Systematic review of the quantity and quality of randomized clinical trials in pancreatic surgery

F. J. Hüttner[1,2], L. Capdeville[1], F. Pianka[1,2], A. Ulrich[1], T. Hackert[1], M. W. Büchler[1], P. Probst[1,2] and M. K. Diener[1,2]

[1]Department of General, Visceral and Transplantation Surgery, and [2]Study Centre of the German Surgical Society, University of Heidelberg, Heidelberg, Germany

Correspondence to: Dr M. K. Diener, Department of General, Visceral and Transplantation Surgery, University of Heidelberg, Im Neuenheimer Feld 110, 69120 Heidelberg, Germany (e-mail: markus.diener@med.uni-heidelberg.de)

Background: **RCTs are considered the reference standard in clinical research. However, surgical RCTs pose specific challenges and therefore numbers have been lower than those for randomized trials of medical interventions. In addition, surgical trials have often been associated with poor methodological quality. The objective of this study was to evaluate the evolution of quantity and quality of RCTs in pancreatic surgery and to identify evidence gaps.**

Methods: **PubMed, CENTRAL and Web of Science were searched systematically. Predefined data were extracted and organized in a database. Quantity and quality were compared for three intervals of the study period comprising more than three decades. Evidence maps were constructed to identify gaps in evidence.**

Results: **The search yielded 8210 results, of which 246 trials containing data on 26 154 patients were finally included. The number of RCTs per year increased continuously from a mean of 2·8, to 5·7 and up to 13·1 per year over the three intervals of the study. Most trials were conducted in Europe (46·3 per cent), followed by Asia (35·0 per cent) and North America (14·2 per cent). Overall, the quality of RCTs was moderate; however, with the exception of blinding, all domains of the Cochrane risk-of-bias tool improved significantly in the later part of the study. Evidence maps showed lack of evidence from RCTs for operations other than pancreatoduodenectomy and for specific diseases such as neuroendocrine neoplasms or intraductal papillary mucinous neoplasms.**

Conclusion: **The quantity and quality of RCTs in pancreatic surgery have increased. Evidence mapping showed gaps for specific procedures and diseases, indicating priorities for future research.**

Paper accepted 27 September 2018

Published online in Wiley Online Library (www.bjs.co.uk). **DOI:** 10.1002/bjs.11030

Introduction

RCTs represent the reference standard of clinical experimental research[1]. Specifically, in the context of evidence-based medicine or evidence-based surgery, RCTs provide the highest level of evidence at the individual-study level and form the basis of high-quality systematic reviews and meta-analyses[2]. Good clinical decision-making for individual patients is based on the available synoptic evidence. However, the number of RCTs assessing surgical interventions has been low since the methodology was introduced in the second half of the 20th century[3]. Furthermore, RCTs of surgical interventions pose specific challenges to researchers, and as a result surgical research has often been associated with low quality of conduct and reporting[4]. Since the identification of these shortcomings, substantial efforts have been made by different institutions to create an infrastructure for surgical research, resulting in larger numbers and higher quality of surgical RCTs[5,6].

Pancreatic cancer is a devastating disease and radical surgery remains the only chance of cure. Benign diseases of the pancreas such as chronic pancreatitis also depend greatly on surgical interventions to assure symptom relief and improve the patient's quality of life. However, pancreatic surgery is technically demanding and complex. Numerous modifications of the standard operative techniques – pancreatoduodenectomy, distal pancreatectomy, duodenum-preserving pancreatic head resection (DPPHR) and others – have been reported, and perioperative management therefore varies widely among institutions and across countries. This exacerbates the challenges for high-quality surgical research in this field.

Evidence mapping is an emerging approach to systematic assessment of quantitative and qualitative aspects in broad fields of research[7]. Although there is as yet no universally applied definition of evidence mapping, its aim is usually to identify gaps in the body of knowledge regarding a specific area of research. In times of scarcity of health system resources and information overload, this approach may enable researchers to define questions that should be prioritized by future research rather than just summarizing the existing evidence. Furthermore, evidence maps can serve to guide the efforts of stakeholders and funding organizations[8].

The aim of this project was to evaluate systematically the quantity and quality of RCTs in pancreatic surgery, and to identify gaps in evidence by construction of evidence maps.

Methods

As far as applicable, all steps of the present study followed the PRISMA recommendations[9] and those of the Cochrane Collaboration.

Literature search

A systematic literature search was conducted in MEDLINE (via PubMed), Cochrane CENTRAL and Web of Science up to May 2018[10]. No restrictions were applied regarding language or publication date.

The search strategy comprised Medical Subject Heading (MeSH) terms and text words (tw) combined with the Boolean operators 'AND' and 'OR'. The final PubMed search strategy was: ((pancreas[MeSH terms] OR pancreas[tw] OR pancreatic[tw])) AND (surgery[tw] OR 'surgical procedures, operative'[MeSH terms] OR ('surgical'[tw] AND 'procedures'[tw] AND 'operative'[tw]) OR ('operative surgical procedures'[tw] OR 'general surgery'[MeSH terms]) OR ('surgery'[tw])) OR (((((((((pancreaticoduodenectomy) OR pancreatoduodenectomy) OR 'pancreatic head resection') OR pancreatectomy) OR 'pancreatic resection') OR pancreaticoduodenectomy[MeSH terms]) OR pancreatectomy[MeSH terms]) OR 'duodenum-preserving pancreatic head resection') OR dpphr) OR 'pancreatic enucleation') AND (randomized controlled trial [pt] OR random*).

