



Universiteit  
Leiden  
The Netherlands

## **National databases for neurosurgical outcomes research: options, strengths, and limitations**

Karhade, A.V.; Larsen, A.M.G.; Cote, D.J.; Dubois, H.M.; Smith, T.R.

### **Citation**

Karhade, A. V., Larsen, A. M. G., Cote, D. J., Dubois, H. M., & Smith, T. R. (2018). National databases for neurosurgical outcomes research: options, strengths, and limitations. *Neurosurgery*, 83(3), 333-344. doi:10.1093/neuros/nyx408

Version: Publisher's Version

License: [Licensed under Article 25fa Copyright Act/Law \(Amendment Taverne\)](#)

Downloaded from: <https://hdl.handle.net/1887/4303208>

**Note:** To cite this publication please use the final published version (if applicable).

# National Databases for Neurosurgical Outcomes Research: Options, Strengths, and Limitations

**Aditya V. Karhade, BE**  
**Alexandra M. G. Larsen, BS**  
**David J. Cote, BS**  
**Heloise M. Dubois, BA**  
**Timothy R. Smith, MD, PhD,**  
**MPH**

Cushing Neurosurgery Outcomes Center,  
 Department of Neurosurgery, Brigham  
 and Women's Hospital, Harvard Medical  
 School, Boston, Massachusetts

#### Correspondence:

Timothy R. Smith, MD, PhD, MPH,  
 Department of Neurosurgery,  
 Brigham and Women's Hospital,  
 15 Francis Street,  
 Boston, MA 02115.  
 E-mail: [trsmith@partners.org](mailto:trsmith@partners.org)

**Received,** November 4, 2016.

**Accepted,** June 21, 2017.

**Published Online,** August 5, 2017.

Copyright © 2017 by the  
 Congress of Neurological Surgeons

**BACKGROUND:** Quality improvement, value-based care delivery, and personalized patient care depend on robust clinical, financial, and demographic data streams of neurosurgical outcomes. The neurosurgical literature lacks a comprehensive review of large national databases.

**OBJECTIVE:** To assess the strengths and limitations of various resources for outcomes research in neurosurgery.

**METHODS:** A review of the literature was conducted to identify surgical outcomes studies using national data sets. The databases were assessed for the availability of patient demographics and clinical variables, longitudinal follow-up of patients, strengths, and limitations.

**RESULTS:** The number of unique patients contained within each data set ranged from thousands (Quality Outcomes Database [QOD]) to hundreds of millions (MarketScan). Databases with both clinical and financial data included PearlDiver, Premier Healthcare Database, Vizient Clinical Data Base and Resource Manager, and the National Inpatient Sample. Outcomes collected by databases included patient-reported outcomes (QOD); 30-day morbidity, readmissions, and reoperations (National Surgical Quality Improvement Program); and disease incidence and disease-specific survival (Surveillance, Epidemiology, and End Results-Medicare). The strengths of large databases included large numbers of rare pathologies and multi-institutional nationally representative sampling; the limitations of these databases included variable data veracity, variable data completeness, and missing disease-specific variables.

**CONCLUSION:** The improvement of existing large national databases and the establishment of new registries will be crucial to the future of neurosurgical outcomes research.

**KEY WORDS:** Neurosurgical outcomes, Databases, Quality Improvement, Epidemiology, Value of care, Healthcare reform

*Neurosurgery* 83:333–344, 2018

DOI:10.1093/neuros/nyx408

[www.neurosurgery-online.com](http://www.neurosurgery-online.com)

**ABBREVIATIONS:** **ACS**, American College of Surgeons; **CBTRUS**, Central Brain Tumor Registry of the United States; **CHA**, Children's Hospital Association; **CMS**, Centers for Medicare and Medicaid Services; **CPT**, Current Procedural Terminology; **ICD**, International Classification of Disease; **KID**, Kids' Inpatient Database; **NCDB**, National Cancer Database; **NCI**, National Cancer Institute; **NIS**, National Inpatient Sample; **N2QOD**, National Neurosurgery Quality and Outcomes Database; **NRD**, Nationwide Readmissions Database; **NSQIP**, National Surgical Quality Improvement Program; **NTDB**, National Trauma Data Bank; **QOD**, Quality Outcomes Database; **SASD**, State Ambulatory Surgery and Services Database; **SID**, State Inpatient Databases; **SEER**, Surveillance, Epidemiology, and End Results

Value-based care delivery,<sup>1</sup> alternative payment models,<sup>2</sup> quality improvement, publicly reported surgeon scorecards,<sup>3</sup> and individualized patient care depend on accurate, complete, and large clinical and financial databases. The development of databases that contain robust outcome, process, and structural measures presents an invaluable opportunity for surgeons to influence future policy changes while improving the quality of individual patient care.<sup>4</sup>

Neurosurgical interventions are among the most expensive and scrutinized services offered by inpatient hospitals and ambulatory centers. Neurosurgeons have been leaders in responding to the value agenda with a dramatic increase

in the number of large, multicenter, national database studies.<sup>4-7</sup> The use of these data streams and creation of additional ones, such as those that include patient-reported outcomes measures, are an integral part of the provider-led effort to demonstrate high-value care by improving outcomes and eliminating unnecessary costs.

Though neurosurgeons are rapidly adapting to the use of national databases for outcomes research, the neurosurgical literature lacks a comprehensive review of large national databases. In this study, we assessed the strengths and limitations of various resources for outcomes research in neurosurgery.

## METHODS

A review of the current literature of PubMed and Web of Science was conducted to identify surgical outcomes studies that used large, national data sets as the source of data. The following databases were included in this review, based on the expertise of the senior author: (1) Quality Outcomes Database (QOD, formerly National Neurosurgery Quality and Outcomes Database [N<sup>2</sup>QOD]), (2) Surveillance, Epidemiology, and End Results (SEER) and SEER-Medicare database, (3) Central Brain Tumor Registry of the United States (CBTRUS), (4) American College of Surgeons (ACS) National Surgical Quality Improvement Program (NSQIP), (5) National Trauma Data Bank (NTDB), (6) National Cancer Database (NCDB), (7) National Inpatient Sample (NIS), (8) Kids' Inpatient Database (KID), (9) Nationwide Readmissions Database (NRD), (10) State Inpatient Databases (SID), (11) State Ambulatory Surgery and Services Database (SASD), (12) Centers for Medicare and Medicaid Services (CMS) Claims Data, (13) Vizient Clinical Data Base and Resource Manager (formerly University HealthSystem Consortium), (14) MarketScan Research Databases, (15) Premier Healthcare Database, (16) PearlDiver, (17) OptumLabs Data Warehouse, and (18) Pediatric Health Information System (PHIS; Table 1).

The PubMed and Web of Science search strategies were based on database names and acronyms and keywords related to "neurosurgery" and "outcomes." Titles and abstracts were screened. Full texts of selected articles were screened for strengths and limitations of each database. The initial search was supplemented by manually searching the references of selected full-text studies. The following files were also used to supplement the PubMed and Web of Science searches: (1) User Guide for the 2014 ACS NSQIP Participant User File,<sup>8</sup> (2) NTDB Annual Report 2015,<sup>9</sup> (3) NCDB Participant User File Data Dictionary,<sup>10</sup> (4) MarketScan Research Databases, Truven Health Analytics documentation,<sup>11</sup> and (5) Agency for Healthcare Research and Quality, Healthcare Cost and Utilization Project documentation for the NIS, KID, SID, NRD, and SASD databases.<sup>12</sup> Information collected included (1) availability of clinical data, (2) availability of demographic data, (3) availability of longitudinal data, (4) years of data collection, (5) strengths, and (6) limitations.

## RESULTS

### Quality Outcomes Database

QOD is a prospective, risk-adjusted, observational registry of NeuroPoint Alliance focused on quality and outcome reporting.<sup>13-18</sup> QOD was formerly known as the N<sup>2</sup>QOD. The American Association of Neurological Surgeons established the

non-for-profit NeuroPoint Alliance in 2008 with the goal of collecting, analyzing, and sharing neurosurgical clinical information for quality improvement. The first database, the Lumbar Spine Registry, was launched in 2012 in order to improve quality of life and decrease disability and readmission rates associated with surgery for degenerative lumbar spinal disorders.<sup>17,18</sup> The current QOD registries available are the Spine Surgery QOD, containing the Lumbar Spine, Spinal Deformity, and Cervical Spine Registries, and the Neurovascular QOD.<sup>19</sup> The tumor QOD registry is in development, with expected launch in 2017. The Lumbar Spine registry has 80 participating centers with over 23 000 patients enrolled, and the Cervical Spine registry has 51 participating centers with over 8500 patients enrolled.<sup>19</sup> The CMS has approved the QOD to identify as a Qualified Clinical Data Registry, with 21 non-Physician Quality Reporting System measures focused on neurosurgery.<sup>6</sup> Asher et al<sup>17</sup> reported in 2014 that the length of follow-up in the Lumbar Spine database is 1 yr, with 30-d morbidity and 3- and 12-mo patient-reported outcomes. The database enrolls 6 patients per week at participating institutions and uses multiple measures to ensure internal validity, including multiple types of audits, ranging from surgeon-led self-audits to random on-site audits.<sup>17</sup> The baseline data completeness in 2014 for the total number of enrolled patients was 98.1%.<sup>17</sup> Limitations of the database include relatively more missing data for the 3- and 12-mo quality measures.<sup>17</sup> Institutions can participate in QOD by submitting a request to NeuroPoint Alliance.

