

Do doctors under-provide, over-provide or do both? Exploring the quality of medical treatment in the Philippines

C.D. JAMES¹, K. HANSON², O. SOLON³, C.J.M. WHITTY² AND J. PEABODY⁴

¹World Health Organization, Western Pacific Regional Office (the majority of this work was undertaken whilst based at the London School of Hygiene and Tropical Medicine), Manila, Philippines, ²London School of Hygiene and Tropical Medicine, London, UK, ³University of the Philippines, Manila, Philippines, and ⁴University of California, San Francisco, CA, USA

Address reprint requests to: C.D. James, WHO Western Pacific Regional Office, PO Box 2932, 1000 Manila, Philippines. Tel: +63-2-5288001; Fax: +63-25-21-10-36; E-mail: chrisdames@fastem.com

Accepted for publication 17 May 2011

Abstract

Objective. To assess the quality of medical treatment by disaggregating quality into components that distinguish between insufficient and unnecessary care.

Design. Randomly selected doctors were asked how they would treat a sick child. Their responses were disaggregated into how much of an evidence-based essential treatment plan was completed and the number of additional non-essential treatments that were given. Key variables included the expected cost, the health consequences of insufficient and unnecessary care and comparisons between public and private physicians. Responses to 160 clinical performance vignettes (CPVs) were analysed.

Setting. Philippines.

Participants. One hundred and forty-three public and private physicians in the Philippines, collected in November 2003–December 2004 and September 2006–June 2007.

Interventions. CPVs administered to physicians.

Main outcome measures. Process quality measures (accounting for the possibility of both over-treatment and under-treatment).

Results. Based on CPVs, doctors gave both insufficient and unnecessary treatment to under-five children in 69% of cases. Doctors who provided the least sufficient care were also the most likely to give costly or harmful unnecessary care. Insufficient care typically had potentially worse health consequences for the patient than unnecessary care, though unnecessary care remains a concern because of overuse of antibiotics (47%) and unnecessary hospitalization (34%).

Conclusions. Quality of care is complex, but over- and under-treatment coexist and, in our analysis physicians that were more likely to under-treat a sick child were also those more likely to over-treat.

Keywords: measurement of quality, quality indicators, appropriateness, under-use and over-use, healthcare system, health policy

Introduction

This research investigates the relationship between the quality and quantity of healthcare, using data from the Philippines. The premise is that doctors can provide too little or too much care, both of which can negatively impact upon healthcare quality [1]. Research elsewhere has already highlighted that some doctors in the Philippines, in common with other

low and medium (and high) income countries, provide a low technical quality of healthcare (e.g. [2]).

Exploration of this quality-quantity relationship adds to such findings by evaluating whether under- or over-provision is likely to be of greater concern to policymakers. By distinguishing between when more healthcare is desirable as compared with more care being unnecessary, it also contributes to the efficiency literature.

To evaluate the extent to which doctors provide too little or too much care, we used clinical performance vignettes (CPVs), whereby physicians were asked how they would care for a range of paediatric cases [3]. These CPV responses were disaggregated into components that distinguish insufficient and unnecessary care. Accordingly, the objective of this research is to provide insights into the different components of sub-optimal quality of care.

Although our study refers only to the Philippines, both insufficient care and unnecessary care are relevant quality concerns worldwide. It is important to note, though, that by assessing how close doctors are to providing a set of actions needed to improve a sick child's health, these CPVs provide a quantitative measure of a physician's *technical*, as opposed to interpersonal, quality of care [4, 5].

Methods

Using CPVs to assess the quality-quantity relationship

In the CPVs, physicians were asked how they would care for a range of paediatric cases. These were administered as written interviews. Their responses were then assessed against a predefined essential treatment plan, which was based on international evidence-based standards that were then reviewed by national (Philippine) experts to ensure they were nationally relevant and reflected best local practice. Further, responses were assessed in terms of which aspects of the predefined treatment plan a doctor failed to give (insufficient care), but also any additional treatments offered that are not part of the predefined treatment plan (unnecessary care). Such disaggregation of CPVs allows the quality-quantity relationship to be explored.

The CPVs used were made up of five sequential domains of care, evaluating a doctor's ability in terms of five technical areas. These relate to: (i) asking questions about the patient's symptoms and medical history; (ii) performing physical examinations; (iii) ordering laboratory tests; (iv) arriving at a diagnosis for the patient; and (v) compiling a recommended treatment plan for the patient.

In this study, analysis focuses on the last domain, namely the doctor's recommended treatment plan. This is because the expected health and cost consequences of over-treatment are more likely to be substantial than over-provision in the other four care domains. For example, unnecessarily giving a child certain medications can be expensive and harmful. In contrast, asking too many questions about a patient's symptoms and medical history, or performing too many physical examinations is only marginally more expensive (by increasing consultation time), and rarely harmful. Further, whilst ordering too many laboratory tests can be as expensive as over-treatment, it is rarely harmful to a patient.

Evaluating the extent and consequences of inappropriate care

For each CPV, essential actions making up an essential treatment plan were defined prior to CPV administration to

physicians (with non-essential actions being those actions not defined). Each doctor's response regarding how they would treat the sick child was disaggregated into how much of the essential treatment plan was completed and the number of additional non-essential treatments given. These measure the extent of *insufficient care* and *unnecessary care* respectively. The expected health and cost consequences of both insufficient care and unnecessary care were evaluated after the sample was collected.

