell D, *et al.* Consumer involvement in topic and outcome selection in the development of clinical practice guidelines. *Health Expect* 2012;**15**:410–23. <http://dx.doi.org/10.1111/j.1369-7625.2011.00676.x>
272. van de Bovenkamp HM, Trappenburg MJ. Reconsidering patient participation in guideline development. *Health Care Anal* 2009;**17**:198–216. <http://dx.doi.org/10.1007/s10728-008-0099-3>
273. van de Bovenkamp HM, Zuiderent-Jerak T. An empirical study of patient participation in guideline development: exploring the potential for articulating patient knowledge in evidence-based epistemic settings [published online ahead of print 2 May 2013]. *Health Expect* 2013. <http://dx.doi.org/10.1111/hex.12067>

274. Clarke A. *The REFER Project: Realistic Effective Facilitation of Elective Referral for Elective Surgical Assessment*. National Institute for Health Research (NIHR) Service Delivery and Organisation programme. Southampton: NIHR; 2010.
275. van der Weijden T, Boivin A, Burgers J, Schunemann HJ, Elwyn G. Clinical practice guidelines and patient decision aids. An inevitable relationship. *J Clin Epidemiol* 2012;**65**:584–9. <http://dx.doi.org/10.1016/j.jclinepi.2011.10.007>
276. O'Connor AM, Wennberg JE, Legare F, Llewellyn-Thomas HA, Moulton BW, Sepucha KR, et al. Toward the 'tipping point': decision aids and informed patient choice. *Health Affairs* 2007;**26**:716–25. <http://dx.doi.org/10.1377/hlthaff.26.3.716>
277. Han PK, Kobrin S, Breen N, Joseph DA, Li J, Frosch DL, et al. National evidence on the use of shared decision making in prostate-specific antigen screening. *Ann Fam Med* 2013;**11**:306–14. <http://dx.doi.org/10.1370/afm.1539>
278. Coiera E. Maximising the uptake of evidence into clinical practice: an information economics approach. *Med J Aust* 2001;**174**:467–70.
279. Mazza D, Chapman A, Michie S. Barriers to the implementation of preconception care guidelines as perceived by general practitioners: a qualitative study. *BMC Health Serv Res* 2013;**13**:36. <http://dx.doi.org/10.1186/1472-6963-13-36>
280. Malik A, Wigney L, Murray S, Gerrand CH. The effectiveness of 'two-week' referrals for suspected bone and soft tissue sarcoma. *Sarcoma* 2007;**2007**:23870. <http://dx.doi.org/10.1155/2007/23870>
281. Jones R, Rubin G, Hungin P. Is the two week rule for cancer referrals working? *BMJ* 2001;**322**:1555–6. <http://dx.doi.org/10.1136/bmj.322.7302.1555>
282. Potter S, Govindarajulu S, Shere M, Braddon F, Curran G, Greenwood R, et al. Referral patterns, cancer diagnoses, and waiting times after introduction of two week wait rule for breast cancer: prospective cohort study. *BMJ* 2007;**335**:288. <http://dx.doi.org/10.1136/bmj.39258.688553.55>
283. Cornford CS, Harley J, Oswald N. The '2-week rule' for suspected breast carcinoma: a qualitative study of the views of patients and professionals. *Br J Gen Pract* 2004;**54**:584–8.
284. Cornford CS, Oswald N. The 2-week rule for patients with suspected breast cancer: what can be learnt by analysing policy documents? *Health Policy* 2004;**68**:263–6. <http://dx.doi.org/10.1016/j.healthpol.2003.07.002>
285. Barwick TW, Murray S, Gerrand C. Impact of 'two-week wait' referrals for suspected bone or soft tissue carcinoma. *Ann R Coll Surg Engl* 2010;**92**:208–11. <http://dx.doi.org/10.1308/147363510X506829>
286. Cookson R, McDaid D, Maynard A. Wrong SIGN, NICE mess: is national guidance distorting allocation of resources? *BMJ* 2001;**323**:743–5. <http://dx.doi.org/10.1136/bmj.323.7315.743>
287. NHS Institute for Innovation and Improvement. *Improving Patient Flows in the NHS: Case Studies in Reducing Delay*. URL: www.carebydesign.org/files/no_delays_achiever_case_studies.pdf (accessed 30 July 2013).
288. Baathe F, Norback LE. Engaging physicians in organisational improvement work. *J Health Organ Manag* 2013;**27**:479–97. <http://dx.doi.org/10.1108/JHOM-02-2012-0043>
289. NHS Institute for Innovation and Improvement. *Building Energy for Change*. URL: www.institute.nhs.uk/tools/energy_for_change/energy_for_change_.html (accessed 30 July 2013).
290. Akkermans H. Renga: a systems approach to facilitating inter-organizational network development. *Syst Dynamics Rev* 2001;**17**:179–93. <http://dx.doi.org/10.1002/sdr.215>

