IDEAL framework for surgical innovation 1: the idea and development stages

Peter McCulloch,[1] Jonathan A Cook,[2] Douglas G Altman,[3] Carl Heneghan,[4] Markus K Diener,[5] On behalf of the IDEAL group

[1]Nuffield Department of Surgical Science, University of Oxford, Oxford, UK
[2]Health Services Research Unit, University of Aberdeen, Aberdeen, UK
[3]Centre for Statistics in Medicine, University of Oxford, UK
[4]Department of Public Health and Primary Care, University of Oxford, UK
[5]Study Centre of the German Surgical Society, Department of General, Visceral, and Transplantation Surgery, Heidelberg University, D-69120 Heidelberg, Germany
Correspondence to: M K Diener
Markus.Diener@med.uni-heidelberg.de
Accepted: 15 March 2013
Cite this as: *BMJ* 2013;346:f3012
doi: 10.1136/bmj.f3012

IDEAL is a framework for evaluations of surgical innovations, which follow a distinct development pathway differing from the approach developed for pharmacological interventions. Many pathway and evaluation challenges are shared by other interventional therapies, requiring individual therapist skills and customisation of treatment to the individual, partly through medical devices. This paper provides an overview of the IDEAL framework and recommendations, and focuses on the first two stages: idea and development.

Box 1 | Recommendations for studies in stages 1 (idea) and 2 (development)

Idea

Mandatory registry for interventions thought to be first in man, with anonymous reporting option

Protocols pre-registered on the above mandatory registry, for planned research programmes on a first-in-man intervention

Development and use of agreed reporting standards and definitions for key outcomes and modifying factors

Development

Prospective development studies, with pre-published protocol and consecutive cases

Publication of findings, including transparent reporting of changes in technique or device design and indication

Introduction

Surgical innovations comprise new techniques, modified strategies, or innovative instruments. The evidence base for many of these approaches—and therefore for much of current surgical practice—is vastly weaker than for most modern drug treatments. Randomised trials of surgical techniques (versus placebo surgery)[1] were conducted within 10 years of the publication of the epochal streptomycin drug trial.[2] Yet, despite rapid growth in recent years, the overall number of randomised controlled trials and systematic reviews in surgical innovations remains small compared with the number of studies evaluating drug treatments. Randomised trials have also been few in number and of poor quality in some other therapeutic specialities, where the success of the intervention depends on the skill and judgment of the individual operator.[3]

The IDEAL Collaboration was born out of a series of conferences between surgeons and methodologists at Balliol College, Oxford,[4 5 6] which was convened to study why high quality trials in surgery were genuinely difficult to conduct, and what could be done to improve the evidence base for surgery. The conclusion was that innovation in surgery inevitably follows a pathway with important differences from that followed by pharmacological developments, and that a different approach to evaluation is therefore needed. It was noted that many non-surgical disciplines had similar problems with evaluation of such treatments (termed as "interventional therapies"), which rely on operator skill and tailoring of the intervention to the patient (for example, cardiac catheterisation, endoscopic techniques, or physiotherapy).

The IDEAL Collaboration developed a framework for the stages in surgical innovation (idea, development, exploration, assessment, and long term study; table) and a set of recommendations on how evaluation should be conducted at each stage (box 1). The collaboration also proposed how the environment for surgical research could be improved by editors, regulators, funders, and professional societies. An open international collaborative group has been developed to explore these issues further.[7] Recent concerns over hip resurfacing techniques[8] and breast implants[9] have raised serious questions about how medical devices are evaluated, and there has been

IDEAL framework

Stage 1: Idea	Stage 2a: Development	Stage 2b: Exploration	Stage 3: Assessment	Stage 4: Long term study
Question				
Can the procedure or device achieve a specific physical or physiological goal?	What is the optimal technique or design, and for which patients does it work best?	What are the outcomes of more widespread use? Can consensus equipoise be reached on a trial question?	How well does the procedure work compared with current standards of care?	What are the long term effects and outcomes of the procedure?
Aim				
Proof of concept	Safety, efficacy	Efficacy	Comparative effectiveness	Quality assurance
Patient base				
Single to few	10s	100s	100s+	100s+
Optimal study design(s)				
First-in-man study; structured case report	Prospective development study	Prospective collaborative observational study (Phase IIS) or feasibility randomised controlled trial (or both)	Randomised controlled trial	Observational study or randomised trial nested within a comprehensive disease based registry
Example of procedure at this stage				
Stem cell based tracheal transplant for tracheal stenosis[7]	Peroral endoscopic myotomy for oesophageal achalasia	Single incision laparoscopy for abdominal surgery	Minimally invasive oesophagectomy	Banding and bypass surgery for morbid obesity