Expected cost consequences. Non-essential treatments and each aspect of the essential treatment plan were classified according to their likely (societal) cost implications. Four broad treatment categories were distinguished. These were hospitalization; medications (drugs and therapies); monitoring of condition by doctor; and advice on homecare and supplements.

Given data limitations, only approximate cost estimates were possible. Treatment categories' expected costs were defined in terms of being low, medium or high using categorizations based on Quality Improvement Demonstration Study (QIDS) data and research by the authors (see below). Note that cost categories assume that changing provider behaviour is inexpensive. Although a detailed analysis of the cost of behavioural change is beyond the scope of this study, other QIDS-related research has shown that at low levels of quality (as measured by CPVs), improvements in quality are linked to lower hospital charges, hypothesized to reflect less provision of unnecessary care [6].

Hospitalization of a patient is relatively costly compared with the other treatment aspects. Therefore it was classified as high cost. For example, facility survey data from 30 hospitals indicate that the total healthcare charges per admission inside a public district hospital were, on average, approximately 1900 Philippine pesos (PHP) (\$39 USD) for pneumonia patients and 1400 PHP (\$29 USD) for diarrhoea patients, in 2006/2007 (unpublished data, QIDS project). Of note is that these are lower-end estimates, as they exclude associated indirect costs such as travel and food costs.

Medications were classified as medium cost. Pharmacy exit data from 29 private pharmacies in the Philippines estimate an average pharmaceutical expenditure of 260 PHP (\$6 USD) in 2007 for customers with a prescription [7]. Monitoring of a patient's condition by the doctor was also classified as medium cost. In these CPVs, monitoring included checking for, and reassessing, vomiting, urine output and heart rate. The main cost was that associated with a doctor's time. Physician survey data indicate an average reported public physician salary (all sources) to be about 24 000 PHP (\$495) per month in 2006/2007 (QIDS dataset). Assuming a doctor works 160 h per month, and that monitoring of a patient's condition takes 1 h, the time cost is approximately 150 PHP (\$3 USD). Other health facility costs included the use of medical instruments and time costs of nurse and other support staff, but there were no data available on these costs.

The treatment category advice on homecare/supplements was classified as low cost. Advice on homecare only adds a few minutes to consultation time, and supplements can be

purchased cheaply from pharmacies (and are sometimes free from primary health centres).

Expected health consequences. To determine the health consequences of a doctor's recommended treatment plan, three physicians were asked to evaluate independently the health consequences of: (i) *not* undertaking different aspects of essential treatment plans and (ii) each non-essential treatment given. The treatments were classified as definitely harmful, probably harmful, possibly harmful, health neutral, possibly beneficial, probably beneficial and definitely beneficial. Categorizations were based on expert clinical opinion. For treatments that were classified as harmful, these were further separated into those that are expected to result in hospitalization or only had moderate/mild adverse effects.

The physicians evaluating expected health consequences were all paediatricians and scorers for the CPVs in the QIDS study. After their independent evaluations of the expected health consequences of insufficient and unnecessary care, results were returned to the physicians with disparities in their scores highlighted. The three physicians then met to discuss these disparities until consensus was reached.

Patterns of inappropriate care

Having evaluated the extent and consequences of insufficient and unnecessary care, the relationship between the two was explored. This provides added insight into the relationship between quantity and quality of care by illustrating different patterns of inappropriate care.

That is, whilst it is evident that reducing the extent of insufficient and unnecessary care improves quality of care—in terms of moving a doctor's healthcare provision closer to best practice—analysing patterns of inappropriate care determines whether doctors more often provide insufficient or unnecessary care.

To better understand the patterns of inappropriate care between public and private practitioners, these were compared in terms of the extent of insufficient care, and also the probability of providing unnecessary care with harmful health effects.

Data

The quality improvement demonstration study. We used data on CPVs from the Philippine Child Health Experiment, known locally as QIDS. The QIDS was a large randomized policy experiment, conducted to evaluate the impact of large-scale government policy interventions on both the delivery of care and long-term health status in children. The QIDS sampling frame consists of 30 districts in the Visayan island group and the northern tip of Mindanao, with data collected in two time periods (see [8] for a comprehensive discussion of the QIDS methodology).

Data collected included CPVs (alongside household, hospital-based patient exit and facility surveys). CPVs were administered to public and private physicians; the following sampling information comes from ref. [9]. For inclusion in the QIDS study sample frame, physicians needed to be

graduates of an accredited medical school, and children had to account for at least a fifth of their clinical practice. Further, public physicians had to work full time in public hospitals; private physicians had to live in the same district as the public hospital and serve the same geographic population. From the sampling frame of doctors who met these conditions, three randomly selected public physicians per public hospital, and two randomly selected private doctors were interviewed in each of the 30 QIDS study districts. These interviews took place between November 2003 and December 2004 and between September 2006 and June 2007.

CPVs selected in this research. In the QIDS study, 15 CPVs were administered to each of these doctors, five each relating to paediatric pneumonia, diarrhoea and dermatological-related cases. These were for common conditions, so clinicians can reasonably be expected to deal with it in practice as well as theory on a regular basis, and hence answers are more likely to be based on real experience than remembered book-learning.

