Identifying research waste from surgical research: a protocol for assessing compliance with the IDEAL framework and recommendations

Jiajie Yu, Fei Shan, Allison Hirst, Peter McCulloch, Youping Li, Xin Sun

ABSTRACT

Introduction Approximately £1130 billion was invested in research worldwide in 2016, and 9.6% of this was on biomedical research. However, about 85% of biomedical research investment is wasted. The Lancet published a series to identify five categories relating to research waste and in 2014. Some categories of research waste in surgery are avoidable by complying with the Idea, Development, Exploration, Assessment, Long-term follow-up (IDEAL) framework for it enables researchers to design, conduct and report surgical studies robustly and transparently. This review aims to examine the extent to which surgical studies adhered to the IDEAL framework and estimate the amount of overall research waste that could be avoided if compliance was improved.

Methods We will search for potential studies published in English and between 1 January 2018 and 31 December 2018 via PubMed. Teams of paired reviewers will screen titles, abstracts and full texts independently. Two researchers will extract data from each paper. Data will be collected about general information and specialised information in each stage, and our IDEAL Compliance Appraisal tool will be used to analyse included studies. Descriptive statistics and χ² or Fisher’s exact tests for comparisons will be presented.

Discussion Our study will provide important information about whether compliance with the specific IDEAL Recommendations has reduced research waste in surgical and therapeutic device studies. And we will identify particular key aspects that are worse and need to focus on improving those in future education.

BACKGROUND

Surgical research is an essential part of clinical research and associated with a number of methodological and practical challenges. The Idea, Development, Exploration, Assessment and Long-term follow-up (IDEAL) Framework and Recommendations were established in 2009, and specify several unique features of clinical studies of surgical interventions, especially for earlier stages of surgical research. The IDEAL pathway integrates proposals regarding the design, conduct and reporting of clinical studies of surgery and has been widely accepted in studies on surgical intervention, medical devices and physical therapy.

According to the report prepared by Elsevier for UK’s Department for Business, Energy & Industrial Strategy, approximately £1130 billion was invested in research worldwide in 2016, and 9.6% of this was on biomedical research. However, it is estimated that about 85% of biomedical research investment is wasted. In 2014, The Lancet published a series to discuss issues relating to research waste. These articles identified five categories of research waste. First, research priorities are not based on the needs of potential users (ie, health professionals, patients, policy-makers). Second, weakness is common in the design, conduct and analysis of health research studies (eg, sample size, randomisation, blinding). Third, full information about studies is not accessible (ie, protocol, analyses plan, full study report and participant level data). Fourth, the regulation and management of the conduct of research are often inefficient. Finally, research waste also arises from inadequate reporting of research studies.

Some categories of research waste in surgery (ie, design and reporting flaws) are avoidable by complying with the IDEAL framework. IDEAL enables researchers and authors to design, conduct and report surgical studies in a robust and transparent way. Studies have explored research waste in IDEAL stage 3 (randomised controlled trial, RCT) and found a considerable burden of research waste in surgical RCTs. RCTs, however, only make up a small proportion of published surgical research. Since the IDEAL recommendations cover studies at all stages of the life cycle, and the majority of published studies of safety, efficacy and effectiveness are not randomised trials, better
adherence to IDEAL recommendations could reduce the total sum of research waste in surgery very significantly. We, therefore, decided to conduct a literature survey to examine the extent to which surgical studies adhered to the IDEAL framework to ensure appropriate, safe and efficient translation of new treatments into surgical practice. This will allow estimation of the amount of overall research waste which could be avoided if compliance was improved.

This study aims to assess the avoidable research waste relating to the design, conduct and reporting of studies evaluating innovation, effectiveness, efficacy or harms of surgical operations or therapeutic devices in recently year. Since the research waste associated with methodological and reporting defects in IDEAL stage 3 (RCT) has already been estimated, we will include papers that report studies in stage 1 (idea), stage 2a (development), stage 2b (exploration) and stage 4 (long-term follow-up) according to IDEAL framework.

**METHODS**

**Eligibility criteria**

Studies will be included in the review if they meet the following inclusion criteria: **participants**: any patient receiving surgical operations or therapeutic devices; **intervention**: at least one patient/group receiving any kind of surgical operations or therapeutic devices; **comparison**: alternative surgical approaches and/or non-surgical treatments; no comparator group in some early phase studies; **outcomes**: technical and/or physiological and/or clinical measures of intended effects and/or safety outcomes (including patient-reported outcomes); **study designs**: non-RCT, prospective cohort study, retrospective cohort study, case-control study, before and after study/interrupted time series, case series and case report.

