

Knee surgery and its evidence base

A Sharma¹, K Hasan², A Carter⁵, R Zaidi¹, S Cro⁴, T Briggs¹, A Goldberg¹

¹Royal National Orthopaedic Hospital NHS Trust, UK

²Barts Health NHS Trust, UK

⁵South Warwickshire NHS Foundation Trust, UK

⁴Medical Research Council, UK

ABSTRACT

INTRODUCTION Evidence driven orthopaedics is gaining prominence. It enables better management decisions and therefore better patient care. The aim of our study was to review a selection of the leading publications pertaining to knee surgery to assess changes in levels of evidence over a decade.

METHODS Articles from the years 2000 and 2010 in *The Knee*, the *Journal of Arthroplasty*, *Knee Surgery, Sports Traumatology, Arthroscopy*, the *Journal of Bone and Joint Surgery (American Volume)* and the *Bone and Joint Journal* were analysed and ranked according to guidelines from the Centre for Evidence-Based Medicine. The intervening years (2003, 2005 and 2007) were also analysed to further define the trend.

RESULTS The percentage of high level evidence (level I and II) studies increased albeit without reaching statistical significance. Following a significant downward trend, the latter part of the decade saw a major rise in levels of published evidence. The most frequent type of study was therapeutic.

CONCLUSIONS Although the rise in levels of evidence across the decade was not statistically significant, there was a significant drop and then rise in these levels in the interim. It is therefore important that a further study is performed to assess longer-term trends. Recent developments have made clear that high quality evidence will be having an ever increasing influence on future orthopaedic practice. We suggest that journals implement compulsory declaration of a published study's level of evidence and that authors consider their study designs carefully to enhance the quality of available evidence.

KEYWORDS

Evidence-based medicine – Knee – Arthroplasty – Review

Accepted 3 September 2015

CORRESPONDENCE TO

Aadhar Sharma, E: aadharsharma0@gmail.com

Global demand for knee surgery has risen steadily. It already represents the largest joint reconstruction market worldwide and is set to expand for the foreseeable future.^{1,2} An ageing population and an obesity epidemic are contributing to an increased number of referrals being made to orthopaedic departments. The number of hip and knee arthroplasties being performed has increased by 300% in ten years.^{3,4} In this period of financial austerity, the UK National Health Service (NHS) is faced with the need to make substantial savings.⁵ A report from 2012 entitled *Getting It Right First Time* followed a review of over 200 hospitals and highlighted significant variations in orthopaedic practice across the UK.⁵ This has resulted in recommendations for major organisational changes to ensure costs are minimised but not at the expense of patient outcome. The adoption of treatments underpinned by a strong evidence base are central to these recommendations.

Evidence-based medicine combines the best available research with clinical expertise so as to optimise patient

care.^{5,6} Having developed from clinical epidemiology, it incorporates scientific methods (particularly the methods of population sciences) into clinical decision making. It has driven a transformation of clinical practice and looks to continue this trend into the future.⁵

Clinical research can be divided broadly into four main categories, namely therapeutic, prognostic, diagnostic and economic, in accordance with the Centre for Evidence-Based Medicine in the UK.⁷ Each of these categories are then subdivided into levels of evidence, from I (high) to V (low). Declaration of a research paper's level of evidence at the time of submission to a journal is on the rise and an increasing number of journals are also declaring articles' evidence ratings at the time of publication.⁸ A high degree of interobserver agreement has been demonstrated when the rating system is applied.⁹ A review of levels of evidence across nine orthopaedic journals in 2005 showed that only 32% of the literature contained high level (I and II) and 68% contained low level (III–V) evidence.¹⁰

In the face of rising demand, treatment underpinned by a strong evidence base is preferred and it is therefore prudent for knee surgeons to provide high quality evidence to support the current best practice in their field so as to optimise patient outcomes. The aim of this study was to investigate in detail the levels of evidence published in five major knee journals across a decade to identify whether researchers have responded to calls for higher levels of evidence. We sought to identify the volume and types of studies, levels of evidence, geographical variation of source of articles and interobserver agreement among the reviewers who rated the articles.