Box 2 | Example of study at idea stage 1

Stem cell tracheal transplant (based on reference 11)

Clinical background at the time of conduct

Loss of an airway is debilitating and replacement is difficult

A stem cell graft embedded in a framework constructed using a de-antigenised collagen matrix may overcome current limitations

Design

Staged programme of multidisciplinary research beginning with detailed preclinical studies

Explicit clinical and data collection protocol written in advance and submitted for ethical review

Detailed informed consent process

Oversight provided by multiple agencies

Findings

Replacement using stem cell tracheal graft is feasible and a good early outcome is achievable

No sign of rejection was achieved without the need for immunosuppressive drugs

considerable interest in applying the IDEAL framework to this problem, since many difficulties in evaluating device innovation mirror those in surgery innovation. In this series of three articles, we explain the problems and discuss proposed solutions put forward in the IDEAL recommendations, using current examples. This first article in the series focuses on the first two stages of the IDEAL framework: idea and development.

Idea (IDEAL stage 1)

Surgical innovations can arise from careful planning and laboratory studies, from necessity created by an emergency, or even by accident. Advances in technology and related devices may make new or substantially different procedures feasible (such as robotic surgery). Planned and unplanned innovations can also occur out of desperation, in situations where the prognosis seems otherwise hopeless (for example, abbreviated "damage control" surgery for major combined vascular and visceral injury[10]). More measured innovation could represent an incremental advance, where the new procedure is a small variant on an older one. Innovations can also be completely novel, and be taken through to clinical trials via a carefully planned research programme, such as the recent successful advances in transplant surgery.[11]

What should a surgeon do if they believe that they have invented or developed something new and different, or if—in the case of industry driven research—they have used a new device in humans for the first time? The answer has two parts: surgeons can report what they have done, and then evaluate the intervention. Because first-in-man studies, by their nature, deal with single cases or a small number of cases, study design considerations might be largely irrelevant; but how they are reported is important. We can develop basic principles using the three pillars of the modern framework of medical ethics: utility, beneficence, and non-maleficence.[12] Surgeons should have the opportunity of learning from each other's experiences, particularly if this helps their patients or avoids them being harmed. Therefore, surgeons have an ethical obligation to share experiences with colleagues. Further, they need to convey sufficient information about what was done and what the consequences were, so that specialist colleagues can understand how to reproduce their success or avoid their failure.

All first-in-man interventions (whether a new procedure or new use of a device) should be included in an open access registry recording key details of the innovation. This registry would facilitate information searches in the published literature, which surgeons should carry out before embarking on a planned first-in-man intervention.[13] Box 2 provides an example of a study at the idea stage. The use of a new innovation, particularly if it is the first use of the innovation at the surgeon's institution, should have some form of independent oversight by those responsible for local clinical governance. Consent for new procedures is important: patients contemplating whether to undergo such procedures must fully understand their experimental nature, and the uncertainty that therefore surrounds any estimates of risk. If patient incapacity or time urgency prevents informed consent, governance authorities and patients' relatives or advocates may need to reach an agreement by discussion, even for retrospective cases. Therefore, hospitals would need systems that allow the right mix of clinical and ethical expertise to be bought to bear rapidly, and outside of normal hours if necessary.

Although surgeons may not need much incentive to report their successful innovations, it is arguably just as important to formally record their unsuccessful ideas or initial failures, to avoid unnecessary repetition by others. For this reason, the IDEAL recommendations include registration of all first-in-man procedures, with the suggestion that anonymous reporting might be permitted. In principle, anonymous reporting of harms or "near misses" might be desirable, but it has serious ethical, practical, and legal difficulties. If reporting is truly anonymous, how can spam or deliberately misleading reports be screened out? On the other hand, if identification of the author is possible in principle, legal discovery attempts and claims for compensation are a near certainty. The unique nature of new procedures might also make it difficult to maintain patient confidentiality. To allow surgeons to report their unsuccessful first-in-man efforts with confidence, a legal framework may be required, supported by the relevant governance and professional bodies to protect surgeons from compensation claims, provided that oversight and informed consent have been satisfactory.