### Surveillance, Epidemiology, and End Results

SEER, originally established by the National Cancer Institute (NCI) in 1973, publishes cancer statistics from population-based cancer registries.<sup>20-22</sup> Statistics available for brain and other nervous system cancers include trends in SEER incidence and US mortality by sex, annual incidence and death rates, incidence and mortality rates by age, 5-yr relative and period survival, relative survival by year of diagnosis, and US prevalence estimates.<sup>23</sup> The database contains 9 675 661 cases from 1973 to 2014.<sup>22,24-34</sup> The database includes information on patient demographics, primary tumor site, stage of cancer at diagnosis, and first course of treatment information. Researchers must request access to the SEER database by submitting an SEER Research Data-Use Agreement form to NCI. Limitations of the SEER database include lack of comprehensive clinical variables, lack of imaging characteristics, incomplete treatment characteristics, lack of postoperative therapies, and missing data for patient comorbidities.<sup>22</sup> SEER data "do not capture information about surgery and radiation past 4 mo of diagnosis, nor is there information about recurrence or metastasis that is detected subsequent to initial diagnosis."<sup>35</sup>

### SEER-Medicare Database

The SEER-Medicare database is a linkage between the NCI SEER database and CMS claims data.<sup>36-39</sup> The purpose is to gather information about patients on Medicare with cancer. The

**TABLE 1. Representative Characteristics by Database**

Dataset	Organization	Example study published with database
QOD	AANS, NeuroPoint Alliance	Comparison of clinical outcomes in the N <sup>2</sup> QOD for open vs minimally invasive transforaminal lumbar interbody fusion (Guan et al). <sup>14</sup>
SEER	NCI	Gross-total resection outcomes in an elderly population with glioblastoma: a SEER-based analysis: clinical article (Noorbakhsh et al). <sup>24</sup>
SEER-Medicare	NCI and CMS	Use of SEER-Medicare data to conduct case-control studies of cancer among the US elderly (Engels et al). <sup>38</sup>
CBTRUS	American Brain Tumor Association	CBTRUS statistical report: primary brain and central nervous system tumors diagnosed in the United States in 2008-2012 (Ostrom et al). <sup>42</sup>
NSQIP	ACS	Length of hospital stay after craniotomy for tumor: a NSQIP analysis (Dasenbrock et al). <sup>47</sup>
NTDB	ACS	Intracranial pressure monitoring in children with severe traumatic brain injury: NTDB-based review of outcomes (Alkhoury et al). <sup>52</sup>
NCDB	ACS and American Cancer Society	Brain tumor survival: results from the National Cancer Data Base (Surawicz et al). <sup>61</sup>
NIS	HCUP-AHRQ	Incidence, hospital costs and in-hospital mortality rates of epidural hematoma in the United States (Bir et al). <sup>65</sup>
KID	HCUP-AHRQ	Pediatric cervical spine and spinal cord injury: a national database study. (Shin et al). <sup>73</sup>
NRD	HCUP-AHRQ	Predictors of hospital readmissions for ulcerative colitis in the United States: a national database study (Poojary et al). <sup>76</sup>
SID	HCUP-AHRQ	Selection of patients for ambulatory lumbar discectomy: results from four US states (Bekelis et al). <sup>79</sup>
SASD	HCUP-AHRQ	Institutional charges and disparities in outpatient brain biopsies in four US States: the State Ambulatory Database (SASD; Bekelis et al). <sup>83</sup>
CMS	CMS	Outcomes after ischemic stroke for hospitals with and without Joint Commission-certified primary stroke centers (Lichtman et al). <sup>87</sup>
Vizient Clinical Data Base	Vizient, Inc.	Smoking is associated with poorer quality-based outcomes in patients hospitalized with spinal disease (Bisson et al). <sup>90</sup>
MarketScan Research Databases	Truven Health Analytics, Inc.	Comparison of bilateral vs staged unilateral deep brain stimulation (DBS) in Parkinson's disease in patients under 70 yr of age (Petraglia et al). <sup>97</sup>
Premier Healthcare Database	Premier, Inc.	Hospital variation in cervical spine imaging of young children with traumatic brain injury (Henry et al). <sup>102</sup>
PearlDiver	PearlDiver Technologies, Inc.	Trends associated with distal biceps tendon repair in the United States, 2007 to 2011 (Wang et al). <sup>107</sup>
OptumLabs Data Warehouse	Optum, Mayo Clinic	External validation of a multivariable claims-based rule for predicting in-hospital mortality and 30-d postpulmonary embolism complications (Coleman et al). <sup>110</sup>
PHIS	CHA	Reinfection following initial cerebrospinal fluid shunt infection (Simon et al). <sup>113</sup>

database contains 17 368 cases of individuals with brain cancer from 1991 to 2013.<sup>40</sup> SEER-Medicare contains information for individuals diagnosed with cancer when enrolled in Medicare part A and part B. The data contain the aspects outlined in the SEER section, including information on first-course treatment such as cancer-directed surgery and radiation, as well as claims files from Medicare inpatient and outpatient visits, home health agencies, hospice, durable medical equipment, and Medicare part D.<sup>38,39</sup> The database also contains patient information from individuals on Medicare without cancer to be used for comparison. Limitations of the SEER-Medicare database include limitations of SEER data, lack of imaging lack of care paid for by entities other than Medicare, and inability to determine disease-free survival.<sup>41</sup>

**Central Brain Tumor Registry of the United States**

CBTRUS is a nonprofit research organization founded in 1992.<sup>42</sup> CBTRUS is a histology-specific, population-based registry of brain tumor data in the United States.<sup>42</sup> CBTRUS receives cancer data from SEER and the Centers for Disease Control and Prevention National Program of Central Registries. The database contains information on primary malignant and nonmalignant tumors of the brain and central nervous system. As of 2016, approximately 100% of the US population is represented in the database through the 50 state cancer registries and District of Columbia registry, which contribute data.<sup>42-44</sup> CBTRUS releases information on incidence (overall, pediatric, pediatric and adolescent, and adolescent and young adult), mortality, lifetime risk, survival, and prevalence.<sup>44</sup> The