Methods

An analysis was performed of articles published in five journals during two time periods separated by ten years: 1 January 2000 to 31 December 2000 and 1 January 2010 to 31 December 2010. Five major journals were selected based on their affiliations with national and international knee societies. The journals were published in the English language, and they had been present in print and online for the whole duration of our included time periods.

The journals selected were: the *Journal of Arthroplasty*, affiliated to the American Association of Hip and Knee Surgeons; *The Knee*, affiliated to the British Association for Surgery of the Knee, the Australian Knee Society and the German Knee Society; and *Knee Surgery, Sports Traumatology, Arthroscopy (KSSTA)*, affiliated to the European Society of Sports Traumatology, Knee Surgery and Arthroscopy. In addition, we included all knee related articles in two journals affiliated to major orthopaedic associations: the *Journal of Bone and Joint Surgery (JBJS)* and the *Bone and Joint Journal (BJJ)*, affiliated to the American Academy of Orthopaedic Surgeons and the British Orthopaedic Association respectively.

In 2013 the *Journal of Bone and Joint Surgery (British Volume)* was renamed the *Bone and Joint Journal*. In this paper, we have decided to simply use 'BJJ', regardless of whether we are referring to the journal before or after its name change.

In order to select the articles to be included in the study, a PubMed search of the aforementioned journals was performed, and all articles relating to knee pathology and its treatment published in the years 2000 and 2010 were selected. Two reviewers (KH and AC) analysed the selected articles independently, and ranked them according to level of evidence and type. Article type was divided into diagnostic, prognostic, therapeutic and economic. The level of evidence was assigned on a scale of I to V. Any disagreements were discussed with the senior authors (TB and AG) and the methodology described by Spindler *et al* was used to reassess such papers.¹¹ A consensus decision was then reached regarding article evidence level or type.

All clinical studies were included. Technical tips, case reports or expert opinion were included but classified as level V evidence. Animal, cadaveric and basic science articles, editorials, surveys, special topics, letters to the editor and correspondence were excluded.

In addition, the lead author (AS) performed a separate search and analysis of level I and II articles from the selected journals for the years 2003, 2005 and 2007. The aim was to further define the trends in high level evidence in the intervening years as well as to help ensure that the results from the main two time periods did not represent anomalies.

Statistical methods

Interobserver agreement was measured using the kappa statistic. Kappa values were assessed using the criteria described by Fleiss.¹²

Fisher's two-sided exact test was used to compare the proportions of the study types and levels of evidence by year of publication as well as by journal. The chi-squared test for trend was used to test for linear trend. Fisher's exact test was also used to examine the proportions of the study types and levels of evidence by year of publication within the five different journals.

All statistical analysis was performed using Stata/IC™ version 12.1 (StataCorp, College Station, TX, US). A *p*-value of <0.05 was considered statistically significant.

Results

Overall, 657 studies were reviewed from 2000 and 2010. After excluding animal, cadaveric and basic science studies (*n*=65), 592 studies were available for analysis.

The kappa value for the interobserver reliability of study type showed excellent agreement between the reviewers ($\kappa=0.932$, $p<0.01$). Interobserver reliability of levels of evidence also showed excellent agreement between the reviewers ($\kappa=0.750$, $p<0.01$).

Additionally, across the intervening years 2003, 2005 and 2007, the five journals included in the study published 1,067 knee surgery related articles.

Level of evidence and study type by year of publication

Between 2000 and 2010, the overall percentage of high level evidence studies (levels I and II) increased and that of low level evidence studies (levels III–V) decreased (Table 1). However, these changes were not statistically significant ($p=0.08$).

There was a significantly higher proportion of level II studies in 2010 than in 2000 (34% vs 24%, $p=0.02$). There was also a significantly lower proportion of level IV studies in 2010 than in 2000 (21% vs 32%, $p=0.01$) (Fig 1).

Table 1 Levels of evidence grouped into high and low evidence by year of publication

	2000	2010	<i>p</i> -value
High level (I–II)	65 (39%)	203 (48%)	0.08
Low level (III–V)	100 (61%)	224 (52%)	
Total	165 (100%)	427 (100%)	

Examination of the intervening years revealed that there was actually a downward trend in high level evidence between 2000 and 2007. A chi-squared test for a linear

trend confirmed this as statistically significant ($p=0.01$). It was not until 2010 that the proportion of high level evidence increased markedly to 48% (Fig 2).