Development (IDEAL stage 2a)

The IDEAL development stage begins once surgeons start to plan a series of procedures using a new technique or device (table). Innovations are especially fluid in this phase; innovations undergo rapid iterative change in the light of experience. Therefore, it is the development stage that most clearly differentiates the pathway for surgery innovation from that for pharmaceutical innovations. In both a scientific and ethical sense, development is the most problematic of the stages, and as a result is often poorly reported.

Experience often makes the need for modification obvious after only a few repetitions, although surgeons are insecure about the logical and ethical justification for making changes on the basis of scarce data that are not definitive. It is therefore tempting for authors to wait until

Box 3 | Common items for which agreed standard definitions are needed

Contextual factors

Grading of patient risk factors
Severity grading of comorbid pathology or general health
Scale of surgical insult
Urgency status of procedure
Environment for surgery (hospital or unit type)

Outcomes

Grading of functional performance
Scope and severity of complications

the development stage has ended, and then report the initial results as if the final version of the technique had been used in all cases. This strategy is adopted in the classic retrospective surgical case series, and is deeply problematic. Obscuring details that authors may not wish to report deprives others of the opportunity to learn from the development process, and can provide a misleading picture of the use of a procedure or device. Authors obscuring changes in eligibility, which naturally occur during development, in order to make it appear predetermined is similarly unhelpful. If the patients undergoing a technique are different at the beginning and end of a reported series, the aggregate outcome might not indicate much about what can be expected from using the final version of the technique in the patient group that trial and error has shown to be most suited to it.

Judgments about success or failure at this stage may be made on the basis of short term outcome measures that might not reflect the most important effects of the procedure, and frequently the data are insufficient to allow any meaningful statistical analysis (even if available data are maximised[14]). A cancer operation or a new artificial joint might seem successful at this stage because recovery from the surgery is quick or complications few, but subsequent data about survival rates or function in the long term may reverse these impressions. One may reasonably question the value of reporting such unreliable figures, but the alternative may be waiting many years for the results of definitive trials—which are unlikely to be undertaken without some pilot data. The pressure to innovate and improve is such that funders, patients, and clinical colleagues expect to be updated on the promise of innovations as rapidly as possible, in order to make decisions about funding, treatment, or use. If such decisions are to be made regardless (which seems a reasonable assumption barring a radical change in how health provision is organised internationally), they should at least be made using the most complete and accurate information available. The key principle, therefore, is transparency.

The IDEAL recommendations recognise that at the development stage, a randomised trial is often operationally undesirable and scientifically of limited use, owing to procedural modifications and varying eligibility. IDEAL supports prospective rather than retrospective studies at this stage, with sequential reporting of all cases and outcomes without omissions, and with clear explanations of when and how technique, design, or indications were changed. Sequential presentation of results might also reveal the effects of operator learning curves, which have an even more important role in the next stage of exploration. To ensure full reporting of relevant outcomes, the prior publication of a protocol at the outset of this type of study would be helpful. The United States Food and Drug Administration, which has been re-evaluating the regulatory framework for implantable devices since the Institute of Medicine report of 2011,[15] has put forward proposals for early studies of innovative devices that closely follow this model, which is encouraging.

The prospective development studies recommended by the IDEAL Collaboration represent a new type of observational study, which will no doubt change and evolve, but examples of this kind of study are now beginning to appear.[16 17] Key elements are a prior protocol, clearly defined objective outcomes, and transparent sequential reporting of cases, showing when changes in indication or technique are made. Data from this type of study will be more reliable and valid than information obtained from retrospective series, although retrospective data require much less effort and planning. We therefore suggest that for techniques and devices in the development stage, journals positively discriminate in favour of prospective studies, and should cease to accept studies based on retrospective data except when it can convincingly be shown that no viable alternative exists.

A much needed, important parallel improvement is the development of international standards for reporting surgical outcomes and contextual factors. Reports that use a common terminology and taxonomy are much more useful than those in which a plethora of definitions of the key data sow confusion and doubt. Groups such as COMET[18] and the Zurich group responsible for the Dindo-Clavien classification of complications[19] have made an important contribution to standardising this language, but further work is still needed. Many specialist endpoints will be best defined by consensus among the specialist community, and specialist societies and journals should work together to standardise terminology in their area of interest. Box 3 shows key outcomes and contextual factors that will need general agreement across the international surgical community. This agreement will need a concerted effort from international societies, national professional bodies, and leading journals, but research funders could also help by insisting on the use of standardised terms in funding applications.