Trial selection

All RCTs of pancreatic surgical interventions or perioperative interventions (medical, nutritional, etc.) directly related to pancreatic surgery were included, as were RCTs comparing surgical with non-surgical interventions (such as endoscopy). Trials evaluating longer-running interventions that did not directly involve the perioperative period, such as neoadjuvant or adjuvant therapy trials, were excluded from analysis. There were no restrictions with regard to underlying diseases, endpoints or duration of follow-up.

Two reviewers independently screened the titles and abstracts of all retrieved references for inclusion. If one reviewer considered an article as potentially eligible, the full text of the publication was obtained and evaluated in detail. In the event of disagreement, consensus was reached by discussion with a third author.

Whenever two or more publications reported on the same RCT, the main publication was included but the full texts of the secondary publications were also obtained and checked for further relevant information.

Data extraction

A standardized electronic extraction sheet (available on request) was used for data extraction. The following predefined data were gathered: first author, year of publication, journal, region of publication (Africa, Asia, Australia/ New Zealand, Europe, North America, South America or other), type of operation (distal pancreatectomy, DPPHR, enucleation, pancreatoduodenectomy, total pancreatectomy or other), type of trial intervention (drug, medical device, nutrition, perioperative management, surgical strategy or other), type of disease (ampullary carcinoma, bile duct carcinoma, chronic pancreatitis, cystic neoplasms, duodenal carcinoma, intraductal papillary mucinous neoplasms, neuroendocrine tumours, pancreatic adenocarcinoma, trauma or other), sample size of trial, duration of follow-up, and quality features as described below.

Assessment of risk of bias

The Cochrane risk-of-bias tool[11] as described in the current version of the Cochrane Handbook[12] was used to assess the methodological quality of the included trials. The tool evaluates the following dimensions: randomization and allocation concealment (selection bias), blinding (performance bias and detection bias), incomplete outcome data (attrition bias), selective reporting (reporting bias) and other sources of bias (baseline imbalances, industry bias and sample size bias). If blinding was not stated, the risk of bias for these dimensions was always judged as high. With regard to selective reporting, trials that did not provide a published protocol or a trial registration were classed as having unclear risk of reporting bias. Whenever obvious discrepancies between the protocol/registration data and the trial publication were present, the trial was judged as high risk. Trials with discrepancies between the methods

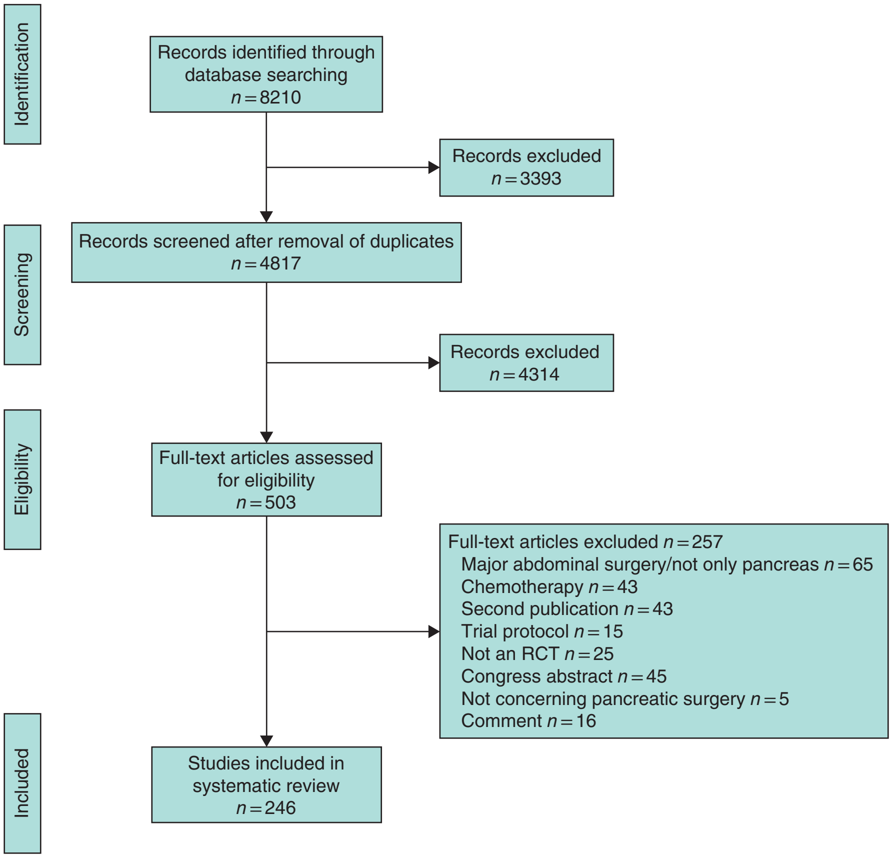

Fig. 1 PRISMA flow diagram showing selection of articles for review

and results section of a manuscript were also considered to be high risk. Lack of information on sample size calculation, shortcomings in statistical analysis (such as lack of intention-to-treat analysis), small sample bias and potential funding/industry bias were classified as other sources of bias. Small sample bias in terms of potential risk of type II error was judged as unclear if the total sample size of the trial comprised fewer than 50 patients, and as high risk if the calculated sample size was not reached.

