Systematic review of the methodological and reporting quality of case series in surgery

R. A. Agha[1,3], A. J. Fowler[4], S.-Y. Lee[7], B. Gundogan[5], K. Whitehurst[5], H. K. Sagoo[6], K. J. L. Jeong[8], D. G. Altman[2] and D. P. Orgill[9]

[1]Balliol College, University of Oxford, and [2]Centre for Statistics in Medicine, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, Departments of [3]Plastic Surgery and [4]Medicine, Guy's and St Thomas' NHS Foundation Trust, [5]University College London Medical School and [6]King's College London GKT School of Medical Education, London, [7]University of Southampton Medical School, Southampton, and [8]Department of Obstetrics and Gynaecology, James Paget University Hospital, Great Yarmouth, UK, and [9]Harvard Medical School, Boston, Massachusetts, USA

Correspondence to: Dr A. J. Fowler, Department of Vascular Surgery, St Thomas' Hospital, Westminster Bridge Road, London SE1 7EH, UK (e-mail: alexjfowler@gmail.com)

Background: Case series are an important and common study type. No guideline exists for reporting case series and there is evidence of key data being missed from such reports. The first step in the process of developing a methodologically sound reporting guideline is a systematic review of literature relevant to the reporting deficiencies of case series.

Methods: A systematic review of methodological and reporting quality in surgical case series was performed. The electronic search strategy was developed by an information specialist and included MEDLINE, Embase, Cochrane Methods Register, Science Citation Index and Conference Proceedings Citation index, from the start of indexing to 5 November 2014. Independent screening, eligibility assessments and data extraction were performed. Included articles were then analysed for five areas of deficiency: failure to use standardized definitions, missing or selective data (including the omission of whole cases or important variables), transparency or incomplete reporting, whether alternative study designs were considered, and other issues.

Results: Database searching identified 2205 records. Through the process of screening and eligibility assessments, 92 articles met inclusion criteria. Frequencies of methodological and reporting issues identified were: failure to use standardized definitions (57 per cent), missing or selective data (66 per cent), transparency or incomplete reporting (70 per cent), whether alternative study designs were considered (11 per cent) and other issues (52 per cent).

Conclusion: The methodological and reporting quality of surgical case series needs improvement. The data indicate that evidence-based guidelines for the conduct and reporting of case series may be useful.

Introduction

A case series is an uncontrolled study that either samples participants with both a specific intervention (exposure) and a specific outcome, or samples participants with a specific outcome of interest regardless of their exposure status[1]. A series sampled only on exposure is a cohort study. Reports of case series are commonly a retrospective review of a string of patients with a unifying feature: exposure (including treatment) or outcome, or both. There has also been significant confusion between case series and a single-group cohort study[2]. Case series are common in the healthcare literature, but are also present within social sciences and the humanities[1]. As with case reports, their value has been debated[3,4]. In the age of evidence-based medicine (EBM), with the RCT as the criterion standard to show the efficacy of a particular treatment, what is their role?

The use of a case series in the recognition of a new disease was exemplified in 1999 by the epidemic of West Nile encephalitis in New York[5]. Historically, case series were important in identifying the impact of maternal drinking and pregnancy outcome[6] and the role of vitamin C in preventing scurvy[7]. More recently, a study by Albrecht and colleagues[8] of case series published in *The Lancet* found

that a high proportion led to follow-up trials and that they were useful in establishing an early evidence base for treatments of rare diseases in which trials would not be feasible. For some specialties, such as in accident and emergency medicine, establishing control groups may be difficult. In the social sciences, many social psychology studies have been case series, for example Yale psychologist Stanley Milgram's seminal work on obedience to authority figures[9].

In a 2005 report, Dalziel and co-workers[10] found that case series were used in 30 per cent of health technology assessments used in the provision and suitability of care. Poor reporting in the case series included in their study, however, severely constrained their analysis and investigation of the hypothesis that findings in case series may be affected by methodological characteristics[10]. Readers need complete, transparent information in all reports of research. Poor reporting of case series undermines critical appraisal, assessment of external validity and whether, for instance, a surgeon should change their practice.