In this paper, we focused on diarrhoea and pneumonia cases, and selected the two simplest cases of each. For diarrhoea, the cases were (i) viral (rotavirus) gastroenteritis with mild dehydration and (ii) acute bacterial gastroenteritis with slight haematochezia. For pneumonia, the two cases were community-acquired pneumonia and acute viral bronchitis. These four vignettes were chosen because over-treatment as well as under-treatment may occur, particularly in relation to unnecessary use of antibiotics and unnecessary hospitalization. Boxes A1–A5 in an Appendix describe in detail each of the four vignettes used.

Sample and pooling validity issues. The final sample was composed of 160 CPVs from 143 different doctors who answered one or more of the four CPVs. This represents all vignette responses for the four CPVs. Data were pooled from the two time periods to increase statistical power, after assessing whether such pooling was valid. Statistical analysis found no significant difference in the extent of insufficient care between the two time periods; and whilst unnecessary care was significantly higher in the second time period than in the baseline, it was still statistically greater than zero in the baseline (based on standard statistical *t*-tests comparing the two time periods). Thus, pooling across the two time periods was deemed valid. Note that this result is in contrast to other QIDS studies evaluating all 15 CPVs across all five dimensions of care [10].

Results

Descriptive statistics

We analysed the complete sample composed of 160 CPVs, of which public doctors answered 60% and private doctors 40%, respectively. The average (mean) physician age was 42 years; female physicians made up 64% of the sample. Table 1 provides further details, disaggregated by the four CPVs used in this study.

Table 1 Sample characteristics, full sample and by CPV

	Full sample	Diarrhoea #1	Diarrhoea #2	Pneumonia #1	Pneumonia #2
Time period					
CPV answered in baseline	112 (70)	27	28	27	30
CPV answered in round 2	48 (30)	13	14	12	9
Doctor's place of work					
Doctors working in private clinic	64 (40)	6	9	24	25
Doctors working in public hospital	96 (60)	34	33	15	14
Gender and age					
CPV answered by female doctor	103 (64)	24	29	25	25
CPV answered by male doctor	57 (36)	16	13	14	14
Average (mean) age of doctor	42	40	42	43	43
Full sample	160 (100)	40	42	39	39

Values in parenthesis are percentage values. *Note:* Two doctors moved from the private to public sector between the baseline and the second round.

Insufficient care

For the majority of doctor–patient interactions, less than half of the recommended essential treatment plan was given. In 30 CPVs (19% of the time), doctors gave <25% of the essential treatment plan; in 86 CPVs (54%), doctors gave <50%. In only eight doctor–patient interactions (5%) was fully sufficient treatment given. The average (mean) percentage of the essential treatment plan given was 50%, the median 44%, with a standard deviation of 24% and a 95% confidence interval of 45.7–53.3%.

Disaggregated by category, results show that the majority of treatment plans were characterized by insufficient advice, monitoring and medication: namely in 111 CPVs (69% of the time) doctors did not give sufficient advice, in 118 cases (74%) doctors did not adequately monitor the patient, and in 116 instances (73%) doctors did not give sufficient medication. In a lower proportion of CPVs (17%), doctors failed to hospitalize a patient when hospitalization was required. See Table A1 in the Appendix for a more detailed presentation of the types of insufficient care not given, for one of the diarrhoea vignettes.

Not giving any individual part of the essential treatment plan always had potentially negative health consequences. Often, this was expected to have serious health consequences. For instance, in 104 of the doctor–patient interactions (65% of the time) part of an essential treatment plan was not given that would cause a ‘severe adverse event’.

After controlling for diarrhoea vs. pneumonia cases, there was no statistical difference in the sufficiency of care given by public vs. private doctors other than for diarrhoea CPV #1, where public doctors were closer to meeting the complete essential treatment plan (38 vs. 24%, P -value = 0.04).

Unnecessary care

Approximately three-quarters of the sample (74%) gave non-essential care. This ranged from 1 to 5 additional treatments,

with a mean of 1.39, median of 1, standard deviation of 1.24 and a 95% confidence interval of 1.20–1.59.

Disaggregating non-essential treatments by category provides some further insights into the expected cost implications of unnecessary care. Of the 118 cases where hospitalization was *not* needed, in 40 CPVs (34%) doctors recommended hospitalization. Doctors also frequently recommended non-essential drugs, particularly antibiotics (47% of the time).

In terms of expected health consequences, for 23 CPVs (14%) doctors gave non-essential treatments that were potentially harmful to patients (e.g. metoclopramide for a paediatric diarrhoea case with viral (rotavirus) gastroenteritis and mild dehydration), although none of these had ‘definitely harmful’ health consequences. Fourteen doctor–patient interactions (9%) were characterized by doctors giving non-essential treatments that had potentially positive health effects (e.g. complete immunization). More often, for 108 CPVs, doctors gave non-essential treatments (69% of the sample) that were health neutral (e.g. decongestants for paediatric pneumonia case with acute viral bronchitis). However, when a sensitivity analysis was conducted—whereby unnecessary antibiotic use and unnecessary hospitalizations were re-categorized from health neutral to potentially harmful—the number of doctors giving potentially harmful treatments increased markedly—to 97 of the CPVs (61% of the sample). See Table A1 in the Appendix for a more detailed presentation of the types of unnecessary care not given, for one of the diarrhoea vignettes.

There were no systematic differences between private and public doctors in terms of unnecessary hospitalization or harmful non-essential treatments. However, more systematic differences between public and private doctors emerge in unnecessary antibiotic use. Public doctors were more likely than private doctors to recommend unnecessary antibiotics to patients for diarrhoea CPV #2 (64 vs. 33%, P -value = .05) and pneumonia CPV #1 (73 vs. 33%, P -value = .01).

