RESEARCH

Reporting of industry funded study outcome data: comparison of confidential and published data on the safety and effectiveness of rhBMP-2 for spinal fusion

OPEN ACCESS

Mark A Rodgers *research fellow*¹, Jennifer V E Brown *research fellow*¹, Morag K Heirs *research fellow*¹, Julian P T Higgins *professor of evidence synthesis*¹, Richard J Mannion *consultant neurosurgeon*², Mark C Simmonds *research fellow*¹, Lesley A Stewart *director and professor of evidence synthesis*¹

¹Centre for Reviews and Dissemination, University of York, UK; ²Addenbrookes Hospital, Cambridge, UK

Abstract

Objective To investigate whether published results of industry funded trials of recombinant human bone morphogenetic protein 2 (rhBMP-2) in spinal fusion match underlying trial data by comparing three different data sources: individual participant data, internal industry reports, and publicly available journal publications and conference abstracts.

Data collection and synthesis The manufacturer of rhBMP-2 products (Medtronic; Minneapolis, MN) provided complete individual participant data and internal reports for all its studies of rhBMP-2 in spinal fusion. We identified publications and conference abstracts through comprehensive literature searches. We compared outcomes provided in the individual participant data against outcomes reported in publications. For effectiveness outcomes, we compared meta-analyses of randomised controlled trials based on each of the three data sources. For adverse events, meta-analysis of the published aggregate data was not possible and we compared the number and type of adverse events reported between data sources.

Results 32 publications reported outcomes from 11 of the 17 existing manufacturer sponsored studies. For individual randomised controlled trials, 56% (9/16) to 88% (15/17) of effectiveness outcomes known to have been collected were reported in the published literature. Meta-analyses of effectiveness data were almost identical for pain outcomes and similar for fusion across the three data sources. A minority of adverse event data known to have been collected were reported in the published literature. Several journal articles reported only "serious," "related," or "unanticipated" adverse events, without defining these terms. Others reported a small proportion of the collected adverse event

categories. Around 23% (533/2302) of the total adverse events collected in published randomised controlled trials have been reported in the literature, with randomised controlled trials evaluating the licensed preparation (Infuse) reporting around 11% (122/1108) of collected adverse events.

Conclusions The published literature only partially represents the total data known to have been collected on the effects of rhBMP-2. This did not lead to substantially different results for meta-analysis of effectiveness outcomes. In contrast, reporting of adverse event data in trial publications was inadequate and inconsistent to the extent that any systematic review based solely on the publicly available data would not be able to properly evaluate the safety of rhBMP-2. Analysis of individual participant data enabled the most complete, detailed, and in-depth analysis and was not more resource intensive than extracting, collating, and analysing aggregate data from multiple trial publications and conference abstracts. Confidential internal reports presented considerably more adverse event data than publications, and in the absence of individual participant data access to these reports would support more accurate and reliable investigation, with less time and effort than relying on incomplete published data.

Introduction

In 2002, the US Food and Drug Administration approved the use of recombinant human bone morphogenetic protein 2 (rhBMP-2, Infuse; Medtronic) in anterior lumbar interbody fusion as an alternative to traditional iliac crest bone graft techniques for spinal fusion surgery. Since then, use of this

Correspondence to: L A Stewart lesley.stewart@york.ac.uk

Extra material supplied by the author (see <http://www.bmj.com/content/346/bmj.f3981?tab=related#webextra>)

Overview of available sources of data for Medtronic rhBMP-2

Outcomes reported in peer reviewed journal articles compared with Medtronic internal reports

Adverse events occurring in Medtronic studies, as reported in published journal articles

Comprehensive reference list

product in spinal surgery—including in “off label” procedures—has increased rapidly.

Published industry sponsored trials of rhBMP-2 reported clinical benefits and stated that there were no (or no unanticipated) adverse events relating to rhBMP-2 treatment.¹⁻¹² However, from 2006, several non-industry published studies reported potential adverse events associated with the use of rhBMP-2. In 2008, the FDA issued a public health notification of potentially life threatening complications associated with swelling of neck and throat tissue after rhBMP-2 use in the cervical spine.¹³ A subsequent review of publicly available data suggested an increased risk of complications and adverse events for patients receiving rhBMP-2 that was 10 to 50 times higher than the original estimates.¹⁴ However, the authors of that review did not have access to either the original study data or the confidential internal reports for these industry funded studies.

In June 2011, a US Senate Finance Committee investigation of publishing practices was initiated, focusing on the relationships between the sponsor (Medtronic) and clinical authors of published articles. Amid this controversy, the Yale University Open Data Access (YODA) project¹⁵ reached a landmark agreement with Medtronic for the manufacturer to provide full individual participant data and internal reports from all their studies of rhBMP-2 in spinal fusion surgery, and importantly to allow unrestricted independent reanalysis of these data.

We conducted one of the two unrestricted reanalyses commissioned by and funded through Yale, details of which are reported elsewhere.^{15 16} YODA acted as an intermediary, ensuring there was no direct contact between Medtronic and the analysis teams, thereby assuring independence. Our systematic review of the Medtronic individual participant data and one additional randomised controlled trial showed that, compared with iliac crest bone graft, rhBMP-2 resulted in a small statistically significant increase in fusion rate and a modest reduction in pain at two years that fell below clinically important thresholds. Although there were signals of other potential harms including sterility (retrograde ejaculation), numbers of events were small giving rise to considerable uncertainty and inconclusive findings. The systematic review also found evidence that rhBMP-2 increases early post-surgical pain. Evidence of an increased incidence of cancer was inconclusive.^{15 16}

An integral part of our research proposal was to additionally investigate whether published results of the rhBMP-2 trials matched the provided individual participant data and to consider if and how this might impact on a standard systematic review. We therefore compared the results of our individual participant data meta-analysis of the data provided by Medtronic with meta-analyses of summary data extracted from the published literature and from the internal reports provided by the manufacturer.

“Individual participant data” refers to the raw data files (supplied as SAS files in this case) and “internal reports” refers to the written internal Medtronic reports of the trial, containing detailed, tabulated data (but not full case histories) for outcomes and adverse event report forms for each patient.