Database

The extracted data were entered into a relational database (Microsoft Access®; Microsoft, Redmond, Washington, USA). Two reviewers independently checked all data entries to exclude transcription errors.

Evidence mapping

To identify evidence gaps in pancreatic surgery, evidence maps were constructed. Bubble plots were created, mapping all RCTs by type of operation against type of intervention and by type of disease against type of intervention. In these plots, each bubble represents an individual trial. The sample size of the trials is represented by bubble size and the geographical regions are colour-coded.

Statistical analysis

The RCTs were divided by date of publication into three subgroups: before 1996 (PI), 1996–2007 (PII) and 2008 onwards (PIII). These intervals were chosen according to relevant milestones regarding conduct and reporting of RCTs; the first version of the CONSORT statement was published in 1996, and mandatory registration of prospective clinical trials was incorporated into the Declaration of Helsinki in 2008.

Categorical variables were analysed as proportions/percentages. The χ^2 test was used for analysis of proportions. GraphPad Prism® version 5 (GraphPad Software, La Jolla, California, USA) and JMP® version 11 (SAS Institute, Cary, North Carolina, USA) were used for statistical analysis. Two-sided $P < 0{\cdot}050$ was considered significant.

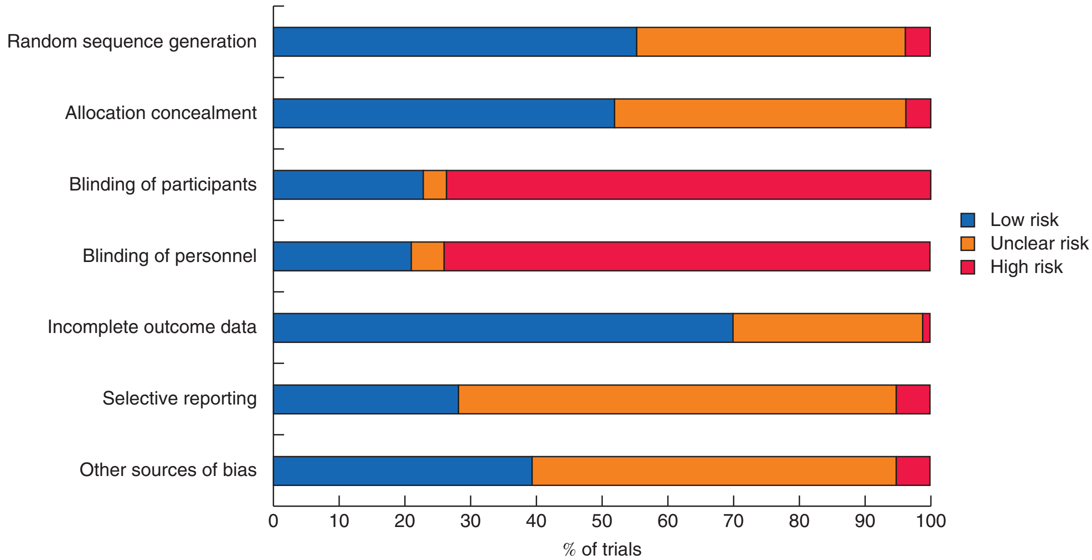

Fig. 2 Risk-of-bias summary

Results

Trial characteristics

A total of 8210 references were identified by systematic search of the three databases. After removal of duplicates, screening of titles and abstracts, and subsequent detailed assessment of potentially relevant full texts (*Fig. 1*), 246 publications of RCTs were included in the database (*Table S1*, supporting information).

Two of the RCTs were conducted in Australia or New Zealand (0·8 per cent), four in South America (1·6 per cent), five in Africa (2·0 per cent), 35 in North America (14·2 per cent), 86 in Asia (35·0 per cent) and 114 in Europe (46·3 per cent). Of the included trials, one was published in Hungarian, one in Korean, one in Russian, two in Italian, two in Spanish, four in Chinese, six in German and 229 in English.

Most of the articles were published in the journals *Annals of Surgery* (56, 22·8 per cent), *BJS* (18, 7·3 per cent), *Surgery* (12, 4·9 per cent), *American Journal of Surgery* (10, 4·1 per cent), *Journal of Gastrointestinal Surgery* (9, 3·7 per cent) and *Hepato-Gastroenterology* (8, 3·3 per cent). Only 12 of 246 RCTs (4·9 per cent) were published in a journal with an impact factor exceeding 10, comprising one trial from Africa, two from North America and nine RCTs from Europe.

The trials included data on a total of 26 154 patients. The sample size ranged from eight to 485 patients, with a median of 70·5. A total of 149 trials (60·6 per cent) provided information on duration of follow-up, whereas 97 publications (39·4 per cent) did not specify follow-up duration. The follow-up period ranged from 2 days to 84 months, with a median of 4 months.

Variation in surgical strategy was assessed in 116 RCTs (47·2 per cent), 44 evaluated drugs (17·9 per cent), 29 investigated nutritional interventions (11·8 per cent), 20 studied perioperative management (8·1 per cent), 17 assessed medical devices (6·9 per cent) and 20 RCTs (8·1 per cent) considered other interventions, such as surgical *versus* medical or endoscopic treatments.