The relationship between insufficient and unnecessary care

We found that most doctors gave insufficient and unnecessary care *simultaneously*. That is, doctors typically replaced needed aspects of the essential treatment plan with non-essential treatments. For 111 CPVs (69% of the cases), doctors gave both insufficient and unnecessary treatment. This compares with 41 (26%) cases of inappropriate care in which doctors gave insufficient treatment only (i.e. did not give unnecessary treatments) and 8 (5%) cases where doctors gave unnecessary treatment only (i.e. did not give insufficient treatment) (Fig. 1).

When insufficient care and unnecessary care were disaggregated, distinct distributions of practice emerged. Doctors who gave less sufficient care were typically more likely to hospitalize patients unnecessarily and to give harmful non-essential treatments. More precisely, the probability of unnecessarily hospitalizing a patient was higher for doctors who gave less than half of the essential treatment plan as compared with those giving more than half (44 vs. 26%, P -value = .04). Similarly, unnecessary hospitalization was more likely for doctors who gave insufficient as compared with sufficient medication (50 vs. 2%, P -value < 0.01). In contrast, there was no significant relationship between the probability of unnecessary hospitalization and the sufficiency of a doctor's advice or monitoring of the patient (Table 2).

The probability of giving harmful non-essential treatments was higher for doctors who gave less than half of the essential treatment plan as compared with those giving more than half (21 vs. 7%, P -value = 0.01). Similarly, the likelihood of giving harmful non-essential treatments was higher for doctors who gave insufficient as compared with sufficient advice (19 vs. 4%, P -value = 0.01), and insufficient vs. sufficient medication (20 vs. 0%, P -value < 0.01). In contrast, there was no significant relationship between the probability of giving harmful non-essential treatments and the sufficiency of a doctor's monitoring of the patient (Table 3).

There were no clear correlations between sufficiency of care and the probability of doctors giving health neutral non-essential treatments. Conversely, the probability of giving beneficial non-essential treatments was positively related to

sufficiency of care. Doctors who gave less than half of the essential treatment plan were less likely to give beneficial non-essential treatments as compared with those giving more than half (5 vs. 14%, P -value = 0.04). Similarly, the likelihood of giving beneficial non-essential treatments was lower for doctors who gave insufficient as compared with sufficient advice (6 vs. 14%, P -value = 0.09), and insufficient vs. sufficient medication (5 vs. 18%, P -value = 0.01).

Discussion

This study characterizes optimal quality healthcare in terms of both the sufficiency and necessity of care, using data from the Philippines. It found that 69% of the time doctors gave *both* insufficient and unnecessary treatment. Further, doctors who provided the least sufficient care were the most likely to give costly and harmful unnecessary care. This was most marked for medication use and unnecessary hospitalization, as opposed to patient advice and monitoring where. For example, doctors who gave insufficient medication were significantly more likely than doctors giving sufficient medication to recommend both unnecessary hospitalization and harmful non-essential treatments.

Results also showed that moving from insufficient to sufficient care can bring large health gains for the patient without much additional expense to the patient. However, as discussed earlier, the total cost of this move to sufficient care will depend on how costly it proves to change provider behaviour. This includes the costs of training and other quality improving strategies which have been shown to be effective. For example, a recent review of studies found that supervision and audit with feedback were two of the more effective ways to improve health worker performance, and that dissemination of treatment guidelines, self-assessment, and training were much more effective when combined with appropriate supervision [11]. However, supervision costs have been considered expensive from studies on supervision that have reported its costs [12].

In contrast, reducing unnecessary care can lead to important cost savings, but it does not always offer substantial health gains. That is, unnecessary care often reflected

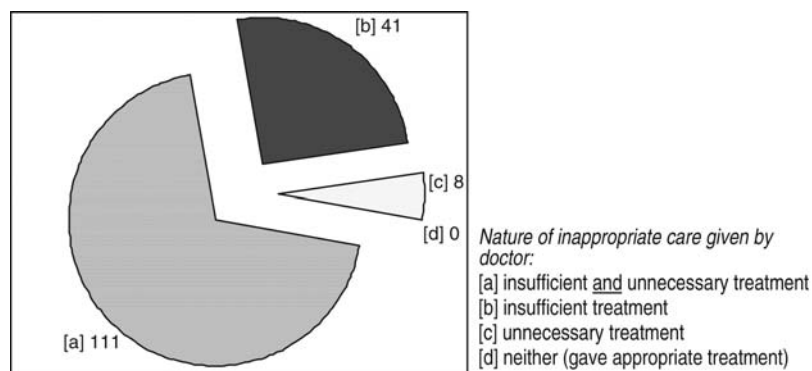


Figure 1 Nature of inappropriate care: too little, too much or both?

Table 2 Sufficiency of care and probability of unnecessary hospitalization

Sufficiency of essential treatment plan given (general and by category)	Was patient hospitalized unnecessarily?			
	Yes (40)	No (78)	Total (118)	Proportion = yes (%)
General				
<50% of essential treatment plan	23	29	52	44
≥50% of essential treatment plan	17	49	66	26
	$\chi^2 = 4.43$; P -value = 0.035			
Advice				
Insufficient	26	45	71	37
Sufficient	14	33	47	30
	$\chi^2 = 0.59$; P -value = 0.443			
Monitoring				
Insufficient	24	56	80	30
Sufficient	16	22	38	42
	$\chi^2 = 1.68$; P -value = 0.194			
Medication				
Insufficient	38	38	76	50
Sufficient	2	40	42	2
	Fisher's exact $P < 0.001$ ($\chi^2 = 24.71$; P -value < 0.001)			

Note: Fisher's exact test, rather than Pearson's chi-squared test, used when sample size of any cell in 2×2 table is < 10 (i.e. for medication 2×2 table).