**TABLE 2. Strengths and Limitations**

Database	Limitations	Strengths
QOD	<ul style="list-style-type: none"> <li>Relatively more missing data for 3-mo and 12-mo time points<sup>17</sup></li> </ul>	<ul style="list-style-type: none"> <li>Patient reported outcomes</li> <li>Multiple registries (spine surgery QOD, neurovascular QOD, and tumor QOD in development)</li> <li>98.1% baseline data completeness in 2014<sup>17</sup></li> <li>&gt; 9 million cases from 1973-2014<sup>22</sup></li> <li>Trends in SEER incidence and US mortality by sex, annual incidence and death rates, incidence and mortality rates by age, 5-yr relative and period survival, relative survival by year of diagnosis, and U.S. prevalence estimates<sup>23</sup></li> </ul>
SEER	<ul style="list-style-type: none"> <li>Not neurosurgery-specific database</li> <li>Lack of imaging characteristics, incomplete treatment characteristics, lack of postoperative therapies, missing data for patient comorbidities<sup>22</sup></li> <li>Lack of information about surgery and radiation past four months after diagnosis and lack of information regarding metastasis or recurrence subsequent to initial diagnosis<sup>35</sup></li> </ul>	<ul style="list-style-type: none"> <li>&gt; 17,000 patients with brain cancer from 1991-2013<sup>40</sup></li> <li>Claims files from Medicare inpatient and outpatient visits, home health agencies, hospice, durable medical equipment, Medicare part D</li> <li>Geographic representation of US population through 50 state cancer registries and District of Columbia<sup>44</sup></li> </ul>
SEER-Medicare	<ul style="list-style-type: none"> <li>Not neurosurgery-specific database</li> <li>Limitations of SEER data</li> <li>Lack of imaging<sup>41</sup></li> <li>Lack of care paid for by entities other than Medicare<sup>41</sup></li> </ul>	<ul style="list-style-type: none"> <li>Trained records reviewer</li> <li>Interobserver disagreement around 2%<sup>8</sup></li> <li>603 NSQIP-participating sites in 2015<sup>48</sup></li> <li>Hospital exclusion if 30-d follow-up rate less than 80%<sup>8</sup></li> </ul>
CBTRUS	<ul style="list-style-type: none"> <li>Not neurosurgery-specific database</li> <li>Record of only initial brain tumor diagnosis, lack of consistently listed WHO grades, potential inaccuracies in assignment of Hispanic ethnicity<sup>43</sup></li> </ul>	<ul style="list-style-type: none"> <li>Largest compilation of trauma data in the US and Canada<sup>50</sup></li> <li>747 facilities with &gt; 860 thousand admission records in 2015<sup>50</sup></li> </ul>
NSQIP	<ul style="list-style-type: none"> <li>Not neurosurgery-specific database</li> <li>Lack of neurosurgery-specific variables and lack of data beyond 30-d<sup>49</sup></li> <li>Wide variability in amount of data for missing data for perioperative variables</li> </ul>	<ul style="list-style-type: none"> <li>Over 34 million records and approximately 70% newly diagnosed cancer cases in the US<sup>58</sup></li> <li>Includes cancer stage, interventions, histological information with public access to more than 9 million cases of the 14 most commonly diagnosed solid tumors<sup>58</sup></li> </ul>
NTDB	<ul style="list-style-type: none"> <li>Not neurosurgery-specific database</li> <li>Lack of inclusive list of risk factors and comorbidity data<sup>57</sup></li> </ul>	<ul style="list-style-type: none"> <li>Largest publicly available database of inpatient stays in the US</li> <li>Total cost and length of hospitalization</li> <li>2014 NIS sampling represented area of 44 states and 96% of the US population<sup>63</sup></li> </ul>
NCDB	<ul style="list-style-type: none"> <li>Not neurosurgery-specific database</li> <li>Changes in classification instructions given to tumor registries and staging rules impacts comparability of data from different years<sup>62</sup></li> </ul>	<ul style="list-style-type: none"> <li>In 2012, weighted sample of 80% of pediatric discharge and 10% of normal newborn discharges from more than 4100 community hospitals in the US<sup>74</sup></li> <li>2014, data from 22 states that covered 49.3% of all US hospitalizations<sup>75</sup></li> <li>Verified confirmation of multiple discharges for individuals</li> </ul>
NIS	<ul style="list-style-type: none"> <li>Not neurosurgery-specific database</li> <li>Lack of participation from several US states, lack of post-discharge information<sup>68</sup></li> </ul>	<ul style="list-style-type: none"> <li>Covers approximately 90% of US hospital discharges and 97% of US community hospital discharges<sup>12</sup></li> <li>Ability to analyze hospitalization patterns within a geographical area containing multiple states, one state, or within a state</li> </ul>
KID	<ul style="list-style-type: none"> <li>Not neurosurgery-specific database</li> <li>Lack of discharge information from rehabilitation hospitals and inability to perform longitudinal analysis across years<sup>69</sup></li> </ul>	<ul style="list-style-type: none"> <li>Ambulatory surgery centers and hospital-based outpatient departments across 35 states<sup>12</sup></li> <li>Identifiers that permit linkage to SID</li> </ul>
NRD	<ul style="list-style-type: none"> <li>Not neurosurgery-specific database</li> <li>Inability to use information for analysis of regional, state, or hospital specific readmission rates<sup>75</sup></li> <li>No clear indication if readmission related to prior admissions<sup>75</sup></li> </ul>	<ul style="list-style-type: none"> <li>Includes part A, part B, service dates, ICD diagnosis, procedure codes, type of healthcare coverage<sup>84</sup></li> <li>Varying privacy levels with research identifiable files containing beneficiary level protected health information</li> </ul>
SID	<ul style="list-style-type: none"> <li>Not neurosurgery-specific database</li> <li>Variability in available data depending on state and year and loss of revisit analysis if patient changed states<sup>79,80</sup></li> </ul>	
SASD	<ul style="list-style-type: none"> <li>Not neurosurgery-specific database</li> <li>Variability in available data depending on state and year and loss of revisit analysis if patient changed states<sup>79,80</sup></li> </ul>	
CMS	<ul style="list-style-type: none"> <li>Not neurosurgery-specific database</li> <li>Variable granularity depending on type of files requested, variable coding of disease specific variables pertinent to the field of investigation<sup>85,89</sup></li> </ul>	

**TABLE 2. Continued**

Database	Limitations	Strengths
Vizient Clinical Data Base	<ul style="list-style-type: none"> <li>Not neurosurgery-specific database</li> <li>Difficulty determining the timing of complications in reference to the date of surgery<sup>90,91</sup></li> </ul>	<ul style="list-style-type: none"> <li>Data representing 97% US academic centers and 160 community hospitals<sup>92</sup></li> <li>Financial and clinical information including length of stay, complications, readmissions, medication use</li> </ul>
MarketScan Research Databases	<ul style="list-style-type: none"> <li>Not neurosurgery-specific database</li> <li>Lack of comprehensive clinical factors relevant to neurosurgery<sup>97</sup></li> </ul>	<ul style="list-style-type: none"> <li>Records for more than 200 million unique patient since 1995<sup>93</sup></li> </ul>
Premier Healthcare Database	<ul style="list-style-type: none"> <li>Not neurosurgery-specific database</li> <li>Lack of comprehensive clinical findings for certain neurosurgical study questions<sup>102</sup></li> </ul>	<ul style="list-style-type: none"> <li>147 million unique patients from over 700 hospitals representing approximately 20% of annual US inpatient discharges<sup>103</sup></li> </ul>
PearlDiver	<ul style="list-style-type: none"> <li>Not neurosurgery-specific database</li> <li>Unequal geographic representation within the US<sup>107</sup></li> </ul>	<ul style="list-style-type: none"> <li>24.6-26.3 million patients from 2007-2011, representing 9% of the US population younger than 65 yr and approx. 13% of US population with private insurance<sup>108</sup></li> </ul>
OptumLabs Data Warehouse	<ul style="list-style-type: none"> <li>Not neurosurgery-specific database</li> <li>Limitations in granularity of diagnosis codes<sup>111</sup></li> <li>Not neurosurgery-specific database</li> </ul>	<ul style="list-style-type: none"> <li>De-identified claims for over 150 million patients<sup>109</sup></li> </ul>
PHIS	<ul style="list-style-type: none"> <li>Missing data on factors of clinical interest and limitations of diagnosis codes<sup>113</sup></li> </ul>	<ul style="list-style-type: none"> <li>Clinical and utilization data from over 45 children's hospitals<sup>114</sup></li> </ul>

QOD, Quality Outcomes Database; SEER, Surveillance, Epidemiology, and End Results; CBTRUS, Central Brain Tumor Registry of the United States; ACS-NSQIP, American College of Surgeons-National Surgical Quality Improvement Program; NTDB, National Trauma Data Bank; NCDB, National Cancer Database; NIS, National Inpatient Sample; KID, Kids' Inpatient Database; NRD, Nationwide Readmissions Database; SID, State Inpatient Database; SASD, State Ambulatory Surgery and Services Database; CMS, Centers for Medicare & Medicaid Services; PHIS, Pediatric Health Information System

epidemiological data can be requested by submitting an online Data Analysis Request through their website. Limitations of CBTRUS include, per Kshetty,<sup>43</sup> records of only the initial brain tumor diagnosis (eg, benign meningioma that later recurred as atypical or anaplastic), lack of consistently listed World Health Organization grade for brain tumors, and potential inaccuracies in assignment of Hispanic ethnicity by use of the North American Association of Central Cancer Registries Hispanic/Latino identification algorithm (Table 2).

**ACS Databases**

*NSQIP Database*

The ACS initiated NSQIP in 2004 in order to decrease 30-d morbidity, mortality, readmissions, and reoperations.<sup>45,46</sup> Reported variables (over 150) in the participant use file include patient demographics, preoperative comorbidities, operative details, and postoperative 30-d complications.<sup>47,48</sup> At each participating institution, a trained records reviewer collects surgical cases on an 8-d cycle, and the associated medical records are reviewed for preoperative, perioperative, and postoperative details up to 30 d following surgery. ACS initiated the NSQIP Pediatric database in 2008 in order to specifically track 30-d outcomes in pediatric surgery. Previous inter-rater reliability audits of the adult NSQIP database have found an overall interobserver disagreement rate around 2%.<sup>8</sup> The 2015 Participant Use Data File contains 885 502 cases from 603 NSQIP-participating

sites.<sup>48</sup> Inclusion criteria are based on Current Procedural Terminology (CPT) code. There are multiple exclusion criteria, such as American Society of Anesthesiologists score of 6 and trauma cases. Hospitals are excluded if their 30-d follow-up rate is less than 80% or if the inter-rater reliability audit disagreement rate is over 5%.<sup>8</sup> Limitations of the NSQIP data include lack of neurosurgery-specific variables, lack of reporting of complications or death after 30 d postoperatively, and wide variability in the amount of missing data for perioperative variables.<sup>49</sup>