Methods

We compared three different sources of study data: individual participant data provided by the manufacturer, confidential internal reports, and publicly available journal publications and conference abstracts.

Identifying the evidence

As part of the YODA agreement, Medtronic provided complete individual participant data for all their clinical studies of rhBMP-2 in patients undergoing spinal fusion surgery for treatment of degenerative disc disease, spondylolisthesis, or other relevant spinal conditions. They also provided internal reports, protocols, and citation details for these studies. Data from redacted clinical study reports produced for FDA premarket approval were only available for three trials (see supplementary table 1). In addition, we identified journal publications and conference abstracts for this set of studies through comprehensive literature searches (details reported elsewhere).^{15 16}

Data collection and extraction

We matched all known published papers and conference abstracts to the relevant studies. For each publication we extracted key study and participant characteristics and all reported data on effectiveness outcomes and adverse events. Where study data were reported in more than one publication, we used the most complete dataset aggregated across all known publications. One researcher extracted data and a second reviewer checked the data extraction. We resolved any discrepancies through consensus or recourse to a third reviewer where necessary.

Data synthesis

Using protocol definitions, we compared the outcomes collected in each study and presented in the individual participant data and internal reports against those reported in each study’s publications and abstracts.

Effectiveness outcomes

We conducted separate meta-analyses of outcome data reported in internal reports and in publications and compared these against meta-analyses of the provided individual participant data. Meta-analyses were performed for pain outcomes (Oswestry disability index, SF-36 physical component score, back pain, leg pain) and fusion success for all Medtronic randomised controlled trials comparing rhBMP-2 with iliac crest bone graft. The definitions of these pain scores and of fusion are detailed elsewhere.^{15 16} We analysed these outcomes at six weeks and at three, six, 12, and 24 months after surgery. Meta-analyses for all three data sources were conducted using DerSimonian-Laird random effects models.¹⁷

For each study we calculated the reported mean differences between treatment arms in the change in score from preoperative values, or the relative risk of successful fusion, and then combined these results across studies in “two stage” meta-analyses. Where standard deviations for outcomes were not reported in the published literature we imputed these values by assuming that the standard deviation of pain scores in the rhBMP-2 or iliac crest bone graft arm was the average of the standard deviations from those studies where this information was reported.

Adverse events

Adverse events were defined in Medtronic trial protocols as “any clinically adverse sign, symptom, syndrome, or illness that occurs or worsens during the operative and postoperative periods of the trial, regardless of causality, that is not otherwise being measured in the trial.” Events were classified using Medtronic’s

own non-standardised “in-house” system developed in partnership with the FDA.

Meta-analysis of adverse events was performed using the individual participant data and these analyses are reported in our clinical paper.¹⁶ However, meta-analysis based on data available in journal publications was not possible as data were too sparse for this to be meaningful. Thus a comparison of meta-analyses arising from the different data sources was not feasible. We therefore compared the total number of adverse events reported across study publications with the total number of events collected in the individual participant data. The reported categories of adverse events were also compared between sources.

Results

We received protocols, internal reports, and individual participant data for all 17 of Medtronic’s relevant rhBMP-2 studies (13 randomised controlled trials, four single arm studies). Three of the randomised controlled trials suspended recruitment earlier than planned (fig 1). Our independent bibliographic searches identified a total of 32 journal publications and conference abstracts relating to the Medtronic studies (see supplementary table 1). We also identified two small randomised controlled trials funded independently of the manufacturer but found no evidence of any additional Medtronic studies.^{18 19}

Ten of the 13 randomised controlled trials and one of the four single arm studies have been published in a scientific journal. Extremely limited data for one further randomised controlled trial (rhBMP-2/BCP Canada pivotal RCT) were available from three short conference abstracts.²⁰⁻²² Nine of the published randomised controlled trials were eligible for our meta-analyses of clinical effectiveness. One randomised controlled trial (MAVERICK) was not eligible for inclusion as it compared rhBMP-2 with an ineligible comparator.

Reporting of effectiveness outcomes

Supplementary table 2 shows which outcomes have been published in peer reviewed journals. No individual abstract or journal article reported all of the clinical outcomes known to have been collected. Studies collected a median of 16 (range 11-18) effectiveness outcomes, with a median of nine (range 1-14) outcomes reported in individual peer reviewed publications. Figure 2 shows that even combining information from all journal publications and conference abstracts could not identify a complete set of outcome data for any study (among published studies, 56% to 88% of known collected outcomes were mentioned across publications and abstracts).

The most commonly defined primary outcomes in Medtronic study protocols were radiographic evidence of fusion and measures of pain/disability (Oswestry disability index; SF-36 physical component score; back, hip, and leg pain scores). Around 70% of these outcomes have been presented somewhere in the published literature. Oswestry disability index, SF-36 physical component score, fusion and additional surgery outcomes were available in some form for all 11 studies that have been published. Other frequently collected outcomes were rarely presented in the literature (for example, overall success 14%, patient perceived effect 15%, doctor’s perception of results, 0%). These latter measures were not analysed in our review of effectiveness as we did not a priori identify them as being of clinical importance. However, we do not rule out the possibility of selective reporting bias among these more sporadically reported, but less important outcomes. Detailed

assessment of all studies using the Cochrane risk of bias tool is available elsewhere.¹⁵

Meta-analyses of effectiveness outcomes

Our review of comparative effectiveness was limited to randomised controlled trials.¹⁶ In all nine published randomised controlled trials comparing rhBMP-2 with iliac crest bone graft, data for the main pain outcomes (Oswestry disability index, or neck disability index, SF-36 physical component score, back and leg pain) were available. Estimates of uncertainty in these pain scores (for example, standard errors) were not commonly reported. We therefore chose to impute standard deviations for studies where these data were not provided, as described previously.

Seven published Medtronic randomised controlled trials presented data on the number of successful fusions at six, 12, and 24 months after surgery. The remaining two published data on successful fusion only at 24 months after surgery.