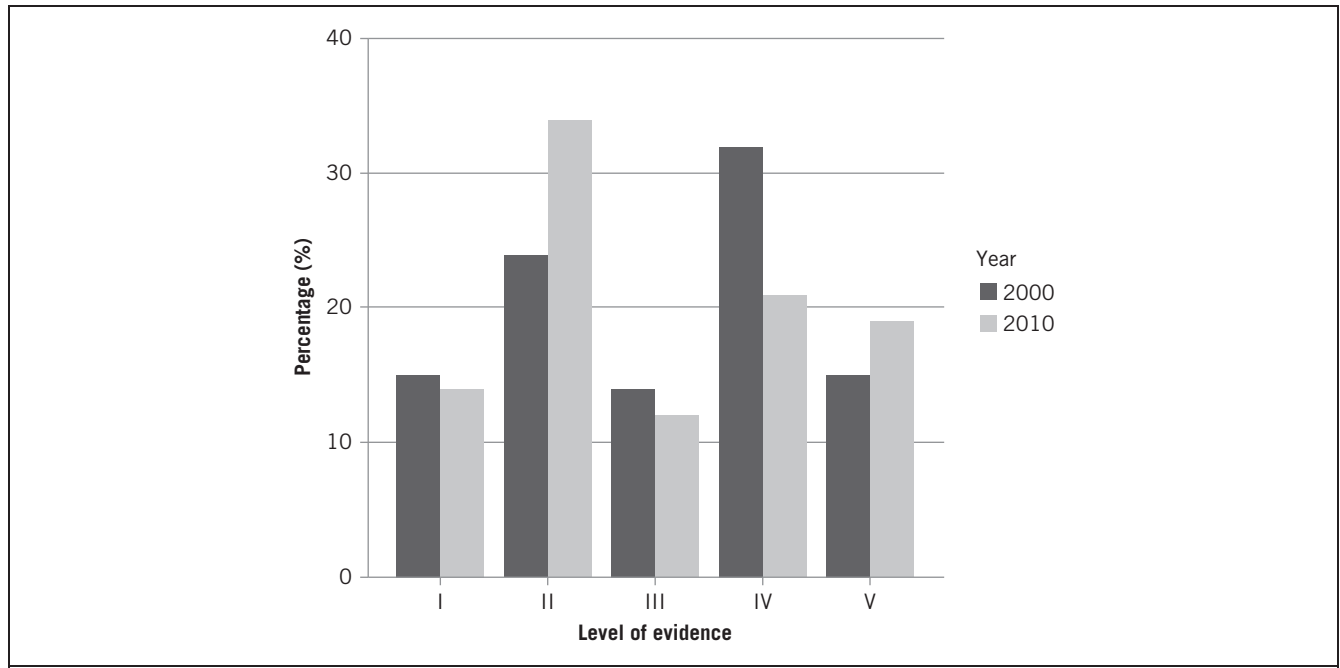


Figure 1 The individual levels of evidence published in the included journals in 2000 and 2010 expressed as a percentage of the total number of articles published in those years