Table 3 Sufficiency of care and probability of giving harmful non-essential treatment

Sufficiency of essential treatment plan given (general and by category)	Doctor gave harmful non-essential treatment?			
	Yes (23)	No (137)	Total (160)	Proportion = yes (%)
General				
<50% of essential treatment plan	18	68	86	21
≥50% of essential treatment plan	5	69	74	7
	Fisher's exact $P = 0.009$ ($\chi^2 = 6.49$; P -value = 0.011)			
Advice				
insufficient	21	90	111	19
Sufficient	2	47	49	4
	Fisher's exact $P = 0.009$ ($\chi^2 = 6.08$; P -value = 0.014)			
Monitoring				
Insufficient	18	100	118	15
Sufficient	5	37	42	12
	Fisher's exact $P = 0.403$ ($\chi^2 = 0.28$; P -value = 0.595)			
Medication				
Insufficient	23	93	116	20
Sufficient	0	44	44	0
	Fisher's exact $P < 0.001$ ($\chi^2 = 10.19$; P -value < 0.001)			

Note: Fisher's exact test, rather than chi-squared test, used when sample size of any cell in 2×2 table is < 10 .

'flat-of-the-curve' medicine and healthcare that are not harmful but equally provide no incremental benefit to the patient [13]. The cost of such unnecessary care can be

substantial: for instance, the World Health Report 2010 estimated that at least 20–40% of total health spending is ineffectively spent. Moreover, it identified overuse of medicines,

healthcare products and services, and medical errors as some of the leading sources of inefficiency [14].

Thus insufficient care is more likely to have worse health consequences for the patient than unnecessary care. An important caveat is that unnecessary care often took the form of doctors recommending antibiotics (i.e. 47% of the time). Although judged by physicians to be mostly health neutral on a case-by-case basis, overuse of antibiotics is a public health concern because it can expose the population as a whole to higher antibiotic resistance, and can be costly [15, 16]. Further, patients were unnecessarily hospitalized 34% of the time. This was also judged to be health neutral, but unnecessary hospitalization increases the risk of individuals acquiring nosocomial infections or having further unnecessary tests and procedures. Finally, we did not find substantive differences in practices between public vs. private practitioners.

The study has some potential limitations. First, CPVs are based on hypothetical behaviour, with some authors arguing that they only measure a doctor's health knowledge [17]. However, previous research has shown CPVs to be a valid measure of process quality in healthcare, outperforming chart abstraction when compared with the gold standard of the 'standardized patient' [3]. This study was repeated in a larger number of settings with a broader range of conditions in a second validation study [18]. The results of the second trial confirmed that performance measured by CPVs came close to (and did not exceed) actual practice. This validation work involved over 200 doctors and 16 different conditions and over 1500 patients. Still, it is recognized that no studies have to date validated CPVs against standardized patients in lower income country settings. Second, CPVs are limited to evaluating the technical quality of healthcare, with no analysis of interpersonal quality. Vignettes also analyse only a selection of conditions. Still, paediatric pneumonia and diarrhoea, analysed here, are common conditions in low- and middle-income countries, the leading causes of death in this population and thus represent a high burden of disease where appropriate clinical care improves a patient's health outcome. These clinical conditions are thus likely to be good 'tracer' conditions in evaluating the broader health system [19, 20], although they are less representative of adult care, particularly chronic conditions.

The design of this analysis also had its own specific limitations. First, only one dimension of care—a doctor's recommended treatment plan—was analysed. Even though inclusion of diagnostic testing would look at additional costs, a doctor's treatment plan is the most relevant dimension for analysing costly and potentially harmful consequences of overprovision, and consequently the relationship between quality and quantity of care. Second, measures of the overall extent of both insufficient and unnecessary care were simple aggregates. Other CPV studies, addressing different issues, have used expert panels [3] or item response theory [21] to weight doctor's responses. Although in this research there was no weighting of individual items as contributors to a single CPV score, the expected health and cost consequences of these individual items were evaluated. More generally, the

study was set in the Philippines. Whilst we believe our general findings about the nature of the quality–quantity relationship are likely to be applicable to other country contexts, we realize that this merits further attention. For instance, studies have consistently shown that provider payment mechanisms influence the quantity of care a doctor provides [22]. Consequently, variation in the exact mix of insufficient and unnecessary care may vary across countries because of different mixes in provider payments.

Notwithstanding these limitations, this research adds to the literature by simultaneously investigating insufficient care and unnecessary care. Existing research has typically evaluated the extent of insufficient care *or* unnecessary care, but not both at the same time. For instance, in other studies that have used CPVs, a doctor's technical quality of care is assessed by analysing whether s/he has provided a comprehensive set of actions needed to improve a patient's health (e.g. [15, 23]), and if not, which actions they did not provide. But no distinction is made in these studies between a doctor failing to recommend a needed treatment (or other action), and a doctor recommending an unnecessary treatment.