*National Trauma Data Bank*

NTDB is the largest compilation of trauma registry data in the United States and Canada, and contains over 7 million records of trauma data. In 2015, 861 888 admission records were entered from 747 facilities.<sup>50,51</sup> The database contains patient demographics, insurance status and type, injury severity score, mechanism of injury, hospital demographics including trauma level, bed size, and geographic location, and outcome information including length of stay, discharge disposition, and complications.<sup>9,52-56</sup> Eligibility for case entry into NTDB is determined by specific International Classification of Disease (ICD)-9 and ICD-10 codes. NTDB data can be purchased online for research purposes with the approval of the ACS. Limitations of NTDB include lack of an inclusive list of risk factors and comorbidity data.<sup>57</sup>

### *National Cancer Database*

The NCDB, established in 1989, is maintained by the American Cancer Society and the ACS.<sup>58</sup> The database contains over 34 million records and approximately 70% of newly diagnosed cancer cases across the United States from over 1500 Commission on Cancer-accredited facilities.<sup>58</sup> The database contains information for all types of cancer and includes patient demographics, cancer stage, and interventions, including first course therapy and type of surgical resection, histological information, and outcomes.<sup>59-61</sup> The database also includes public access to more than 9 million cases of the 14 most commonly diagnosed solid tumors in the United States from 2003 to 2014, including brain cancer.<sup>58</sup> Limitations include changes in the classification instructions given to tumor registries and changes in staging rules that impact comparability of data over the years.<sup>62</sup>

### **Agency for Healthcare Research and Quality and Healthcare Cost and Utilization Project (AHRQ-HCUP) Databases**

All databases are available for purchase through the Healthcare Cost and Utilization Project Central Distributor.

### *National Inpatient Sample*

The NIS is the largest publicly available database of inpatient stays in the United States, containing information from all payer types. The NIS represents a 20% stratified sample of discharges from US community hospitals, excluding long-term acute care and rehabilitation hospitals from 2012 onward. The 2014 NIS sampling represented an area of 44 states and more than 96% of the US population.<sup>63</sup> The database contains patient demographics, total cost and length of hospitalization, status at discharge, diagnosis and procedural codes, hospital characteristics, and severity and comorbidity information.<sup>64-67</sup> Limitations of the database include lack of participation from several US states and lack of post-discharge information.<sup>63,68</sup>

### *Kids' Inpatient Database*

KID is the largest pediatric inpatient database for patients younger than 21 yr in the United States. It includes all payer types and contains approximately 3 million pediatric discharges per year. The database contains information in 3-yr intervals from 1997 to 2012 and 2016.<sup>69</sup> Collected variables include patient demographics, hospital characteristics, procedural and diagnostic codes, total charges and length of stay, type of payment, and severity and comorbidity measures.<sup>70-73</sup> In 2012, the database contained a weighted sample of 80% of pediatric discharges and 10% of normal newborn discharges from more than 4100 community hospitals in the United States.<sup>74</sup> The large sample size allows for analysis of rare neurosurgical conditions and treatments. Limitations of the database include lack of discharge information from rehabilitation hospitals and inability to perform longitudinal analysis across years.<sup>69</sup>

### *Nationwide Readmissions Database*

NRD focuses on the analysis of hospital readmissions, with data from approximately 15 million discharges each year drawn from the SID. In 2014, the data came from 22 states that covered 49.3% of all US hospitalizations.<sup>75</sup> The database contains patient demographics, payment source, total hospitalization cost, ICD procedure and diagnostic codes, and specific variables that pertain to readmission, including verified confirmation of multiple discharges for individuals, length of inpatient stay, and timing between admissions for a specific patient.<sup>75,76</sup> In order to track patients using different hospitals within a state, individual patients are given a linkage number in the SID. A main limitation of the database is the inability to use the information for the analysis of regional-, state-, or hospital-specific readmission rates and no clear indication if the readmission was related to prior admissions.<sup>75</sup>

### *State Inpatient Database*

SID is an administrative claims database that contains records starting in 1990 for state inpatient hospital discharges from an area that covers approximately 90% of all US hospital discharges and 97% of all US community hospital discharges.<sup>12,77</sup> The database contains patient demographics, diagnosis and procedure codes, payment sources, total costs of hospitalization, admission and discharge status, and length of hospital stay.<sup>12</sup> Strengths of the data include ability to analyze hospitalization patterns within a geographical area containing multiple states, 1 state, or within a state.<sup>78</sup> Limitations include variability in available data depending on state and year, and loss of revisit analysis if the patient changed states.<sup>79,80</sup>

### *State Ambulatory Surgery and Services Database*

SASD contains records for state ambulatory surgical cases from ambulatory surgery centers and hospital-based outpatient departments across 35 states.<sup>79-83</sup> The database contains patient demographics, total cost of procedure, payment source, diagnosis and procedure codes, and identifiers that permit linkage to the SID.<sup>82</sup> Limitations include variability in available data depending on state and year, and loss of revisit analysis if the patient changed states.<sup>79,80</sup>

### **CMS Claims Data**

CMS databases through the State Data Resource Center contain patient information derived from claims data. The information available for request includes part A (inpatient, outpatient, skilled nursing facility, home health, and hospice) and part B (carrier and durable medical equipment), patient demographics, service dates, ICD diagnosis and procedural codes, and type of health coverage (Medicare, Medicaid).<sup>84,85</sup> Claims data are available in varying privacy levels with Research Identifiable Files, which contain beneficiary-level protected health information; Limited Data Sets, which contain beneficiary-level health information with selected variables "blacked, encrypted,

or ranged”; and Public Use Files, which contain “aggregated summary level health information.”<sup>86</sup> Data contained vary by file type; the inpatient Limited Data Sets contain procedural ICD codes, Medicare Severity Diagnosis Related Group, reimbursement amount, hospital provider number, dates of services, and patient demographic information.<sup>86</sup> The data are available for purchase through the CMS website. Limitations of the database include variable granularity depending on the type of files requested and variable coding of disease-specific variables pertinent to the field of investigation.<sup>87-89</sup>

## Private Databases

### *Vizient Clinical Data Base and Resource Manager*

Vizient, formerly the University HealthSystem Consortium, contains data representing over 97% of US academic medical centers and 160 community hospitals.<sup>90-92</sup> The database contains demographic, financial, and clinical information including the length of stay, complication rates, readmission rates, morbidity and mortality rates, and medication use.<sup>90,91</sup> The database provides the ICD codes for each hospital discharge that can be used for research inclusion or exclusion criteria. Limitations of the data include lack of specific individual patient characteristics and inability to determine the timing of complications in reference to the date of surgery.<sup>90,91</sup>

### *MarketScan Research Databases*

Truven Health Analytics MarketScan Research Databases include patient information from commercial insurance, hospital drug, primary care electronic medical records, and multi-state Medicaid and Medicare Supplemental databases.<sup>93-96</sup> The databases include demographic information (age, gender, state, and employment status), insurance features (plan type and coverage limitations), financial information (total admission payments and net payments), inpatient and outpatient medical information (length of stay, primary/secondary diagnosis, service ICD, and CPT codes), and drug information (therapeutic class, generic product identification, and days supplied).<sup>93,94,97,98</sup> Data also include information on claims related to hospital drugs, mortality, primary care electronic medical records, benefit plan design, lab values, health and productivity management, health risk appraisals, and dental health.<sup>94,99,100</sup> The databases have records for more than 200 million unique patients since 1995.<sup>95,96</sup> Data are de-identified from individual patients; however, each patient is assigned a unique identification number that can be used to identify patients across modalities within each database. Petraglia et al’s<sup>97</sup> reported limitations of the MarketScan database include lack of comprehensive clinical factors relevant to neurosurgery.

### *Premier Healthcare Database*

Premier Healthcare Database (Premier, Inc., Charlotte, North Carolina), established in 2000, is a US-based administrative database that contains billed items from hospital discharge as

well as hospital-based outpatient encounters.<sup>101</sup> The database contains over 147 million unique patients with 550 million outpatient visits and 80 million inpatient admissions from greater than 700 hospitals, representing approximately 20% of annual US inpatient discharges.<sup>101-103</sup> The database includes information on patient demographics, procedure and diagnosis codes (CPT, ICD), hospital costs and charges, physician characteristics, outcomes (length of stay and readmissions), laboratory results, drug utilization, and hospital characteristics such as teaching status, geographic location, and number of beds.<sup>103</sup> Each patient is given a unique identifier to track admissions across the inpatient and outpatient setting. Limitations of the Premier Healthcare Database include lack of comprehensive clinical findings for certain neurosurgical study questions.<sup>102</sup>

### *PearlDiver*

PearlDiver Technologies, Inc. (Warsaw, Indiana) established the PearlDiver Patient Record Database as a commercially available national database of public and private insurance records. PearlDiver contains patient records from the PearlDiver Private Payer Database of the UnitedHealth Group, Medicare Carrier Files, and Medicare Standard Analytical Files.<sup>104</sup> PearlDiver has partnered with Comprehensive Health Insights to integrate Humana’s claims database, which represents 22 million patients from 2007 to Q2 2016 in the PearlDiver database.<sup>105-108</sup> Wang et al<sup>107</sup> reported that “from 2007 to 2011, the database (PearlDiver) captured 24.6–26.3 million patients, representing approximately 9% of the US population younger than 65 yr and approximately 13% of the US population with private insurance.” PearlDiver includes diagnosis and procedure codes (ICD, CPT, and Diagnosis Related Group), patient demographics, discharge status, length of stay, postoperative complications, physician specialty, medication information, and hospital geographic information.<sup>108</sup> Wang et al<sup>107</sup> reported that PearlDiver has unequal geographic representation within the United States, as UnitedHealth group has increased coverage in the south relative to other regions of the country.