A surgical operation is defined as the procedure where access is gained via an incision, natural orifice or percutaneous puncture or one that involves a device used inside the body. Radiological procedures performed without additional surgical intervention are not considered to be within the definition of a surgical study. Procedures testing injection, acupuncture or diagnostic interventions are not eligible, and those that tested effects of medical therapies in patients receiving surgical procedures, such as antibiotics or adjuvant treatments in cancer, will also not be considered. A therapeutic device is defined as a medical device used on humans with therapeutic benefits generally caused through a physical or mechanical effect on the body.

**Literature search**

We will search for potential publications via PubMed and the search strategy will be developed with the assistance of a specialist librarian. Search terms for surgical intervention and medical therapeutic devices will use those in our published study, and study designs will refer to search filters in Cochrane and the BMJ (online supplementary appendix 1). We will further limit our search to literature published in English and between 1 January 2018 and 31 December 2018. This will allow us to study a very large population of recent surgical studies, and the interval chosen will allow sufficient time for MeSH indexing to be completed before the search for published manuscripts. Citations will be collected by Endnote reference management software.

**Screening and selection procedure**

An inclusion/exclusion form will be used to screen titles and abstracts. Teams of paired reviewers, trained in the IDEAL framework, will use pilot-tested, standardised forms to screen titles and abstracts independently. We will conduct a pilot study before formal screening and data extraction, during which we will iteratively adjust the protocol and definitions used by reviewers to improve consistency between reviewers, and regular sampling with analysis of the interclass correlation coefficient (ICC) will be used to assess the progress of this process. Any conflicts not resolved by discussion between these reviewers will be referred to the study team for discussion. Full-text versions of papers will be obtained for inclusion in the study. The flow chart for retrieval, screening and exclusions are shown in figure 1.

**Data extraction**

Data will be extracted independently by at least two assessors for each study. We will randomly select 10% of included papers and evaluate the ICC between reviewers for these papers for both data extraction results and IDEAL compliance tool results. The following information will be collected for each eligible study:

**General information**

1. Author, date of publication, country of study (correspondent author), journal of publication; number of arms; recruitment setting (institute, national or international); number of centres; sample size.
2. Intervention type (surgical operation or therapeutic devices).
3. Trial design (non-RCT, prospective cohort study, retrospective cohort study, case-control study, before and after study/interrupted time series, cross-sectional study, case series, case report).
4. The study stage (stage 1, stage 2a, stage 2b and stage 4). This will be determined in term of the purpose/research questions of the study with the updated IDEAL Framework and recommendations. The key research questions (purpose) in each stage are: what is the new treatment concept and why is it needed (proof of concept/stage 1)? Has the intervention reached a state of stability sufficient to allow replication by others (development of procedure/stage 2a)? Have the questions that might compromise the chance of conducting a successful RCT been addressed (achieving consensus between surgeons and centres/stage 2b)? Are there any long-term or rare adverse effects or changes in indications or delivery quality over time (surveillance/stage 4).
5. Details of the protocol (reference, Digital Object Identifier (DOI), registration number, supplement).
6. Whether informed consent was obtained or not.
7. Funding source and/or conflict of interest.
8. Whether the authors state they were trying to use the IDEAL recommendations in their study.

Specialised information in each stage
The study will use our IDEAL Compliance Appraisal tool to analyse included studies. Details of the development and validation of the tool will be published separately, and the results of this study will provide further evidence to confirm its validity and reliability. Every domain in each stage is assessed as ‘yes/green’ for a positive result, ‘no/red’ for a negative result or ‘partial yes/yellow’ where neither of these appears appropriate. The details of the appraisal tool see in Table 1. We will include a brief narrative report on observed tendencies and trends within the data, including analysis of the areas of best and worst compliance overall, and a comment on any notable differences between the surgery and device subgroups.