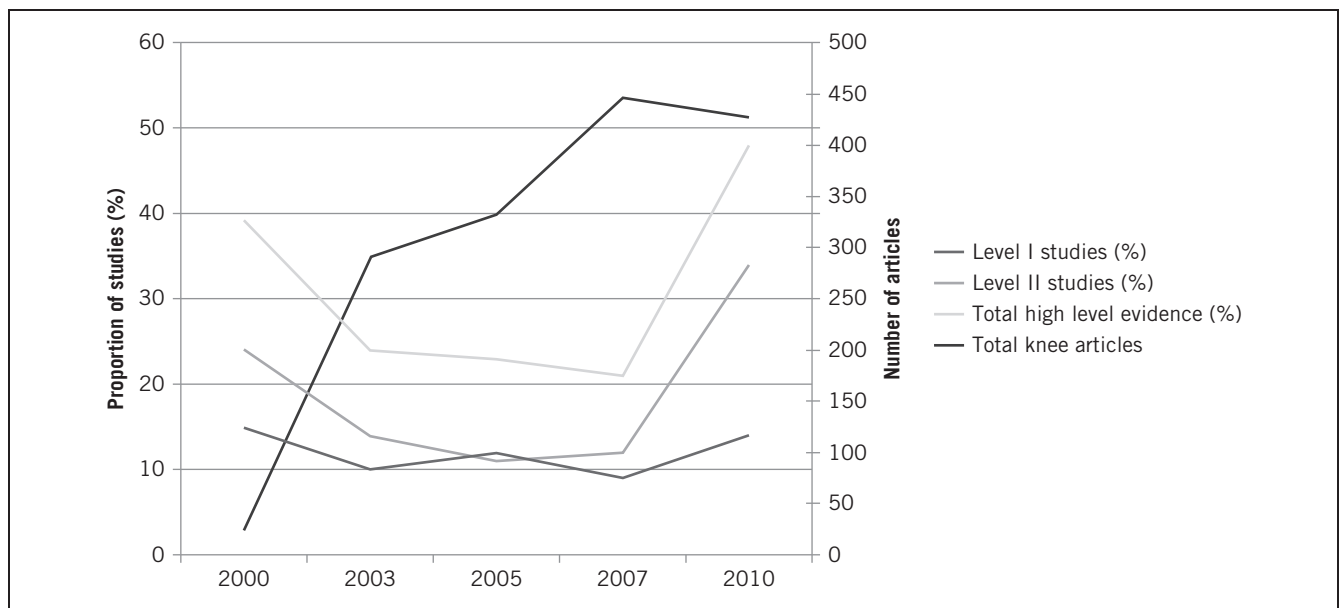


Figure 2 The trends between 2000 and 2010 of published high level evidence as well as the total number of articles published in the included journals

	2000	2010	p-value
Diagnostic	17 (10%)	28 (7%)	0.16
Prognostic	8 (5%)	14 (3%)	0.34
Therapeutic	139 (84%)	382 (89%)	0.09
Economic	1 (1%)	3 (1%)	0.99
Total	165 (100%)	427 (100%)	

No significant differences in the proportions of study types were found between 2000 and 2010 (Table 2). Therapeutic studies were by far the most common.

Level of evidence and study type by year and journal

No statistically significant differences were found in the study types or levels of evidence between 2000 and 2010 for *BJJ* or *KSSTA*. Significant differences occurred in the levels of evidence and study types in the *Journal of Arthroplasty*, *JBJS* and *The Knee*.

In 2010 the *Journal of Arthroplasty* had a greater proportion of level II studies than in 2000 (38% vs 19%, $p=0.04$). *The Knee* also saw a rise in the proportion of level II studies between 2000 and 2010 (from 5% to 28%, $p<0.01$). In 2010, *JBJS* had a significantly lower proportion of level IV studies than in 2000 (10% vs 47%, $p=0.01$). Figure 3 illustrates the contributions of each journal to the number of

high and low level evidence articles in 2000 and 2010 as well as each journal’s contributions to the number of articles as a whole.

Geographical variation

The total number of papers included from the five journals in the study increased from 165 in 2000 to 427 in 2010 (Table 1). The combined contribution from the UK and US was 50% of the total volume in 2000. This figure dropped significantly to 37% for 2010 ($p=0.01$). Conversely, the proportion of papers from countries other than those named individually in Figure 4 increased from 13% to 23% ($p=0.02$). The proportion of total papers from the remaining countries did not differ significantly.

Discussion

This study shows an overall increase in high levels of evidence in the knee literature over the course of a decade although this did not reach statistical significance ($p=0.08$). The rise in the levels of evidence occurred in the latter part of this decade, with evidence levels actually showing a significant downward trend between 2000 and 2007 ($p=0.01$).