291. Pieters A, Akkermans H, Franx A. E pluribus unum: using group model building with many interdependent organizations to create integrated health-care networks. *Adv Health Care Manag* 2011;**10**:321–44. [http://dx.doi.org/10.1108/S1474-8231\(2011\)0000010025](http://dx.doi.org/10.1108/S1474-8231(2011)0000010025)
292. Rouwette EAJA, Vennix JAM, van Mullekom T. Group model building effectiveness: a review of assessment studies. *Syst Dynamics Rev* 2002;**18**:5–45. <http://dx.doi.org/10.1002/sdr.229>
293. Rouwette EJA, Vennix JM, Felling AA. On evaluating the performance of problem structuring methods: an attempt at formulating a conceptual model. *Group Decis Negot* 2009;**18**:567–87. <http://dx.doi.org/10.1007/s10726-007-9100-z>
294. Dewey J. *Logic: The Theory of Inquiry*. New York, NY: Henry Holt and Company; 1939.
295. Cruz AM. Chatter. *People*, 29 December 2003.
296. Oxman A, Sackett DL, Chalmers I, Prescott TE. A surrealistic mega-analysis of reorganization theories. *J Royal Soc Med* 2005;**98**:563–8. <http://dx.doi.org/10.1258/jrsm.98.12.563>
297. Stone JR. *The Routledge Book of World Proverbs*. Abingdon: Routledge; 2006.
298. Reinink A. Don't Cut Out the Middleman – Become One. *Entrepreneur Magazine*, 15 September 2009.
299. Austin J. *Mansfield Park*. London: Thomas Egerton; 1814. <http://dx.doi.org/10.1093/oseo/instance.00080851>
300. Lask T. Stars from Britain: Donald Pleasence. *New York Times*, 10 December 1961.
301. Nancarrow Sa, Borthwick AM. Dynamic professional boundaries in the healthcare workforce. *Social Health Illn* 2005;**27**:897–919. <http://dx.doi.org/10.1111/j.1467-9566.2005.00463.x>
302. Popper K. *Conjectures and Refutations*. London: Routledge and Keagan Paul; 1963.
303. Akbari A, Mayhew A, Al-Alawi M, Grimshaw J, Winkens R, Glidewell E, *et al*. Interventions to improve outpatient referrals from primary to secondary care. *Cochrane Database Syst Rev* 2008;**4**:CD005471. <http://dx.doi.org/10.1002/14651858.cd005471.pub2>
304. Jack EP, Powers TL. A review and synthesis of demand management, capacity management and performance in health services. *Int J Manag Rev* 2009;**11**:149–74. <http://dx.doi.org/10.1111/j.1468-2370.2008.00235.x>
305. NHS Evidence Adoption Centre East of England. *Effectiveness and Cost-Effectiveness of Targeted Interventions to Reduce Unnecessary Referrals and Improve the Quality of Referrals from Primary to Secondary Care*. Cambridge: NHS Evidence Adoption Centre East of England; 2010.
306. Thomas R, Grimshaw J, McClinton S. Cluster randomised controlled trial of a guideline based open access urological investigation service. *Fam Pract* 2003;**20**:646–54. <http://dx.doi.org/10.1093/fampra/cm605>
307. Sibbald B, McDonald R, Roland M. Shifting care from hospitals to the community: a review of the evidence on quality and efficiency. *J Health Serv Res Policy* 2007;**12**:110–17. <http://dx.doi.org/10.1258/135581907780279611>

Appendix 1 Search and selection strategies

This appendix provides further details of the methods used in searching and selecting the primary sources, which provide the platform for the synthesis. Let us put the question very directly. As can be seen, the report utilises several hundred of primary inquiries, some of which are singled out for particularly detailed analysis. The question, especially for those unfamiliar with theory-driven inquiry, is how were these studies sought and selected? Why these and not a thousand others?

A more familiar approach to searching for the primary studies relevant to a review is to begin with the PICO formula (www.york.ac.uk/inst/crd/pdf/Systematic_Reviews.pdf) and we begin with a brief comment on why this method is inapplicable in reviews of complex social interventions. In reviewing clinical interventions, the researcher would normally begin the search process with close operational definitions of the intervention (I), the population (P) targeted, the intended outcomes (O) and the comparative interventions (C), if any, against which it may be assessed. All of these features resist tight pre-specification in complex service delivery interventions. In the case of demand management, not only are there a prolific number of different programmes, but each one is subject to and works through constant negation and adaptation. The inescapable discovery of the review is that we are dealing with responsive interventions under constant adaptation. Nor are demand management interventions aimed at populations. They are targeted from the top to the bottom of entire regimes – organisational strategies to everyday practices, policy-makers to practitioners to patients. Programme outcomes can indeed be pre-specified; this is an inquiry into demand efficiencies. However, it should also be quite clear from the report that such gains or losses depend entirely on underlying processes and surrounding circumstances and it is material on these which needs to be identified in any search strategy. And, finally, comparators? Demand management interventions are embedded in services, which are bombarded simultaneously with an endless stream of policy imperatives and in regimes which already have quite different histories of referral management. There is no 'usual care' or 'usual practice' against which to compare.

Realist synthesis thus begins with a different conception of the purpose of a review and different search strategy. The starting point is to begin with programme theory, the ideas underlying a family of interventions, as it is these theories which are tested in the course of the review. Our exploration of demand management theories is covered in *Chapter 3*. As our focus is on underlying policy thinking, we directed our attention, self-evidently, to the 'ideas literature'. In other words we searched for sources in the so-called grey literature (planning documents, guidance materials, discussion documents, proposals, rationales, policy expositions, professional journals and critical debate). Here it suffices to use rather simple search terms that have been identified in our research brief and to borrow those terms and synonyms used previous reviews or demand management, which identify the core approaches such as 'referral management centres', 'guidelines', 'feedback', 'GPs with special interests', 'direct access to test results' and so on.

The immediate task is to extract and compile the underlying programme theories on the basis a close reading of this documentation and it will be recalled that we distinguished different and substantial currents of strategic, administrative, role, procedural and motivational thinking. It is worth emphasising the exact status of our claim to have captured the 'theories of demand management' here, for this is also the warrant for the initial search strategy. The idea is not that we compile an exhaustive list of every ambition and every conjecture that has ever been forwarded in the name of demand and referral management, only that we capture a *key subset of ideas in play from the key players*. This is the claim against which our selection of programme theories must be judged. To be sure, some demand management interventions have been mounted with overtly political ambitions to expand or shrink medical power hierarchies; some will have been applied in furtherance of the careers of key players; some will have been mounted as an adjunct to broader health service reforms; and so on. Such matters were detected in our original search but its purpose and its justification is to place on the table some of the clearly influential programme theories.

Influential ideas usually turn out to be contentious and it is the key points of contention that are then put to further analysis.

This brings us to the second phase of our search, namely to unearth primary studies that enable us to test the intervention theories. These inquiries provide the foundation for refining the respective programme ideas by providing evidence on their implementation and targeting – in what guises, in what respects and in what circumstance have the programme theories come to fruition? The primary materials of interest here are evaluative inquiries and so the stock-in-trade materials take the form of formal research ‘reports’ as well as papers and commentary from the many health-care journals. Theory testing can make use of findings that emerge from any form of quantitative or qualitative inquiry; there is no hierarchy of evidence and thus no search restrictions on that basis.

The preliminary trawl for these empirical studies is conducted on orthodox lines. Having identified the broad families of interventions (as in *Chapters 4–7*), searches commence utilising key terms and synonyms associated with each intervention and employing the standard databases for health service research. Examples of the initial search strategies and decision flow charts for three of our interventions are reported below.

