Evaluation of intervertebral body implant performance using active surveillance of electronic health records

Edward A Frankenberger,1 Frederic S Resnic,1,2 Henry Ssemaganda,1 Susan Robbins,1 Melissa R Dunbar,1,3 Paul Coplan,4 Shumin Zhang,4 Cortney Bruno,5 Mitchell Maltenfort,5 Jillian B Benedetti,6 Michael E Matheny,7 Zoher Ghogawala1,3,8

ABSTRACT

Objectives To assess the feasibility of using electronic health record (EHR) derived clinical data within an active surveillance setting to evaluate the safety of a novel intervertebral body implant (IVBI) stabilization device.

Design Retrospective, longitudinal observational cohort study comparing clinical outcomes for patients seen through 1 year following spinal fusion surgery.

Setting Lahey Health network, which includes academic tertiary hospitals, outpatient clinics, and independent provider offices in the New England region of the USA.

Participants All spine surgery patients aged 18 or older who underwent thoracic or lumbar spinal arthrodesis surgeries were included.

Main outcome measures The clinical outcomes of patients treated with the CONCORDE Bullet (CB) interbody spine system (DePuy) between April 2015 and December 2018 were compared with those patients receiving alternative spine stabilization interbody device implants. The primary endpoint was reoperation rate at 1 year, with secondary endpoints including the requirement for blood transfusion during index hospitalization, 1 year rate of any cause hospitalization, 1 year rate of surgical site infection, and mortality at 1 year.

Results Among the 606 patients undergoing thoracic or lumbar spinal fusion surgery during the study period, 136 received only the CB. In comparison with patients who did not receive the CB, no significant differences were found in the rate of reoperation at 1 year or the rates of secondary safety outcomes.

Conclusions Data derived from the EHR can be successfully leveraged to assess the safety of IVBI devices, in this case demonstrating no significant differences in the rates of risk-adjusted safety endpoints between patients undergoing spinal surgery with the CB as compared with alternative spinal implants.

INTRODUCTION

Intervertebral body stabilization device implants (IVBI) have been used with increasing frequency in an effort to achieve more predictable surgical results during spinal fusion surgeries.1,2 The high costs of spinal fusion implants along with limited effectiveness data highlight the importance of developing data collection systems to address the safety and comparative effectiveness of interbody fusion implants. However, like most medical devices, confirmation of the safety and effectiveness of IVBI have been hampered by the small size of pre-market studies and the inadequacy of voluntary reporting of adverse events through the submission of medical device adverse event reports to the Food and Drug Administration (FDA).3–5

Active surveillance of electronic health record (EHR) derived clinical data has been proposed as a strategy that is complementary to the existing pre-market and post-market evaluation tools, which have primarily relied on the evaluation of adverse event reports.6 We have previously developed and validated an open-source, active surveillance software.
system called DELTA (Data Extraction and Longitudinal Trend Analysis) to leverage data from high quality clinical data sources for the purposes of comparative safety assessment for a variety of high-risk medical devices. However, to date, DELTA has not been applied to EHR-derived clinical data for the purpose of active surveillance by automated analysis and re-analysis of accumulating clinical data.

To foster the development of optimized post-approval evaluation of medical devices, the FDA founded the National Evaluation System for health Technologies coordinating center (NESTcc) in 2018. Johnson & Johnson proposed a safety surveillance study for the CONCORDE Bullet (CB) IVBI in the first series of NESTcc test case studies in an effort to validate the strategy of applying clinical surveillance methods to EHR-derived data. In addition, the sponsor sought high quality clinical information for submission to European Union regulators in order to meet recent changes in requirements for medical device post-market monitoring, but importantly, there was no data suggesting a safety concern with the CB device prior to the initiation of the study.

The CB DELTA study was therefore designed to evaluate the use of active surveillance tools applied to EHR-derived data, from a single healthcare system, to assess the safety of a commonly used IVBI used in posterior spinal fusion procedures.

**METHODS**

**Study design and oversight**

A comprehensive written protocol was developed, and plans for interim data reviews and a study oversight committee were established with representation from NESTcc, PEDSnet, Johnson & Johnson, and Lahey Medical Center, which reviewed and approved the final study protocol prior to any data analysis. The institutional review boards of Lahey Medical Center and Children’s Hospital of Philadelphia reviewed and cleared the study protocol prior to the review of any study data.