### *OptumLabs Data Warehouse*

OptumLabs Data Warehouse, founded as a partnership between Optum and Mayo Clinic in 2013, contains de-identified medical and pharmacy claims for over 150 million patients.<sup>109,110</sup> Data can be queried using ICD-9-CM diagnosis codes, CPT procedure codes, and Healthcare Common Procedure Coding System procedure codes, site of service codes, and provider specialty codes.<sup>111</sup> The database contains sociodemographic and clinical information (patient vital signs, diagnoses and treatments from electronic health records, physician text notes, pathology reports, operative reports, and imaging notes), diagnosis and treatment information from all claims both inpatient and outpatient, and pharmacy prescriptions filled.<sup>109</sup> Limitations of that database include the use of nonspecific diagnosis codes.<sup>111</sup>

### *Pediatric Health Information System*

The Children's Hospital Association (CHA) maintains the PHIS with data from more than 45 children's hospitals.<sup>112-115</sup> PHIS contains clinical data, including procedures and diagnosis (ICD) codes and utilization data for inpatient, emergency department, ambulatory surgery, and observation unit patient encounters.<sup>113-115</sup> Simon et al<sup>113,114</sup> reported that the limitations of PHIS data include missing data on factors of clinical interest and limitations of patient identification based on ICD-9 codes.

## DISCUSSION

### Unmet Investigative and Clinical Need

Limitations of surgical outcomes research often cited in the literature include selection bias, relatively small sample sizes, and the lack of effectiveness of randomized controlled trials for surgical interventions, which together result in a relatively lower quality of evidence to support the outcomes-related conclusion at hand.<sup>116-122</sup> Historically, much of the neurosurgical outcomes literature has been driven by single-center cohort studies of specific procedures, which typically have a relatively small sample size, lack adequate controls, and are retrospective in nature.

Therefore, the purpose of large national databases includes meeting both an unmet investigative and clinical need. In terms of investigation, national databases provide externally validated, nationally or internationally representative outcomes. Clinically, the data and resulting studies can improve neurosurgical practice by providing predictive risk scores and clinical risk calculators. Eventually, with the aid of novel technologies such as machine learning and predictive analytics, these databases may be able to improve clinical encounters by providing analytics in the office. Nevertheless, the databases listed in this study are imperfect, and there remain many obstacles to successfully integrating big data in clinical research and clinical care. A careful discussion of strengths and limitations and the possible effects of successfully integrating these data can shed light on the progress that must be made in the coming years.

### Strengths and Limitations

Large data sets offer the advantage of a large population size and multi-institutional sampling that often cover a diverse patient population with respect to race, gender, socioeconomic status, and insurance coverage. These data are often collected from both academic and community medical centers, which may result in more representative data sets. The databases contain critical patient information, including demographics, preoperative variables, medical and surgical procedures, inpatient and postoperative complications, and mean cost and length of hospitalization. Together, the large number of patients included and the useful variables collected can be used to provide convincing analyses of postoperative outcomes.

Problems with the use of these large databases for epidemiological or clinical research include the veracity of data coded in

administrative databases, degree of completeness, and standardization of the information.<sup>39,123-135</sup> Other limitations include lack of variables needed to comprehensively identify predictive factors suggested by the literature.

If a neurosurgeon has a particular clinical question, they are interested in addressing using a large national database such as those described here, there are several barriers to a successful investigation. First, the researcher must identify the data set most useful for answering the clinical question at hand, by considering the data and number of patient records needed to adequately power the study, and the costs and other limitations associated with each database. Once the proper database is identified, the researcher must then access it, either by applying for access or purchasing access, depending on the database selected. Financial barriers likely limit research in this arena, but administrative barriers may be just as important—applying for data is often a prolonged process, and it can take months from the time a project is conceived until the data arrive for analysis. Finally, the data must then be accurate, complete, useful, and analyzed in a meaningful way, leading to an answer to the original clinical question. These obstacles, though obvious, are nevertheless the most pressing logistical issues in the use of the valuable data described in this study. These databases, and those that will be established in the future, must make every effort to limit barriers to access.

### Clinical Decision-Making

For neurosurgeons, the ultimate purpose of establishing and studying databases is to provide meaningful insights that aid clinical decision-making. Establishing uniform data veracity, data completeness, and validated clinical and patient-reported measures is a prerequisite to this goal. Disease-specific variables that are not formally recorded or collected cannot be optimized. As demonstrated by the lack of these core concepts in many of the databases studied here, much work remains to be done in the construction and redesign of new and existing databases. At present, the strengths of large patient populations, integration of multiple data streams from different databases, and multi-institutional sampling have allowed for the creation of predictive risk scores and clinical risk calculators. However, if these limitations of present databases are successfully addressed, the role of databases can continue to expand significantly in clinical decision-making. Personalized medicine, as it relates to databases in neurosurgery, will be achieved when structured and unstructured data from the clinical encounter are seamlessly translated from health records to both individual site-based and national registries. The power of predictive analytics, integration of continuous patient physiological sensors, and continuous subjective self-report surveys from mobile devices will then become the next frontier for patient databases and personalized medicine in neurosurgery.

Health care reform has provided the financial incentives for health care systems, providers, and payers to work together in

reducing unnecessary costs and improving the quality of care. The analysis of the databases mentioned in this review has demonstrated a growth in research focused on surgical outcomes, has helped identify risk factors for adverse events, and has drawn attention to areas where quality improvement is required. The ultimate goal of outcomes research is to identify ways to supplement and improve clinical decision-making, with the purpose of improving patient care at the level of the individual and at the level of the institution, region, state, and nation. The future of national health care dialogue will be informed by the mobilization and analysis of big data, including the improvement of existing databases and the continued creation of new registries.

## CONCLUSION

Limitations of large national databases for neurosurgery include lack of granularity, variable data accuracy in administrative databases, and variable data completeness. Strengths of multi-institutional registries include greater external validity, large numbers of rare pathologies, and comparative health analytics. The improvement of existing data streams and the creation of new registries will inform the future of neurosurgical practice.

## Disclosure

The authors have no personal, financial, or institutional interest in any of the drugs, materials, or devices described in this article.