No standardized reporting criteria have been developed within a robust methodological framework for case series. The aim of the present study was to close this gap and produce a reporting guideline for case series that is methodologically robust, easy to use, and accepted internationally across a broad range of specialties and disciplines. Following guidance on guideline development, the early steps in this process require an analysis of previous literature to identify previous guidance (if any) and to analyse relevant evidence on the quality of reporting of published research articles within the domain of interest[11]. This phase of the study involves a systematic review of the reporting within published surgical case series.

Methods

This systematic review was conducted according to the recommendations outlined in the Cochrane Handbook version 5.1.0 for Systematic Reviews[12] and reported in line with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement[13]. A protocol was developed and registered on the National Institute of Health Research (NIHR) – PROSPERO database (CRD42015016145). There were no deviations from the protocol during the conduct of the study[14].

Criteria for selecting studies

The search criteria were devised to locate studies specifically pertaining to the reporting quality of case series within surgery and to inform the development of recommendations for reporting such studies (*Table 1*).

Table 1 Search criteria

Criterion	Description
Types of study/material	Research articles and systematic reviews that highlight reporting issues within case series
Types of participant	Human participants undergoing surgery
Types of intervention	Any surgical intervention
Types of comparator	Case series typically have no comparator or control group. Nothing was specified here within the search criteria

Outcomes

Specified reporting deficiencies identified in the articles relating to case series were categorized under the following five headings: failure to use standardized definitions (for example, for outcomes and complications); missing or selective data (such as failing to document loss to follow-up and omitting whole cases or only presenting certain important variables and not others); lack of transparency or incomplete reporting (for example, failure to describe the patient population, intervention or outcomes in sufficient detail); whether alternative study designs were considered; and other issues (any other reporting deficiencies of note that did not come under 4 items above).

Search methods for identification of studies

Electronic searches

The following electronic databases were searched from their inception to 5 November 2014: MEDLINE, Embase, Cochrane Methods Register and Science Citation Index restricted to the English language. In addition, as part of the 'grey' literature search, the Conference Proceedings Citation Index was also searched.

Search terms and keywords

The search strategy was developed through consultation with an information specialist based at the Bodleian Library, University of Oxford. Its aim was to locate papers related specifically to the reporting quality of case series (*Table 1*). The search was performed on 24 September 2014 from inception of the database to this date. This search utilized English-language keywords combined with Boolean logical operators. The search was restricted to the English language and tailored to the idiosyncrasies of each database.

An example of a search strategy for the MEDLINE database is shown in *Table S1* (supporting information).