Indeed, most studies measuring the technical quality of healthcare can be understood as focusing on the extent of insufficient care, with no direct analysis of unnecessary care. Structural quality measures can (at best) assess whether doctors are likely to be constrained in their attempts to provide comprehensive care (e.g. [24]). Related QIDS research analysed the impact of structural factors on quality of care in the same study area as our paper. They found that staffing levels, medical supplies and other structural factors had little impact on quality of care [25]. However, structural quality measures provide no information on the potential for over-provision. Studies using other process quality measures—such as chart abstraction [26], direct observation [17] and standardized patient approaches [27]—can also be interpreted in the same way. That is, they compare a doctor's healthcare provision against a checklist of required actions, with the focus being on which aspects of this checklist the doctor failed to complete. Outcome measures could ultimately provide the most accurate measure of healthcare quality. However, they cannot easily separate out the impact (positive or negative) of individual aspects of a doctor's treatment plan on a patient's health.

In contrast, the literatures on health provider efficiency and supplier-induced demand assess unnecessary care but not insufficient care. The sole focus of the supplier-induced demand literature is on whether, and if so how, doctors can influence patients to utilize more healthcare than is clinically necessary [28]. The efficiency literature has shown that hospitals (as a whole) have some costs that are due to waste or poor decision-making. But most of these efficiency studies implicitly assume adequate quality [29]. Some efficiency studies do account for quality, and consequently the possibility of insufficient care [30–32]. Nevertheless, these studies concentrate on identifying when quantity of care can be reduced without negatively impacting upon healthcare quality, rather than on quality directly.

To conclude, this research shows that the relationship between the quality and quantity of care cannot be collapsed to a question of whether doctors provide too little or too much care, since doctors typically do both simultaneously. One important solution is greater use of and adherence to standardized, evidence-based guidelines. This has been demonstrated in a range of settings to improve the quality of service provision [33–35]. Further, insufficient care was shown to be more likely to have adverse health effects than unnecessary care. But unnecessary care remains a concern since it can be costly for the patient and society, and often involves unnecessary use of antibiotics. Moreover, doctors that provide the least sufficient care are the most likely to give harmful and costly unnecessary care, and thus both over- and under-treatment need to be tackled together.

Acknowledgements

The paper utilized data from the Philippine Child Health Experiment, which provided technical and logistical support to the lead author. The authors would like to thank Palmera I. Baltazar, Evelyn Duron and Aura Mar G. Mathew. We would also like to thank the editor and two anonymous reviewers of the manuscript for their insightful comments.

Funding

This work was supported by a joint Economic and Social Research Council/Medical Research Council scholarship, which funded the lead author's PhD. This work was also indirectly supported by the U.S. National Institute of Child Health and Human Development, which funded the Philippines Child Health Experiment.

Appendix

Boxes A1–A4: Description of the four vignettes used in the study

A1: Summary description of Diarrhoea Vignette #1: viral (rotavirus) gastroenteritis with mild dehydration

- A mother comes to the clinic with her daughter, an 8-month old baby. She states that her daughter has had diarrhoea and is vomiting.
- *Symptoms/medical history:* The diarrhoea started 2 days ago at the same time as the vomiting. The baby has had very loose, watery stools, without blood or mucus, in her diaper about six to seven times throughout the day and night. She has had a low-grade fever and has not eaten very much. She vomited three or four times yesterday but only twice today. The child has been almost weaned

from breast milk, breastfeeding twice a day for the past month. Yesterday, she drank a little water from a cup but would not breast-feed. Today she has breast-fed once and has been drinking some diluted mango juice. She has urinated once today, about 6 h ago. The little girl's older sister, aged 2, had a similar problem about 1 week ago but the symptoms lasted only a day. She has no prior history of similar episodes, any known drug allergies or other medical problems. Her mother reports that the delivery was normal.

- *Physical examinations:* The child is alert and interactive, but tearful and irritable. The pulse is strong at 170 beats per min. Temperature is 39°C. Eyes are sunken. Mucus membranes are somewhat dry. The skin pinch goes back in 1 s. The fontanelle is not depressed. The head is normal without nuchal rigidity, the abdomen is mildly tender but there is no guarding, rigidity or rebound. Bowel sounds are normal. Her capillary refill time is ~3 s. Faeces in the diaper are negative for blood. Weight is 7.5 kg and the length is 68 cm.
- *Laboratory tests:* All laboratory tests are normal.

A2: Summary description of Diarrhoea Vignette #2: acute bacterial gastroenteritis with slight haematochezia

- A 3-year-old boy is brought to the clinic by his mother. She states that her son has diarrhoea and vomiting.
- *Symptoms/medical history:* The diarrhoea started 2 days ago at the same time as the vomiting. The stool was described as loose and watery without blood or mucus. The diarrhoea episodes occur seven to eight times throughout the day and night. He vomited three to four times yesterday and today; there is no blood in the vomit. He had a low-grade fever and has not eaten well. His mother offered some water but he refuses and drinks only a small sip. He has not been given any medication nor has he ingested any new foods or foods that might be contaminated. His mother does not know when he last urinated. No one else has been ill in the (his) family. He has no previous history of similar symptoms, any known drug allergies or other medical problems. His mother reports the delivery was normal.
- *Physical examinations:* The boy appears calm and he is lethargic. He weighs 10 kg and his height is 96 cm. He is afebrile, the pulse rate is 150 beats/min, his blood pressure is 70/35 and his respiration 55. The mucus membranes are dry and the eyes are sunken. The skin pinch returns to normal in 3 s. The chest examination is normal. The abdomen is soft with no guarding, rigidity or rebound and the bowel sounds appear increased by otherwise normal. Stool/rectal examination is negative for blood.