**Study environment, data source and data element extraction**

Epic Systems (Verona, Wisconsin, USA) has been used as the EHR system at all Lahey Health hospitals and clinics to support all clinical activities. Data elements were extracted from Lahey’s clinical data warehouse derived from the Epic EHR, containing demographic, clinical, laboratory, and claims data for all patient encounters in any Lahey facility. The universal device identifier (UDI) system was not implemented within the Lahey EHR until after the study period, and therefore device implants were identified through manufacturer model numbers recorded in the EHR surgical log. Of note, all patients undergoing spine surgery at Lahey are routinely followed with in-person or telehealth visits at 30 days and 12 months postoperatively, thereby minimizing the risk of the loss to follow-up in the first year after spine surgery at the medical center.

**Patient eligibility, device exposures, and endpoint definitions**

Patients 18 years or older, undergoing spinal arthrodesis surgery between April 1, 2015, and December 31, 2018, at a Lahey Health System hospital were included in the study (see online supplemental figure S1). Patients were excluded from the analysis if their index surgical procedure included the cervical spine or if the patient had any spine surgery performed in the 12 months prior to their qualifying surgery.

Among the patients eligible for inclusion, treatment with the device of interest (the CB), as well as prespecified alternative devices (the DePuy Synthes OPAL and the Medtronic CAPSTONE PEEK) was identified by matching surgical log device implant records to device model numbers available from the manufacturers.

The primary safety outcome was the proportion of patients undergoing spine reoperation for any cause at 1 year. Secondary outcomes included mortality, hospitalization for surgical site infection, any hospitalization within 1 year, and requirement for blood transfusion during index hospitalization. The safety endpoints and clinical covariate definitions used for this study are provided in online supplemental tables S1 and S2 of the Supplemental content.

**Covariate and data validation**

Each demographic, clinical, and procedural covariate and outcome was validated through manual chart review of a random 5% sample of patients included in each cohort. Discrepancies found in manual chart review were used to refine covariate filter definitions, and the random 5% manual chart review was repeated until there was 100% agreement between the extraction filters and domain expert chart review for all covariates and outcomes.

**Risk adjustment methods**

Multivariable adjusted logistic regression models were developed to estimate the probability of being treated with the CB. The model included risk factors for the adverse outcomes of interest, as well as factors considered to influence the selection of IVBI. A total of 12 demographic, clinical and procedural variables were included in the final propensity score model (variable definitions available in online supplemental table S6 of the Supplemental content) including age, gender, preoperative body mass index, history of coronary artery disease, heart failure, diabetes, active smoking and history of any prior spine surgery. Procedural covariates included surgeon specific annual spine surgical volume, emergency surgery status, American Society of Anesthesiology physical status classification, and number of spinal levels stabilized during index surgery.

The propensity-matched comparison group was selected on the basis of a non-parsimonious propensity model and matched in a 1:1 manner with CB cases matched within a caliper width of 0.6 of the SD of the logit of the propensity score. Missing data were handled using univariate rules, assuming absence of a condition
for dichotomous variables, and using the median value for continuous variables. The relative imbalance between the CB and comparator groups was assessed using absolute standardized mean difference in covariate means and proportions, with values greater than 0.10 considered suboptimally balanced.\(^{14}\)

Because the number of patients successfully matched could be low and lead to an underpowered analysis, we also performed inverse probability of treatment weighting (IPTW) analysis. IPTW included all patients in both the treatment and control groups, weighted based on the probability of treatment with the device of interest, and therefore minimizes information loss through case exclusion. Weights were trimmed at 5% and 95% to avoid excessive influence of patients with extremes of propensity scores.\(^ {15}\)

A ‘Falsification Hypothesis’ analysis was also prespecified in order to assess the possibility of significant residual confounding after propensity matching or weighting. For this analysis, patients were evaluated for the development of late postoperative renal dysfunction after postoperative day 30, defined as an increase in serum creatinine by at least 50%. Late kidney dysfunction was not thought to be plausibly related to the implantation of a particular IVBI, and was therefore considered an appropriate endpoint for use in Falsification Hypothesis testing.

For the propensity match analysis, a significant difference was considered present if the confidence intervals (CI) between two independent proportions, as measured by the Wilson method, did not cross zero when using an alpha of 0.05. For IPTW, adjusted odds ratio (OR\(_{adj}\)) were considered significant for increased risk of adverse events if the OR\(_{adj}\) was greater than 1.0, with a 95% CI excluding 1.0. All analyses were performed within the DELTA application, relying on R-based statistical packages.\(^{7-9}\)

