This is a repository copy of *Variation in descriptors of patient characteristics in randomized clinical trials of peptic ulcer repair: a systematic review*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/160865/

Version: Published Version

Variation in descriptors of patient characteristics in randomized clinical trials of peptic ulcer repair: a systematic review

M. J. Lee[1,2] , **P. O. Coe[2], R. O'Donoghue[3], M. Peirson[4] and A. Saha[5]**

[1]Department of Oncology and Metabolism, University of Sheffield, and [2]Academic Directorate of General Surgery, Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, [3]Mater Infirmorum, Belfast, [4]Department of General Surgery, Ysbyty Gwynedd Hospital, Bangor, and [5]Department of General Surgery, Calderdale and Huddersfield NHS Foundation Trust, Huddersfield, UK

Correspondence to: Dr M. J. Lee, Office FU32, F Floor, The Medical School, University of Sheffield, Sheffield S10 2RX, UK (e-mail: m.j.lee@sheffield.ac.uk; 🐦 @wannabehawkeye, @petecoe1, @rodonoghue91, @arinsaha6)

Background: The ability to compare findings across surgical research is important. Inadequate description of participants, interventions or outcomes could lead to bias and inaccurate assessment of findings. The aim of this study was to assess consistency of description of participants using studies comparing laparoscopic and open repair of peptic ulcer as an example.

Methods: This systematic review is reported in line with the PRISMA checklist. Searches of MEDLINE and Embase databases were performed to identify studies comparing laparoscopic and open repair of perforated peptic ulcer in adults, published in the English language. Manuscripts were dual-screened for eligibility. Full texts were retrieved and dual-screened for inclusion. Data extracted from studies included descriptors of participants in studies from tables and text. Descriptors were categorized into conceptual domains by the research team, and coverage of each domain by study was tabulated.

Results: Searches identified 2018 studies. After screening, 37 full texts were retrieved and 23 studies were included in the final synthesis. A total of 76 unique descriptors were identified. These were classified into demographics (11 descriptors), vital signs (9 descriptors), disease-specific characteristics (10 descriptors), presentation and pathway factors (4 descriptors), risk factors (8 descriptors), laboratory tests (14 descriptors) and baseline health (28 descriptors). The number of descriptors in a single study ranged from three to 31. All studies reported at least one demographic descriptor. Laboratory tests was the least frequently described domain.

Conclusion: Study participants are described inconsistently in studies of a single example surgical condition.

Introduction

The purpose of scientific enquiry is to expand the realms of knowledge and understanding. One of the key ways in which this can be achieved is through standardization or control of a range of variables to eliminate bias. The rigorous application of this approach allows us to add weight to the findings. It can also help to address some of the current issues around replication of findings, the so-called 'replication crisis'. The other way of assessing the consistency of findings is through systematic review and meta-analysis. This allows us to identify whether treatments work, and for whom. Reviews are, however, necessarily limited by the quality of the studies entered into them.

There are three main components within any surgical study: the subjects who were studied (in terms of both inclusion and exclusion criteria and unreported characteristics); the intervention or treatment(s) compared; and the outcomes reported. A mismatch between studies in any of these components has the potential to introduce heterogeneity into an analysis, potentially leading researchers to an incorrect conclusion. This is a particular risk as there has been a proliferation of systematic reviews and meta-analyses in recent years, some of which have been completed with methodological flaws[1,2]. The potential error introduced by incomplete or inconsistent reporting of population characteristics could have several effects. One of these might be preventing the identification of

characteristics associated with favourable or unfavourable outcomes. Inadequate characterization of patient populations also presents challenges to the external validity of studies, as it impairs the comparison of a study to real-world clinical populations.

Previous work[3–5] has shown a range of outcomes reported in surgical studies, with limited ability to compare across them. This has led to recent attempts to develop 'core outcome sets' across a number of clinical settings[6–8]. This is an established methodology[9] which aims to rationalize outcomes to an agreed set that can be reported and compared consistently. It does not prevent researchers from reporting additional outcomes.

It is not clear whether there are issues relating to the description of patients entered into surgical studies, although it is recognized that selective inclusion of patients could alter study findings[10]. The aim of this analysis was to explore variation in reporting of baseline descriptors. This was conducted by assessing the literature comparing laparoscopic with open treatment of perforated peptic ulcer (PPU) as an example model, and exploring and quantifying any variation between these studies in their description of participants. This condition was selected as a subject of interest to the authors' research team.

Methods

This systematic review was performed with reference to the Cochrane Handbook[11], and is reported in line with PRISMA guidance[12]. It was not registered prospectively. Systematic searches of the MEDLINE and Embase databases were performed using a predefined search strategy, adapted from the previous Cochrane systematic review[13]. Search strategy is presented in *Appendix S1* (supporting information).

Manuscripts reporting comparison of short-term clinical outcomes (up to 90 days) between laparoscopic and open repair of PPU, published in the English language at any time were eligible for inclusion. Papers reporting longer-term outcomes or non-clinical outcomes (such as health economic evaluation) were not eligible for inclusion. Single-arm (non-comparative) studies and case series were excluded. Abstracts were screened against selection criteria for eligibility independently by two reviewers, with conflicts assessed by a third reviewer. This process was repeated for full-text assessments.