These searches provide only the raw materials for the extended syntheses and here we describe how the search for and selection of further studies proceeds on an iterative and purposive basis. Greenhalgh *et al.*'s review conducted along realist lines found that the majority of ‘hits’ were discovered using these progressive, snowballing searches and we illustrate the same process with an example from the present exercise.⁶ In the review of role change interventions (see *Chapter 5*), we discovered on reading several papers that ensuing demand efficiencies depended considerably on whether the new role involved ‘substitution’ or ‘support’ (i.e. concepts not used in the original searches). Repeated sightings of these terms then lead back to key papers by, for instance, Nancarrow and Borthwick,³⁰¹ which explored the ‘permeability’ and ‘specialisation’ of different roles. This iterative process provides us with both an additional armoury of research terms and new resources, namely article bibliographies and citation indices, with which to pursue further sources to interrogate our developing programme theory (e.g. Google Scholar lists 194 citations for the Nancarrow and Borthwick paper). The best primary inquiries have exactly the same structure. They develop in a process of cross-referencing. Study A develops a conjecture B, which is partially refuted by study C, which leads refined theory D, which is partially refuted by study E, which leads to emerging conjecture E, and so on ad infinitum.³⁰² It is this mutual interrogation which is the source of scientific progress and which is imitated in realist search strategies.

We now reach both the opportunity and the dilemma of theory-led inquiry. We have the opportunity to follow finer and finer lines of inquiry, while at the same time recognising that there is a multiplication in the number of potential lines of inquiry. This predicament changes the underlying dynamics of the search strategy. The field of cognate studies become infinite and the objective of the search strategy transforms to that of finding the most *relevant* primary studies to test a *promising* line of inquiry. At this stage judgement must be applied to restrict and direct the searchlight; searching becomes purposive. For instance, in the study of direct access to tests (see *Chapter 6*) an examination of outcome patterns led us to develop the theory that the function of the test – ‘ruling out’, ‘differential diagnosis’, ‘clinical indicator’ – shaped the opportunities for demand reduction. This then guided the search towards studies that allowed us to test out the subtle outcome differences across these three categories. Other conjectures fell by the wayside. We did not pursue potentially important theories about the speed and number of handovers in relaying test results. We did not pursue the 1962 *Lancet* hypothesis about the need for physical redesign of hospitals with diagnostic services at the entry post.²⁰³

It is quite possible, of course, that our judgement might be questioned. For instance, in the analysis of the effectiveness of guidelines in reducing demand (see *Chapter 7*), we lighted on the idea that local adaptation of guidelines was a necessary but not sufficient condition to propel them into usage. Our emerging theory was that guidance remains paper guidance unless it is embodied in the division of labour

responsible for its uptake. A journal paper by Capdenat Saint-Martin was used to as evidence to show that guidelines work most effectively when the personnel responsible for adapting the guidelines are the same people who organise the implementation of the new guidelines.²⁶⁰ It could be argued, of course, that this is a hospital-based study and not about guidelines produced to govern the interface between primary and secondary care. Our conjecture is that same principle will apply. It is open to refutation.

Searching the literature for primary investigations that yield explanations is quite different from seeking out studies that identify populations, interventions, comparators and outcomes.

Search strategy for Chapter 4

Database: Ovid MEDLINE(R)

Date range searched: 1946 to week 3, November 2013.

Date of search: 25 November 2013.

1. ("referral management" adj3 (centre* or center* or gateway* or clinic or system)).tw. (15)
2. "Referral and Consultation"/ (51,778)
3. ("multi-disciplinary clinic*1" or "multidisciplinary clinic*1").tw. (347)
4. 2 and 3 (32)
5. *"Referral and Consultation"/ (19,005)
6. exp *Primary Health Care/ or *general practice/ or *family practice/ (87,730)
7. (referral centre* or referral center*).tw. (11,992)
8. 5 and 6 and 7 (17)
9. 1 or 4 or 8 (64)
10. ("referral" adj3 (centre* or center* or gateway* or clinic or system)).tw. (15,221)
11. 5 and 6 and 10 (89)
12. "clinical assessment service*".tw. (17)
13. "clinical assessment cent*".tw. (2)
14. 12 or 13 (19)
15. "one stop service*".tw. (22)
16. "single point of access".tw. (19)
17. 15 or 16 (41)
18. (triage or referral*).tw. (72,934)
19. 17 and 18 (8)
20. (Triage and referral*).tw. (716)
21. 5 and 20 (120)
22. Triage/ (8072)
23. 5 and 22 (175)
24. 21 or 23 (226)
25. 14 or 19 or 24 (253)
26. 14 or 19 or 24 [final extended referral mgt centre search] (253)

This search was repeated in EMBASE and identified 565 citations.