The proportion of high level articles in our study was greater than that identified by Obremskey *et al*, who examined a cross-section of the orthopaedic literature over a six-month period in 2003.¹⁰ A third (32%) of the studies constituted high levels of evidence. The only knee journal analysed in that study was the *Journal of Arthroplasty*,

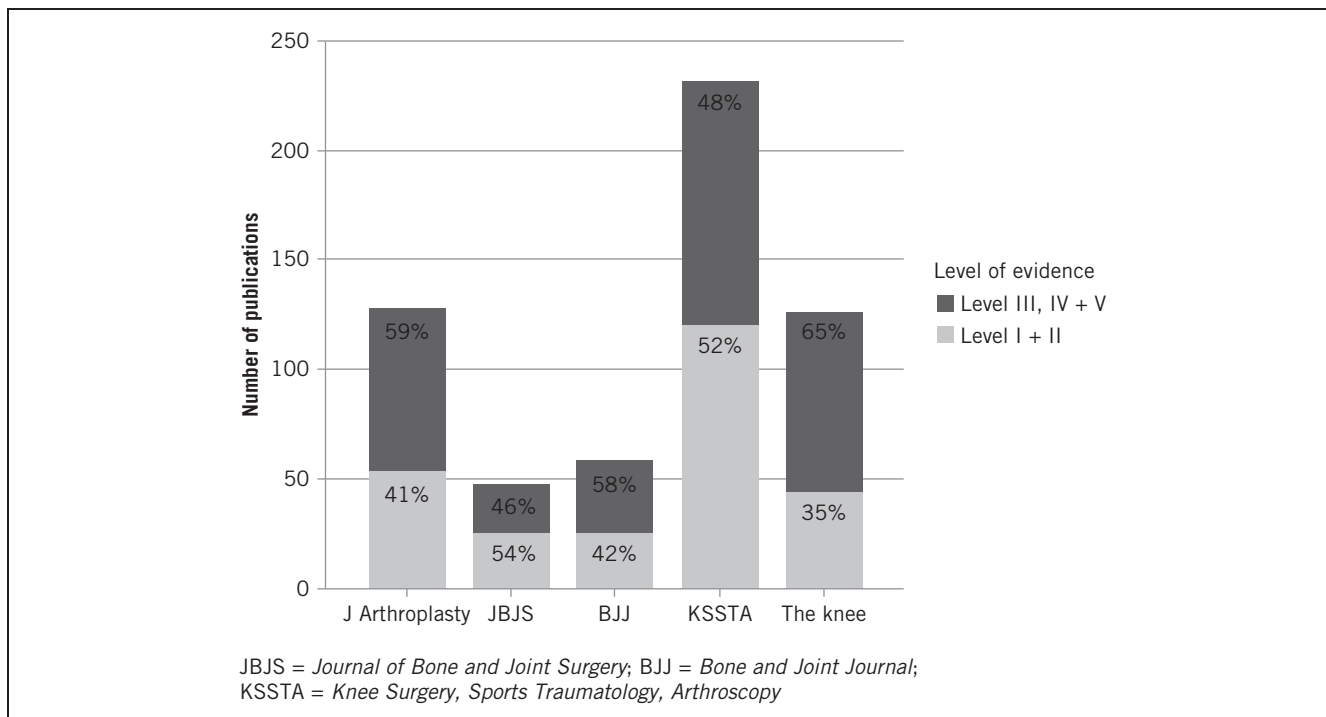


Figure 3 The percentage of high level evidence (I and II) versus low level evidence (III, IV and V) published in each journal across 2000 and 2010. This chart also shows the number of published studies from each journal across our main calendar years of 2000 and 2010.