**RESULTS**

**Unadjusted outcomes**

A total of 612 patients underwent thoracic or lumbar spinal fusion between April 1, 2015, and December 31, 2018, at a Lahey facility, of which 6 patients were excluded due to prior surgery on the thoracic or lumbar spine within the prior 12 months. The clinical characteristics of the remaining 606 patients identified for analysis are shown in online supplemental table S1 of the Supplemental content, along with the subgroups receiving CB, those receiving an alternative IVBI, and those receiving both CB and an alternative device. A total of 10 patients (1.65%) underwent reoperation within 1 year of the index surgery, 3 patients (0.5%) developed a surgical site infection, 105 patients (17.5%) were hospitalized within 1 year of their index surgery, and 3 patients (0.5%) died within 12 months of surgery (see online supplemental table S2 of the Supplemental content). In addition, 66 patients (10.9%) received blood transfusions immediately following surgery. The 158 patients who received both a CB as well as one of the alternative IVBI were excluded from subsequent safety analyses as outcomes could not be definitively attributed to a specific device type.

**Prespecified propensity analyses**

Patients treated with the CB IVBI, alternative IVBI, and those patients treated with neither during qualifying thoracic and lumbar spine fusion surgeries were identified from EHR derived data (see online supplemental figure S1 in the Supplemental content). Among the 136 patients receiving only CB during their qualifying surgery, 69 (51%) were successfully matched with patients in the comparator group (receiving an alternative IVBI) after performing a propensity score match. As shown in table 1, post-matching standardized differences were less than 0.10 for 13 of the 18 covariates, with 5 covariates having post-matching standardized differences ranging from 0.12 to 0.19, indicating suboptimal covariate balance. Thus, we also performed inverse probability of treatment weighting (IPTW) analysis.

There were no significant differences observed between the clinical outcomes in the CB treated patients as compared with the patients treated with alternative intervertebral body device implantation in the propensity matched analysis, as demonstrated in table 2. Among the 69 patients treated with the CB in the matched cohort, only 1 patient required reoperation within 12 months, as compared with no patients among the matched control patients.

Online supplemental figure S2 in the Supplemental content provides the results of the prospective active surveillance analysis, demonstrating stability of the results by repeated analysis of the primary endpoint outcome, by calendar quarter, throughout the study period.

As shown in table 3, the OR for the risk of reoperation at 12 months was 1.01 (95% CI: 0.96 to 1.03) as estimated by IPTW methodology, indicating no significantly increased risk for CB relative to the prespecified alternative IVBI. An extended analysis of the primary endpoint of reoperation out to 2 years maintained these null findings (OR 0.99, 95% CI: 0.96 to 1.03, p value 0.91). In addition, there were no significantly increased risks for any of the secondary outcomes at 12 months.

**Additional sensitivity analyses**

The eligible patient population was expanded in order to perform several post-hoc sensitivity analyses. Patients without a specific comparator device of interest, including patients with no device implanted at the time of index surgery (identified by procedure codes in online supplemental tables S7 and S8), as well as those patients undergoing complex repair (covering more than two spine levels) at the time of their index surgery were included in the analyses. Propensity score matching and IPTW analyses were performed on these additional groups (see online supplemental figure S3 of the Supplemental content).
and the results were consistent with the primary analysis, with no increased risk of adverse events associated with the use of the CB.

Finally, the results of the Falsification Hypothesis comparing CB and alternative IVBI patients demonstrate a relative risk of 0.67 (95% CI: 0.11 to 3.87, p value 0.017) of developing renal dysfunction in the year following index surgery, indicating no evidence of significant residual confounding in the cohorts analyzed. Residual confounding was also absent in the post-hoc sensitivity analysis given the observed relative risk of developing renal dysfunction (0.25 with 95% CI: 0.15 to 2.75, p value 0.472) for the expanded pool of patients in the CB group compared with patients not exposed to CB.

### Table 1 Covariate distributions prior to the propensity score match, after the match, as well as in the remaining unmatched CONCORDE Bullet treated patients