Discussion

Early evaluations of a surgical innovation face common challenges; however, these difficulties must not prevent such studies being conducted. Current practices of study design and reporting are suboptimal and need upgrading. In particular, meaningful reporting of first-in-man cases should become routine (irrespective of the findings). Studies in the development stage need to be prospective, based on consecutive case reporting, and need to be open about the changes in indication, technique, and use of equipment that occur as experience is gained. Studies in the development stage also need a rapid, flexible, and expert system of governance to make decisions about whether to permit new procedures or devices to go ahead, and

SUMMARY POINTS

Innovations in surgery have several features that make scientific evaluation challenging, such as an early phase of rapid modification, learning curve, and strong therapist preferences

The IDEAL framework describes five stages of development and evaluation for surgical and interventional innovations: idea, development, exploration, assessment, and long term study

The IDEAL recommendations identify design and reporting ideas that could help in dealing with specific problems at each stage in the framework

At stage 1 (idea), accuracy, transparency, and completeness of reporting are key elements. Recommendations include standardisation of reports and development of an open access database for lodging reports of first-in-man procedures

At stage 2a (development), innovations are in a state of flux, undergoing modifications and changes in indication. Prospective studies with comprehensive sequential reporting of changes to technique and indication are recommended, together with standardisation of terminology

particularly about whether these new interventions can be modified, as is typical during this stage. Research ethics committees and device regulatory bodies could help by requiring a declaration of the IDEAL stage that the investigators feel the device or procedure has reached, with supporting evidence. Innovations in the idea stage would then be expected to lead to proposals for a prospective development study.

Modifications of interventions are a problem because of their likely frequency and the need for a rapid and ethical response. This task could be delegated to hospitals and universities, because more centralised bodies would be unable to gather information and respond appropriately in a realistic timescale. It would be sensible for existing structures to take on this role, in addition to their other functions, rather than to set up a new infrastructure. In the United Kingdom, trusts each have a committee to review proposals for procedures that are new to the trust, which would be the obvious body to prepare such a response. It is essential that the body responsible for providing an ethical opinion abides by certain principles to maintain an appropriate balance between fostering innovation and protecting current patients. Such considerations include developing sensible standards for documentation, reporting, and patient consent procedures; ensuring oversight of the committee itself; and ensuring access to suitable expert advice. Professional societies and bodies, healthcare institutions, and national regulatory agencies can all contribute towards better surgical research.

Summary

Early evaluations of a surgical innovation (whether an operation, invasive procedure, or use of a medical device) face a common set of difficulties, related principally to the need to modify and redefine the intervention and indication during evaluation. The IDEAL recommendations propose abandonment of retrospective case series, and instead recommend adoption of mandatory registration of first-in-man reporting, standardisation of reporting, and use of prospective study designs. Current regulatory and ethical governance structures need to be refined to facilitate an appropriate balance between fostering innovation and protecting patients.

The Health Services Research Unit is core funded by the Chief Scientist Office of the Scottish Government Health Directorates. Views expressed are those of the authors and do not necessarily reflect the view of the Chief Scientist Office or the funders.

Contributors: JAC and PM formulated the IDEAL series to which this paper belongs. PM wrote the first draft of this paper and MKD, JAC, CH, and DGA all commented on the draft. All authors approved the final version, and PM is the guarantor. The papers were informed by discussions during the IDEAL group in December 2010.

IDEAL workshop participants (December 2010): Doug Altman, Jeff Aronson, David Beard, Jane Blazeby, Bruce Campbell, Andrew Carr, Tammy Clifford, Jonathan Cook, Pierre Dagenais, Philipp Dahm, Peter Davidson, Hugh Davies, Markus K Diener, Jonothan Earnshaw, Patrick Ergina, Shamiram Feinglass, Trish Groves, Sion Glyn-Jones, Muir Gray, Alison Halliday, Judith Hargreaves, Carl Heneghan, Jo Carol Hiatt, Sean Kehoe, Nicola Lennard, Georgios Lyratzopoulos, Guy Maddern, Danica Marinac-Dabic, Peter McCulloch, Jon Nicholl, Markus Ott, Art Sedrakyan, Dan Schaber, Frank Schuller, Bill Summerskill.