Year of publication and study design were collated for each included manuscript. Data extracted from studies included all baseline descriptors presented in the study demographics table (typically Table 1), as well as any descriptors presented in the results paragraph. Baseline

descriptors were those characteristics present at baseline and not impacted by the treatment decision, intervention or operative approach (such as age, sex and peritoneal contamination). Descriptors were classed as unique even if they reported the same measure (such as age or blood biochemistry measures) but used different cut-off levels, as these would affect the ability to interpret across studies. The number and nature of unique descriptors in each study was identified and aggregated across studies.

Baseline descriptors were presented to the research team as a long list. The team grouped descriptors according to the major concept they were thought to be measuring. After they had been grouped, a descriptive name was attached to each category. No bias assessment was performed, as this study reflects collation of content and not assessment of substance or results.

Descriptive reporting of the number of items reported was performed. With the hypothesis that reporting of descriptors may have improved in recent years, correlation between number of descriptors reported and year of publication was explored using Spearman's correlation. Significance was set at $P = 0.050$ *a priori*, and analyses were
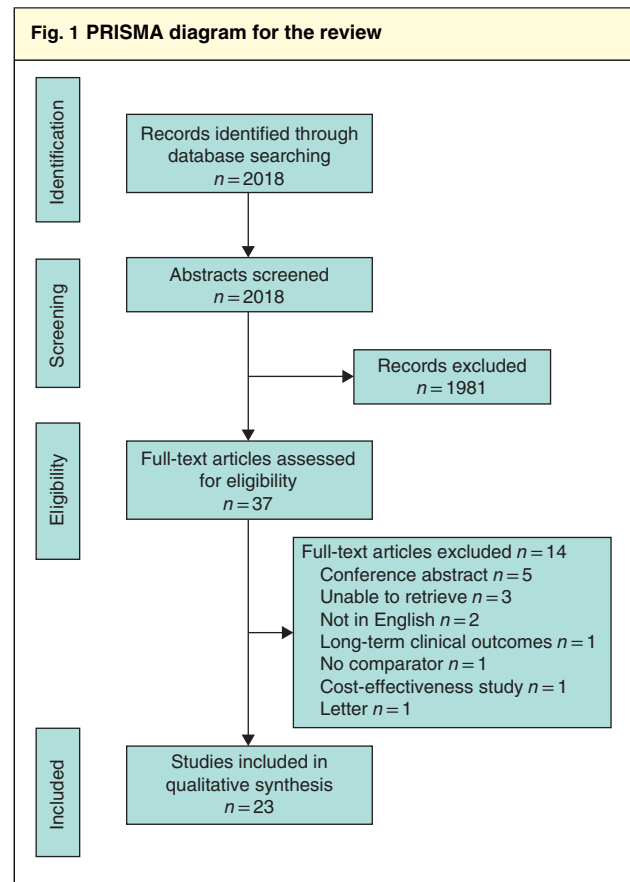

Fig. 1 PRISMA diagram for the review

Table 1 Summary of study design and coverage of included domains in the 23 studies

			No. of descriptors reported according to domain							
Reference	Design	No. of patients in study	Demographics	Risk factors	Disease-specific characteristics	Presentation and pathway factors	Vital signs	Laboratory tests	Baseline health status	Total no. of measures used
14	RCT	101	3	0	1	1	4	1	1	11
22	Retrospective cohort	50	1	2	0	0	0	0	0	3
23	Case−control*	527	4	0	0	0	1	6	5	16
24	Retrospective cohort	163	2	2	0	0	0	0	1	5
25	Retrospective cohort*	4210	2	1	2	0	2	4	2	13
15	RCT	119	2	2	1	0	1	1	2	9
19	Prospective cohort	27	4	0	3	0	0	0	2	9
26	Retrospective cohort†	2909	3	1	0	0	1	0	1	6
16	RCT	103	2	4	0	1	1	0	2	10
27	Retrospective cohort	148	2	0	1	1	2	4	1	11
28	Retrospective cohort‡	13 022	2	0	0	0	0	0	2	4
20	Prospective cohort	77	2	2	0	0	0	0	0	4
29	Retrospective cohort	184	2	0	2	0	1	0	1	6
30	Retrospective cohort	34	2	0	2	0	0	1	0	5
31	Retrospective cohort	74	2	4	0	1	0	0	1	8
17	RCT	50	2	0	2	0	0	0	0	4
32	Retrospective cohort	131	2	0	4	1	1	1	7	16
18	RCT	130	2	0	3	1	1	0	0	7
33	Retrospective cohort	109	3	1	0	1	2	0	1	8
21	Prospective cohort	61	4	0	0	0	0	0	0	4
34	Retrospective cohort	114	3	2	1	0	0	0	2	8
35	Case−control*	2462	4	3	0	0	3	8	13	31
36	Retrospective cohort§	726	3	0	0	0	0	0	5	8

*National Surgical Quality Improvement Project; †Ministry of Health, Labour and Welfare, Japan; ‡Hospital Episode Statistics, UK; §Danish Clinical Register of Emergency Surgery.

performed using R for statistics (The R Foundation for Statistical Computing, Vienna, Austria).