- *Laboratory tests:* The only laboratory result that is available is CBC which shows haemoglobin of 12.5 and WBC of 6. Fecalysis reveals no RBCs or excessive WBC in the stool. All other laboratory tests are pending.

A3: Summary description of Pneumonia Vignette #1: community-acquired pneumonia

- A mother brings her 6-month old baby to the clinic. She states that her daughter has had cough and fever.
- *Symptoms/medical history:* The baby's condition started 1 week ago with cough and colds with whitish nasal discharge. She later developed moderate- to high-grade fever temporarily relieved by paracetamol. She is active, cries easily but is consolable but irritable and the mother reports that she continues to feed. The mother does not report any difficulty breathing or any episodes of cyanosis, convulsion or any rashes. The baby was breastfed for the first 2 months and formula fed thereafter. No solid food has been introduced yet. The baby has received the following immunization through the local health centre: BCG, DPT (three doses) and OPV (three doses). The mother denies any history of asthma in the family or the patient having been with a respiratory disease in the past.
- *Physical examinations:* The baby weighs 8 kg and is 68 cm long. She is awake but crying. She is febrile with temperature of 38.5°C. Her heart rate is 140 beats/min; respiratory rate is increased at 42/min but she has no circumoral cyanosis. There is no tonsillopharyngeal congestion. No stridor is noted. There are supraclavicular and intercostals retractions but there does not appear to be any lower chest indrawing. There are crepitant rales on all lung fields, bilateral. There is no wheezing and there are no cardiac thrills or murmurs. The nailbeds are pinkish. There is no evidence of dehydration.
- *Laboratory tests:* The CBC showed an elevated white blood count of 14 000 with predominance of polymorphonuclear leukocytes. No chest X-ray is available.

A4: Summary description of Pneumonia Vignette #2: acute viral bronchitis

- A mother brings her 3-year-old boy to the clinic. She states that her son has fever and cough.
- *Symptoms/medical history:* The mother states that her son was previously well until 1 week ago when he started to have fever and cough. The fever is low to moderate grade and occurs intermittently. She also reports that he had some sneezing and a runny nose. It is temporarily relieved by paracetamol. The cough is noted to be getting somewhat worse since its onset and it is productive of moderate amounts of clear-white phlegm. He did not have any episode of cyanosis or difficulty breathing however. His appetite is fine and the child is not extremely

thirsty. His immunizations are updated. His 8-month-old sister also had similar symptoms and has just been discharged from the hospital 1 week ago. He has no other medical problems, and has not been sick like this in the past. The child has no known allergies and is on no other medications.

- *Physical examinations:* The boy weighs 13.5 kg and is 90 cm tall. He is awake and active. He is febrile with temperature of 38.7°C. The cardiac rate is 120 beats/min and respiratory rate is 33/min. There is no circumoral cyanosis, the nailbeds are pink and no tonsillopharyngeal congestion. There is no intercostal retraction or lower chest wall indrawing. There are no rales or wheezes but there is occasional ronchi. There are no thrills or murmurs.
- *Laboratory tests:* The CBC showed white blood cell count of 11 000. The haemoglobin is 13.2 g %. The mother refused to have the chest X-ray done.

Table A1 Example of classification of insufficient and unnecessary care: Diarrhoea #1: viral (rotavirus) gastroenteritis with mild dehydration

	Expected health consequence ^a	Expected severity
Essential treatments NOT offered		
Oral rehydration salts (in correct dosage, frequency)	Definitely harmful (-3)	Hospitalization (A)
Monitor for vomiting, urination, normalization of heart rate	Definitely harmful (-3)	Hospitalization (A)
Reassess child after 3-4 h and treat accordingly (details further specified)	Definitely harmful (-3)	Hospitalization (A)
Antipyretic for fever (paracetamol)	Probably harmful (-2)	Moderate adverse effect (B)
Once qualified for discharge, advice on when mother should return (details further specified)	Definitely harmful (3)	Hospitalization (A)
Advise mother on continued homecare	Probably harmful (-2)	Moderate adverse effect (B)
Recommendations on breastfeeding	Probably harmful (-2)	Moderate adverse effect (B)
Recommendations on supplements	Possibly harmful (-1)	Moderate adverse effect (B)
Additional treatments offered		
Ampicillin and gentamycin	Probably harmful (-2)	Moderate adverse effect (B)
Cotrimoxazole	Possibly harmful (-1)	Moderate adverse effect (B)

(continued)

Table A1 Continued

	Expected health consequence ^a	Expected severity
Unspecified antibiotics: assumed to be amoxicillin	Health neutral (0)	–
Hospitalization	Health neutral (0)	–
Metoclopramide	Probably harmful (–2)	Hospitalization (A)
Dilute milk formula or change to non-lactose preparation	Possibly harmful (–1)	Moderate adverse effect (B)

^aAnswers on expected health consequence and severity are final answers following discussion between the three Philippine paediatricians on any score disparities.