<table>
<thead>
<tr>
<th>Covariate</th>
<th>Total study population</th>
<th>After propensity match</th>
<th>Unmatched exposures</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CONCORDE Bullet (N=136)</td>
<td>Alternate IVBI (N=98)</td>
<td>Std. diff</td>
</tr>
<tr>
<td>Female</td>
<td>52.94%</td>
<td>52.04%</td>
<td>0.018</td>
</tr>
<tr>
<td>Age—median (IQR)</td>
<td>61.84±12.43</td>
<td>58.87±13.16</td>
<td>0.232</td>
</tr>
<tr>
<td>Body mass index—median (IQR)</td>
<td>30.03±5.69</td>
<td>29.29±5.80</td>
<td>0.130</td>
</tr>
<tr>
<td>Diabetes</td>
<td>13.97%</td>
<td>10.20%</td>
<td>0.116</td>
</tr>
<tr>
<td>Heart failure</td>
<td>0.74%</td>
<td>1.02%</td>
<td>0.031</td>
</tr>
<tr>
<td>Coronary artery disease</td>
<td>4.41%</td>
<td>4.08%</td>
<td>0.016</td>
</tr>
<tr>
<td>Smoker (active)</td>
<td>13.24%</td>
<td>8.16%</td>
<td>0.165</td>
</tr>
<tr>
<td>Preoperative opioid Rx</td>
<td>19.12%</td>
<td>27.55%</td>
<td>0.200</td>
</tr>
<tr>
<td>ASA class 1–2</td>
<td>68.38%</td>
<td>72.45%</td>
<td>0.089</td>
</tr>
<tr>
<td>ASA class 3–5</td>
<td>31.62%</td>
<td>27.55%</td>
<td>0.089</td>
</tr>
<tr>
<td>Prior spine surgery</td>
<td>0.74%</td>
<td>0.00%</td>
<td>–</td>
</tr>
<tr>
<td>Single-level</td>
<td>40.44%</td>
<td>44.90%</td>
<td>0.090</td>
</tr>
<tr>
<td>Two-level</td>
<td>0.00%</td>
<td>1.02%</td>
<td>–</td>
</tr>
<tr>
<td>Complex repair (&gt;2 levels)</td>
<td>40.44%</td>
<td>6.12%</td>
<td>0.889</td>
</tr>
<tr>
<td>No level identified</td>
<td>19.12%</td>
<td>47.96%</td>
<td>0.642</td>
</tr>
<tr>
<td>Surgeon volume low</td>
<td>5.15%</td>
<td>27.55%</td>
<td>0.636</td>
</tr>
<tr>
<td>Surgeon volume medium</td>
<td>72.79%</td>
<td>64.29%</td>
<td>0.184</td>
</tr>
<tr>
<td>Surgeon volume high</td>
<td>22.06%</td>
<td>8.16%</td>
<td>0.396</td>
</tr>
<tr>
<td>ASA class, American Society of Anesthesiology physical status classification</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>IVBI, intervertebral body implant; Std. diff, standard difference.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 2 Clinical outcomes of the propensity match analysis comparing patients treated with the CONCORDE Bullet with matched patients treated with alternative therapies

<table>
<thead>
<tr>
<th>Outcomes: Propensity match analysis</th>
<th>CONCORDE Bullet (%) (n=69)</th>
<th>Alternative IVBI (%) (n=69)</th>
<th>Relative risk</th>
<th>95% CI</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary endpoint:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reoperation at 1 year</td>
<td>1</td>
<td>1.4</td>
<td>0</td>
<td>n/a</td>
<td>0.316</td>
</tr>
<tr>
<td>Secondary endpoints:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mortality at 1 year</td>
<td>1</td>
<td>1.4</td>
<td>0</td>
<td>n/a</td>
<td>0.316</td>
</tr>
<tr>
<td>Any hospitalization at 1 year</td>
<td>15</td>
<td>21.7</td>
<td>7</td>
<td>10.1</td>
<td>2.14</td>
</tr>
<tr>
<td>Blood transfusion (postoperative)</td>
<td>7</td>
<td>10.1</td>
<td>3</td>
<td>4.3</td>
<td>2.33</td>
</tr>
<tr>
<td>Surgical site infection</td>
<td>0</td>
<td>0.0</td>
<td>0</td>
<td>n/a</td>
<td>n/a</td>
</tr>
</tbody>
</table>

IVBI, intervertebral body implant.
DISCUSSION

The CB DELTA study was designed to assess the feasibility of applying active surveillance tools to accumulating clinical data derived from the EHR to evaluate the safety of a commonly used IVBI. We compared the safety profile of different IVBI devices, through the application of rigorous, validated, prospective, active surveillance methods tools using clinical data collected during the routine course of care. During the study period of 34 months, 136 patients underwent thoracic or lumbar spine surgery with the implantation of the device of interest, while 212 patients were treated exclusively with alternative spinal stabilization devices. Overall, our analysis showed similar 12-month rates of reoperation, mortality and any cause rehospitalization in the CB and alternative implant groups. The findings were consistent among subgroups analyzed and were confirmed through IPTW analysis, in addition to propensity score matching.