Results

Initial searches identified 2018 unique papers, of which 37 were selected for full-text assessment. Of these, 14 were excluded, leaving 23 papers for inclusion in the study (*Fig. 1*). Included studies were published between 1996 and 2019. Five were RCTs[14–18], three were prospective cohort studies[19–21], and the other 15 studies used retrospective cohort or case–control designs[22–36] (*Table 1*). All compared outcomes of laparoscopic *versus* open peptic ulcer repair.
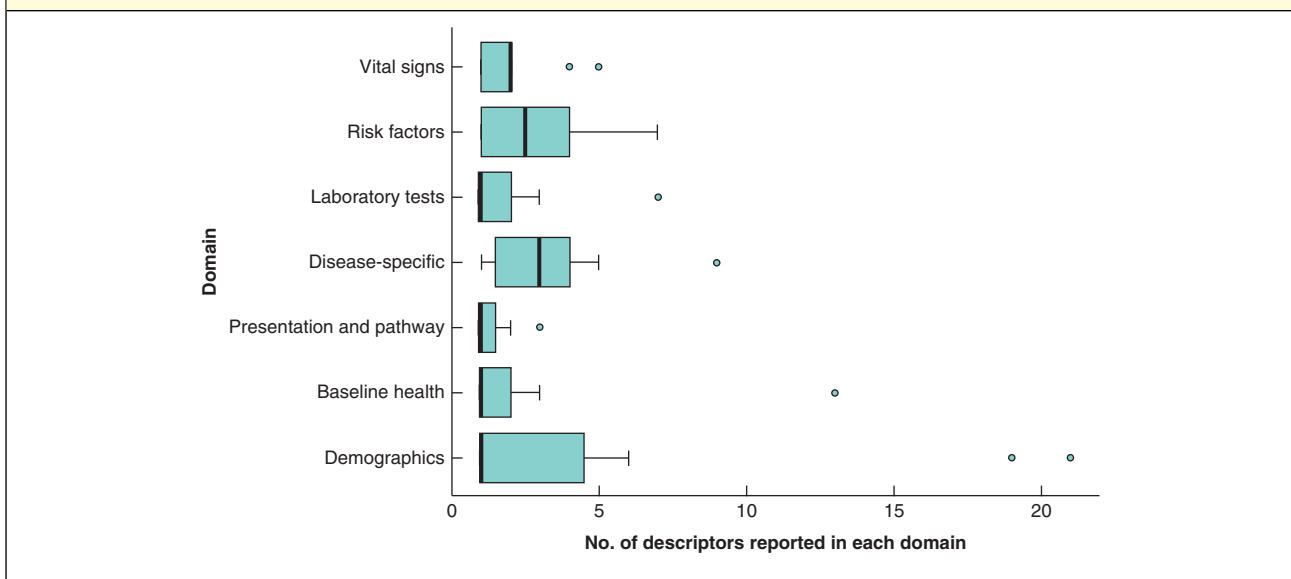
Participant descriptor domains

After longlisting of the identified domains, the items were categorized into seven conceptual groups: demographics; measures of baseline health; laboratory tests; risk factors for development of PPU; vital signs; disease-specific characteristics; and presentation and pathway factors. The longlist of

Table 2 Potential core data required for perforated peptic ulcer studies

Domain	Measures
Demographics	Age
	Sex
Baseline health	Composite co-morbidity measure (e.g. Charlson Co-morbidity Index)
Risk factors	Smoking status
Vital signs	BP on arrival
	Heart rate on arrival
Laboratory tests	Creatinine level (renal function)
	White cell count
Disease-specific data	Location of perforation
	Size of perforation (mm)
	Mannheim Peritonitis Index
Presentation factors	Time since onset of pain

descriptors and frequency of reporting is shown in *Table 2*. A summary of the number of descriptors reported per study in each domain is presented in *Fig. 2*.

Fig. 2 Frequency of reporting of items in each domain per study

Median values, interquartile ranges and ranges, excluding outliers (O), are denoted by vertical bars, boxes and error bars respectively.

Participant descriptor summary

A total of 76 unique descriptors were identified. The median number of descriptors published in a study was 11 (range 4–22). The most frequently reported descriptors were age, sex, ASA grade, site of perforation, smoking status, white blood cell count, BMI, preoperative shock, Boey score, alcohol use, and previous peptic ulcer disease. A further 46 descriptors were reported in just one study each.

Descriptor reporting by study design was explored. A median of 14 (range 8–18) descriptors were reported in RCTs, 8 (7–9) in prospective cohorts, 11 (4–17) in retrospective cohorts, and 14·5 (7–22) in case–control designs. Owing to the relatively small number of studies identified, a formal statistical analysis of variation between study designs was not performed. *Fig. S1* (supporting information) shows general overlap in the ranges of number of descriptors reported in each study type. There was no correlation between year of publication and number of descriptors reported ($r_s = 0.73$). A scatter plot of number of descriptors by year of publication is presented in *Fig. S2* (supporting information).

Demographics

Eleven demographic measures were identified with variable levels of reporting. Sex of participants was reported in 22 (96 per cent) of the 23 studies, and age as a continuous measure in 20 (87 per cent). Patient age was presented as a non-continuous measure in three studies (13 per cent),

with variable age groupings used. The characteristics of minority ethnicity and socioeconomic status by quartile were reported in one study (4 per cent) each (*Fig. 3a*).