## REFERENCES

- Porter ME, Lee TH. Why strategy matters now. *N Engl J Med*. 2015; 372(18):1681-1684.
- Kamath AF, Courtney PM, Bozic KJ, Mehta S, Parsley BS, Froimson MI. Bundled payment in total joint care: survey of AAHKS membership attitudes and experience with alternative payment models. *J Arthroplasty*. 2015;30(12):2045-2056.
- McCarthy M. Controversial online "scorecard" shows complication rates of 17,000 US surgeons. *BMJ*. 2015;351:h3873.
- Bekelis K, McGirt MJ, Parker SL, et al. The present and future of quality measures and public reporting in neurosurgery. *Neurosurg Focus*. 2015; 39(6):E3.
- Menger RP, Guthikonda B, Storey CM, Nanda A, McGirt M, Asher A. Neurosurgery value and quality in the context of the Affordable Care Act: a policy perspective. *Neurosurg Focus*. 2015;39(6):E5.
- Parker SL, McGirt MJ, Bekelis K, et al. The National Neurosurgery Quality and Outcomes Database Qualified Clinical Data Registry: 2015 measure specifications and rationale. *Neurosurg Focus*. 2015;39(6):E4.
- Michael L, Klimo P. Outcomes research in neurosurgery: do administrative databases hold the answers? *World Neurosurg*. 2015;84(5): 1193-1195.
- Program ACoNSQL. *User Guide for the 2014 ACS NSQIP Participant Use Data File (PUF)*. 2015. Available at: [https://www.facs.org/~media/files/qualityprograms/nsqip/nsqip\\_puf\\_userguide\\_2014.ashx](https://www.facs.org/~media/files/qualityprograms/nsqip/nsqip_puf_userguide_2014.ashx). Accessed March 4, 2017.
- Nance ML. *National Trauma Data Bank 2015 Annual Report*. 2015. Available at: [https://www.facs.org/~media/files/qualityprograms/trauma/ntdb/ntdb\\_annual\\_report\\_2015.ashx](https://www.facs.org/~media/files/qualityprograms/trauma/ntdb/ntdb_annual_report_2015.ashx). Accessed March 16, 2016.
- National Cancer Data Base Participant Use Data File (PUF) Data Dictionary. 2014. Available at: <https://www.facs.org/~media/files/quality%20programs/cancer/ncdb/puf%20data%20dictionary%20version%20puf%202014.ashx>. Accessed March 16, 2016.
- Hansen LG, Chang S. White Paper. Health Research Data for the Real World: the MarketScan Databases. 2011. Available at: [http://truvenhealth.com/portals/0/assets/PH\\_11238\\_0612\\_TEMP\\_MarketScan\\_WP\\_FINAL.pdf](http://truvenhealth.com/portals/0/assets/PH_11238_0612_TEMP_MarketScan_WP_FINAL.pdf). Accessed March 4, 2017.
- Databases H. *Healthcare Cost and Utilization Project (HCUP)*. 2016. Available at: <http://www.hcup-us.ahrq.gov/sidoverview.jsp>. Accessed February 16, 2016.
- McGirt MJ, Parker SL, Hilibrand A, et al. Lumbar surgery in the elderly provides significant health benefit in the US health care system: patient-reported outcomes in 4370 patients from the N2QOD registry. *Neurosurgery*. 2015;77(suppl 4):S125-S135.
- Guan J, Bisson EF, Dailey AT, Hood RS, Schmidt MH. Comparison of clinical outcomes in the national neurosurgery quality and outcomes database for open versus minimally invasive transforaminal lumbar interbody fusion. *Spine (Phila Pa 1976)*. 2016;41(7):E416-E421.
- Azad TD, Kalani M, Wolf T, et al. Building an electronic health record integrated quality of life outcomes registry for spine surgery. *J Neurosurg Spine*. 2016;24(1):176-185.
- Ziewacz JE, McGirt MJ, Chewning SJ. Adverse events in neurosurgery and their relationship to quality improvement. *Neurosurg Clin N Am*. 2015;26(2):157-165.
- Asher AL, Speroff T, Dittus RS, et al. The National Neurosurgery Quality and Outcomes Database (N2QOD): a collaborative North American outcomes registry to advance value-based spine care. *Spine*. 2014;39(22 suppl 1):S106-S116.
- McGirt MJ, Speroff T, Dittus RS, Harrell FE, Jr, Asher AL. The National Neurosurgery Quality and Outcomes Database (N2QOD): general overview and pilot-year project description. *Neurosurg Focus*. 2013;34(1):E6.
- Alliance N. Current QOD Registries Available. 2017. Available at: [http://www.neuropoint.org/NPA\\_N2QOD\\_Current\\_Modules.aspx/](http://www.neuropoint.org/NPA_N2QOD_Current_Modules.aspx/). Accessed March 4, 2017.
- Hankinson TC, Dudley RW, Torok MR, et al. Short-term mortality following surgical procedures for the diagnosis of pediatric brain tumors: outcome analysis in 5533 children from SEER, 2004-2011. *J Neurosurg Pediatr*. 2016;17(3):289-297.
- Smoll NR, Weber DC. SEER analysis and statistical judgment. *J Neurosurg*. 2014;120(2):578-579.
- Sampson JH, Lad SP, Herndon JE, 2nd, Starke RM, Kondziolka D. SEER insights. *J Neurosurg*. 2014;120(2):297-298.
- Institute NC. *SEER Cancer Statistics Review 1975-2013*. 2016. Available at: [https://seer.cancer.gov/csr/1975\\_2013/browse\\_csr.php](https://seer.cancer.gov/csr/1975_2013/browse_csr.php). Accessed March 4, 2017.
- Noorbakhsh A, Tang JA, Marcus LP, et al. Gross-total resection outcomes in an elderly population with glioblastoma: a SEER-based analysis: clinical article. *J Neurosurg*. 2014;120(1):31-39.
- Mukherjee D, Chaichana KL, Parker SL, Gokaslan ZL, McGirt MJ. Association of surgical resection and survival in patients with malignant primary osseous spinal neoplasms from the Surveillance, Epidemiology, and End Results (SEER) database. *Eur Spine J*. 2013;22(6):1375-1382.
- Jeswani S, Nuño M, Folkerts V, Mukherjee D, Black KL, Patil CG. Comparison of survival between cerebellar and supratentorial glioblastoma patients: Surveillance, Epidemiology, and End Results (SEER) analysis. *Neurosurgery*. 2013;73(2):240-246.
- Stessin AM, Schwartz A, Judanin G, et al. Does adjuvant external-beam radiotherapy improve outcomes for nonbenign meningiomas? A Surveillance, Epidemiology, and End Results (SEER)-based analysis: clinical article. *J Neurosurg*. 2012;117(4):669-675.
- Selvanathan SK, Hammouche S, Smethurst W, Salminen HJ, Jenkinson MD. Outcome and prognostic features in adult pineoblastomas: analysis of cases from the SEER database. *Acta Neurochir (Wien)*. 2012;154(5):863-869.
- Mukherjee D, Chaichana KL, Gokaslan ZL, Aaronson O, Cheng JS, McGirt MJ. Survival of patients with malignant primary osseous spinal neoplasms: results from the Surveillance, Epidemiology, and End Results (SEER) database from 1973 to 2003: clinical article. *J Neurosurg Spine*. 2011;14(2):143-150.
- Mukherjee D, Chaichana KL, Adogwa O, et al. Association of extent of local tumor invasion and survival in patients with malignant primary osseous spinal neoplasms from the surveillance, epidemiology, and end results (SEER) database. *World Neurosurg*. 2011;76(6):580-585.
- Selvanathan SK, Hammouche S, Salminen HJ, Jenkinson MD. Outcome and prognostic features in anaplastic ganglioglioma: analysis of cases from the SEER database. *J Neurooncol*. 2011;105(3):539-545.
- Goodwin TL, Sainani K, Fisher PG. Incidence patterns of central nervous system germ cell tumors: a SEER Study. *J Pediatr Hematol Oncol*. 2009;31(8):541-544.