Identification and selection of articles

Studies identified by the electronic search strategy were listed. Results including citation, title and abstracts were populated into an Excel® database (Microsoft, Redmond,
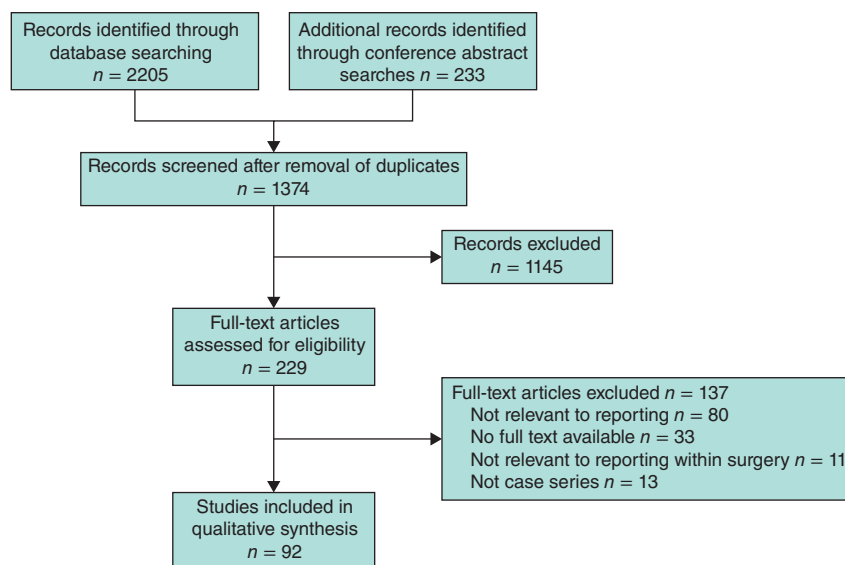
```
┌─────────────────────────┐   ┌─────────────────────────┐
│ Records identified through │   │ Additional records identified │
│   database searching     │   │ through conference abstract │
│       n = 2205           │   │    searches n = 233       │
└─────────────────────────┘   └─────────────────────────┘
            │                             │
            └──────────────┬──────────────┘
                           ▼
            ┌─────────────────────────────────────┐
            │ Records screened after removal of duplicates │
            │            n = 1374                  │
            └─────────────────────────────────────┘
                           │
                           │          ┌─────────────────────┐
                           ├─────────▶│ Records excluded     │
                           │          │    n = 1145          │
                           ▼          └─────────────────────┘
            ┌─────────────────────────┐
            │ Full-text articles       │
            │ assessed for eligibility │
            │       n = 229            │
            └─────────────────────────┘
                           │          ┌──────────────────────────────────────────────┐
                           │          │ Full-text articles excluded n = 137           │
                           ├─────────▶│ Not relevant to reporting n = 80              │
                           │          │ No full text available n = 33                 │
                           │          │ Not relevant to reporting within surgery n = 11 │
                           │          │ Not case series n = 13                        │
                           ▼          └──────────────────────────────────────────────┘
            ┌─────────────────────────┐
            │ Studies included in      │
            │ qualitative synthesis    │
            │       n = 92             │
            └─────────────────────────┘
```

Fig. 1 Flow diagram for the study

Washington, USA) and duplicates removed. Titles and abstracts were screened independently by two teams of authors for issues relating to the reporting quality of case series. Articles selected after title and abstract screening had their full text downloaded, and a further assessment was made of eligibility. Once articles had been selected for inclusion, data extraction took place. Any conflicts in either the eligibility of articles or the data extracted from them, not resolvable between the two teams, were referred to a single senior author for resolution, selected on the basis of recent experience with multiple other published systematic reviews.

Data extraction and management

Data were extracted independently by two teams of authors using a standardized data extraction form. Disagreements were resolved by discussion. Where resolution was not achieved, the same senior author as in the preceding section made a final decision. Data were then entered into a Microsoft Excel® 2011 database.

Data synthesis and statistical analysis

Outcomes were tabulated, with descriptive statistics used to determine frequently missing types of data within reports of case series.

Subgroup analysis

A sensitivity analysis was performed whereby results from studies whose primary aim was to assess the methodological and reporting quality of multiple case series (such as a research paper assessing reporting quality of case series) were considered separately from those articles that simply mentioned an issue in passing in their discussion (such as a single case series or a systematic review related to a particular clinical condition/treatment).

Results

The searches identified 1374 records. Through the process of screening and eligibility assessments, 92 articles, published over the period 1990–2014, were selected for inclusion (*Fig. 1*).

Within the two independent reviewing teams after record screening, there were discrepancies over whether 46 papers should be considered for full-text assessment. After discussion, 45 were excluded and one was included in the final list to go forward to full-text assessment. When it came to data extraction, there were discrepancies over 105 (22·8 per cent) of the 460 points that could not be resolved by discussion between the two teams and that were resolved in accordance with the protocol.

Results are summarized in *Table 2*.

Subgroup analysis

There were three systematic reviews and eight articles where the aim was specifically to investigate the methodological and reporting quality of case series (*Table 3*).