Measures of baseline health

Twenty-eight measures were related to baseline or chronic health (*Fig. 3b*). ASA grade was the most frequently reported descriptor, in 13 studies (57 per cent). Charlson Co-morbidity Index, presence of chronic obstructive pulmonary disease or diabetes mellitus were reported in three studies each (13 per cent). The descriptor 'co-morbidities present', reported as a binary 'yes' or 'no' status, was used in two studies (9 per cent), as was the presence of congestive cardiac failure, hypertension and renal disease (which was itself variably defined or undefined).

Risk factors for development of perforated peptic ulcer

Eight risk factors for PPU disease were reported in studies (*Fig. 3c*). These were smoking status (7 studies, 30 per cent), alcohol use (4 studies, 17 per cent), previous peptic ulcer disease (4 studies, 17 per cent), steroid use (3 studies, 13 per cent) and non-steroidal anti-inflammatory use (2 studies, 9 per cent). The remaining characteristics of use of ulcerogenic drugs, use of aspirin and 'patients with prognostic factors' were reported in one study (4 per cent) each.

Fig. 3 Reporting frequency of descriptors in each domain

a Demographic descriptors
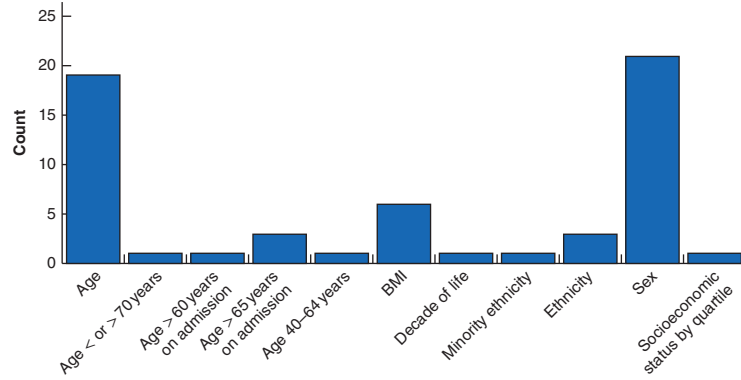
b Baseline health descriptors
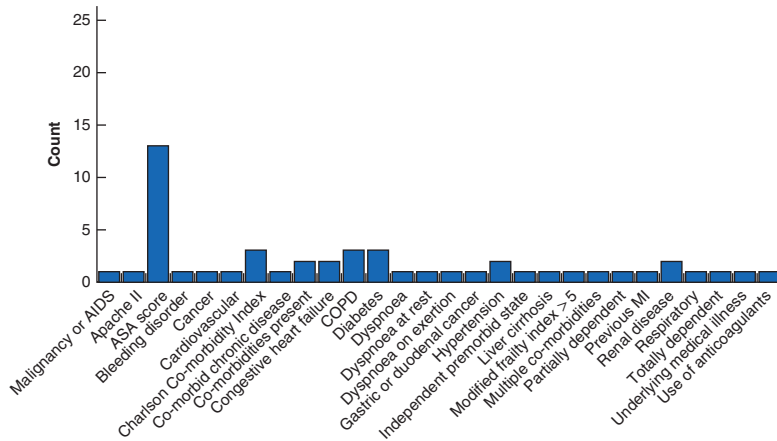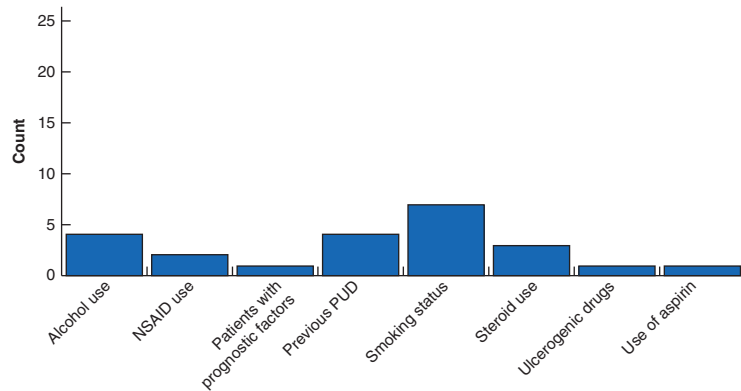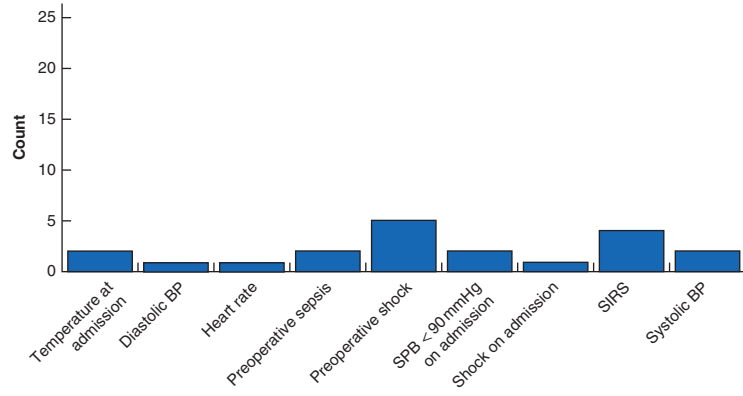
c Risk factor descriptors