Funding: The IDEAL group meeting in December 2010 was funded by the National Institute for Health Research's Health Technology Assessment programme, Johnson & Johnson, Medtronic, and Zimmer (all unrestricted grants). JAC holds a Medical Research Council Methodology Fellowship (G1002292).

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: PM received financial support from the National Institute for Health Research's Health Technology Assessment programme, Johnson & Johnson, Medtronic, and Zimmer for the IDEAL collaboration and for a workshop; PM and JC received support for travel to the US from the FDA to attend a seminar on IDEAL; no other financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Peer review and provenance: Not commissioned; externally peer reviewed.

1 Cobb LA, Thomas GI, Dillard DH, Meredino KA, Bruce RA. An evaluation of internal mammary artery ligation by a double-blind technic. *N Engl J Med* 1959;260:1115-8.
2 Streptomycin in Tuberculosis Trials Committee. Streptomycin treatment of pulmonary tuberculosis. A Medical Research Council investigation. *BMJ* 1948;2:769-82.
3 Wente MN, Seiler CM, Uhl W, Büchler MW. Perspectives of evidence-based surgery. *Dig Surg* 2003;20:263-9.
4 Barkun JS, Aronson JK, Feldman LS, Maddern GJ, Strasberg SM; Balliol Collaboration, et al. Evaluation and stages of surgical innovations. *Lancet* 2009;374:1089-96.
5 Ergina PL, Cook JA, Blazeby JM, Boutron I, Clavien PA, Reeves BC, et al. Challenges in evaluating surgical innovation. *Lancet* 2009;374:1097-104.
6 McCulloch P, Altman DG, Campbell WB, Flum DR, Glasziou P, Marshall JC, et al. No surgical innovation without evaluation: the IDEAL recommendations. *Lancet* 2009;374:1105-12.
7 IDEAL Collaboration. Home page. 2013. www.ideal-collaboration.net/.
8 Graves SE, Rothwell A, Tucker K, Jacobs JJ, Sedrakyan A. A multinational assessment of metal-on metal bearings in hip replacement. *J Bone Joint Surg Am* 2011;93(suppl 3):43-7.
9 O'Dowd A. French women to have PIP breast implants removed for free. *BMJ* 2011;343:d8329.
10 Rotondo MF, Schwab CW, McGonigal MD, Phillips GR 3rd, Fruchterman TM, Kauder DR, et al. 'Damage control': an approach for improved survival in exsanguinating penetrating abdominal injury. *J Trauma* 1993;35:375-82.
11 Macchiarini P, Jungebluth P, Go T, Asnaghi MA, Rees LE, Cogan TA, et al. Clinical transplantation of a tissue-engineered airway. *Lancet* 2008;372:2023-30.
12 Gillon R. Doctors and patients. *Br Med J (Clin Res Ed)* 1986;292:466-9.
13 Clarke M, Hopewell S, Chalmers I. Reports of clinical trials should begin and end with up-to-date systematic reviews of other relevant evidence: a status report. *J R Soc Med* 2007;100:187-90.
14 Lilford RJ, Thornton JG, Braunholtz D. Clinical trials and rare diseases: a way out of a conundrum. *BMJ* 1995;311:1621-5.
15 Institute of Medicine. Medical devices and the public's health: the FDA's 510(k) clearance process at 35 years. National Academies Press, 2011.
16 Blazeby JM, Blencowe NS, Titcomb DR, Metcalfe C, Hollowood AD, Barham CP. Demonstration of the IDEAL recommendations for evaluating and reporting surgical innovation in minimally invasive oesophagectomy. *Br J Surg* 2011;98:544-51.
17 Ahmed HU, Hindley RG, Dickinson L, Freeman A, Kirkham AP, Sahu M, et al. Focal therapy for localised unifocal and multifocal prostate cancer: a prospective development study. *Lancet Oncol* 2012;13:622-32.
18 Comet Initiative. Home page. 2013. www.comet-initiative.org/.
19 Dindo D, Demartines N, Clavien PA. Classification of surgical complications: a new proposal with evaluation in a cohort of 6336 patients and results of a survey. *Ann Surg* 2004;240:205-13.