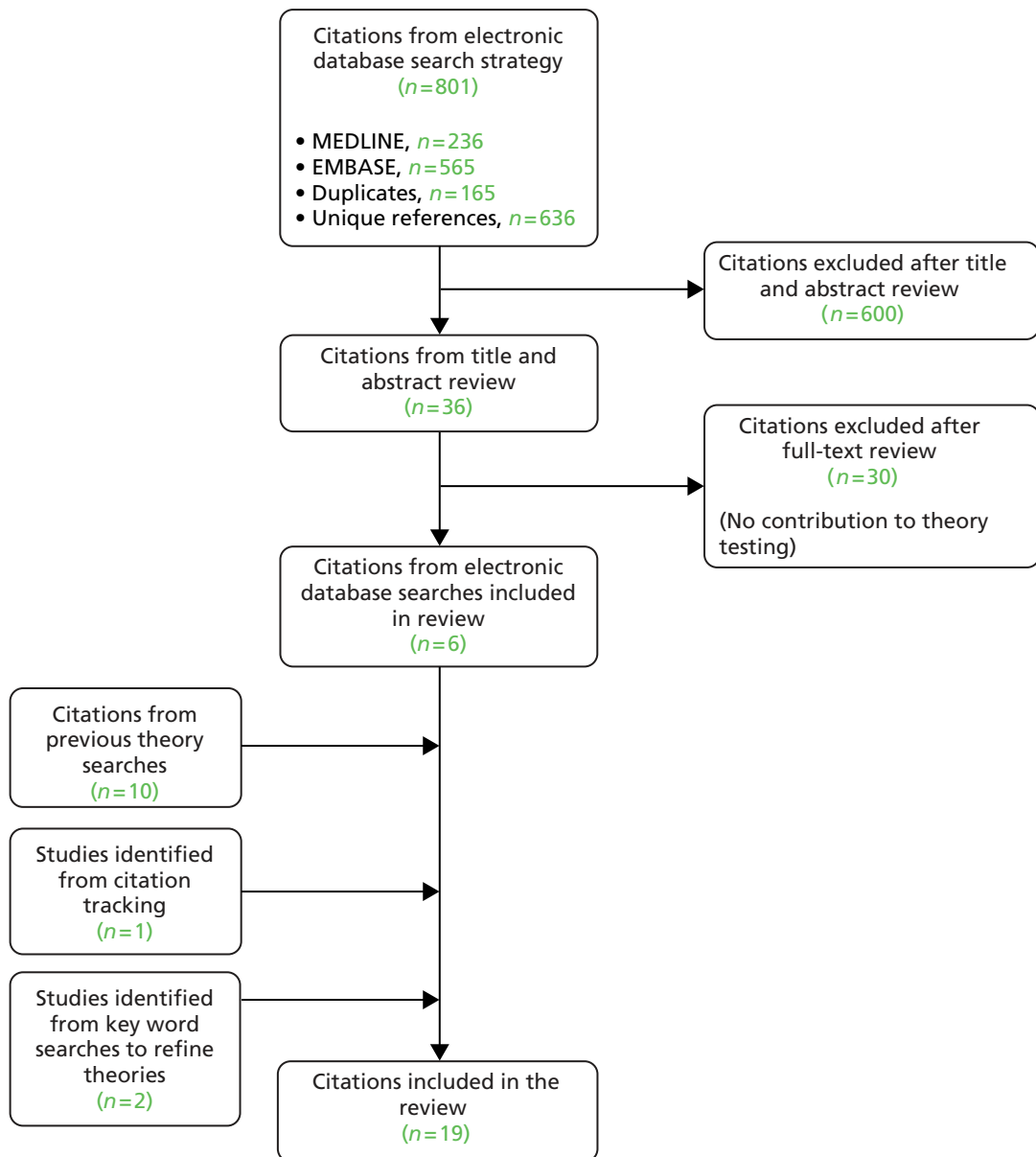


FIGURE 34 Flow chart of inclusion of studies into review (RMC).

Search strategy for Chapter 5

Database: Ovid MEDLINE(R)

Date range searched: 1946 to week 5, April 2014.

Date of search: 5 May 2014.

1. GPSI.tw. (11)
2. GP with a Special Interest.tw. (6)
3. Specialist GP.tw. (15)
4. (General Practitioner adj3 special interest).tw. (19)
5. 1 or 2 or 3 or 4 (49)
6. GPSI.tw. (11)
7. GPwSI.tw. (13)
8. 5 or 6 or 7 (59)
9. limit 8 to yr="1902-2013" (59)

EMBASE

Database: EMBASE Classic + EMBASE

Date range searched: 1947 to 7 May 2014.

Date of search: 13 January 2014.

1. GPSI.tw. (16)
2. GP with a Special Interest.tw. (12)
3. Specialist GP.tw. (18)
4. (General Practitioner adj3 special interest).tw. (28)
5. 1 or 2 or 3 or 4 (72)
6. GPSI.tw. (16)
7. GPwSI.tw. (39)
8. 5 or 6 or 7 (101)
9. limit 8 to yr="1902 - 2013" (99)

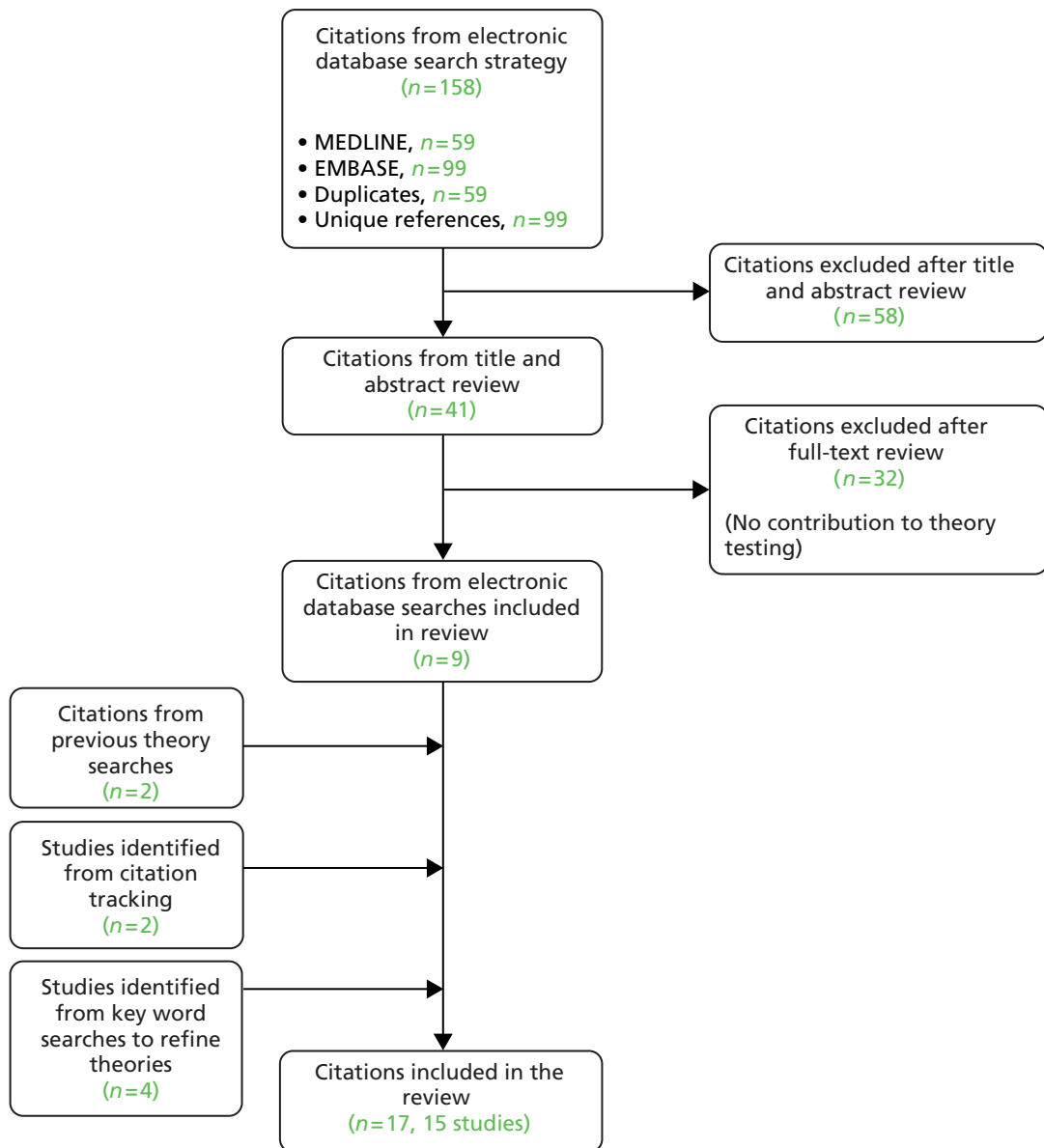


FIGURE 35 Flow chart of inclusion of studies into review (GPwSI).