Fig. 3 Continued

d Vital sign descriptors

Bar chart (y-axis: Count, 0–25) with categories: Temperature at admission, Diastolic BP, Heart rate, Preoperative sepsis, Preoperative shock, SPB < 90 mmHg on admission, Shock on admission, SIRS, Systolic BP.

e Laboratory test descriptors

Bar chart (y-axis: Count, 0–25) with categories: Albumin, Albumin < 30 g/dl, ALP, Bilirubin, BUN, BUN > 40 mg/dl, Creatinine, Creatinine > 1·2 mg/dl, Haematocrit, INR, SGOT, Sodium, Sodium > 135 mmol/l, White cell count.
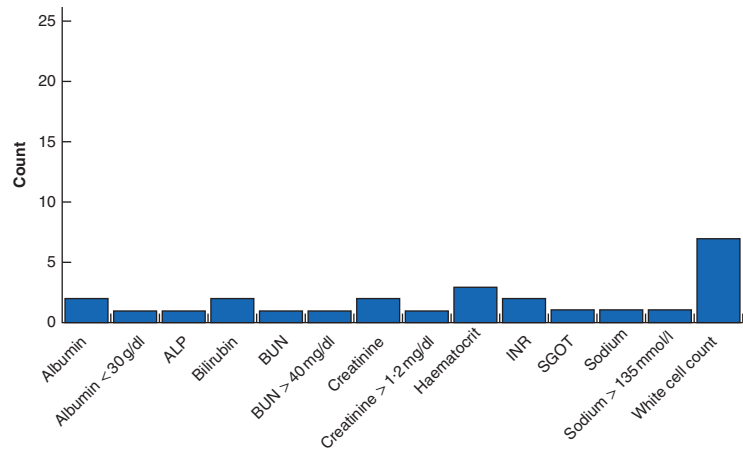
f Disease-specific descriptors

Bar chart (y-axis: Count, 0–25) with categories: Boey score, Concurrent bleeding, Mannheim peritonitis index, Mannheim peritonitis index > 27, Mean size of defect, Perforation > 1 cm, Preoperative transfusion, Preoperative ventilation, Site of perforation, Volume of transfusion.

Fig. 3 Continued

g Presentation and pathway related descriptors

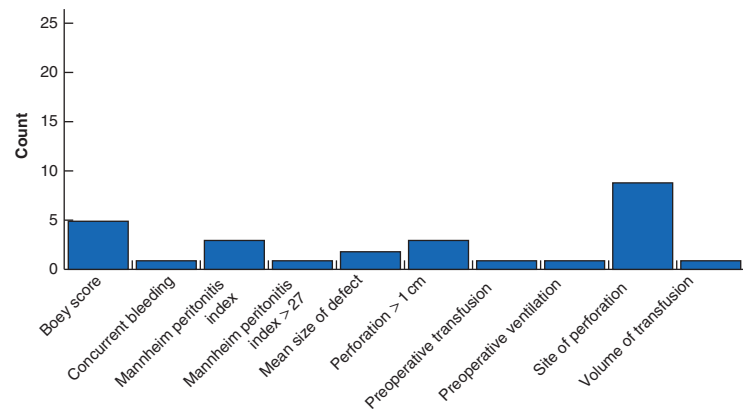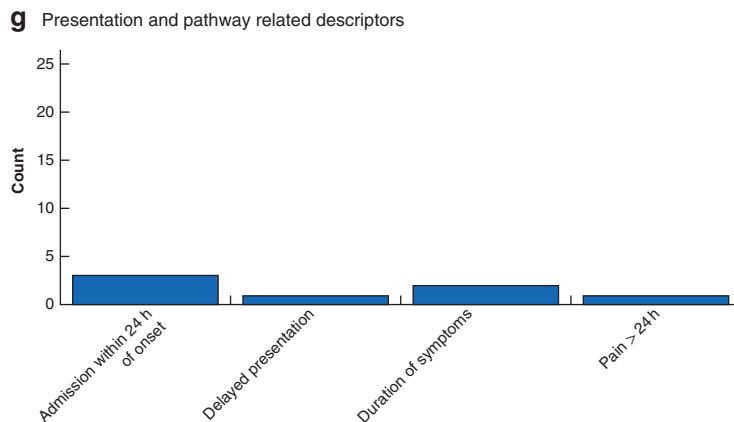**a** Demographics, **b** baseline heath, **c** risk factors for peptic ulcer development, **d** vital signs, **e** laboratory tests, **f** disease-specific characteristics, **g** presentation and pathway factors. AIDS, acquired immune deficiency syndrome; COPD, chronic obstructive pulmonary disease; MI, myocardial infarction; NSAID, non-steroidal anti-inflammatory drug; PUD, peptic ulcer disease; SIRS, systemic inflammatory response syndrome; ALP, alkaline phosphatase; BUN, blood urea nitrogen; INR, international normalized ratio; SGOT, serum glutamic–oxaloacetic transaminase.

Vital signs

There was considerable variability in the reporting of vital signs (*Fig. 3d*). Overall, nine descriptions of vital signs were recorded. Five of these were related to haemodynamic variables such as heart rate, systolic or diastolic BP. Preoperative shock was reported in five studies (22 per cent), mean systolic BP in two studies (9 per cent) and systolic BP of less than 90 mmHg on admission in two studies (9 per cent). Systemic inflammatory response syndrome was reported in four studies and temperature at admission in two studies. Shock on admission, diastolic BP on admission and heart rate were reported in one study (4 per cent) each.