The main 'other issues' identified in this cohort include failure to define clearly the patient population under investigation, selection bias, insufficient follow-up time, and

Table 2 Summary of results

	n	Failure to use standardized definitions	Missing or selective data	Lack of transparency or complete reporting	Were alternative study designs considered?	Other issues
Articles	46	25	29	34	4	22
Systematic reviews	32	25	23	26	6	15
Conferences	14	2	9	4	0	11
Total	92	52 (57)	61 (66)	64 (70)	10 (11)	48 (52)

Values in parentheses are percentages.

Table 3 Summary of subgroup analysis

	n	Failure to use standardized definitions	Missing or selective data	Lack of transparency or incomplete reporting	Were alternative study designs considered?	Other issues
Articles	8	6	8	6	3	8
Systematic reviews	3	1	3	2	1	3
Total	11	7	11	8	4	11

need for validated outcomes (even though the ones used were well defined).

Discussion

The results show that surgical case series suffer from methodological and reporting issues. These can essentially be broken down along the five main lines of enquiry anticipated with high percentage frequencies across all areas studied. The other issues can be further segmented with sample size calculation, patient population definition, length of follow-up and whether outcomes are validated.

The value of case reports and case series has been questioned in the EBM era. Hoffman[15] has stated that 'more often than not' new ideas from such work are not sustained on further research. The focus of EBM is in finding the 'best' available research evidence to answer a given clinical question. The 'best' will have the least bias and is more likely to approach the truth of a given clinical question. However, is the poor reputation of the surgical case series the 'fault' of the concept, rather than its methodological and reporting execution? This systematic review clearly shows that those assessing case series often highlight areas that could have been improved through better conduct and reporting within the construct of the case series design.

Problems with the reporting of surgical case series were documented in a recent systematic review of autologous fat grafting for breast reconstruction[16]. In this study, 25 of 31 included studies were case series. Failure correctly to define the patient population under investigation, their demographic details and previous treatments is important, yet 20 per cent did not mention the age of the participants and 48 per cent did not indicate whether the participants

had been treated with radiotherapy, an important prognostic factor.

It has been noted elsewhere[17] that there are few formal assessments of how often the conclusions based on cases and case series are actually correct. This was highlighted in an investigation assessing side-effects reports, where 35 of 47 anecdotal reports were qualified as 'clearly correct'. Hence the predictive record of unstructured observations may be valuable[18]. Furthermore, two modelling exercises[19,20] have shown that case reports are likely to pick up true associations, for either rare diseases or more common diseases with a high relative risk. Indeed, such types of association led to the detection of 'flock-workers lung'[21].

So, when should a case series be performed? Vandenbroucke[4] argued in defence of case reports and case series, and listed their potential roles as: the recognition and description of new diseases, the detection of drug side-effects (adverse or beneficial), study of the mechanisms of disease, medical education and audit, and the recognition of rare manifestations of disease.

For surgical case series specifically, the following can also be advocated: rare diseases or rare circumstances (such as emergencies), new diseases – their description, natural history and management, studying the mechanism of disease and studying the impact of established procedures. In addition, late or delayed effects following surgical interventions, such as biliary malignancy after biliodigestive anastomosis, could be collated into a case series.

Where a new technique or device has been conceived and requires development and assessment, the IDEAL (Idea, Development, Exploration, Assessment and Long-term follow-up) framework is recommended[22].

In the ongoing drive to improve the evidence base for clinical practice, a number of tools have been developed to improve the quality of reporting research. For example, publication of the CONSORT statement has seen the quality of articles in some fields improve significantly[23,24]. The CONSORT statement has also been used to highlight and raise awareness of poor compliance in some fields[25−29]. The same is true of observational studies using the STROBE guideline[30].