Internal reports and individual participant data were available for the nine relevant published randomised controlled trials plus two further unpublished randomised controlled trials (INTERFIX ALIF pilot, BCP Canada). By contrast with the published evidence, internal reports provided complete data on pain and function scores (including measures of variability) and fusion status at six, 12, and 24 months after surgery. Individual participant data contained some additional follow-up data, extra demographic data, and details on missing outcomes reporting.

Figure 3 shows the results of meta-analyses of published pain and function outcomes against data from internal reports and individual participant data corresponding to the same set of published studies and from complete available data from the individual participant data and internal reports—that is, incorporating additional data from unpublished trials. Where observations were missing for patients in the individual participant data we used multiple imputation to impute plausible observations. This was not possible when analysing the tabulated data in the publications or internal reports as they lacked data on which patients had missing observations. This different handling of missing observations led to the slightly different results seen across data sources. Otherwise, results are almost identical for all four pain and function outcomes at all follow-up times. On this basis, there does not seem to be evidence of a systematic bias in the reporting of pain outcomes in the published literature.

Figure 4 shows the results of meta-analyses for fusion success across the different sources of outcome data. While the results for the different data sources are broadly similar some differences exist. Including unpublished data in the internal report and individual participant data analyses increases relative risks at six months because of one unpublished trial (rhBMP-2/BCP Canada pivotal), which indicated a strongly beneficial result at this time point, with the increase in heterogeneity widening the confidence intervals. This variation was less noticeable at 24 months. The results, however, are generally consistent, with estimates derived from all sources having overlapping confidence intervals. There does not appear to be evidence of a systematic bias in the reporting of fusion in the published literature.

Reporting of adverse event categories

Table 1 shows the adverse event categories as reported in the summary tables of internal reports (one single arm study and 10 randomised controlled trials, of which nine informed our clinical effectiveness analyses). After the initial pilot study

(LT-CAGE pilot), it seems that a common set of adverse event categories were specified across all subsequent studies. Later studies (AMPLIFY, MASTERGRAFT pilot, MAVERICK) reported additional adverse event categories.

For published studies, only a minority of categories that were recorded and reported in the internal reports have been reported in the published literature (table 1). This is summarised in figure 5.

Publications relating to the two most recent randomised controlled trials^{10 23} published in 2009 and 2011 reported a comprehensive table of adverse events similar to that provided in the internal report. Of the earlier trials, three randomised controlled trials reported that there were no “unanticipated” events related to treatment. Although they limited reporting to “serious,” “related,” or “unanticipated” adverse events, these terms were not defined in either the publications or the corresponding clinicaltrials.gov protocols^{3 8 24} (see supplementary table 3). The remaining publications limited reporting to a small proportion of the adverse events noted in the internal reports. There was no apparent consistency in the selection of adverse event categories.

In some cases the way in which adverse events were reported also differed between sources reporting the same type of event. For example, a total of six retrograde ejaculation events occurred in the LT-CAGE open pivotal randomised controlled trials. The journal publication⁵ failed to report that five of these six events occurred in the rhBMP-2 treatment group, whereas this fact was noted in the internal report²⁵ and captured in the individual participant data. The publication also reported this important adverse event as a proportion of all participants (6/146; 4.1%) rather than as a proportion of male participants (6/78; 7.7%).

Several observations identified by other authors as being of serious concern (for example, heterotopic bone formation, osteolysis, radiculitis) were not included as adverse event categories in the main individual participant data datasets or internal reports, even though such observations were considered important enough to suspend recruitment in one study.²⁶ The manufacturer informed us (mediated through YODA) that this was because adverse events were classified on the basis of clinical signs and symptoms rather than by the possible cause of event. However, some journal publications described occurrences that, although not labelled as adverse events, could independently be interpreted as evidence of osteolysis²⁷ or heterotopic bone formation.^{26 28}

For example, Haid et al¹² (INTERFIX PLIF randomised controlled trials) reported a statistically significantly higher rate of “new bone formation extending outside the disc space and into the spinal canal or neuroforamina” in rhBMP-2 treated patients than in controls (24 v 4 events, $P<0.001$), which we interpreted as heterotopic bone formation. This was mentioned in the internal report (box) but not recorded in the summary adverse event table. The “study progress” section of the same report states that “No unanticipated adverse device events have been reported . . . The reports of posterior bone formation are not considered unanticipated adverse device events since this was a possible adverse event listed in the risk analysis and informed consent form.” Medtronic provided the relevant individual participant data from the outset and later confirmed that these data were the result of investigation done outside of the protocol that looked at this occurrence in both arms of the study and were subsequently reported in the Haid et al paper.¹² They confirmed that heterotopic bone formation was systematically reviewed on computed tomography scans by an investigator surgeon and, based on protocol definitions at the

time the studies were run, these radiographic observations of heterotopic bone formation may not have been reported as an adverse event.

Similarly a description of “14 patients with transient localised areas of bone remodelling in the vertebral body adjacent to the allograft dowel” in a bone dowel pivotal randomised controlled trials publication⁷ could be interpreted as possible evidence of osteolysis. Corresponding individual participant data were provided from the outset, again as data beyond the protocol. Medtronic confirmed that adverse events were classified on the basis of signs and symptoms, so potential adverse events such as heterotopic bone formation, osteolysis, and radiculitis not originally listed as adverse event categories in study protocols do not appear in summary tables in any internal reports. They suggested that any clinical consequences of such events may have been classified as “neurological” or as “back and/or leg pain.”

Total number of reported adverse events

We attempted to compare the total number of adverse events that could be obtained from publications with those observed in the individual participant data. However, we note that since the same event may have been reported in more than one publication, our summation of events across publications may overestimate the actual number of reported events. Table 2 shows that, despite this likely over-counting, far fewer total adverse events were reported across publications than were collected in the individual participant data. Around 18.5% of adverse events recorded in the individual participant data across all randomised controlled trials (published and unpublished) have been reported somewhere in the published literature (19% of events with rhBMP-2 and 18% with iliac crest bone graft). Among randomised controlled trials comparing the licensed Infuse preparation of rhBMP-2 against iliac crest bone graft, around 10.5% of adverse events collected in the individual participant data have been reported in the published literature (12% with Infuse, 9% with iliac crest bone graft). Restricting the analysis to published trials alone did not substantially change the proportion of reported individual participant data adverse events (23% across all published studies; 11% across published Infuse studies). Figure 6 shows the total number of adverse events reported in individual participant data and publications for each of these randomised controlled trials.