Two hundred trials focused on only one specific type of operation. In most instances this was pancreatoduodenectomy (188, 76·4 per cent), followed by distal pancreatectomy in 19 trials (7·7 per cent), DPPHR in five (2·0 per cent), enucleation in one (0·4 per cent) and other types of operation in 33 RCTs (13·4 per cent), including necrosectomy in acute pancreatitis and bypass procedures in unresectable malignant disease. Forty-six RCTs (18·7 per cent) assessed more than one type of pancreatic operation; the second type of surgery was distal pancreatectomy in 11 studies, DPPHR in seven, total pancreatectomy in four, pancreatoduodenectomy in one RCT and others in 23 trials.

Most of the trials (160, 65·0 per cent) did not focus on an individual disease. However, pancreatic ductal adenocarcinoma was the most frequently occurring disease in 158 trials (64·2 per cent), whereas 23 trials (9·3 per cent) focused primarily on patients with chronic pancreatitis. Other diseases studied included ampullary carcinoma, bile duct carcinoma, intraductal papillary mucinous neoplasms, neuroendocrine tumours, and other lesions leading to an indication for surgery.

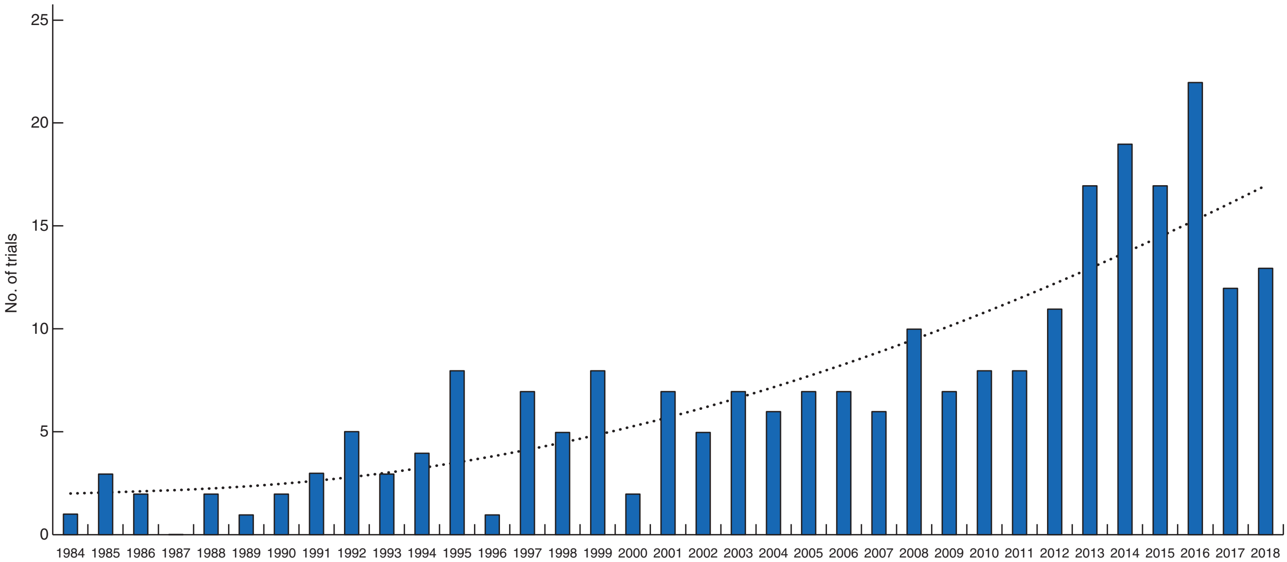

Fig. 3 Time trend of number of RCTs in pancreatic surgery. RCTs were divided by date of publication: before 1996 (PI), 1996-2007 (PII) and 2008 onwards (PIII)