Search strategy and decision flow chart for direct access to tests, Chapter 6

Database: Ovid MEDLINE(R)

Date range searched: 1946 to week 3, November 2013.

Date of search: 13 January 2014.

1. exp *Primary Health Care/ or *general practice/ or *family practice/ (87,730)
2. "direct access".tw. (1932)
3. 1 and 2 (86)
4. access.ti. (26,521)
5. 1 and 4 (838)
6. exp "diagnostic techniques and procedures"/ or diagnostic tests, routine/ (5,681,971)
7. 1 and 4 and 6 (93)
8. 3 or 7 (159)

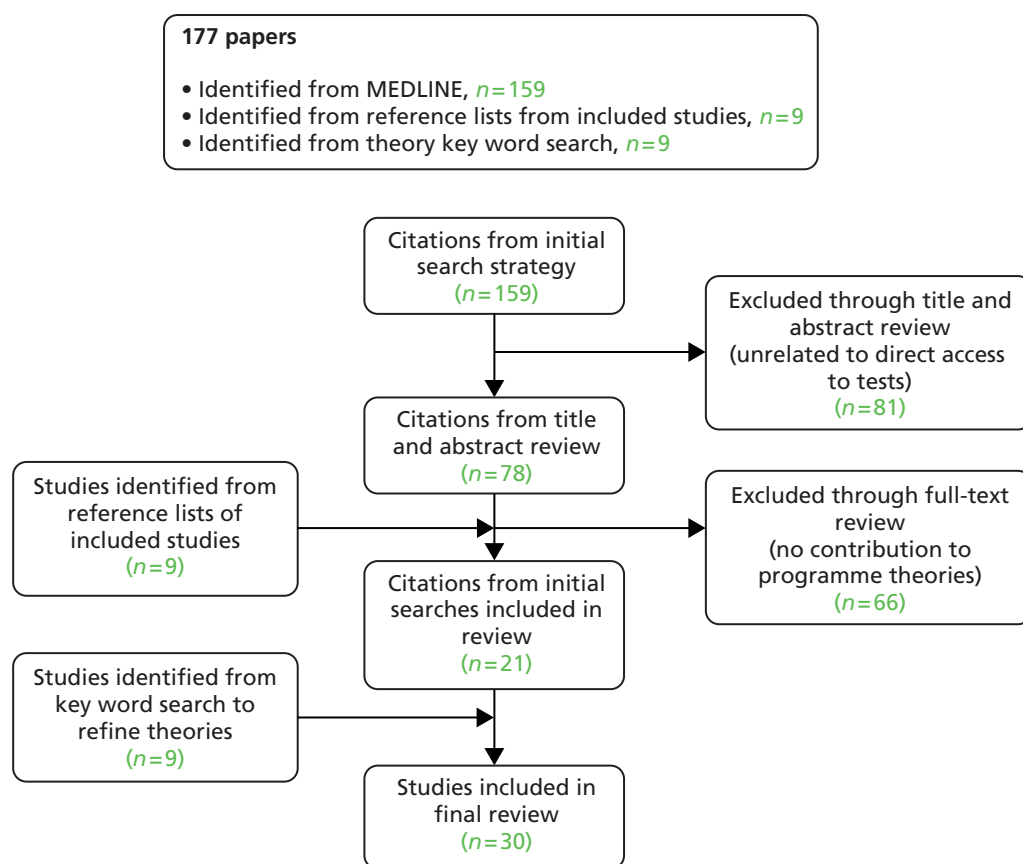


FIGURE 36 Flow chart of studies included in the review (direct access to tests).

Appendix 2 Reviewing the field

Text in this appendix has been reproduced from Pawson and colleagues under the terms of the Creative Commons Licence.⁹²

This appendix provides a brief examination of previous attempts to review the field of demand management intervention in health-care provision. The problem in a nutshell is that all reviews on this topic have to confront what only can be conceived as a rapidly expanding policy domain; there is a continuous evolution of new species and subspecies of programmes. This situation is light years away from the original idea that systematic review examined *all* studies on *one* type of intervention. The actual task is to review *many* studies of *several* overlapping and constantly adapting interventions.

This is a daunting undertaking and most of the previous attempts to review demand management have opted for an approach which might be termed the 'typological review'. The basic format is summarised in the matrix in *Table 14*. First, a decision is made on the significant types and subtypes of the broad intervention under review (types A–E in the table). Most complex interventions have multiple ambitions and the second element of the review structure is to decide on the various outputs and outcomes in terms of which the different family members will be assessed and compared. The schema is completed in the cells of the matrix where the findings are arrayed. The subject matter being complex adaptive interventions, relatively few RCTs are to be found in the primary research and so the review findings are predominantly discursive. The expectation, nevertheless, is that the review will call on a variety of evidence to provide a 'scorecard' or 'dashboard summary' of each species in terms of whether the evidence is broadly positive (✓), broadly negative (✗) or absent/indecisive (?). The overall objective of the typological review is to distinguish the most effective subtypes (depicted notionally in the diagram as type C).

We have identified five previous demand management/referral management reviews.^{93,95,303–305} Roland *et al.*⁹⁵ employ a 'scoping review' covering all types of primary study designs; Akbari *et al.*³⁰³ is a Cochrane review; Jack and Powers international review uses narrative description;³⁰⁴ The Evidence Adoption Centre compile a review of reviews;³⁰⁵ and Imison and Naylor combine their review with case studies.⁹³ We offer no comprehensive review of reviews here, noting only that neophyte typological reviews clearly allow for many variations on the theme. Our analysis concentrates on the common challenges they face, most particularly in respect of our core thesis about the constantly shifting ground of intervention theory.