This suggests any independent systematic review based only on data presented in journal publications and conference abstracts (that is, the public sources available to standard systematic reviews) would not be able to fully evaluate the safety of rhBMP-2 in spinal fusion.

Discussion

Controversy around use of recombinant human bone morphogenetic protein 2 (rhBMP-2) in spinal fusion began with suggestions that inadequate peer review and editorial oversight were responsible for an apparent absence of adverse events from early peer reviewed publications of industry sponsored studies. A subsequent investigation led by the editor in chief of *The Spine Journal*¹⁴ found fewer reported adverse events in academic publications of industry sponsored studies than in related study data available in FDA summaries and public meeting documents. The FDA materials available for that investigation appear to be partial outcome data from a subset of industry funded studies evaluating Infuse/LT-CAGE, Infuse/MASTERGRAFT, and AMPLIFY rhBMP-2 matrix preparations. Our investigation was able to explore this issue in greater depth having individual

Excerpt from Infuse/INTERFIX PLIF internal report

Briefly, in December 1999, Medtronic Sofamor Danek suspended enrolment in this clinical trial. This stoppage occurred mainly because of the results of two patients who developed posterior bone over the cages. The exact cause of the bone formation in these cases is still unknown. Possible contributing factors have been speculated to be the less than ideal placement of the cages, the lack of countersinking performed the use of ADCON-L material during the surgery of one patient, the use of hemostatic agents such as gel foam sponges, and patient selection.

In April 2001, a group of spinal surgeons, including several investigators, evaluated the radiographs of the investigational patients in the clinical trial. They considered six of these patients to potentially have a clinically significant amount of posterior bone formation. The reviewers judged 17 patients to have no or slight bone formation and nine patients to have posterior bone formation but not to a possibly clinically significant amount. A few months later, a group of spinal surgeons reviewed the investigational patients' radiographs again. At this time they noted similar results, with 10 patients having posterior bone formation with some degree of foraminal or canal stenosis, nine patients having posterior bone formation without foraminal stenosis, and 14 patients having no or greater than the expected posterior bone formation.

participant data on all recorded effectiveness and safety outcomes as well as internal reports provided for all Medtronic funded studies of rhBMP-2 for spinal fusion. Given their previous confidentiality, high degree of detail, and availability for all known studies, we consider the combination of individual participant data and internal reports to be the most comprehensive and trustworthy source of Medtronic trial evidence available to date.

Incomplete reporting of trial outcome data

We found that the published scientific literature only partially reported the total outcome data known to have been collected on the effects of rhBMP-2 in spinal surgery in manufacturer sponsored studies. This is partly because some studies were not published or were reported only in brief conference abstracts. For studies that were published, not all collected outcomes were reported. Despite being incomplete, meta-analyses for main effectiveness outcomes using data reported in publications and abstracts did not give substantially different results from meta-analyses using the individual participant data or the internal reports. The reporting of these results does not therefore seem to be substantially biased.

In contrast, we found adverse events to be incompletely and inadequately described in the study publications. In general, published papers provided substantially less information than was available in the internal reports or in the individual participant data. The way in which the adverse event data were presented in the published literature was inconsistent and the rationale for presenting some adverse events and not others was rarely clear. Although broad, non-specific definitions of "related" and "unanticipated" adverse events appeared in confidential Medtronic protocols, it is not clear how these definitions were operationalised and the meaning of these terms was not given in any publication. Brief statements that simply noted "no unanticipated adverse events related to the use of Infuse" were uninformative and, in our opinion, manuscripts making such statements, without supporting evidence, should not have been considered acceptable for publication. Although the proportion of reported events was similar across treatment arms, incomplete reporting in publications precluded a comparative meta-analysis across sources. The academic literature alone is therefore an inadequate resource for evaluating the safety of rhBMP-2. This is not unique to these Medtronic studies and several authors have noted poor reporting of adverse events from studies in the academic literature more widely.²⁹⁻³²

Some potential adverse events that have caused concern in the post-marketing period (for example, heterotopic bone formation, radiculitis, osteolysis) were not routinely collected or reported in the original Medtronic studies. In one internal report, even though apparent heterotopic bone formation led to study suspension—with substantially more cases in the rhBMP-2 arm—this was noted but not formally labelled as an adverse event in the internal report or subsequent academic publication.¹²

When this was queried with the manufacturer, the reasons given were that such events were not unanticipated and that adverse data collection was based on signs and symptoms. Again, this type of approach to reporting may not be unusual, but we suggest that in general any events of sufficient importance to suspend a trial or warrant additional analyses should be openly discussed in trial publications and reports.

Under-reporting of outcome data may not be intentionally biased, but failure to report outcomes fully can undermine confidence in the reported findings and leaves study authors open to accusations of deliberate concealment. Given the large number of adverse events observed in most trials, study authors often discuss only those events considered to be of greatest clinical importance. However, modern electronic publishing methods enable these discussions to be supplemented by complete data on adverse events. We do not expect authors to discuss every single event or type of event; rather we would still expect them to draw the reader's attention to the most important events (supported by a clearly stated rationale) within this summary data. Going further and routinely allowing independent scrutiny of trial data would not only protect sponsors and authors from accusations of concealment, but might also lead to earlier detection of harms.

Our investigation cannot shed light on the decisions or motivations of the authors of rhBMP-2 studies; we can conclude only that trials collected a substantial amount of adverse event data that were not reported in the journal publications. A recently published US Senate Finance Committee investigation,³³ of which Medtronic strongly rejected many conclusions, was critical of publication practices. Assertions that "Medtronic was heavily involved in drafting, editing, and shaping the content of medical journal articles authored by its physician consultants who received significant amounts of money through royalties and consulting fees from Medtronic" underscore the importance of full disclosure of author contributions and full declaration of conflict of interest when reporting all clinical trials.