33. Davis FG, Freels S, Grutsch J, Barlas S, Brem S. Survival rates in patients with primary malignant brain tumors stratified by patient age and tumor histological type: an analysis based on Surveillance, Epidemiology, and End Results (SEER) data, 1973-1991. *J Neurosurg.* 1998;88(1):1-10.
34. Institute NC. *SEER Data 1973–2013.* 2013. Available at: <https://seer.cancer.gov/data/>. Accessed March 4, 2017.
35. Healthcare Delivery Research Program. *Data, Tools, and Initiatives. Publicly Accessible Data. SEER-Medicare Linked Database. The SEER-Medicare Data Files. SEER Program & Data.* Available at: <https://healthcaresdelivery.cancer.gov/seermedicare/aboutdata/program.html>. Accessed June 2, 2017.
36. Nuño M, Ly D, Mukherjee D, Ortega A, Black KL, Patil CG. Quality of surgical care and readmission in elderly glioblastoma patients. *Neurooncol Pract.* 2014;1(2):33-39.
37. Halasz LM, Weeks JC, Neville BA, Taback N, Punglia RS. Use of stereotactic radiosurgery for brain metastases from non-small cell lung cancer in the United States. *Int J Radiat Oncol Biol Phys.* 2013;85(2):e109-e116.
38. Engels EA, Pfeiffer RM, Ricker W, Wheeler W, Parsons R, Warren JL. Use of surveillance, epidemiology, and end results-medicare data to conduct case-control studies of cancer among the US elderly. *Am J Epidemiol.* 2011;174(7):860-870.
39. Warren JL, Klabunde CN, Schrag D, Bach PB, Riley GF. Overview of the SEER-Medicare data: content, research applications, and generalizability to the United States elderly population. *Med Care.* 2002;40(8 suppl):IV-3-18.
40. Institute NC. *SEER-Medicare: Number of Cases for Selected Cancers Appearing in the Data.* 2017. Available at: <https://healthcaresdelivery.cancer.gov/seermedicare/aboutdata/cases.html>. Accessed March 12, 2017.
41. Healthcare Delivery Research Program. *Data, Tools, Initiatives. Publicly Accessible Data. SEER-Medicare Linked Database. Analytic Support for Researchers. Measures that are Limited or not Available in the Data.* Available at: <https://healthcaresdelivery.cancer.gov/seermedicare/considerations/measures.html>. Accessed June 2, 2017.
42. Ostrom QT, Gittleman H, Fulop J, et al. CBTRUS statistical report: primary brain and central nervous system tumors diagnosed in the United States in 2008-2012. *Neuro Oncol.* 2015;17( suppl 4):iv1-iv62.
43. Kshetty VR, Ostrom QT, Kruchko C, Al-Mefty O, Barnett GH, Barnholtz-Sloan JS. Descriptive epidemiology of World Health Organization grades II and III intracranial meningiomas in the United States. *Neuro Oncol.* 2015;17(8):1166-1173.
44. States CBTRotU. *2016 CBTRUS Fact Sheet.* 2016. Available at: <http://www.cbtrus.org/factsheet/factsheet.html>. Accessed March 4, 2017.
45. Raval MV, Dillon PW, Bruny JL, et al. American College of Surgeons National Surgical Quality Improvement Program Pediatric: a phase 1 report. *J Am Col Surg.* 2011;212(1):1-11.
46. Longo WE, Cheadle W, Fink A, et al. The role of the Veterans Affairs Medical Centers in patient care, surgical education, research and faculty development. *Am J Surg.* 2005;190(5):662-675.
47. Dasenbrock HH, Liu KX, Devine CA, et al. Length of hospital stay after craniotomy for tumor: a National Surgical Quality Improvement Program analysis. *Neurosurg Focus.* 2015;39(6):E12.
48. American College of Surgeons National Surgical Quality Improvement Program. *User Guide for the 2015 ACS NSQIP Participant Use Data File (PUF).* Available at: [https://www.facs.org/~media/files/quality%20programs/nsqip/nsqip\\_puf\\_user\\_guide\\_2015.ashx](https://www.facs.org/~media/files/quality%20programs/nsqip/nsqip_puf_user_guide_2015.ashx). Accessed March 16, 2016.
49. Karhade AV, Cote DJ, Larsen AM, Smith TR. Neurosurgical infection rates and risk factors: a national surgical quality improvement program analysis of 132,000 patients, 2006-2014. *World Neurosurg.* 2017;97:205-212.
50. American College of Surgeons. *National Trauma Data Bank. 2016. Annual Report.* 2016. Available at: <https://www.facs.org/~media/files/quality%20programs/trauma/ntdb/ntdb%20annual%20report%202016.ashx>. Accessed March 16, 2016.
51. Nembhard CE, Hwabejire JO, Obirieze AC, et al. Trauma associated splenic artery aneurysms: an analysis of the National Trauma Data Bank. *Am J Surgery.* 2016;211(4):739-743.
52. Alkhoury F, Kyriakides TC. Intracranial pressure monitoring in children with severe traumatic brain injury: National Trauma Data Bank-based review of outcomes. *JAMA Surg.* 2014;149(6):544-548.
53. Chen CM, Yi HY, Yoon YH, Dong C. Alcohol use at time of injury and survival following traumatic brain injury: results from the National Trauma Data Bank. *J Stud Alcohol Drugs.* 2012;73(4):531-541.
54. Majidi S, Siddiq F, Qureshi AI. Prehospital neurologic deterioration is independent predictor of outcome in traumatic brain injury: analysis from National Trauma Data Bank. *Am J Emerg Med.* 2013;31(8):1215-1219.
55. Sweeney TE, Salles A, Harris OA, Spain DA, Staudenmayer KL. Prediction of neurosurgical intervention after mild traumatic brain injury using the national trauma data bank. *World J Emerg Surg.* 2015;10:23.
56. Teixeira PG, Karamanos E, Okoye OT, et al. Splenectomy in patients with traumatic brain injury: protective or harmful? A National Trauma Data Bank analysis. *J Trauma Acute Care Surg.* 2013;75(4):596-601.
57. Auer R, Riehl J. The incidence of deep vein thrombosis and pulmonary embolism after fracture of the tibia: an analysis of the National Trauma Databank. *J Clin Orthop Trauma.* 2017;8(1):38-44.
58. Surgeons ACo. *National Cancer Database.* 2016. Available at: <https://www.facs.org/quality%20programs/cancer/ncdb>. Accessed March 16, 2016.
59. Surgeons ACo. *Public Access to Cancer Data.* 2016. Available at: <https://www.facs.org/quality%20programs/cancer/ncdb/publicaccess>. Accessed March 16, 2016.
60. Greenleaf EK, Hollenbeak CS, Wong J. Trends in the use and impact of neoadjuvant chemotherapy on perioperative outcomes for resected gastric cancer: evidence from the American College of Surgeons National Cancer Database. *Surgery.* 2016;159(4):1099-1112.
61. Surawicz TS, Davis F, Freels S, Laws ER, Menck HR. Brain tumor survival: results from the National Cancer Data Base. *J Neurooncol.* 1998;40(2):151-160.
62. Cooper JS, Porter K, Mallin K, et al. National Cancer Database report on cancer of the head and neck: 10-year update. *Head Neck.* 2009;31(6):748-758.
63. (HCUP) HCAUP. *HCUP Databases.* 2016. Available at: <http://www.hcup-us.ahrq.gov/nisoverview.jsp>. Accessed March 16, 2016.
64. Attenello FJ, Wen T, Huang C, Cen S, Mack WJ, Acosta FL. Evaluation of weekend admission on the prevalence of hospital acquired conditions in patients receiving thoracolumbar fusions. *J Clin Neurosci.* 2015;22(8):1349-1354.
65. Bir SC, Maiti TK, Ambekar S, Nanda A. Incidence, hospital costs and in-hospital mortality rates of epidural hematoma in the United States. *Clin Neurol Neurosurg.* 2015;138:99-103.
66. Kalakoti P, Missios S, Maiti T, et al. Inpatient outcomes and postoperative complications after primary versus revision lumbar spinal fusion surgeries for degenerative lumbar disc disease: a national (nationwide) inpatient sample analysis, 2002-2011. *World Neurosurg.* 2016;85:114-124.
67. Singla A, Fargen K, Blackburn S, et al. National treatment practices in the management of infectious intracranial aneurysms and infective endocarditis. *J Neurointerv Surg.* 2016;8(7):741-746.
68. Alluri RK, Leland H, Heckmann N. Surgical research using national databases. *Ann Transl Med.* 2016;4(20):393.
69. Project HCAU. *HCUP Databases: KID.* 2016. Available at: <https://www.hcup-us.ahrq.gov/kidoverview.jsp>. Accessed February 20, 2016.
70. Attenello FJ, Christian E, Wen T, et al. Reevaluating the weekend effect on patients with hydrocephalus undergoing operative shunt intervention. *J Neurosurg Pediatr.* 2016;17(2):156-162.
71. Lin Y, Pan IW, Mayer RR, Lam S. Complications after craniostomy surgery: comparison of the 2012 Kids' Inpatient Database and Pediatric NSQIP Database. *Neurosurg Focus.* 2015;39(6):E11.
72. Salazar JH, Goldstein SD, Yang J, et al. Regionalization of pediatric surgery: trends already underway. *Ann Surg.* 2016;263(6):1062-1066.
73. Shin JI, Lee NJ, Cho SK. Pediatric cervical spine and spinal cord injury: a national database study. *Spine.* 2016;41(4):283-292.
74. Distributor HC. *Introduction to the HCUP KIDS' Inpatient Database (KID).* 2015. Available at: [https://www.hcup-us.ahrq.gov/db/nation/kid/KID\\_2012\\_Introduction.pdf](https://www.hcup-us.ahrq.gov/db/nation/kid/KID_2012_Introduction.pdf). Accessed March 16, 2016.
75. Quality AfHRA. *NRD Overview.* 2015. Available at: <https://www.hcup-us.ahrq.gov/nrdoverview.jsp>. Accessed February 16, 2016.
76. Poojary P, Saha A, Chauhan K, et al. Predictors of hospital readmissions for ulcerative colitis in the United States: a national database study. *Inflamm Bowel Dis.* 2017;23(3):347-356.
77. Schairer WW, Nwachukwu BU, McCormick F, Lyman S, Mayman D. Use of hip arthroscopy and risk of conversion to total hip arthroplasty: a population-based analysis. *Arthroscopy.* 2016;32(4):587-593.
78. Reports HSR. *Healthcare Cost and Utilization Project.* 2015. Available at: <https://www.hcup-us.ahrq.gov/db/state/sidrelatedreports.jsp>. Accessed February 16, 2016.