If our previous analysis is correct, it would suggest that the typological review, because of the very way it is constructed, will face a number of severe methodological challenges. Basically, our claim is that demand management interventions have a natural tendency to hybridise. In order to work, strategic innovations will need to be buttressed over time with administrative reforms; changes in everyday duties are more likely to bed down if supported by role realignment and a motivated workforce. . . and so on. Such cross-fertilisation suggests formidable challenges ahead for any attempt to lay down a classification system in the expectation that a review will then go on to encounter pure empirical instances of each type. In particular, we might anticipate stumbling blocks in terms of (1) disagreement in the definitions of the

TABLE 14 The typological review

Intervention	Output 1	Output 2	Output 3	Outcome 4	Outcome 5
Type A	✓	✗	✗	?	✗
Type B	✗	✓	✗	✗	?
Type C	✓	?	✓	✓	✓
Type D	✗	✓	✓	✗	?
Type E	✓	✗	✓	?	?

operational types and subtypes, and (2) indeterminacy in being able to assess and differentiate clear outcomes corresponding to each category.

We explore the first of these conjectures in *Table 15*, which summarises the different schemata employed in the five reviews. An immediate observation is to note both promising similarities and unsettling differences between the schemas. Many of the same categories and subcategories are discernible. There is, however, no consensus on how to subdivide the field. The coverage, number, naming and content of each type and subtype differs from review to review. In terms of coverage, Jack and Powers stretch the field massively to cover demand, capacity and performance management schemes,³⁰⁴ whereas Roland *et al.*⁹⁵ review only those schemes involved in establishing specialist outpatient services within the primary care environment. Even when there is overlap in the broad categories, the 'same' interventions are classed differently. For example, 'in house second opinion by a GP' is conceptualised as a 'professional behaviour change' intervention by Roland *et al.*,⁹⁵ an 'organisational intervention' by Akbari,³⁰³ a 'managerial intervention' by the Evidence Adoption Centre³⁰⁵ and a 'clinical triage and assessment' intervention by Imison and Naylor.⁹³

One frequently proposed method for improving discernment in a typology is to identify subtypes within each domain. As can be seen in the bracketed descriptions in *Table 15*, this tactic is employed in all but the Jack and Powers review.³⁰⁴ Such subcategorisation is achieved by moving down a level of abstraction. Rather than using generic titles like 'organisational' or 'behavioural' interventions, the subschemes are then identified using practitioner- or domain-specific NHS terminology (e.g. 'consultation liaison psychiatry', 'PCT-led referral management systems'). The Akbari *et al.* review employs this strategy in the extreme with the subtypes or subsubtypes being represented by individual or very small batches of interventions (small, incidentally, because in this the interventions need to have been researched in primary studies that meet Cochrane methodological requirements).³⁰³

Does the fragmentation of a classification system resolve the problem of the intermixing of intervention theory? We think not, for the very idea of typological review is to serve evidence-based management by differentiating approaches and choosing between them. The more the subtype identified is specific to local implementation and contextual conditions, the longer become the odds of reproducing it elsewhere. It becomes much more difficult to identify a transferable formula when the methodological entreaty is to adhere to 'type 4, subtype 5'. For instance, one of Akbari *et al.*'s³⁰³ successful subtypes turns out to be 'a multi-faceted intervention involving educational meetings, a new referral and reply sheet, new staff and changes in equipment and facilities'.³⁰⁶ What type or subtype is under consideration here? How could one follow that template? The paradox remains: typologies are simplifications and accentuations – interventions are complex and adaptive.

We turn now to our second hypothesis in relation to the findings of the typological review. If programme typologies fail to recognise interdependencies between programme theories, if they mask the depth ontology of organisation change, what is the fate of the outcome 'scorecards'? Somewhat against our contention that they will end in indeterminacy, let us begin with instances from our set of reviews that claim to identify winners and losers. In the latter category we have the most persistent finding from the reviews, namely that the 'passive use of guidelines' is ineffective in controlling demand through changing GPs' referral behaviour.

The explanation here is relatively straightforward. If referral guidelines are simply issued to GPs they have little support or incentive to follow them – given the scores of other demands of the working day. Note, however, the provenance of this miniature explanation. It is a perfect exemplification of our overall thesis that successful innovations depend on action up and down the organisations strata described in *Figure 9*. To quote Akbari *et al.* guidelines may be effective if 'local secondary care providers are involved in dissemination activities, structured referral sheets are used, secondary care management is responsive to changes in primary care behaviour as a result of the guidelines and if they reflect local circumstances and address local barriers'.³⁰³ In the right conditions the all-time loser can become a winner.

TABLE 15 'Types' identified in demand management reviews

Roland <i>et al.</i> (2006) ⁹⁵	Jack and Powers (2009) ³⁰⁴	Akbari <i>et al.</i> (2008) ³⁰³	Evidence Adoption Centre (2010) ³⁰⁵	Imison and Naylor (2010) ⁹³
Transfer of services to primary care (surgical/medical clinics in primary care, GPwSIs, changing outpatient follow-up, GP direct access to diagnostic tests or services)	Demand modelling	Professional education interventions (passive dissemination of local guidelines, guidelines with structured referral sheets, secondary care provider-led educational strategies)	Professional education interventions (structured referral sheets, educational activities by secondary care providers, passive dissemination of referral guidelines)	RMCs
Relocation of services to primary care (outpatient clinics held in primary care, telemedicine, attachment of specialists to primary care teams)	Health maintenance organisations	Organisational interventions (attachment of specialists to primary care teams, in-house second opinion, changes to appointment systems based on size of practice)	Managerial interventions (in-house second opinion, primary care clinics for chronic disease, changing outpatient follow up, GP direct access to diagnostic test, GP direct access to services, GPwSIs, referral management schemes)	Clinical triage and assessment (e.g. in-house second opinion, community based specialist clinics, specialists attached to primary care teams, photo triage)
Joint working between primary and secondary care (shared care management of chronic disease, consultation liaison psychiatry)	Vertical/horizontal integration	Financial interventions (change in remuneration system, GP fundholding, change in patient charging systems)	Financial interventions (incentives to GPs)	Peer review and feedback (e.g. peer review of referral letters, educational outreach visits by specialists)
Professional behaviour change (structured referral sheets, educational outreach by specialists, in-house second opinion by a GP, financial incentives)	Capacity management			Financial incentives (e.g. use of referral standards in contracts, PCTs led referral management system)
	Workforce management			Guidelines plus other support (e.g. structured referral sheets)
	Resource utilisation			Passive use of guidelines
	Subcontracting			
	Information technology			
	Quality of care			
	Technical efficiency			
	Financial incentives			