Laboratory tests

Fourteen different laboratory measures were used in the studies (*Fig. 3e*). Of these, three were related to inflammatory parameters: white blood cell count at admission (7 studies, 30 per cent), albumin at admission (2 studies, 9 per cent) and albumin concentration below 30 g/dl (1 study, 4 per cent). Renal function was assessed using six different measures: creatinine (2 studies, 9 per cent), blood urea nitrogen (1 study, 4 per cent), blood urea nitrogen level above 40 mg/dl (1 study, 4 per cent), serum creatinine level above 1·2 mg/dl (1 study, 4 per cent), serum sodium (1 study, 4 per cent) and serum sodium level below 135 mmol/l (1 study, 4 per cent).

There were four descriptors related to liver function: international normalized ratio (2 studies, 9 per cent), bilirubin (2 studies, 9 per cent), alkaline phosphatase (1 study, 4 per cent) and serum glutamic–oxaloacetic transaminase (1 study, 4 per cent). Haematocrit level was reported in three studies (13 per cent).

Disease-specific characteristics

Ten measures of disease-specific characteristics were identified. These included factors related to the ulcer including site of perforation (9 studies, 39 per cent), defect size greater than 1 cm (3 studies, 13 per cent) and mean size of defect (2 studies, 9 per cent). Measures of items relating to degree of peritoneal contamination were also reported: the Boey score was reported in five studies (22 per cent), the Mannheim Peritonitis Index in three studies (13 per cent), and patients with a Mannheim Peritonitis Index score above 27 was reported in one study (4 per cent). Need for blood transfusion, volume of blood transfused and need for preoperative ventilation was reported in one study (4 per cent) each.

Presentation and pathway factors

Four factors were reported relating to timing of presentation. These were admission within 24 h (3 studies, 13 per cent) and duration of symptoms (2 studies, 9 per cent); delayed presentation and pain for more than 24 h were reported in one study (4 per cent) each.

Coverage of domains by study

Reporting of factors in each domain was assessed both as presence or absence and as the proportion of identified characteristics in that domain (*Table 1*). Laboratory tests was the least frequently reported domain, covered in eight studies. Eleven studies addressed risk factors, 12 reported preoperative vital signs, and 11 reported disease-specific characteristics. Seventeen studies reported baseline health descriptors, and all 23 studies reported some aspects of patient demographics.

Discussion

This study reviewed reporting of baseline characteristics of patients included in studies comparing laparoscopic with open repair of PPU. It demonstrated variable reporting of characteristics, with a range of measures used. These measures were, in turn, reported variably, with continuous variables frequently presented as categorical data. This poses challenges for the comparison of outcomes across studies.

The variation in descriptors between studies may in part be explained by variation in the prognosticators of clinical outcome that are reported by the literature. Yet, as an example, a putative predictor of mortality is systolic BP at admission. In the present systematic review it was found that this baseline characteristic was reported in five different ways, making comparison across studies difficult. Many studies used scores or measurements that were composites or indirect descriptors, for example BMI or Boey score. Where these were used, summaries of the constituent data were not presented. This hints that descriptive data were collected but not reported. This may be for reasons such as brevity of report, or may relate to journal style or policies.

Although it is recognized that the evidence base for emergency surgery should be improved by way of high-quality RCTs[37,38], there is a long-standing issue of unrepresentative samples contained within clinical research[39]. This affects both interpretation and comparison of results. These data show, through the example intervention of surgery for PPU, that rigorous and consistent characterization may be lacking in the acute surgical setting. This, in turn, may impact the external validity of studies, and prevent the identification of interventions that will truly improve outcomes for specific patient groups. A further benefit of standardizing descriptors for conditions under study might be to reduce heterogeneity of the populations under study, and thereby reduce heterogeneity in meta-analyses. Specifically, this might permit reliable comparisons of individual-level data in individual patient data meta-analysis in a way not currently achievable.

There are some limitations to this study that necessarily affect the strength of the conclusions. The study was not registered prospectively and no bias assessment has been performed, in common with other studies that longlist outcomes[3,4]. However, the review employed standard systematic review techniques with dual assessors of studies for inclusion. The search terms were robust, which should ensure the widest possible sampling of the literature. It is plausible that poor characterization of the underlying patient groups impedes comparison of studies. Where populations with varied and unreported characteristics are compared, this may contribute to heterogeneity[40]. To deal with the related issue around outcomes, the development of core outcome sets has occurred. These are likely to reduce heterogeneity attributable to variable outcome reporting[41]. Equally, improving descriptions of characteristics of interventions will likely address some of the heterogeneity attributable to surgical procedures[42]. The logical extension of this is to attempt to standardize the reporting of at least key descriptors in studies. A review of the literature identified only one further study[43] that addressed the question of variable descriptor reporting. This looked at the issue within the context of lower back pain, and highlighted the same issues around barriers to comparison and appreciating clinical applicability.

The noted variation in baseline descriptors is unlikely to be limited to this setting alone. Other avenues for investigation might be the characterization of patients in benign *versus* malignant conditions, emergency *versus* chronic or elective conditions, and also quality of descriptor reporting according to study design. This is likely to be an issue in other acute surgical conditions, although to what extent remains to be defined. Further work is required to define this across other conditions and interventions.