A wide variety of reporting guidelines are now available across different research study types, from case reports to systematic reviews, but case series are a notable exception[31,32]. Surgery has the additional complexity of learning curves. The surgical technique selected is not the sole factor affecting outcome. Patients need to be selected carefully, appropriately investigated, and the technique has to be implemented meticulously in an appropriate setting with relevant safe anaesthesia and with an appropriate postoperative setting/regimen. The entire package has to be delineated and documented in case series in order to be reproducible by others.

Strengths of this review include conduct by a group with significant interest and experience in this area of methodological and reporting quality[25,26,30,33,34]. Limitations include restriction to the English language, although it has been estimated that 80–90 per cent of papers in scientific journals are written in English[35]. No synonyms for 'case series' were used, as none of the authors was aware of any. The potential to have missed relevant articles in the search or scored the articles incorrectly was hopefully minimized by having two teams that independently selected and scored articles.

The group's focus will now shift towards developing a guideline for the conduct and reporting of a case series. This has been named PROCESS (Preferred Reporting Of CasE Series in Surgery), and has been registered on the EQUATOR Network site, a repository for reporting guidelines[36]. This systematic review has now provided the initial items for an expert panel to consider through a Delphi consensus exercise.

Acknowledgements

References

1 Dekkers OM, Egger M, Altman DG, Vandenbroucke JP. Distinguishing case series from cohort studies. *Ann Intern Med* 2012; **156**: 37–40.

2 Esene IN, Ngu J, El Zoghby M, Solaroglu I, Sikod AM, Kotb A *et al.* Case series and descriptive cohort studies in neurosurgery: the confusion and solution. *Childs Nerv Syst* 2014; **30**: 1321–1332.

3 Agha R, Rosin DR. Time for a new approach to case reports. *Int J Surg Case Report* 2010; **1**: 1–3.

4 Vandenbroucke JP. In defense of case reports and case series. *Ann Intern Med* 2001; **134**: 330–334.

5 Exotic diseases close to home. *Lancet* 1999; **354**: 1221.

6 Landesman-Dwyer S. Maternal drinking and pregnancy outcome. *Appl Res Ment Retard* 1982; **3**: 241–263.

7 Lind J. *A Treatise on the Scurvy*. A. Millar: London, 1753.

8 Albrecht J, Meves A, Bigby M. Case reports and case series from Lancet had significant impact on medical literature. *J Clin Epidemiol* 2005; **58**: 1227–1232.

9 Milgram S. Behavioral study of obedience. *J Abnormal Psychol* 1963; **67**: 371–378.

10 Daziel K, Round A, Stein K, Garside R, Castelnuovo E, Payne L. Do the findings of case series studies vary significantly according to methodological characteristics? *Health Technol Assess* 2005; **9**: iii–iv, 1–146.

11 Moher D, Schulz KF, Simera I, Altman DG. Guidance for developers of health research reporting guidelines. *PLoS Med* 2010; **7**: e1000217.

12 Higgins JPT, Green S. *Cochrane Handbook for Systematic Reviews of Interventions*, version 5.1.0. http://handbook.cochrane.org [accessed 6 June 2015].

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

14 Agha R, Fowler A, Lee SY, Gundogan B, Whitehurst K, Sagoo H *et al. A Systematic Review Protocol for Reporting Deficiencies within Surgical Case Series*. http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42015016145#.VWI1redYWvk [accessed 24 May 2015].

15 Hoffman JR. Rethinking case reports. *West J Med* 1999; **170**: 253–254.

16 Agha RA, Fowler AJ, Herlin C, Goodacre T, Orgill DP. Use of autologous fat grafting for breast reconstruction: a systematic review and meta-analysis. *J Plast Reconstr Aesthet Surg* 2015; **68**: 143–161.

17 Chalmers I. Evaluating the effects of care during pregnancy and childbirth. In *Effective Care in Pregnancy and Childbirth*, Chalmers I, Enkin M, Keirse MJ (eds). Oxford University Press: New York, 1989; 3–38.