Encouragingly, a complete list of adverse events similar to the tabulated adverse event data reported in the corresponding internal report has been published for the two largest rhBMP-2 trials, key publications of which also disclose financial conflicts of interest.^{10 23} Release of adverse event data is important because in practice clinicians and patients do not currently have access to information made available for regulatory purposes and so must make decisions weighing benefits and harms of interventions based on what is publicly available in the peer reviewed literature. We hope that this approach to disclosure will become routine in future.

Improving access to trial data

Increasing access to clinical trial data is currently the topic of intense debate. The *BMJ* has stated that it will no longer publish any trial of drugs or devices where the authors do not commit to making the relevant anonymised patient level data available,

and the AllTrials initiative (<http://www.alltrials.net/>) currently has over 53 000 signatures. There have also been moves toward greater openness and transparency from the European Medicines Agency and by the pharmaceutical multinational GlaxoSmithKline, while in the UK a House of Common's Science and Technology Committee inquiry on clinical trials is ongoing.

As part of this discussion, Medtronic should be commended for taking the unprecedented step of voluntarily submitting their data to independent scrutiny, and for their commitment, beyond the two independent investigations commissioned by Yale University Open Data Access (YODA), to making their data available for further independent analysis. While our findings suggest that trial reporting practices have historically been inadequate, by allowing such an analysis the manufacturer has made a welcome move towards greater openness and transparency.

With suitable safeguards to protect patient privacy and transparency around vested and potential conflict of interest, access to participant level data for research purposes should be the ultimate goal of initiatives to increase access to trial data. Contrary to common expectation, our analysis of the individual participant data provided proved to be no more resource intensive than collecting, extracting, and synthesising data from multiple overlapping journal publications and conference abstracts for each trial. We note that here we did not have to negotiate provision of data from trial investigators, which can be a time consuming first step for systematic reviews of individual participant data done under more usual circumstances. The individual participant data were relatively straightforward to check¹⁶ and also enabled us to provide a more complete, detailed, and in-depth analysis than was possible from any other source. For example, individual participant data allowed us to investigate treatment by subgroup interactions across all included trials, whereas this was precluded by the summary and/or incomplete nature of the other data sources. However, we cannot entirely exclude the possibility of errors in the provided individual participant data; a "forensic" examination of data collection beyond standard data checking techniques was outside the aims and resources of this particular investigation.

The findings of this and other reviews³⁴ suggest that even in the absence of individual participant data, access to complete study protocols and the reports that are produced by manufacturers for regulatory or other purposes also produce more accurate, reliable, and robust findings with less time and effort than relying on publicly available data alone. Release of internal reports could therefore be an important and potentially almost immediately achievable first step toward increasing transparency of trials.

The debate around disclosure and data release will undoubtedly continue with further discussion of if, how, and to whom information and data should be released. The YODA initiative has been "proof of concept" showing that manufacturers can release data and that the systematic reviews based on these data are more complete and reliable than systematic reviews restricted to data presented in journal publications.

Contributors: MAR contributed to the development of the study protocol, eligibility screening, data extraction, and assessment of risk of bias, and conducted analyses, including adverse event comparisons. He drafted and edited the manuscript. JVEB contributed to eligibility screening, data extraction, assessment of risk of bias, and the mapping of the adverse events publications. She contributed to the writing of the manuscript. MKH contributed to the management and checking of data

provided by the manufacturer. She was involved in the eligibility screening, data extraction, and risk of bias assessment processes, and contributed to the writing of the manuscript. JPTH contributed to the development of the proposal and protocol, undertook some of the analyses for safety, provided methodological advice, and contributed to the writing of the manuscript. RJM contributed to the development of the proposal and protocol, provided advice on clinical significance and context, and contributed to the writing of the manuscript. MCS was responsible for the management and checking of the data provided by the manufacturer. He conducted the effectiveness comparisons and contributed to the writing of the manuscript. LAS devised the project and developed the proposal with input from co-applicants. She was responsible for the overall direction and supervision of the project, contributed to the writing of the manuscript, and is the guarantor.

Funding: This study was supported by the Yale University Open Data Access project. The funders had no role in study design, analysis, decision to publish, or preparation of the manuscript.

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: grant funding from Yale University for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Ethical approval: Not required.

Data sharing: As part of the broader YODA project, Medtronic have agreed to provide clinical research data on rhBMP-2 to external investigators on request. Details of the data release policy are available from the YODA website at http://medicine.yale.edu/core/projects/yodap/medtronic_data/medtronic_data.aspx.