The development of a core descriptor set should be part of the formal process for development of a core outcome set, as the two are intrinsically entwined. The development of such a methodology will require work with stakeholders to establish common definitions of included characteristics. Implementation of such an intervention would also require scrutiny. Specifically, at what number of descriptors would such a set become a burden that is ignored by researchers? Opinion is also required on which descriptors may bring particular challenges in data collection such that missing data could become a problem. The present study suggests that around 11 descriptors seems to be a broadly acceptable number.

It might be anticipated that key characteristics could be defined that map to the domains identified here. These are proposed in *Table 2*. The exact measurement of each domain is subject to debate. For example, renal function

is measured preferentially using creatinine concentration in some settings, and blood urea nitrogen levels in others. There is also the potential that a key prognostic factor might not be presented here. Further work is required to establish a definitive core descriptor set for this condition that can be used across multiple health systems and study designs, with minimal burden to researchers.

Disclosure

The authors declare no conflict of interest.

References

1 O'Connor C. Meta-analysis bouillabaisse. *JACC Heart Fail* 2014; **2**: 201–202.

2 Pussegoda K, Turner L, Garritty C, Mayhew A, Skidmore B, Stevens A *et al.* Systematic review adherence to methodological or reporting quality. *Syst Rev* 2017; **6**: 131.

3 Mellor K, Hind D, Lee MJ. A systematic review of outcomes reported in small bowel obstruction research. *J Surg Res* 2018; **229**: 41–50.

4 Chapman SJ, Thorpe G, Vallance AE, Harji DP, Lee MJ, Fearnhead NS *et al.* Systematic review of definitions and outcome measures for return of bowel function after gastrointestinal surgery. *BJS Open* 2019; **3**: 1–10.

5 Fish R, Sanders C, Williamson PR, Renehan AG. Core outcome research measures in anal cancer (CORMAC): protocol for systematic review, qualitative interviews and Delphi survey to develop a core outcome set in anal cancer. *BMJ Open* 2017; **7**: e018726.

6 Haywood K, Whitehead L, Nadkarni VM, Achana F, Beesems S, Böttiger BW *et al.* COSCA (Core Outcome Set for Cardiac Arrest) in adults: an advisory statement from the international liaison committee on resuscitation. *Circulation* 2018; **137**: e783–e801.

7 Allin B, Ross A, Marven S, Hall N, Knight M. Development of a core outcome set for use in determining the overall success of gastroschisis treatment. *Trials* 2016; **17**: 360.

8 Sahnan K, Tozer PJ, Adegbola SO, Lee MJ, Heywood N, McNair AGK *et al.* Developing a core outcome set for fistulising perianal Crohn's disease. *Gut* 2019; **68**: 226–238.

9 Kirkham JJ, Davis K, Altman DG, Blazeby JM, Clarke M, Tunis S *et al.* Core Outcome Set-STAndards for Development: the COS-STAD recommendations. *PLoS Med* 2017; **14**: e1002447.

10 Page MJ, McKenzie JE, Kirkham J, Dwan K, Kramer S, Green S *et al.* Bias due to selective inclusion and reporting of outcomes and analyses in systematic reviews of randomised trials of healthcare interventions. *Cochrane Database Syst Rev* 2014; (1)MR000035.

11 Higgins JPT, Green S. *Cochrane Handbook for Systematic Reviews of Interventions*. John Wiley: Chichester, 2011.

12 Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Int J Surg* 2010; **8**: 336–341.

13 Sanabria A, Villegas MI, Morales Uribe CH. Laparoscopic repair for perforated peptic ulcer disease. *Cochrane Database Syst Rev* 2013; (2)CD004778.

14 Bertleff MJOE, Halm JA, Bemelman WA, van der Ham AC, van der Harst E, Oei HI *et al.* Randomized clinical trial of laparoscopic *versus* open repair of the perforated peptic ulcer: the LAMA trial. *World J Surg* 2009; **33**: 1368–1373.

15 Ge B, Wu M, Chen Q, Chen Q, Lin R, Liu L *et al.* A prospective randomized controlled trial of laparoscopic repair *versus* open repair for perforated peptic ulcers. *Surgery* 2016; **159**: 451–458.

16 Lau WY, Leung KL, Kwong KH, Davey IC, Robertson C, Dawson JJ *et al.* A randomized study comparing laparoscopic *versus* open repair of perforated peptic ulcer using suture or sutureless technique. *Ann Surg* 1996; **224**: 131–138.

17 Shah FH, Mehta SG, Gandhi MD, Saraj. Laparoscopic peptic ulcer perforation closure: the preferred choice. *Indian J Surg* 2015; **77**: 403–406.

18 Siu WT, Leong HT, Law BKB, Chau CH, Li ACN, Fung KH *et al.* Laparoscopic repair for perforated peptic ulcer: a randomized controlled trial. *Ann Surg* 2002; **235**: 313–319.

19 Johansson B, Hallerbäck B, Glise H, Johnsson E. Laparoscopic suture closure of perforated peptic ulcer. *Surg Endosc* 1996; **10**: 656–658.

20 Mehendale VG, Shenoy SN, Joshi AM, Chaudhari NC. Laparoscopic *versus* open surgical closure of perforated duodenal ulcers: a comparative study. *Indian J Gastroenterol* 2002; **21**: 222–224.