18 Venning GR. Validity of anecdotal reports of suspected adverse drug reactions: the problem of false alarms. *Br Med J (Clin Res Ed)* 1982; **284**: 249–252.

19 Tubert P, Bégaud B, Péré JC, Haramburu F, Lellouch J. Power and weakness of spontaneous reporting: a probabilistic approach. *J Clin Epidemiol* 1992; **45**: 283–286.

20 Bégaud B, Moride Y, Tubert-Bitter P, Chaslerie A, Haramburu F. False positives in spontaneous reporting: should we worry about them? *Br J Clin Pharmacol* 1994; **38**: 401–404.

21 Kern DG, Crausman RS, Durand KT, Nayer A, Kuhn C III. Flock worker's lung: chronic interstitial lung disease in the nylon flocking industry. *Ann Intern Med* 1998; **129**: 261–272.

22 Hirst A, Agha RA, Rosin D, McCulloch P. How can we improve surgical research and innovation? The IDEAL framework for action. *Int J Surg* 2013; **11**: 1038–1042.

23 Turner L, Shamseer L, Altman DG, Weeks L, Peters J, Kober T *et al.* Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals. *Cochrane Database Syst Rev* 2012; (11)MR000030.

24 Turner L, Shamseer L, Altman DG, Schulz KF, Moher D. Does use of the CONSORT Statement impact the completeness of reporting of randomised controlled trials published in medical journals? A Cochrane review. *Syst Rev* 2012; **1**: 60.

25 Lee S, Teoh P, Camm CF, Agha RA. Compliance of randomised controlled trials in trauma surgery with the CONSORT Statement: a systematic review. *J Trauma Acute Care Surg* 2013; **75**: 562–572.

26 Agha RA, Camm CF, Doganay E, Edison E, Siddiqui MR, Orgill DP. Randomised controlled trials in plastic surgery: a systematic review of reporting quality. *Eur J Plast Surg* 2013; **37**: 55–62.

27 Camm CF, Agha RA, Edison E. CONSORT adherence in journals is still far from perfect. *Ann Surg* 2015; **261**: e38.

28 Khajuria A, Agha R. CONSORT compliance in surgical randomized trials: possible solutions. *Ann Surg* 2015; **261**: e135.

29 Carthy ER, Agha RA, Gray R, Sullivan M, Altman DG, Gordon-Weeks AN. Adherence of trials of operative intervention to the CONSORT statement. *Ann R Coll Surg Engl* 2014; **96**: 86.

30 Agha R, Lee S, Jeong KJ, Fowler A, Orgill DP. Reporting quality of observational studies in plastic surgery needs improvement: a systematic review. *Ann Plast Surg* 2016; **76**: 585–589.

31 Gagnier J, Kienle G, Altman DG, Moher D, Sox H, Riley DS; CARE Group. The CARE guidelines: consensus-based clinical case report guideline development. *J Clin Epidemiol* 2014; **67**: 46–51.

32 Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009; **6**: e1000097.

33 Agha RA, Camm CF, Edison E, Orgill DP. The methodological quality of randomized controlled trials in plastic surgery needs improvement: a systematic review. *J Plast Reconstr Aesthet Surg* 2013; **66**: 447–452.

34 Yao A, Khajuria A, Camm C, Edison E, Agha R. The reporting quality of parallel randomised controlled trials in ophthalmic surgery in 2011: a systematic review. *Eye (Lond)* 2014; **28**: 1341–1349.

35 Montgomery S. Of towers, walls, and fields: perspectives on language in science. *Science* 2004; **303**: 1333–1335.

36 Equator Network. *Preferred Reporting Of CasE SerieS (PROCESS) Checklist.* http://www.equator-network.org/library/reporting-guidelines-underdevelopment/#23 [accessed 10 June 2015].

Supporting information

Additional supporting information may be found in the online version of this article:

Table S1 Search strategy for the MEDLINE database (Word document)