- 1 Baskin DS, Ryan P, Sonntag V, Westmark R, Widmayer M. A prospective, randomized, controlled cervical fusion study using recombinant human bone morphogenetic protein-2 with the CORNERSTONE-SR allograft ring and the ATLANTIS anterior cervical plate. *Spine (Phila Pa 1976)* 2003;28:1219-24.
- 2 Boakye M, Mummaneni PV, Garrett M, Rodts G, Haid R. Anterior cervical discectomy and fusion involving a polyetheretherketone spacer and bone morphogenetic protein. *J Neurosurg Spine* 2005;2:521-5.
- 3 Boden SD, Kang J, Sandhu H, Heller J. Use of recombinant human bone morphogenetic protein-2 to achieve posterolateral lumbar spine fusion in humans: a prospective, randomized clinical pilot trial: 2002 Volvo Award in clinical studies. *Spine (Phila Pa 1976)* 2002;27:2662-73.
- 4 Boden SD, Zdeblick T, Sandhu H, Heim S. The use of rhBMP-2 in interbody fusion cages. Definitive evidence of osteoinduction in humans: a preliminary report. *Spine (Phila Pa 1976)* 2000;25:376-81.
- 5 Burkus J, Gornet M, Dickman C, Zdeblick T. Anterior lumbar interbody fusion using rhBMP-2 with tapered interbody cages. *J Spinal Disord Tech* 2002;15:337-49.
- 6 Burkus J, Heim S, Gornet M, Zdeblick T. Is INFUSE bone graft superior to autograft bone? An integrated analysis of clinical trials using the LT-CAGE lumbar tapered fusion device. *J Spinal Disord Tech* 2003;16:113-22.
- 7 Burkus J, Sandhu HS, Gornet MF, Longley MC. Use of rhBMP-2 in combination with structural cortical allografts: clinical and radiographic outcomes in anterior lumbar spinal surgery. *J Bone Joint Surg Am* 2005;87:1205-12.
- 8 Burkus KJ, Transfeldt E, Kitchel S, Watkins R, Balderston R. Clinical and radiographic outcomes of anterior lumbar interbody fusion using recombinant human bone morphogenetic protein-2. *Spine (Phila Pa 1976)* 2002;27:2396-408.
- 9 Dawson E, Bae H, Burkus J, Stambough J, Glassman S. Recombinant human bone morphogenetic protein-2 on an absorbable collagen sponge with an osteoconductive bulking agent in posterolateral arthrodesis with instrumentation. A prospective randomized trial. *J Bone Joint Surg Am* 2009;91:1604-13.
- 10 Dimar J, Glassman S, Burkus J, Pryor P, Hardacker J, Carreon L. Clinical and radiographic analysis of an optimized rhBMP-2 formulation as an autograft replacement in posterolateral lumbar spine arthrodesis. *J Bone Joint Surg Am* 2009;91:1377-86.
- 11 Glassman SD, Dimar J, Burkus K, Hardacker JW, Pryor PW, Boden SD, et al. The efficacy of rhBMP-2 for posterolateral lumbar fusion in smokers. *Spine (Phila Pa 1976)* 2007;32:1693-8.
- 12 Haid R, Branch C, Alexander J, Burkus J. Posterior lumbar interbody fusion using recombinant human bone morphogenetic protein type 2 with cylindrical interbody cages. *Spine J* 2004;4:527-38.
- 13 Center for Devices and Radiological Health. Public Health Notifications (Medical Devices)—FDA Public Health Notification: life-threatening complications associated with recombinant human bone morphogenetic protein in cervical spine fusion. www.fda.gov/MedicalDevices/Safety/AlertsandNotices/PublicHealthNotifications/ucm062000.htm.
- 14 Carragee EJ, Hurwitz EL, Weiner BK. A critical review of recombinant human bone morphogenetic protein-2 trials in spinal surgery: emerging safety concerns and lessons learned. *Spine J* 2011;11:471-91.
- 15 Yale School of Medicine. YODA project: Yale University Open Data Access (YODA) project. 2013. <http://medicine.yale.edu/core/projects/yodap/index.aspx>.
- 16 Simmonds M, Brown J, Heirs M, Higgins JPT, Mannion RJ, Rodgers MA, et al. The safety and effectiveness of recombinant human bone morphogenetic protein-2 (rhBMP-2) for

What is already known on this topic

Concern is increasing about non-reporting of clinical study outcomes, particularly from industry funded studies

Much controversy around the use of recombinant human bone morphogenetic protein 2 (rhBMP-2) for spinal fusion has stemmed from this problem

Although systematic reviews are widely regarded as providing best evidence to inform decision making, they rely on the integrity of the available information and data from included studies in reaching reliable conclusions

What this study adds

This investigation compared a systematic review based on publications from industry funded rhBMP-2 studies with systematic reviews of their complete underlying study data

Although published outcome data were incomplete, this did not lead to substantially different results for meta-analysis of effectiveness outcomes, but adverse event data were under-reported and did not permit meta-analysis

While availability of individual participant data for research purposes remains the ideal, access to industry funded study protocols and reports could be an important first step toward increasing transparency of trials conducted for regulatory purposes

spinal fusion: An individual participant data meta-analysis. *Ann Intern Med* 2013;158:877-89.

- 17 DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986;7:177-88.
- 18 Shapiro S, Rodgers RB, Sloan R, Altstadt T, Miller J. A randomized trial of instrumented posterior lumbar interbody fusion using machined cortical wedges/local bone with or without rhBMP2 in the treatment of degenerative lumbar spondylolisthesis with stenosis. *Neurosurgery* 2005;57:398.
- 19 Glassman SD, Carreon L, Dimar JR. Outcome of lumbar arthrodesis in patients sixty-five years of age or older. *J Bone Joint Surg Am* 2010;92A:77-84.
- 20 Abraham E, Alexander D, Bailey S, Hurlbert J, McBroom R, Mahood J, et al. A long-term radiographic and clinical evaluation of a new rhBMP-2 formulation in a prospective randomized lumbar posterolateral spine fusion study. *Can J Surg* 2008;51(3 Suppl):S6-7.
- 21 Alexander D, Bailey S, Hurlbert RJ, Abraham E, McBroom R, Mahood J, et al. A long-term radiographic and clinical evaluation of a new rhBMP-2 formulation in a prospective randomized lumbar posterolateral spine fusion study. *Neurosurgery* 2007;61:200-1.
- 22 Abraham EP, Hurlbert J, Alexander D, Bailey S, Fisher C. Evaluation of an rhBMP-2 formulation in 2-level posterolateral lumbar spine arthrodesis. *Spine J* 2010;10(9 Suppl 1):105S.
- 23 Gornet MF, Burkus JK, Dryer RF, Pelozo JH. Lumbar disc arthroplasty with MAVERICK disc versus stand-alone interbody fusion: a prospective, randomized, controlled, multicenter investigational device exemption trial. *Spine (Phila Pa 1976)* 2011;36:E1600-11.
- 24 Burkus JK, Gornet MF, Schuler TC, Kleeman TJ, Zdeblick TA. Six-year outcomes of anterior lumbar interbody arthrodesis with use of interbody fusion cages and recombinant human bone morphogenetic protein-2. *J Bone Joint Surg Am* 2009;91:1181-9.
- 25 Medtronic. A clinical investigation of recombinant human bone morphogenetic protein-2 and absorbable collagen sponge with the NOVUSTM L T device for anterior lumbar interbody fusion in patients with symptomatic degenerative disc disease [Final report] [received from Medtronic 1 Dec 2011], 2002.
- 26 Medtronic. A prospective, randomized clinical investigation of recombinant human bone morphogenetic protein-2 and absorbable collagen sponge with the INTER FIX™ device for posterior lumbar interbody fusion in patients with degenerative disc disease [Final report] [Received from Medtronic 1 Dec 2011], 2002.