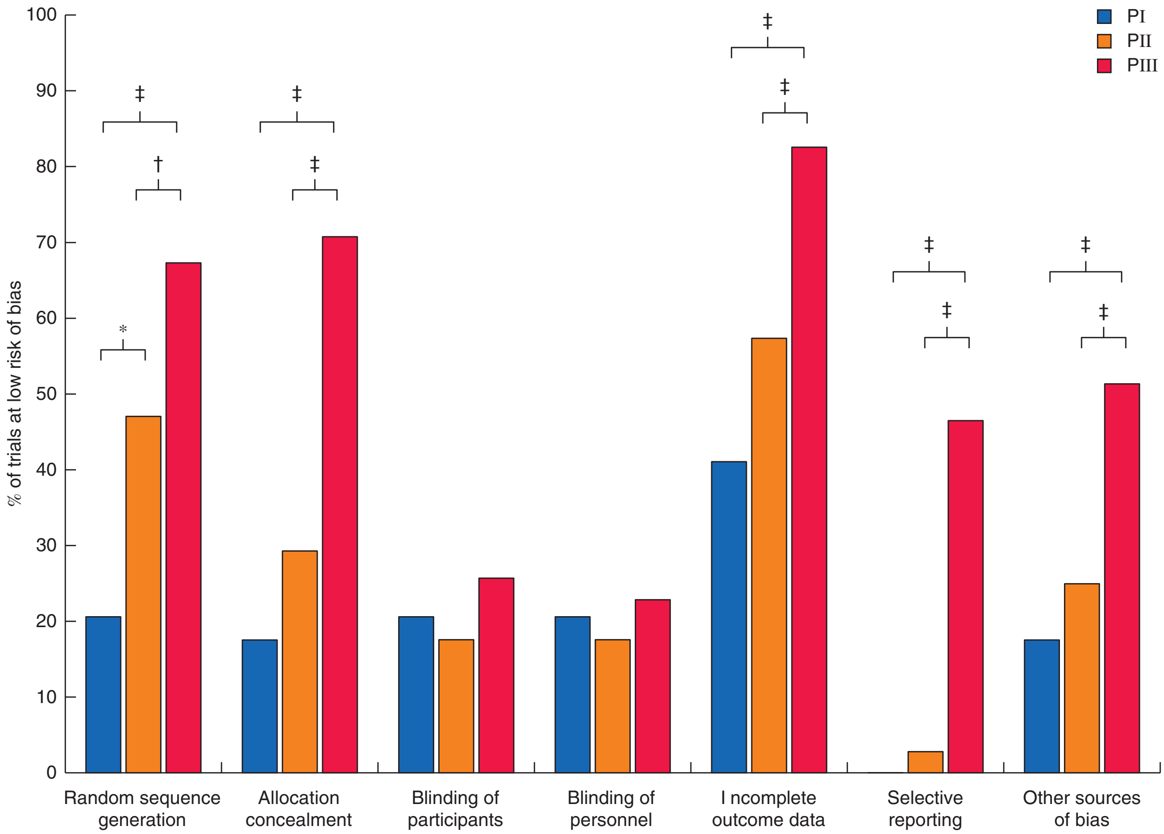

Fig. 4 Time trend of trial quality over the three parts of the study period: before 1996 (PI), 1996-2007 (PII) and 2008 onwards (PIII). Two-sided $^{*}P < 0{\cdot}050$, $^{\dagger}P < 0{\cdot}010$, $^{\ddagger}P < 0{\cdot}001$ (χ^2 test)