References

- Institute of Medicine. *Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington, DC: National Academy Press, 2001.
- Peabody JW, Liu A. A cross-national comparison of the quality of clinical care using vignettes. *Health Policy Plan* 2007;**22**:294–302.
- Peabody JW, Luck J, Glassman P *et al*. Comparison of vignettes, standardized patients, and chart abstraction: a prospective validation study of 3 methods for measuring quality. *JAMA* 2000;**283**:1715–22.
- Brook RH, McGlynn EA, Shekelle PG. Defining and measuring quality of care: a perspective from US researchers. *Int J Qual Health Care* 2000;**12**:281–95.
- Donabedian A. The quality of care. How can it be assessed? *JAMA* 1988;**260**:1743–8.
- Peabody JW, Florentino J, Shimkhada R *et al*. Quality variation and its impact on costs and satisfaction: evidence from the QIDS study. *Med Care* 2010;**48**:25–30.
- James CD, Peabody J, Solon O *et al*. An unhealthy public-private tension: pharmacy ownership, prescribing, and spending in the Philippines. *Health Aff* 2009;**28**:1022–33.
- Shimkhada R, Peabody JW, Quimbo SA *et al*. The quality improvement demonstration study: an example of evidence-based policy-making in practice. *Health Res Policy Syst* 2008;**6**:5.
- Quimbo SA, Peabody JW, Shimkhada R *et al*. Should we have confidence if a physician is accredited? A study of the relative impacts of accreditation and insurance payments on quality of care in the Philippines. *Soc Sci Med* 2008;**67**:505–10.
- Peabody J, Shimkhada R, Quimbo S *et al*. Financial incentives and measurement improved physicians' quality of care in the Philippines. *Health Aff* 2010;**30**:773–81.
- Rowe AK, de Savigny D, Lanata CF *et al*. How can we achieve and maintain high-quality performance of health workers in low-resource settings? *Lancet* 2005;**366**:1026–35.
- Bosch-Capblanch X, Garner P. Primary health care supervision in developing countries. *Trop Med Int Health* 2008;**13**:369–83.
- Fuchs VR. More variation in use of care, more flat-of-the-curve medicine. *Health Aff* 2004; **Oct 2004**:VAR104–7.
- World Health Organization. *Health Systems Financing: The Path to Universal Coverage*. Geneva: World Health Organization, 2010.
- Coast J, Smith RD. Antimicrobial resistance: cost and containment. *Expert Rev Anti Infect Ther* 2003;**1**:241–51.
- Kunin CM. Resistance to antimicrobial drugs—a worldwide calamity. *Ann Intern Med* 1993;**118**:557–61.
- Das J, Hammer J, Leonard K. The quality of medical advice in low-income countries. *J Econ Perspect* 2008;**22**:93–114.
- Peabody JW, Luck J, Glassman P *et al*. Measuring the quality of physician practice by using clinical vignettes: a prospective validation study. *Ann Intern Med* 2004; **141**:771–80.
- Kessner DM, Kalk CE, Singer J. Assessing health quality—the case for tracers. *N Engl J Med* 1973;**288**:189–94.
- Neuhauser D. Assessing health quality: the case for tracers. *J Health Serv Res Policy* 2004;**9**:246–7.
- Das J, Hammer J. Which doctor? Combining vignettes and item response to measure clinical competence. *J Dev Econ* 2005;**78**:348–83.
- Gosden T, Forland F, Kristiansen IS *et al*. Impact of payment method on behaviour of primary care physicians: a systematic review. *J Health Serv Res Policy* 2001;**6**:44–55.
- Das J, Hammer J. Money for nothing: the dire straits of medical practice in Delhi, India. WB Policy Research Working Paper 3669, 2005.
- Barber SL, Gertler PJ. *Child Health and the Quality of Medical Care*. Berkeley: University of California, 2002.
- Solon O, Woo K, Quimbo SA *et al*. A novel method for measuring health care system performance: experience from QIDS in the Philippines. *Health Policy Plann* 2009;**24**:167–74.
- Iezzoni LI. Assessing quality using administrative data. *Ann Intern Med* 1997;**127**(8 Pt 2):666–74.
- Beullens J, Rethans JJ, Goedhuys J *et al*. The use of standardized patients in research in general practice. *Fam Pract* 1997;**14**:58–62.
- McGuire TG. Physician agency. In: Culyer AJ, Newhouse JP (eds). *Handbook of Health Economics*. Amsterdam: Elsevier, 2000, 461–536.
- Hussey PS, de Vries H, Romley J *et al*. A systematic review of health care efficiency measures. *Health Serv Res* 2009;**44**: 784–805.
- McKay NL, Deily ME. Cost inefficiency and hospital health outcomes. *Health Econ* 2008;**17**:833–48.
- Mutter RL, Rosko MD, Wong HS. Measuring hospital inefficiency: the effects of controlling for quality and patient burden of illness. *Health Serv Res* 2008;**43**: 1992–2013.

32. Zuckerman S, Hadley J, Iezzoni L. Measuring hospital efficiency with frontier cost functions. *J Health Econ* 1994;**13**:255–80; discussion 335–40.
33. Kelly CS, Andersen CL, Pestian JP *et al.* Improved outcomes for hospitalized asthmatic children using a clinical pathway. *Ann Allergy Asthma Immunol* 2000;**84**:465–560.
34. Kotagal UR, Robbins JM, Kini NM *et al.* Impact of a bronchiolitis guideline: a multisite demonstration project. *Chest* 2002;**121**:1789–97.
35. Muething S, Schoettker PJ, Gerhardt WE *et al.* Decreasing overuse of therapies in the treatment of bronchiolitis by incorporating evidence at the point of care. *J Pediatrics* 2004;**144**: 693–872.