With this possible exception, our claim is that typological scorecards tend to the indeterminate. We begin by noting a familiar limiting feature of review methodology: the stunting of firm conclusions by study selection. Demand management/referral management interventions are colossal in number and yet they are represented in Akbari *et al.*'s review³⁰³ by 17 studies and in Roland *et al.*'s review⁹⁵ by 119 inquiries. The former uses the customary Cochrane restriction to primary studies conducted by RCTs and a fair proportion of its discussion is given over to the further methodological weaknesses of the subset that has passed initial muster. Here, then, is the entirely typical root of 'indeterminacy mark I'. Reviews proceed rigorously forward but by passing down responsibility to poorly executed primary studies it becomes difficult to draw firm conclusions.

The second root of indeterminacy is the proper caution expressed by authors as a result of the mixed outcomes that typically emerge in typological reviews. To appreciate this we examine a real scorecard (Table 16) extracted from analysis derived from Roland *et al.*'s review,⁹⁵ and compare it with the idealised framework with its notional 'winners' and 'losers' as in Table 14. This table covers three different demand management schemes involving the relocation of services to primary care.³⁰⁷ As a first and overall impression, one notes the equivocation expressed in many of the cells. The authors are able to point to partial gains here and losses there but the majority of the cells report 'insufficient evidence', 'no change' and 'variable outcomes'. Also noticeable but hard to quantify is a counterbalancing tendency, whereby a gain on one outcome is met by loss in another quarter.

There are a couple of reasons for this mixed picture. 'Indeterminacy mark II' is rooted in the classic problems of under-reporting and non-equivalence of analytic frameworks in primary studies. In total Roland's team seek to review four major intervention modes, broken down into 15 subtypes, analysed on six different outcomes ('equity', 'quality', 'hospital impact', 'GP impact', 'cost' and 'implementation issues'). Typological high ambition is met with paucity in the coverage and content of primary studies. There is no reason why a one-off evaluation should employ the generic framework, resulting in the inevitable absence of information on some cells.

This brings us to the all-pervasive 'indeterminacy mark III'. If one examines closely the summary finding deposited in most cell entries, they are not in fact verdicts but, quite properly, conditional statements. To use an obvious illustration, the first 'results cell' in Table 16 tells us that relocating an outpatient clinic will improve access and equity *if* it is located in the appropriate (badly served) area. In another example it transpires that telemedicine is more likely to be cost-efficient *if* utilised in remote areas. What is being discovered, unsurprisingly, is that there are no universal panaceas and that conditionality is the norm. This feature is most evident in the final column of the evidence matrix, where sits a brief assessment of 'implementation issues'.³⁰⁷ In each and every one there is a 'condition' that for the subscheme to work, other resources need to be brought to bear and other working practices must be changed to fall in line with the developing programme . . . scheme X requires expansion of a specialist workforce, scheme Y depends on good communication between actors, scheme Z requires major revision in working practices, and so on.

Inexorably, we are returned to our core thesis that, whatever the initial orientation of a demand management scheme, for it to work requires the co-ordination of strategies, rules, roles, routines and individual support. Alas, there has always been considerable antithesis to producing contingent findings in evaluative inquiry and research reviews. And this distaste results in a final characteristic of the scorecard summaries of the typology review, namely their free use of the model auxiliary verb. Of the various approaches it is frequently concluded that 'it *may* improve the referral process', 'it *may* not improve the quality of referrals', 'it *appears* a promising approach', 'it represents a *plausible* strategy', 'it is *generally* attractive'. Because they cannot fulfil the grand ambition of sorting intervention wheat from chaff, scorecards have to retreat from probabilities to possibilities. We suppose that 'possibilities' are a poor guide to policy-making. The proper answer to the 'what works' question is always 'it depends', and the job of research synthesis, we submit, is to deepen our understanding of the contingencies. And this is what we have attempted in the main body of the review.

TABLE 16 Summary of findings: relocation of services to primary care

Intervention	Outcomes	Quality/health	Hospital impact	General practice impact	Health service cost	Implementation issues
Shifted outpatient clinic	Improved access: potential to improve equity if located to populations with poor access to secondary care	Insufficient evidence: in theory, quality should be unchanged	Insufficient evidence on outpatient use: some patients will require added outpatient visit because primary care lacks diagnostic facilities	Insufficient evidence on workload: no gains in GP knowledge or skills	Clinics serving urban advantaged populations are not cost-effective due to loss of economies of scale	Requires expansion in specialist workforce to compensate for loss of economies of scale
Telemedicine	Improved access for remote populations; potential to improve equity of located to populations with poor access to secondary care	Insufficient evidence on health outcomes; diagnosis more difficult for some specialities (i.e. dermatology) but may improve with advances in technology	Insufficient evidence on outpatient use; some patients will require added outpatient visit because primary care lacks diagnostic facilities	Insufficient evidence but likely to increase primary care workload	Cost-effectiveness is highly context dependent but generally better when telemedicine clinics are located in remote areas where patient travel costs to outpatient clinics are high	Requires substantial investment in equipment and training of clinicians
Attachment of specialist	Improved access	No change	Variable: reduced outpatient referrals in some specialities (physiotherapy) but not others	No change	Variable: appears cost-effective for some specialities but not others	Requires expansion if specialist workforce and deployment to primary care teams

Reproduced with permission from Roland M, McDonald R, Sibbald B. Outpatient Services and Primary Care: A scoping review of research into strategies for improving outpatient effectiveness and efficiency. Final Report to National Institute for Health Research Service Delivery and Organisation programme 2006.⁹⁵

Appendix 3 Stakeholder and patient involvement

Throughout the review we sought advice on our theories and synthesis from a range of stakeholders including those within our project team, individual and group meetings with key stakeholders within the local health economy, a specially convened stakeholder group meeting and three meetings with our patient group, chaired by Laurence Wood. Their views helped us to develop and revise our programme theories, assess whether our findings were of use to their own decision-making and identify appropriate methods of dissemination.