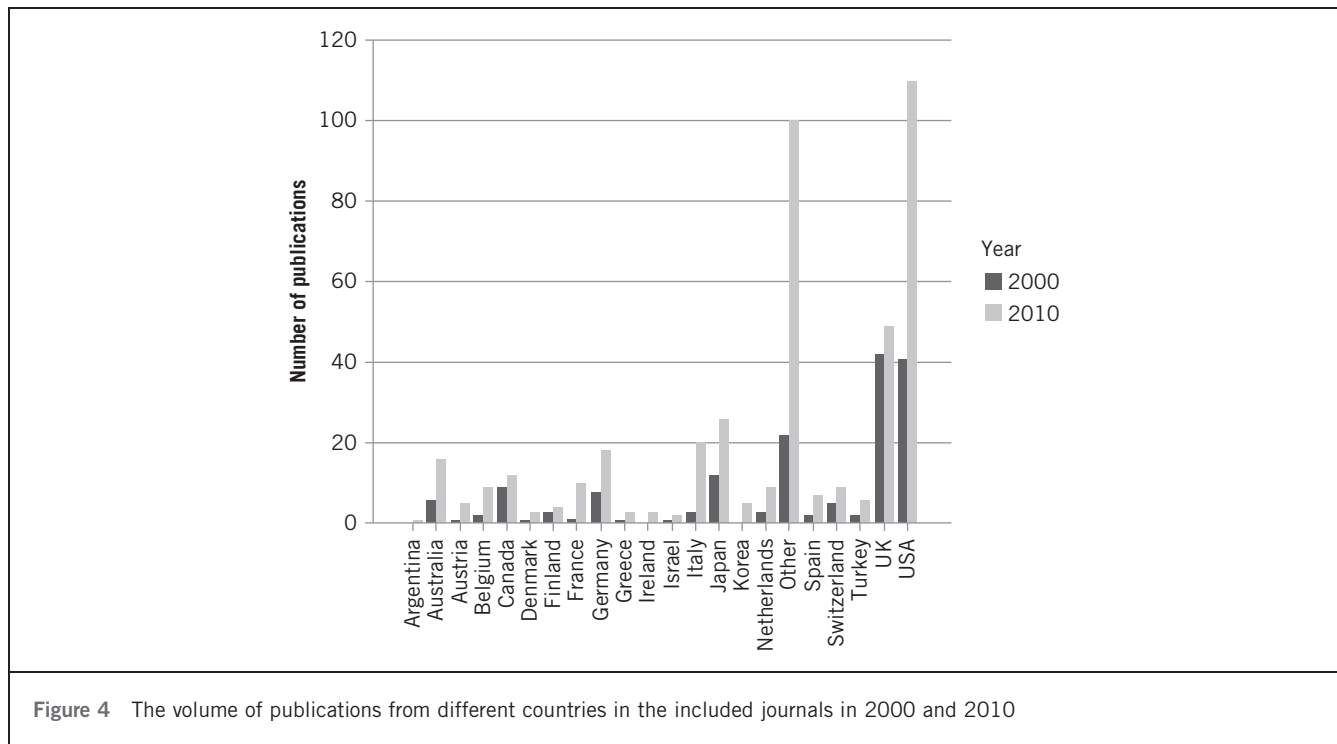


Figure 4 The volume of publications from different countries in the included journals in 2000 and 2010

which had 20% of articles categorised as high levels of evidence. The authors excluded level V evidence articles, however, which were included in our study. In response to this, after excluding level V studies from our analysis, the percentage of high levels of evidence was 46% and 59% in 2000 and 2010 respectively although our data included all papers from the full calendar years.

A study assessing papers on spinal surgery found that 38.4% of papers had high levels of evidence but this also excluded level V studies.¹⁵ The levels of evidence found in our study therefore compare favourably not only with this study but also with similar studies in foot and ankle and paediatric surgery.^{14,15} We believe that the inclusion of level V articles (which represent a large volume of the literature) provides a much more realistic cross-sectional overview of the proportion of evidence types and as a result, it was considered important to include them.

Additionally, the levels of evidence in the literature on knee surgery compare favourably with those for other surgical disciplines. For example, only 5.9% of studies on plastic surgery in 2005 were reported as having a high level of evidence and in 1993, the proportion was 5.3%.¹⁶

The fact that a high percentage of articles were level IV and V (the most common being a therapeutic level IV study) concurs with previous reports on the subject.^{8,17} There are a number of reasons cited for these findings, including ethical barriers limiting the use of sham procedures in surgical research.¹⁸ These can be overcome by proper design and methodology, enabling randomised trials with sham controls to be performed.¹⁹

Another commonly cited barrier to conducting randomised controlled trials (RCTs) in surgical practice is the

variation of presentation and treatment, surgeons fearing that the application of a rigid methodology could undermine the effectiveness of a surgical intervention and impact negatively on the patient's outcome. It is well known that residents are required to publish regularly to achieve steady career progression.¹⁷ Level IV studies are easier to conduct, less expensive and less time consuming than RCTs, and can be performed within the constraints of time and resources available to a trainee. This may also explain the high number of these studies being performed.