21 Sreeramulu PN, Venkatachalapathy TS, Supreet CS, Prathima S. A comparative study of laparoscopic *vs* open surgery for the management of duodenal ulcer perforation. *World J Laparosc Surg* 2013; **6**: 11–14.

22 Bhatti S, Muzaffar N, Jalloun H. Managing perforated peptic ulcer – a comparison between different approaches. *Mortality* 2014; **12**: 15.

23 Byrge N, Barton RG, Enniss TM, Nirula R. Laparoscopic *versus* open repair of perforated gastroduodenal ulcer: a national surgical quality improvement program analysis. *Am J Surg* 2013; **206**: 957–962.

24 Critchley AC, Phillips AW, Bawa SM, Gallagher PV. Management of perforated peptic ulcer in a district general hospital. *Ann R Coll Surg Engl* 2011; **93**: 615–619.

25 Davenport DL, Ueland WR, Kumar S, Plymale M, Bernard AC, Roth JS. A comparison of short-term outcomes between laparoscopic and open emergent repair of perforated peptic ulcers. *Surg Endosc* 2019; **33**: 764–772.

26 Kuwabara K, Matsuda S, Fushimi K, Ishikawa KB, Horiguchi H, Fujimori K. Community-based evaluation of laparoscopic *versus* open simple closure of perforated peptic ulcers. *World J Surg* 2011; **35**: 2485–2492.

27 Lee DJK, Ye M, Sun KH, Shelat VG, Koura A. Laparoscopic *versus* open omental patch repair for early presentation of

perforated peptic ulcer: matched retrospective cohort study. *Surg Res Pract* 2016; **2016**: 8605039.

28 Leusink A, Markar SR, Wiggins T, Mackenzie H, Faiz O, Hanna GB. Laparoscopic surgery for perforated peptic ulcer: an English national population-based cohort study. *Surg Endosc* 2018; **32**: 3783–3788.

29 Mirabella A, Fiorentini T, Tutino R, Falco N, Fontana T, De Marco P *et al.* Laparoscopy is an available alternative to open surgery in the treatment of perforated peptic ulcers: a retrospective multicenter study. *BMC Surg* 2018; **18**: 78.

30 Miserez M, Eypasch E, Spangenberger W, Lefering R, Troidl H. Laparoscopic and conventional closure of perforated peptic ulcer. *Surg Endosc* 1996; **10**: 831–836.

31 Næsgaard JM, Edwin B, Reiertsen O, Trondsen E, Færden AE, Rosseland AR. Laparoscopic and open operation in patients with perforated peptic ulcer. *Eur J Surg* 1999; **165**: 209–214.

32 Siow SL, Mahendran HA, Wong CM, Hardin M, Luk TL. Laparoscopic *versus* open repair of perforated peptic ulcer: improving outcomes utilizing a standardized technique. *Asian J Surg* 2018; **41**: 136–142.

33 Smith RS, Sundaramurthy SR, Croagh D. Laparoscopic *versus* open repair of perforated peptic ulcer: a retrospective cohort study. *Asian J Endosc Surg* 2019; **12**: 139–144.

34 Thorsen K, Glomsaker TB, von Meer A, Søreide K, Søreide JA. Trends in diagnosis and surgical management of patients with perforated peptic ulcer. *J Gastrointest Surg* 2011; **15**: 1329–1335.

35 Vakayil V, Bauman B, Joppru K, Mallick R, Tignanelli C, Connett J *et al.* Surgical repair of perforated peptic ulcers:

laparoscopic *versus* open approach. *Surg Endosc* 2019; **33**: 281–292.

36 Wilhelmsen M, Møller MH, Rosenstock S. Surgical complications after open and laparoscopic surgery for perforated peptic ulcer in a nationwide cohort. *Br J Surg* 2015; **102**: 382–387.

37 EMSurg Collaborators. Methodological overview of systematic reviews to establish the evidence base for emergency general surgery. *Br J Surg* 2017; **104**: 513–524.

38 Milton A, Drake TM, Lee MJ. The current status of clinical trials in emergency gastrointestinal surgery: a systematic analysis of contemporary clinical trials. *J Trauma Acute Care Surg* 2019; **86**: 524–531.

39 Altman DG. The scandal of poor medical research. *BMJ* 1994; **308**: 283–284.

40 Gagnier JJ, Moher D, Boon H, Beyene J, Bombardier C. Investigating clinical heterogeneity in systematic reviews: a methodologic review of guidance in the literature. *BMC Med Res Methodol* 2012; **12**: 111.

41 Webbe J, Sinha I, Gale C. Core outcome sets. *Arch Dis Child Educ Pract Ed* 2018; **103**: 163–166.

42 Blencowe NS, Mills N, Cook JA, Donovan JL, Rogers CA, Whiting P *et al.* Standardizing and monitoring the delivery of surgical interventions in randomized clinical trials. *Br J Surg* 2016; **103**: 1377–1384.

43 Wertli MM, Schöb M, Brunner F, Steurer J. Incomplete reporting of baseline characteristics in clinical trials: an analysis of randomized controlled trials and systematic reviews involving patients with chronic low back pain. *PLoS One* 2013; **8**: e58512.

Supporting information

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