- 27 Medtronic. A clinical investigation of recombinant human bone morphogenetic protein-2 and absorbable collagen sponge with allograft bone dowel for anterior lumbar interbody fusion in patients with symptomatic degenerative disc disease [Final report] [Received from Medtronic 1 Dec 2011], 2004.
- 28 Medtronic. A pilot prospective, randomized, clinical investigation of recombinant human bone morphogenetic protein-2 and absorbable collagen sponge with the CORNERSTONE-SR™ allograft ring and the ATLANTISTM anterior cervical plate system for the treatment of patients with degenerative cervical disc disease [Final report] [Received from Medtronic 1 Dec 2011], 2003.
- 29 De Vries T, van Roon E. Low quality of reporting adverse drug reactions in paediatric randomised controlled trials. *Arch Dis Child* 2010;95:1023-6.
- 30 Pitrou I, Boutron I, Ahmad N, Ravaud P. Reporting of safety results in published reports of randomized controlled trials. *Arch Intern Med* 2009;169:1756-61.
- 31 Sinha S, Sinha S, Ashby E, Jayaram R, Grocott M. Quality of reporting in randomized trials published in high-quality surgical journals. *J Am Coll Surg* 2009;209:565-71.
- 32 Breau R, Gaboury I, Scales CJ, Feserman S, Watterson J, Dahm P. Reporting of harm in randomized controlled trials published in the urological literature. *J Urol* 2010;183:1693-7.
- 33 Baucus-Grassley investigation into medtronic reveals manipulated studies, close financial ties with researchers. 2012. www.finance.senate.gov/newsroom/chairman/release/?id=b1d112cb-230f-4c2e-ae55-13550074fe86.
- 34 Doshi P, Jefferson T, Del Mar C. The imperative to share clinical study reports: recommendations from the Tamiflu experience. *PLoS Med* 2012;9:e1001201.

Accepted: 18 June 2013

Cite this as: *BMJ* 2013;346:f3981

This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 3.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/3.0/>.

Tables

Table 1 | Reporting of adverse events in internal reports and journal articles for published Medtronic studies

Adverse event as reported	Study (date of first patient enrolment)										
	LT-CAGE pilot RCT (Jan 1997)	Bone dowel pilot RCT (Apr 1998)	LT-CAGE open RCT (Aug 1998)	LT-CAGE laparoscopic (Nov 1998)	Interfix PLIF RCT (Mar 1999)	BCP US RCT (May 1999)	CORNERSTONE pilot RCT (Sep 1999)	Bone dowel pivotal RCT (Sep 2000)	AMPLIFY RCT (Mar 2002)	MASTERGRAFT pilot RCT (Apr 2003)	MAVERICK RCT (Apr 2003)
Wound dehiscence	Yes	—	—	—	—	—	—	—	—	—	—
Back pain	Yes	—	—	—	—	—	—	—	—	—	Yes
Trauma	Yes	No	No	No	No	No	No	No	Yes	No	Yes
Back strain	No	—	—	—	—	—	—	—	—	—	—
Endcap event	No	—	—	—	—	—	—	—	—	—	—
Urinary retention	Yes	—	—	—	—	—	—	—	—	—	—
Fracture	No	—	—	—	—	—	—	—	—	—	—
Pseudoarthrosis	No	—	—	—	—	—	—	—	—	Pub*	—
Shortness of breath	No	—	—	—	—	—	—	—	—	—	—
Numbness/SLA	No	—	—	—	—	—	—	—	—	—	—
Drug use	No	—	—	—	—	—	—	—	—	—	—
Facet joint pain	No	—	—	—	—	—	—	—	—	—	—
Rectal bleeding/haemorrhoids	No	—	—	—	—	—	—	—	—	—	—
Anatomical/technical difficulty	—	No	Yes	Yes	—	No	—	No	Yes	—	Yes
Back and/or leg pain	—	No	No	No	No	Yes†	No	No	Yes	No	Yes
Cancer	—	No	No	No	—	No	—	No	Yes	No	Yes
Cardiovascular	—	No	No	No	No	No	No	No	Yes	No	Yes
Death	—	No	No	No	No	No	No	—	Yes	—	Yes
Dural injury	—	No	No	No	Yes	Yes†	—	—	Yes	Yes	—
Gastrointestinal	—	No	No	No	No	No	No	No	Yes	No	Yes ‡
Graft site related	—	No	No	No	Yes	No	—	No	Yes	No	—
Implant displacement/loosening	—	No	Yes	Yes	—	No	—	No	Yes	No	Yes§
Infection	—	No	Yes	No	No	No	No	No	Yes	No	Yes
Malpositioned implant	—	No	Yes	No	—	No	—	—	Yes	Yes	Yes§
Neurological	—	No	No	No	Yes	No	No	No	Yes	No	Yes
Non-union	—	No	Yes	No	No	No	—	No	Yes	No	Yes
Other	—	No	No	No	No	No	No	No	Yes	No	Yes
Other pain	—	No	No	No	No	No	No	No	Yes	No	Yes
Respiratory	—	No	No	No	No	No	No	No	Yes	No	Yes
Retrograde ejaculation	—	No	Yes	No	—	No	—	—	—	—	Yes
Spinal event	—	No	No	No	No	No	No	No	Yes	No	Yes
Subsidence	—	No	Yes	Yes	—	No	—	No	—	—	Yes
Urogenital	—	No	No	No	No	No	No	No	Yes	—	Yes
Vascular intra-op	—	No	Yes	No	No	—	No	—	—	Yes	—
Vertebral fracture	—	No	No	—	—	No	—	—	Yes	—	Yes
Neck and arm pain	—	—	—	—	—	—	No	—	—	No	Yes
Dysphagia/dysphonia	—	—	—	—	—	No	No	—	—	—	—
Headache	—	—	—	—	—	—	No	—	—	—	—
Accidental injury/muscle strain	—	—	—	—	—	—	—	—	—	No	—

Table 1 (continued)