Table 1 IDEAL Compliance Appraisal tool

<table>
<thead>
<tr>
<th>Stage</th>
<th>Domains</th>
<th>Option</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage 1 Idea</td>
<td></td>
<td>□ Yes □ Partial Yes □ No</td>
</tr>
<tr>
<td>Item 1</td>
<td>Was the patient selection process fully described.</td>
<td>□ Yes □ Partial Yes □ No</td>
</tr>
<tr>
<td>Item 2</td>
<td>Did the study fully describe the procedure of interest.</td>
<td>□ Yes □ Partial Yes □ No</td>
</tr>
<tr>
<td>Item 3</td>
<td>Did the study clearly state the rationale for the innovation.</td>
<td>□ Yes □ Partial Yes □ No</td>
</tr>
<tr>
<td>Item 4</td>
<td>Did the study appropriately report relevant outcome measures.</td>
<td>□ Yes □ Partial Yes □ No</td>
</tr>
<tr>
<td>Stage 2a Development</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 1</td>
<td>Did the study fully describe the patient selection process as used in the first case in the 2a study.</td>
<td>□ Yes □ Partial Yes □ No</td>
</tr>
<tr>
<td>Item 2</td>
<td>Did the study describe the procedure of interest as it was performed in the first case in the 2a study.</td>
<td>□ Yes □ Partial Yes □ No</td>
</tr>
<tr>
<td>Item 3</td>
<td>Did the study report all changes in the design, use of the device or the way of performing the procedure.</td>
<td>□ Yes □ Partial Yes □ No</td>
</tr>
<tr>
<td>Item 4</td>
<td>Did the study report key outcome measures sequentially (case by case) to allow evaluation of change.</td>
<td>□ Yes □ No</td>
</tr>
<tr>
<td>Stage 2b Assessment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 1</td>
<td>Did the study state if the innovation stable enough to move to stage 2b.</td>
<td>□ Yes □ Partial Yes □ No</td>
</tr>
<tr>
<td>Item 2</td>
<td></td>
<td>□ Yes □ Partial Yes □ No</td>
</tr>
<tr>
<td>Item 3</td>
<td>Did the study define the measures used to evaluate the quality of the procedure or device use.</td>
<td>□ Yes □ No</td>
</tr>
<tr>
<td>Item 4</td>
<td>Did the study provide information about the learning curve assessment.</td>
<td>□ Yes □ Partial Yes □ No</td>
</tr>
<tr>
<td>Item 5</td>
<td>Did the study provide a qualitative evaluation of the stakeholder values.</td>
<td>□ Yes □ Partial Yes □ No</td>
</tr>
<tr>
<td>Item 6</td>
<td>Did the study describe a pre-planned review of results and discussion to achieve the agreement on the stage 3 trial.</td>
<td>□ Yes □ Partial Yes □ No</td>
</tr>
<tr>
<td>Stage 4 Long-term follow-up.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 1</td>
<td>Did the study report the use of the primary outcome effect estimate to perform power calculations for possible RCTs.</td>
<td>□ Yes □ No</td>
</tr>
<tr>
<td>Item 2</td>
<td></td>
<td>□ Yes □ Partial Yes □ No</td>
</tr>
<tr>
<td>Item 3</td>
<td>Did the study describe the data source.</td>
<td>□ Yes □ No</td>
</tr>
<tr>
<td>Item 4</td>
<td>Did the study clearly and precisely describe the definition of the dataset.</td>
<td>□ Yes □ Partial Yes □ No</td>
</tr>
<tr>
<td>Item 5</td>
<td>Did the study conduct appropriate analysis of data.</td>
<td>□ Yes □ Partial Yes □ No</td>
</tr>
<tr>
<td>Item 6</td>
<td>Did the study report the missing data of the dataset.</td>
<td>□ Yes □ Partial Yes □ No</td>
</tr>
</tbody>
</table>

RCT, randomised controlled trial.
of search literature for screening and data extraction. A sample of studies will be selected randomly in a 1:1 ratio between studies of operations and therapeutic devices. If the number of eligible studies is insufficient after screening and data extraction, sampling of the identified studies will continue, until 100 of each type is selected.

**Statistical analysis**

We will estimate the avoidable research waste in each sample by comparing the actual number of ‘green’ item assessments versus the total number of items. We will estimate the potential contribution to reducing overall research waste if IDEAL were perfectly complied with by considering the percentage of all research registered on PubMed represented by the outcome studies selected by our search, and the proportion of this represented by different subgroups (eg, device vs procedure studies; prospective vs retrospective design).

For all descriptive analyses, we will use frequencies (and percentages) for dichotomous variables, and **χ²** (and percentages) for binary variables, and Spearman’s or Fisher’s exact tests will be used for the comparison between studies in surgical intervention and medical therapeutic devices. All analyses will be performed by IBM SPSS Statistics V.26.

**DISCUSSION**

The Lancet published one of its most provocative Commentaries in 1996 and labelled ‘surgical research’ as ‘comic opera’. This critique was based largely on surgical studies in the early stages. Our study will provide important information about whether compliance with the specific IDEAL Recommendations has reduced research waste in surgical and therapeutic device studies, especially those studies in earlier stages of surgical research. Additionally, we will identify particular aspects of study design, conduct and reporting that need improvement, usefully directing the focus of future education.

To our knowledge, this study is the first study assessing research waste in non-randomised studies and will include all specialties and all stage of the life cycle (from innovation to long-term follow-up), which should lead to better generalisability of the findings.

There are some limitations in this study. First, of the five key elements of research waste in the Lancet series, we are not able to assess waste due to failure to identify the research priority, research management issues and lack of fully accessible research, because published reports contain no information about these problems. Our methodological study, therefore, focuses on avoidable design, conduct and reporting errors. Second, our study is based on published information only, and may not reflect the actual implementation of surgical research.

**REFERENCES**