The findings of this study are significant for two reasons. First, this NESTcc feasibility study further demonstrates the utility of integrating routinely collected EHR data into prospective, active safety surveillance activities for medical devices. Using well-structured clinical data available in our hospital system, we were able to validate the device exposure, key clinical covariates and relevant outcomes, suggesting that safety surveillance of spinal implants is feasible using routinely collected EHR data. Second, there is limited comparative safety data of spinal surgery techniques, and this analysis is among the first reports comparing the safety of a commonly used IVBI with that of risk-adjusted comparator treatments. As the number of options for interbody devices expands, the ability to generate clinical effectiveness and safety data regarding fusion rates and reoperation rates from EHR data will allow providers to better assess the comparative safety, effectiveness, and costs of different treatment strategies.

Limitations

There are several important limitations of this study. The patients studied were treated within a single healthcare system, by a small group of experienced spine surgeons, limiting the generalizability to different environments. Also, the sample size of patients treated with the CB and alternative spine implants were limited, thereby reducing the power of the analysis to identify small differences in adverse outcome rates. However, the application of IPTW methods to reduce the loss of information by including all treated patients, suggests that the sample size did not materially impact the findings. In addition, additional patient groups were explored as comparators, with similar findings, further supporting the primary results. In addition, the residual imbalance in clinical risk factor distribution noted between the treated and comparator groups limit the accuracy of the estimated differences in adverse event rates.

Additional limitations include the risk of residual confounding, which we sought to minimize through robust risk adjustment using propensity matching and propensity weighting. In support of the effectiveness of the methods, the prespecified Falsification Hypothesis analyses found that late renal dysfunction, thought unrelated to choice of surgery, occurred with similar frequency between the matched populations, thereby indicating that there was no evidence for significant residual confounding in the cohorts analyzed.

This study was supported through a competitive research grant from NESTcc in accordance with Federal research regulations, and the study steering committee included regulatory and quality representatives from Johnson & Johnson, the manufacturer of the device studied.

Author affiliations

1Comparative Effectiveness Research Institute, Lahey Hospital and Medical Center, Burlington, Massachusetts, USA
2Department of Cardiovascular Medicine, Tufts University School of Medicine, Boston, Massachusetts, USA
3Department of Neurosurgery, Lahey Hospital and Medical Center, Burlington, Massachusetts, USA
4Johnson & Johnson Medical Devices, New Brunswick, New Jersey, USA
5University of Pennsylvania College of Medicine, Philadelphia, Pennsylvania, USA
6Department of Pediatrics, Lurie Children’s Hospital of Chicago, Chicago, Illinois, USA
7Vanderbilt University Medical Center, Nashville, Tennessee, USA
8Department of Neurosurgery, Tufts University School of Medicine, Boston, Massachusetts, USA

Acknowledgements The authors would like to thank Heather M Colvin, Director of Medical Device Regulatory Affairs Evidence and Outcomes Policy, Johnson & Johnson Corporation, for her thoughtful review of the original protocol, and for her advice of the study throughout its execution.
Contributors EAF performed data curation, validation of data curation methods and output, utilized DELTA statistical software to perform formal analysis of the data, as well as helping to write, review and edit the manuscript. FSR, along with ZG, helped to conceptualize the project, provided clinical insight and leadership related to research activities, assisted in the preparation of original manuscript materials and act as guarantor for overall content. HS participated in data curation, validation, developed the DELTA statistical software and statistical analysis plan, and assisted in writing and review of the manuscript. SR and MRD assisted in validation of data, and coordinated responsibilities for research planning and execution. SZ and PC provided guidance on methodology, and CB, MM, JBB and MEM assisted in the development and validation of DELTA as an active surveillance tool.

Funding This research was supported through a research grant from the National Medical Device Evaluation System Coordination Center (NESTcc) (FDA U01FD006292, MDC/Nest Sub-Award Number 6292-2019-R1TCb-001). MEM efforts were additionally supported, in part, by VA HSR&D IIR 13-152, and NIH NIDDK 5R01 DK113201-02 and 1U01 FD004493-01. Competing interests PC and SZ are representatives from Johnson & Johnson, the manufacturer of the device studied.

Patient consent for publication Not applicable.

Ethics approval The Lahey Medical Center Institutional Review Board (IRB) determined this project to be Exempt Research as set forth in 45 CFR 46.104(d) (4)(ii) (IRB review #141615). Informed consent requirements were waived on the basis of minimal risk to individual patients through secondary research use of their de-identified clinical data.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. De-identified data supporting the findings of this study may be available from the corresponding author (EAF) upon reasonable request.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) 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, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID iDs Edward A Frankenberger http://orcid.org/0000-0003-4951-4312 Frederic S Resnic http://orcid.org/0000-0002-4105-4529

REFERENCES