Adverse event as reported	Study (date of first patient enrolment)											
	LT-CAGE pilot RCT (Jan 1997)	Bone dowel pilot RCT (Apr 1998)	LT-CAGE open RCT (Aug 1998)	LT-CAGE laparoscopic (Nov 1998)	Interfix PLIF RCT (Mar 1999)	BCP US RCT (May 1999)	CORNERSTONE pilot RCT (Sep 1999)	Bone dowel pivotal RCT (Sep 2000)	AMPLIFY RCT (Mar 2002)	MASTERGRAFT pilot RCT (Apr 2003)	MAVERICK RCT (Apr 2003)	
Allergic reaction	—	—	—	—	—	—	—	—	—	No	Yes	
Cervical spinal event	—	—	—	—	—	—	—	—	Yes	No	Yes	
Electrolyte imbalance	—	—	—	—	—	—	—	—	—	No	—	
Endocrine	—	—	—	—	—	—	—	—	—	No	—	
Incision related	—	—	—	—	—	—	—	—	—	No	Yes	
Lower back pain not of back etiology	—	—	—	—	—	—	—	—	—	No	—	
Spinal event at other lumbar level(s)	—	—	—	—	—	—	—	—	Yes	No	Yes	
Depression	—	—	—	—	—	—	—	—	—	—	Yes	
Elevated temperature without infection	—	—	—	—	—	—	—	—	—	—	Yes	
Arthritis/bursitis	—	—	—	—	—	—	—	—	Yes	—	—	
Carpal tunnel	—	—	—	—	—	—	—	—	Yes	—	—	
Leg pain	—	—	—	—	—	Pub†	—	—	—	—	Yes	
Lower extremity pain	—	—	—	—	—	—	—	—	—	—	Yes	
Pain management control	—	—	—	—	—	—	—	—	—	—	Yes	
Peritoneal tear	—	—	—	—	—	—	—	—	—	—	Yes	
Rash	—	—	—	—	—	—	—	—	—	—	Yes	
Spinal event at target level	—	—	—	—	—	—	—	—	—	—	Yes	
Spinal event at thoracic level	—	—	—	—	—	—	—	—	Yes	No	Yes	
Upper extremity pain	—	—	—	—	—	—	—	—	—	No**	Yes	
Vascular event post-op	—	—	—	—	—	—	—	—	—	—	Yes	
Spinal event at other cervical level	—	—	—	—	—	—	—	—	Yes	—	—	
Wound infection	—	—	—	—	—	—	—	—	—	Pub*	—	
Total reported in internal reports	13	23	23	22	15	24	15	18	26	27	38	

RCT=randomised controlled trial; Yes=reported in internal report and at least one full academic journal publication; No=only reported in internal report (clinical studies report) table of adverse events; Pub=only reported in a publication (specific events may have been subsumed under other categories in summary reports—for example, wound infection under “infections,” pseudarthrosis under “non-union (pending).”

LT-CAGE pilot reports fracture, other studies report vertebral fracture.

Three Medtronic randomised controlled trials (INTERFIX ALIF pilot, Cornerstone pivotal, BCP Canada) and three non-randomised studies (Telamon, CRM Two-level, BCP Mexico) have not been published to date.

*Dawson et al⁹ report pseudarthrosis in two patients, and three wound infections (two surgical site, one graft donor site).

†Boden et al³ report one transient leg pain after surgery, one epidural haematoma, and one persistent low back/leg pain among patients receiving rhBMP-2.

‡Split ilium and other.

§Grouped together.

¶Haid et al¹² report one deep vein thrombosis as a vascular complication.

**Not of neck cause.

Table 2| Comparison of number of adverse events reported in Medtronic individual participant data versus Medtronic publications

Publications	All studies (n=17)		All published studies (n=11)		Infuse studies (n=12)		Published Infuse studies (n=9)	
	IPD events	Published events* (% of IPD)	IPD events	Published events* (% of IPD)	IPD events	Published events* (% of IPD)	IPD events	Published events* (% of IPD)
Single arm studies†:								
rhBMP-2	733	169 (23.1)	588	169 (28.7)	661	169 (25.5)	588	169 (28.7)
Randomised controlled trials:								
rhBMP-2	1517	288 (19.0)	1222	288 (23.6)	621	75 (12.1)	592	75 (12.7)
ICBG	1367	245 (17.9)	1080	245 (22.7)	541	47 (8.7)	516	47 (9.1)

IPD=individual participant data; rhBMP-2=recombinant human bone morphogenetic protein 2; ICBG=iliac crest bone graft.

†Includes rhBMP-2 (Infuse) comparison arm from the MAVERICK study.

*"Published" values reflect maximum possible number of reported events summed across publications; some events may have been double counted.

When calculating number of adverse events reported across publications: total number of adverse events were counted, not just "device related" or other subsets; potential adverse events (for example, heterotopic bone formation, osteolysis) were counted if they were reported, regardless of whether they were explicitly referred to as "adverse events" or not; number of events or the number of participants experiencing events were counted, whichever number was larger; where two publications reported on exactly the same study population but the reporting of adverse events differed, these values from each publication were summed; publications reporting on subgroups of the same study population were not counted if there was a publication also available that reported adverse events for the entire study population; where the number of events was not reported according to treatment arm, they were assumed to be distributed evenly between arms.

Figures

Fig 1 Medtronic studies included in analysis

Fig 2 Proportion of measured effectiveness outcomes available from journal publications and conference abstracts

rhBMP-2=recombinant human bone morphogenetic protein 2; ICBG=iliac crest bone graft

Fig 3 Meta-analyses of pain and function across three different data sources. rhBMP-2=recombinant human bone morphogenetic protein 2; ICBG=iliac crest bone graft

Fig 4 Meta-analyses of successful fusion across three data sources. rhBMP-2=recombinant human bone morphogenetic protein 2; ICBG=iliac crest bone graft

Fig 5 Number of adverse event categories reported in Medtronic internal reports and journal publications by date of first patient enrolment

Fig 6 Total number of adverse events reported in Medtronic randomised controlled trials comparing recombinant human bone morphogenetic protein 2 (rhBMP-2) with iliac crest bone graft (ICBG): individual participant data (IPD) versus publications