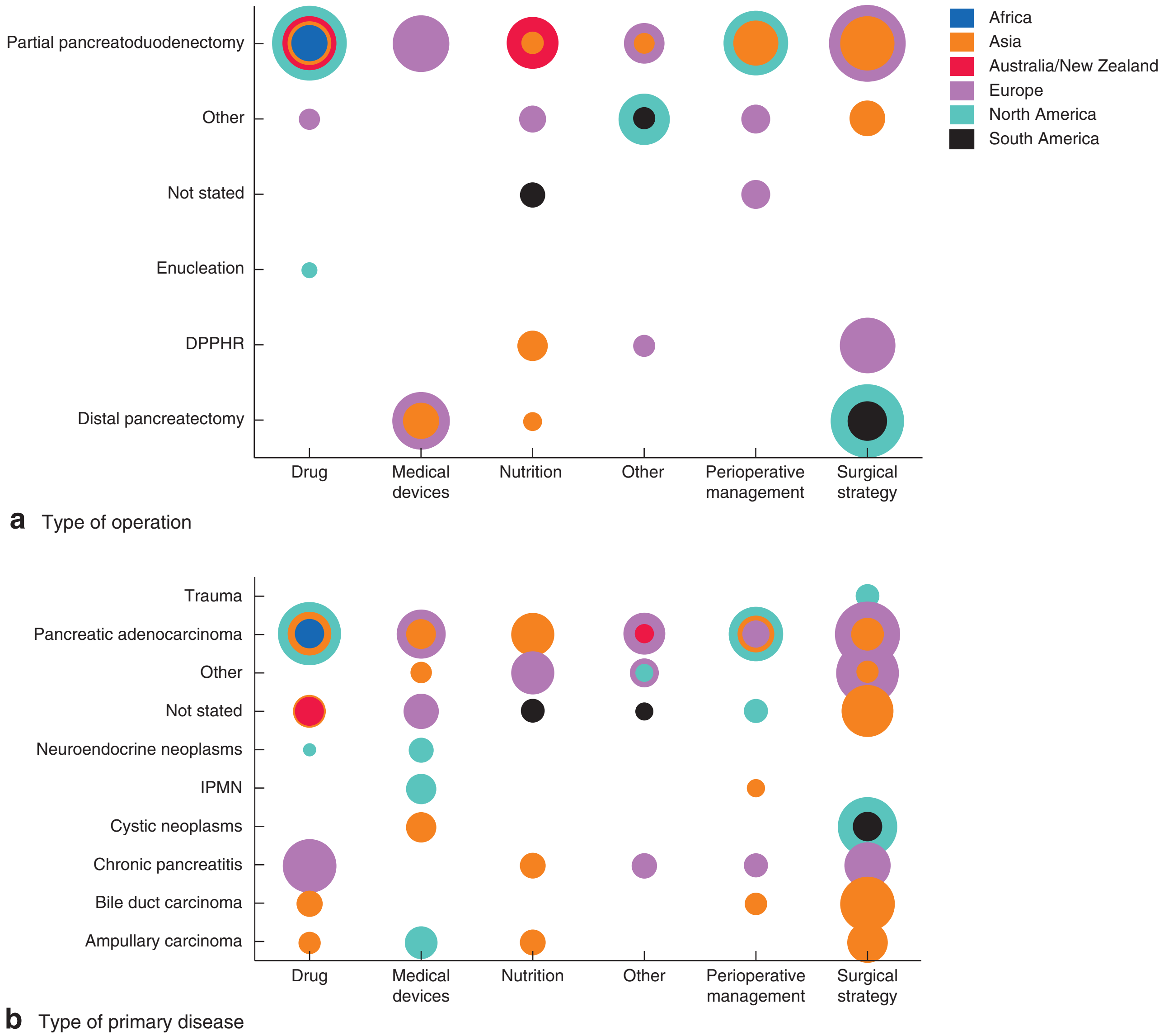

Fig. 5 Evidence maps showing types of operation and disease in relation to type of intervention, by geographical region. **a** Type of operation and **b** type of primary disease investigated against type of intervention. Each bubble represents an individual trial, and the trial sample size is represented by bubble size. DPPHR, duodenum-preserving pancreatic head resection; IPMN, intraductal papillary mucinous neoplasm.

Quality assessment

A total of 136 trials (55·3 per cent) clearly reported an adequate method of random sequence generation and thus were judged to be at low risk of bias in this domain; 101 RCTs (41·1 per cent) were judged as unclear risk, and nine trials (3·7 per cent) had a high risk of bias. Concerning allocation concealment, 128 of the publications (52·0 per cent) described an adequate method, whereas 109 (44·3 per cent) were at unclear risk and nine (3·7 per cent) at high risk of bias in this domain (*Fig. 2*).

Blinding of participants and personnel was reported infrequently. Fifty-six RCTs (22·8 per cent) reported adequate blinding of participants, whereas 52 (21·1 per cent) reported sufficient blinding of personnel. Most of the trials were at high risk of bias concerning blinding (participants: 181 of 246, 73·6 per cent; personnel: 182 of 246, 74·0 per cent); the remainder had an unclear risk of performance bias (blinding of participants: 9 of 246, 3·7 per cent; blinding of personnel: 12 of 246, 4·9 per cent).

Concerning incomplete outcome data, 172 trials (69·9 per cent) presented a flow chart or reported all dropouts and losses to follow-up adequately. However, in 71 RCTs (28·9 per cent) the risk of attrition bias was unclear and in three articles (1·2 per cent) a high risk was detected.

For most of the RCTs (164, 66·7 per cent), no trial registration or trial protocol was available, so these trials were judged to be at unclear risk of selective reporting. For 69 trials (28·0 per cent), a registration with sufficient information or a publicly accessible protocol was available and predefined outcomes were reported as planned, leading to a judgement of low risk in this domain. Thirteen publications (5·3 per cent) showed differences between predefined and finally reported outcomes, and were therefore at high risk of reporting bias.

An unclear risk of potential other sources of bias was found in 136 RCTs (55·3 per cent), mainly associated with missing sample size calculation, small sample bias or potential industry bias. In 97 trials (39·4 per cent) no other sources of bias were present, while 13 trials (5·3 per cent) had a high risk of other bias.

Most trials (129 of 246, 52·4 per cent) did not state the funding source; 95 (38·6 per cent) were funded by non-industry organizations and 22 (8·9 per cent) reported financing by industry.

Evolution of quantity and quality over time

The number of RCTs in pancreatic surgery increased steadily over the study period. Only 34 trials were published in PI, whereas 68 appeared in PII and 144 in PIII (*Fig. 3*). The mean(s.d.) number of RCTs per year was 2·8(2·0) in PI, 5·7(2·1) in PII and 13·1(4·8) in PIII.

With regard to random sequence generation, the frequency of low risk of bias was significantly higher in PII than PI ($P = 0{\cdot}017$). PIII had the highest rate of low risk of bias in this domain with a statistically significant difference compared with PI ($P < 0{\cdot}001$) and PII ($P = 0{\cdot}007$). Low risk of bias for allocation concealment was present more frequently in PIII than in PI and PII (both $P < 0{\cdot}001$) (*Fig. 4*). Blinding of participants and personnel did not show any statistically significant differences between the study intervals.

Frequency of low risk of bias for incomplete outcome data was higher in PIII than in both previous intervals (both $P < 0{\cdot}001$). Similarly, PIII had the highest rate of low risk of bias for selective reporting ($P < 0{\cdot}001$ *versus* PI and *versus* PII). Finally, low risk of other bias was more frequent during PIII than during PI and PII (both $P < 0{\cdot}001$).

Evidence mapping

Two evidence maps were created, one with the type of operation plotted against the type of trial intervention, and one with type of disease plotted against the type of trial intervention (*Fig. 5*). Fields with only small bubbles or no bubble at all had evidence gaps in the scientific literature on pancreatic surgery. Concerning the type of operation, there was wide evidence for partial pancreatoduodenectomy, but several ‘white spots’ were identified for distal pancreatectomy and DPPHR; no trial assessed total pancreatectomy in particular. Regarding individual diseases, neuroendocrine neoplasms, intraductal papillary mucinous neoplasms and pancreatic trauma in particular showed several evidence gaps, whereas pancreatic adenocarcinoma and chronic pancreatitis were covered most comprehensively.