A RCT naturally represents the gold standard in clinical research.¹¹ A study looking at the feasibility of conducting RCTs in surgical practice found that at least 40% of clinical research questions were potentially answerable with a RCT²⁰ and a study published in *The Lancet* in 1995 showed that 53% of clinical decisions in acute internal medicine were backed by a RCT.²¹ These findings refute some of the commonly cited obstacles to performing RCTs in clinical practice.

Many of the articles assessed in this study could have been upgraded by the use of a control group and by performing the study prospectively, often without much difficulty to the researching team. This could enable a level IV study to become a level II study. It is therefore vital for surgeons to consider the potential design of their clinical research project carefully. In many cases, it may be easier to perform a high level study than is originally perceived. On occasion, however, circumstances may dictate that level IV is the only research design possible. If well designed, such studies can still be invaluable in answering certain research questions. This also applies to technical tips and case reports of rare pathologies, which can be useful methods for the dissemination of important information.

In a given year, the impact factor of a journal is the average number of citations received per article published in that journal during the preceding two years. It is used as a measure of the importance of a journal in its field. The correlation between high levels of evidence and a journal's impact factor has already been shown,^{15,22} and the results of our study support this notion. According to the Web of Science™ (Thomson Reuters, New York, US), *KSSTA* and *JBJS* had the highest impact factors of the journals analysed in our study. They also produced the majority of the articles with high levels of evidence.

The bulk of the articles in our study fell into the therapeutic category. There were only four articles (1%) in the economic group. The lack of economic studies is a common finding in orthopaedic literature although they can wield considerable influence on resource allocation and healthcare commissioning, particularly in times of financial austerity.⁸

The highest producer of knee articles was *KSSTA*, which along with *JBJS* contributed the highest percentage of articles with high levels of evidence. Both journals have an international readership. The increase in volume of articles over the ten years appears to have come from outside the US and UK, reflecting a wider international readership. Previously, researchers outside the US and UK may have been more likely to publish in local journals; the increase in the volume of publications from these countries reflects the globalisation of the knee literature.

Study limitations

A limitation to this study is that only two main time periods were selected (ie the calendar years 2000 and 2010) and not each intervening year. In a study from 2012 examining the foot and ankle literature, three timepoints were chosen: the years 2000, 2005 and 2010.¹⁵ Choosing an interim timepoint at 2005 did not alter the thrust of the overall message. Similarly, in our study, while additional analysis of level I and II articles for the interim years 2003, 2005 and 2007 helped to further define trends, it did not result in significant changes to our overall findings.

A further limitation of our study is that the reviewers were not blinded to the journal source; this created a potential source of detection bias. In addition, the inclusion of level V articles is likely to have increased the interobserver agreement because such studies are easy to identify. The purpose of this study, however, was to obtain a clear understanding of the levels of evidence and not to establish the validity of interobserver reliability, which has already been verified.⁹ We recommend a future study that ideally includes all journals to give a wider view of the literature. Despite these weaknesses, to our knowledge, this study gives the most detailed assessment of the levels of evidence in the knee literature to date.

The expansion of knee surgery as a subspecialty is reflected in an increase in the volume of published research between 2000 and 2010 (Fig 2). Faced with rising demand and increasing financial pressures, it is important for knee surgeons to have high quality evidence available to optimise levels of care. A study from 2013 has shown that knee

arthroplasty performed using unproven implants with a poor evidence base leads to a higher rate of early revision; this can be expensive for health services and traumatic for patients.⁴ Evidence-based orthopaedics results in better patient outcomes and a reduction in costly variations of practice.⁵

Entering all implants into a database such as a joint registry (eg the National Joint Registry for England, Wales, Northern Ireland and the Isle of Man) will help to ensure that eventually, only those with a strong evidence base are used.⁵ Consequently, the levels of evidence underpinning various procedures and implants will be having a greater impact on their uptake in the near future.

Conclusions

Between 2000 and 2010, the overall trend was towards higher levels of evidence in the knee literature. However, the differences did not reach statistical significance. As the major increase in evidence levels occurred in the latter part of the decade, a further study is recommended in due course to investigate whether this progress has been maintained. In order to stimulate change at a more rapid rate, it is also advised that knee journals consider making the publication of levels of evidence mandatory, and that surgeons take responsibility for developing and designing higher quality studies in the future.