In this appendix we provide a brief summary of how their views have shaped our review.

We presented our initial programme theories outlined in *Chapter 3* to a group of nine local NHS stakeholders at a 3-hour workshop and our patient group at a separate 3-hour workshop. The aim of the workshops was to validate, prioritise and refine our theories. A number of key issues emerged from these workshops, which informed our approach to evidence review:

- All demand management interventions operate within a complex system; unless change is directed at all levels of the system, the intervention will not be successful.
- Although the literature has attempted to categorise and analyse demand management interventions are separate interventions – in practice, different configurations of these interventions are often implemented together – for example, a reform in personnel such as the introduction of GPWSIs will often be implemented alongside the use of a procedural change such as the usage of guidelines to review referrals.

On the basis of this workshop we developed an overall framework to structure the remainder of our review. All demand management interventions attempt to influence the referral process through bringing about organisational change. Therefore, our basic framework for co-ordinating our review has drawn on the organisational change literature. Organisations are layered and change is mediated by intercommunication between these layers. In the existing review literature, demand management interventions are often characterised as bringing about change at a specific level of the organisation. However, substantive theories of organisational change tell us that what is provoked at one level may be prevented at another. Accordingly, researching demand management interventions requires an evaluation of the 'programme theory' around the entire system. With this in mind our research synthesis has explored the interconnectivity of different demand management schemes through the different organisational strata.

The workshops also identified other key issues that we explored during our review:

- Both patients and stakeholders expressed concern that RMCs may have the unintended consequence of 'de-skilling GPs' because if GPs refer to a RMC they are not required to think as carefully about the most appropriate management plan for the patient. This was a theory we tested in our review of RMCs.
- Both patients and stakeholders questioned how applicable standardised guidelines were to patients with comorbidity and whether or not following one guideline may disagree with the recommendations of another guideline. This was a 'system strain' that we explored in our guidelines review.

Patients also questioned the extent to which guidelines, RMCs and internal peer review of referrals would interfere with the relationship they had with their doctor. They questioned if their trust in the doctor and faith in their expertise would be eroded if they felt that their GP had to ask for a second opinion before referring. They also questioned how guidelines would fit with doctors offering them a choice of treatment. This was an issue we explored in our guidelines review.

Appendix 4 Dissemination events and outputs

During our project we have engaged in a number of local and national dissemination activities, detailed in the table below.

Presentations

Date	Title and location	Content/focus	Audience
November 2012	<i>Demand Management: A Realist Synthesis</i> . Presented at the Leeds Institute for Health Sciences Seminar series, University of Leeds, Leeds, UK	We presented our initial theories about the causes of demand and capacity problems in the NHS for discussion and feedback	Academics, GPs
March 2013	Referral Management Stakeholder project meeting, University of Leeds, Leeds, UK	We presented our initial programme theories of how demand management interventions are intended to work to a group of local NHS stakeholders	GPs, commissioners, NHS managers, patients
June 2013	<i>Pills and Pachyderms: Dealing with Complexity in Research Synthesis</i> . Paper presented at the Health Services Network conference, Nottingham, UK	The presentation discussed the challenges of complexity in reviewing demand management interventions, provided a critique of previous reviews of demand management interventions and presented a systems theory of organisational change, discussed in <i>Chapter 3</i> of our report. This was then written up as a paper for <i>Social Science and Medicine</i> ⁹²	Academics, GPs, clinicians, NHS managers, commissioners, patients
November 2013	<i>Comprehending the Literature on Healthcare Guidelines: A Realist Approach</i> . Presented and discussed at a visit to RAND, Cambridge, UK	The presentation provided a critique of previous approaches to reviewing the guidelines literature, presented an alternative approach of reviewing the 'system strains' in implementing guidelines, provided two examples of findings from a realist synthesis of guidelines. A summary of the findings discussed in <i>Chapter 7</i> of this report	Academics, GPs
February 2014	<i>Comprehending the Literature on Healthcare Guidelines: A Realist Approach</i> . Presented at the Health Policy, Politics and Organisation Group (HiPPO) seminar series, Centre for Primary Care, University of Manchester, Manchester, UK	The paper provided a critique of previous approaches to reviewing the guidelines literature, presented an alternative approach of reviewing the 'system strains' in implementing guidelines, provided two examples of findings from a realist synthesis of guidelines. A summary of the findings discussed in <i>Chapter 7</i> of this report. This has now been written up and has been submitted to the <i>Journal of Health Organisation and Management</i> , currently under revision	Academics, GPs, NHS managers
March 2014	<i>Realist Synthesis of Demand Management for Planned Care</i> . Presented at the School of Health and Related Research (SchARR) seminar series	The presentation provided an overview of the process of realist synthesis and presented findings from realist synthesis of GPwSIs and RMCs, discussed in <i>Chapters 4 and 5</i> of the report	Academics

Date	Title and location	Content/focus	Audience
June 2014	<i>Changing Mandates for Referral Decision-Making: Substitution, Support, or Short-circuit?</i> Poster presented at the Health Services Research Network Conference, Nottingham, UK	The poster provided an overview of the findings from our synthesis of substitution at the organisational level (RMCs), individual level (GPwSIs) and operational level (GP direct access to tests)	Academics, GPs, clinicians, NHS managers, commissioners, patients
June 2014	<i>Optimising Referral Management: Tailoring Solutions to Local Circumstances.</i> Stakeholder dissemination event, Weetwood Hall, Leeds, UK	Presented an overview of the process of the review, its findings and practical recommendations for NHS decision-makers	GPs, clinicians, NHS managers, commissioners, patients
October 2014	<i>Dealing with complexity in whole system change.</i> A presentation and round table discussion in collaboration with Allan Best, Centre for the Advancement for Realist Evaluation and Synthesis Conference, Liverpool, UK	This paper provided two presentations exploring approaches to evaluating system change (realist synthesis and systems theory) and a round table discussion	Academics

A decorative graphic consisting of numerous thin, parallel green lines that curve from the left side of the page towards the right, creating a sense of movement and depth.

**EME
HS&DR
HTA
PGfAR
PHR**

Part of the NIHR Journals Library
www.journalslibrary.nihr.ac.uk

This report presents independent research funded by the National Institute for Health Research (NIHR). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health

Published by the NIHR Journals Library