Discussion

This analysis of the development of RCTs in pancreatic surgery over three decades found that the quantity and quality of trials increased for most domains. The only item that did not show significant improvement was blinding, which is more difficult in surgical trials than in drug trials[13,14]. The significant improvement is probably driven by efforts such as the implementation of the CONSORT statement to improve the reporting quality of RCTs[15].

Even though the quality of RCTs in pancreatic surgery has improved significantly over the past few decades, there are still substantial shortcomings regarding the whole body of evidence in this field. Only slightly more than half of all trials reported their randomization process adequately and could thus be judged as having a low risk of bias in this respect, and more than 70 per cent did not present a published protocol or sufficient trial registration. In addition, the reports of the majority of the included trials did not provide information on prerandomization exclusions; it is unclear to what extent selection of the trial cohorts occurred before randomization. This might impair the generalizability of the trial results, and highlights the need for clear and transparent reporting of patient flow in future clinical research. Furthermore, more than half of the trials yielded some sort of ‘other bias’, such as lack of sample size calculation or potential industry bias. Another aspect that raises concerns about the reliability of the evidence is the fact that the median sample size was 70·5, meaning that a large proportion of trials may be prone to some sort of small sample bias[16,17]. Furthermore, the median duration of follow-up was only 4 months, so data on the long-term effects of specific interventions are mostly lacking. These drawbacks show plainly that so-called evidence-based decisions in current guidelines and clinical practice that have been taken on the basis of this body of research may not be as reliable as thought. In common with the present report, Kaido's[18] analysis of RCTs on pancreatoduodenectomy found substantial shortcomings in trial design, conduct and reporting, which might impair the internal and external validity of such trials.

A recent analysis of the volume and quality of surgical RCTs in general, which compared trials from 1999 and 2009, also showed that the quantity and quality of surgical RCTs has increased[19]. Furthermore, the analysis revealed remarkable quality differences between different regions. The impact of region on methodological quality was not analysed in the present study. However, the finding that most RCTs published in high-impact journals were conducted in Europe and none in Asia is in line with the findings that the highest proportion of low-risk trials was from Europe, whereas the lowest was from Asia[19]. This interpretation is further corroborated by another study from the same group of authors[20] indicating that publication in a high-impact journal was independently associated with improved methodological quality. Another analysis focusing on RCTs in laparoscopic surgery also demonstrated that volume and quality have improved in recent years[21]. However, the authors stressed that there is still substantial room for improvement for several domains of trial conduct and reporting; for example, they showed improvement in reporting of dropouts and details of the analysis, whereas no improvement was seen in terms of blinding and reporting of funding[21]. These findings are in line with the present analysis.

The majority of trials analysed in the present study (52·4 per cent) did not report any information on funding. As industry bias is known to lead to exaggerated positive reporting of outcomes, this finding emphasizes the need for full disclosure of funding and potential conflicts of interest in the future[22]. Great efforts have already been made by the International Committee of Medical Journal Editors (ICMJE)[23,24] and others, but this work must be continued and the recommendations of the ICMJE need to be adopted comprehensively by journals around the globe.

The majority of RCTs in pancreatic surgery were conducted in Europe or Asia; only 14·2 per cent of trials took place in North America. Similar findings have been demonstrated for surgical RCTs in general and for laparoscopic RCTs[19,21]. Another analysis of trends in surgical RCTs revealed that the volume of surgical RCTs from the USA has stagnated and has been overtaken by most European countries at least in per capita publications[25]. There are several potential explanations for this in the context of trials in pancreatic surgery. First, American clinical research currently seems to be focused more on retrospective analysis of correlations in large quantities of data (big data analyses) than on investigating causalities by means of prospective randomized trials. Second, the clinical and scientific infrastructure in the USA complicates the conduct of multicentre RCTs, whereas in Europe several surgical trial networks have been established. Finally, the Institute of Medicine's *Initial National Priorities for Comparative Effectiveness Research*[26] included only a small number of surgical research priorities and discarded pancreatic disorders from the final list, even though 'malignant neoplasm of the pancreas' was among the top 20 causes of death in the USA at the time. Others[24] have postulated a generally diminished allocation of funding in surgery as a potential reason for the stagnation of surgical research in the USA.

The continuously increasing amount of evidence demands consistent compilation of the available data within any clearly defined medical field. The evidence mapping approach enables visualization of evidence gaps, which itself forms the basis for future prioritization of relevant research endeavours. This approach must adopt the methods of qualitative systematic review, because existence of data does not necessarily equate to existence of useful evidence. Evidence mapping is an adjunct to other methods of evidence-based research that may help to develop relevant research questions and guide future research projects. Prioritization will be key to assessing the effectiveness of interventions, especially in times of increasing costs in healthcare accompanied by scarcity of resources for research funding. On the other hand, this might in fact reduce healthcare costs by sorting out ineffective interventions. Therefore, such evidence mapping procedures will be of interest not only to researchers, but also to stakeholders and funding organizations.