References

1. Kurtz SM, Ong KL, Lau, E, Bozic KJ. Impact of the economic downturn on total joint replacement demand in the United States: updated projections to 2021. *J Bone Joint Surg Am* 2014; **96**: 624–630.
2. Knee Replacement Market Bends in the Right Direction, Despite Dwindling Revenue Growth, says GlobalData. GlobalData. <http://healthcare.globaldata.com/media-center/press-releases/medical-devices/knee-replacement-market-bends-in-the-right-direction-despite-dwindling-revenue-growth-says-globaldata> (cited January 2016).
3. Briggs TW. *Getting It Right First Time*. Stanmore: TWR Briggs; 2012.
4. Ng Man Sun S, Gillott E, Bhamra J, Briggs T. Implant use for primary hip and knee arthroplasty: are we getting it right first time? *J Arthroplasty* 2013; **28**: 908–912.
5. Sackett DL, Rosenberg WM, Gray JA *et al*. Evidence based medicine: what it is and what it isn't. *BMJ* 1996; **312**: 71–72.
6. Belsey J. *What Is Evidence-based Medicine?* 2nd edn. Newmarket: Hayward Medical Communications; 2009.
7. OCEBM Levels of Evidence. Centre for Evidence-Based Medicine. <http://www.cebm.net/index.aspx?o=5653> (cited January 2016).
8. Wright JG, Swiontkowski MF, Heckman JD. Introducing levels of evidence to the journal. *J Bone Joint Surg Am* 2003; **85**: 1–3.
9. Bhandari M, Swiontkowski MF, Einhorn TA *et al*. Interobserver agreement in the application of levels of evidence to scientific papers in the American volume of the *Journal of Bone and Joint Surgery*. *J Bone Joint Surg Am* 2004; **86**: 1,717–1,720.
10. Obrebsky WT, Pappas N, Attallah-Wasif E *et al*. Level of evidence in orthopaedic journals. *J Bone Joint Surg Am* 2005; **87**: 2,632–2,638.
11. Spindler KP, Kuhn JE, Dunn W *et al*. Reading and reviewing the orthopaedic literature: a systematic, evidence-based medicine approach. *J Am Acad Orthop Surg* 2005; **13**: 220–229.
12. Fleiss JL. *Statistical Methods for Rates and Proportions*. 2nd edn. New York: John Wiley; 1981.
13. Wupperman R, Davis R, Obrebsky WT. Level of evidence in Spine compared to other orthopedic journals. *Spine* 2007; **32**: 388–393.
14. Cashin MS, Kelley SP, Douziech JR *et al*. The levels of evidence in pediatric orthopaedic journals: where are we now? *J Pediatr Orthop* 2011; **31**: 721–725.
15. Zaidi R, Abbasian A, Cro S *et al*. Levels of evidence in foot and ankle surgery literature: progress from 2000 to 2010? *J Bone Joint Surg Am* 2012; **94**: e112.

16. Loiselle F, Mahabir RC, Harrop AR. Levels of evidence in plastic surgery research over 20 years. *Plast Reconstr Surg* 2008; **121**: 207e–211e.
17. Hanzlik S, Mahabir RC, Baynosa RC, Khiabani KT. Levels of evidence in research published in The Journal of Bone and Joint Surgery (American Volume) over the last thirty years. *J Bone Joint Surg Am* 2009; **91**: 425–428.
18. Stirrat GM. Ethics and evidence based surgery. *J Med Ethics* 2004; **30**: 160–165.
19. Moseley JB, O'Malley K, Petersen NJ *et al*. A controlled trial of arthroscopic surgery for osteoarthritis of the knee. *N Engl J Med* 2002; **347**: 81–88.
20. Solomon MJ, McLeod RS. Should we be performing more randomized controlled trials evaluating surgical operations? *Surgery* 1995; **118**: 459–467.
21. Ellis J, Mulligan I, Rowe J, Sackett DL. Inpatient general medicine is evidence based. *Lancet* 1995; **346**: 407–410.
22. Hoeffel C. Journal impact factors. *Allergy* 1998; **53**: 1,225.