In the present review, it was possible to identify white spots on the evidence maps that represent areas of insufficient knowledge in the field of pancreatic surgery. For example, although there is a substantial amount of evidence for all types of interventions in pancreatoduodenectomy, knowledge is almost completely lacking for most fields in DPPHR or distal pancreatectomy. However, the evidence maps cannot provide specific causes for this. With further improvement in the quality of clinical research, the work presented here may help to form a benchmark of the current status and may guide future research prioritization in pancreatic surgery.

Acknowledgements

P.P. and M.K.D. contributed equally to this study.
Disclosure: The authors declare no conflict of interest.

References

1 Bothwell LE, Greene JA, Podolsky SH, Jones DS. Assessing the gold standard – lessons from the history of RCTs. *N Engl J Med* 2016; **374**: 2175–2181.
2 Oxford Levels of Evidence Working Group. *The Oxford 2011 Levels of Evidence*; 2011. http://www.cebm.net/index.aspx?o=5653 [accessed 27 July 2018].
3 Wente MN, Seiler CM, Uhl W, Büchler MW. Perspectives of evidence-based surgery. *Dig Surg* 2003; **20**: 263–269.
4 Horton R. Surgical research or comic opera: questions, but few answers. *Lancet* 1996; **347**: 984–985.
5 Hamdy FC. 'Comic opera' is dead … surgical research is alive and well. *NCRI Cancer Conference*, Liverpool, 2015.
6 Hüttner FJ, Doerr-Harim C, Probst P, Tenckhoff S, Knebel P, Diener MK. Study methods in evidence-based surgery: methodological impediments and suggested approaches for the creation and transfer of knowledge in surgery. *Eur Surg Res* 2014; **53**: 86–94.
7 Bragge P, Clavisi O, Turner T, Tavender E, Collie A, Gruen RL. The Global Evidence Mapping Initiative: scoping research in broad topic areas. *BMC Med Res Methodol* 2011; **11**: 92.
8 Miake-Lye IM, Hempel S, Shanman R, Shekelle PG. What is an evidence map? A systematic review of published evidence maps and their definitions, methods, and products. *Syst Rev* 2016; **5**: 28.
9 Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP *et al.* The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ* 2009; **339**: b2700.
10 Goossen K, Tenckhoff S, Probst P, Grummich K, Mihaljevic AL, Büchler MW *et al.* Optimal literature search for systematic reviews in surgery. *Langenbecks Arch Surg* 2018; **403**: 119–129.
11 Higgins JP, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD *et al.* The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011; **343**: d5928.
12 Higgins JPT, Green S (eds). *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0* [updated March 2011]. The Cochrane Collaboration, 2011. http://handbook.cochrane.org [accessed 17 October 2018].
13 Cook JA. The challenges faced in the design, conduct and analysis of surgical randomised controlled trials. *Trials* 2009; **10**: 9.
14 Speich B. Blinding in surgical randomized clinical trials in 2015. *Ann Surg* 2017; **266**: 21–22.
15 Schulz KF, Altman DG, Moher D; CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010; **340**: c332.
16 Button KS, Ioannidis JP, Mokrysz C, Nosek BA, Flint J, Robinson ES *et al.* Power failure: why small sample size undermines the reliability of neuroscience. *Nat Rev Neurosci* 2013; **14**: 365–376.
17 Ioannidis JP. Why most published research findings are false. *PLoS Med* 2005; **2**: e124.
18 Kaido T. Recent randomized controlled trials in pancreaticoduodenectomy. *Pancreas* 2006; **33**: 228–232.
19 Ahmed Ali U, van der Sluis PC, Issa Y, Habaga IA, Gooszen HG, Flum DR *et al.* Trends in worldwide volume and methodological quality of surgical randomized controlled trials. *Ann Surg* 2013; **258**: 199–207.
20 Ahmed Ali U, Reiber BMM, Ten Hove JR, van der Sluis PC, Gooszen HG, Boermeester MA *et al.* Journal impact factor and methodological quality of surgical randomized controlled trials: an empirical study. *Langenbecks Arch Surg* 2017; **402**: 1015–1022.
21 Antoniou SA, Andreou A, Antoniou GA, Koch OO, Köhler G, Luketina RR *et al.* Volume and methodological quality of randomized controlled trials in laparoscopic surgery: assessment over a 10-year period. *Am J Surg* 2015; **210**: 922–929.
22 Probst P, Knebel P, Grummich K, Tenckhoff S, Ulrich A, Büchler MW *et al.* Industry bias in randomized controlled trials in general and abdominal surgery: an empirical study. *Ann Surg* 2016; **264**: 87–92.
23 Drazen JM, Van Der Weyden MB, Sahni P, Rosenberg J, Marusic A, Laine C *et al.* Uniform format for disclosure of competing interests in ICMJE journals. *Lancet* 2009; **374**: 1395–1396.
24 International Committee of Medical Journal Editors (ICMJE). International Committee of Medical Journal Editors (ICMJE): Uniform Requirements for Manuscripts Submitted to Biomedical Journals: writing and editing for biomedical publication. *Haematologica* 2004; **89**: 264.
25 Forrester JA, Forrester JD, Wren SM. Trends in country-specific surgical randomized clinical trial publications. *JAMA Surg* 2018; **153**: 386–388.
26 Institute of Medicine. *Initial National Priorities for Comparative Effectiveness Research*. National Academies Press: Washington, 2009.

Supporting information

Additional supporting information can be found online in the Supporting Information section at the end of the article.