79. Bekelis K, Missios S, Kakoulides G, Rahmani R, Simmons N. Selection of patients for ambulatory lumbar discectomy: results from four US states. *Spine J*. 2014;14(9):1944-1950.
80. Missios S, Rahmani R, Bekelis K. Spinal cord stimulators: socioeconomic disparities in four US states. *Neuromodulation*. 2014;17(5):451-455; discussion 455-456.
81. Fischer JP, Fox JP, Nelson JA, Kovach SJ, Serletti JM. A longitudinal assessment of outcomes and healthcare resource utilization after immediate breast reconstruction-comparing implant- and autologous-based breast reconstruction. *Ann Surg*. 2015;262(4):692-699.
82. Databases H. *Healthcare Cost and Utilization Project*. 2015. Available at: <https://www.hcup-us.ahrq.gov/sasdooverview.jsp>. Accessed February 16, 2016.
83. Bekelis K, Missios S, Roberts DW. Institutional charges and disparities in outpatient brain biopsies in four US States: the State Ambulatory Database (SASD). *J Neurooncol*. 2013;115(2):277-283.
84. Center SDR. *Using and Requesting Medicare Data for Medicare-Medicaid Care Coordination and Program Integrity*. 2016. Available at: [http://www.statedataresourcecenter.com/assets/files/SDRC\\_FAQ.pdf](http://www.statedataresourcecenter.com/assets/files/SDRC_FAQ.pdf). Accessed March 4, 2017.
85. Lichtman JH, Leifheit-Limson EC, Goldstein LB. Centers for medicare and medicaid services medicare data and stroke research: goldmine or landmine? *Stroke*. 2015;46(2):598-604.
86. Research Data Assistance Center. *CMS Data Request Center*. Available at: <https://www.resdac.org/cms-data/request/cms-data-request-center>. Accessed March 8, 2017.
87. Lichtman JH, Jones SB, Wang Y, Watanabe E, Leifheit-Limson E, Goldstein LB. Outcomes after ischemic stroke for hospitals with and without Joint Commission-certified primary stroke centers. *Neurology*. 2011;76(23):1976-1982.
88. Lichtman JH, Leifheit-Limson EC, Jones SB, Wang Y, Goldstein LB. 30-day risk-standardized mortality and readmission rates after ischemic stroke in critical access hospitals. *Stroke*. 2012;43(10):2741-2747.
89. Hammill BG, Hernandez AF, Peterson ED, Fonarow GC, Schulman KA, Curtis LH. Linking inpatient clinical registry data to Medicare claims data using indirect identifiers. *Am Heart J*. 2009;157(6):995-1000.
90. Bisson EF, Bowers CA, Hohmann SF, Schmidt MH. Smoking is associated with poorer quality-based outcomes in patients hospitalized with spinal disease. *Front Surg*. 2015;2:20.
91. Fennell VS, Martirosyan NL, Palejwala SK, Lemole GM, Jr, Dumont TM. Morbidity and mortality of patients with endovascularly treated intracerebral aneurysms: does physician specialty matter? *J Neurosurg*. 2016;124(1):13-17.
92. Vizient. *Clinical Analytics and Benchmarking*. 2017. Available at: <https://www.vizientinc.com/What-we-do/Patient-Impact/Clinical-analytics-and-benchmarking>. Accessed March 12, 2017.
93. Analytics TH. *Putting Research Data Into Your Hands With the MarketScan Databases*. 2017. Available at: <http://truvenhealth.com/markets/life-sciences/products/data-tools/marketscan-databases>. Accessed March 4, 2017.
94. Lam S, Harris D, Rocque BG, Ham SA. Pediatric endoscopic third ventriculostomy: a population-based study. *J Neurosurg Pediatr*. 2014;14(5):455-464.
95. Ackerman SJ, Polly DW, Jr, Knight T, Schneider K, Holt T, Cummings J Jr. Comparison of the costs of nonoperative care to minimally invasive surgery for sacroiliac joint disruption and degenerative sacroiliitis in a United States commercial payer population: potential economic implications of a new minimally invasive technology. *Clinicoecon Outcomes Res*. 2014;6:283-296.
96. Ackerman SJ, Polly DW, Jr, Knight T, Holt T, Cummings J. Management of sacroiliac joint disruption and degenerative sacroiliitis with nonoperative care is medical resource-intensive and costly in a United States commercial payer population. *Clinicoecon Outcomes Res*. 2014;6:63-74.
97. Petraglia FW, 3rd Farber SH, Han JL, et al. Comparison of bilateral vs. staged unilateral deep brain stimulation (DBS) in Parkinson's disease in patients under 70 years of age. *Neuromodulation*. 2016;19(1):31-37.
98. Connolly ID, Cole T, Veeravagu A, Papat R, Ratliff J, Li G. Craniotomy for resection of meningioma: an age-stratified analysis of the MarketScan longitudinal database. *World Neurosurg*. 2015;84(6):1864-1870.
99. Huang KT, Hazzard M, Thomas S, et al. Differences in the outcomes of anterior versus posterior interbody fusion surgery of the lumbar spine: a propensity score-controlled cohort analysis of 10,941 patients. *J Clin Neurosci*. 2015;22(5):848-853.
100. Adogwa O, Huang K, Hazzard M, et al. Outcomes after cervical laminectomy with instrumented fusion versus expansile laminoplasty: a propensity matched study of 3185 patients. *J Clin Neurosci*. 2015;22(3):549-553.
101. Larsen MD, Cars T, Hallas J. A minireview of the use of hospital-based databases in observational inpatient studies of drugs. *Basic Clin Pharmacol Toxicol*. 2013;112(1):13-18.
102. Henry MK, Zonfrillo MR, French B, Song L, Feudtner C, Wood JN. Hospital variation in cervical spine imaging of young children with traumatic brain injury. *Acad Pediatr*. 2016;16(7):684-691.
103. Premier I. *Premier Research Services Creates Insights Through Data*. 2016. Available at: <https://www.premierinc.com/transforming-healthcare/healthcare-performance-improvement/premier-research-services/>. Accessed February 16, 2016.
104. Savage JW, Kelly MP, Ellison SA, Anderson PA. A population-based review of bone morphogenetic protein: associated complication and reoperation rates after lumbar spinal fusion. *Neurosurg Focus*. 2015;39(4):E13.
105. Bedard NA, Pugely AJ, Elkins JM, et al. The John N. Insall Award: do intra-articular injections increase the risk of infection after TKA? *Clin Orthop Relat Res*. 2017;475(1):45-52.
106. Puvanesarajah V, Jain A, Shimer AL, Shen FH, Hassanzadeh H. The effect of coagulopathies on perioperative complications and clinical outcomes in patients treated with posterior lumbar fusions. *Spine (Phila Pa 1976)*. 2016;41(17):E1063-E1068.
107. Wang D, Joshi NB, Petrigliano FA, et al. Trends associated with distal biceps tendon repair in the United States, 2007 to 2011. *J Shoulder Elbow Surg*. 2016;25(4):676-680.
108. PearlDiver. *Health Research: Research Capabilities*. 2016. Available at: <http://www.pearliverinc.com/researchinfo.html>. Accessed March 4, 2017.
109. Wallace PJ, Shah ND, Dennen T, Bleicher PA, Crown WH. Optum Labs: building a novel node in the learning health care system. *Health Aff (Millwood)*. 2014;33(7):1187-1194.
110. Coleman CI, Peacock WF, Fermann GJ, et al. External validation of a multi-variable claims-based rule for predicting in-hospital mortality and 30-day post-pulmonary embolism complications. *BMC Health Serv Res*. 2016;16(1):610.
111. Maradit Kremers H, Schilz SR, Van Houten HK, et al. Trends in utilization and outcomes of hip arthroscopy in the United States between 2005 and 2013. *J Arthroplasty*. 2017;32(3):750-755.
112. Ruth A, McCracken C, Hall M, Fortenberry J, Hebbard K. 1102: advanced technologies in pediatric severe sepsis: findings from the PHIS database. *Crit Care Med*. 2013;41(12):A278-A279.
113. Simon TD, Hall M, Dean JM, Kestle JR, Riva-Cambrin J. Reinfection following initial cerebrospinal fluid shunt infection. *J Neurosurg Pediatr*. 2010;6(3):277-285.
114. Simon TD, Hall M, Riva-Cambrin J, et al. Infection rates following initial cerebrospinal fluid shunt placement across pediatric hospitals in the United States. Clinical article. *J Neurosurg Pediatr*. 2009;4(2):156-165.
115. Pan IW, Kuo GM, Luerssen TG, Lam SK. Impact of antibiotic prophylaxis for intrathecal baclofen pump surgery in pediatric patients. *Neurosurg Focus*. 2015;39(6):E10.
116. Ingraham AM, Richards KE, Hall BL, Ko CY. Quality improvement in surgery: the American College of Surgeons National Surgical Quality Improvement Program approach. *Adv Surg*. 2010;44:251-267.
117. Bruny JL, Hall BL, Barnhart DC, et al. American College of Surgeons National Surgical Quality Improvement Program Pediatric: a beta phase report. *J Pediatr Surg*. 2013;48(1):74-80.
118. Anderson JE, Chang DC. Using electronic health records for surgical quality improvement in the era of big data. *JAMA Surg*. 2015;150(1):24-29.
119. Mathias B, Lipori G, Moldawer LL, Efron PA. Integrating "big data" into surgical practice. *Surgery*. 2016;159(2):371-374.
120. Nouraei SA, Hudovsky A, Frampton AE, et al. A study of clinical coding accuracy in surgery: implications for the use of administrative big data for outcomes management. *Ann Surg*. 2015;261(6):1096-1107.
121. Patel AA, Singh K, Nunley RM, Minhas SV. Administrative databases in orthopaedic research: pearls and pitfalls of big data. *J Am Acad Orthop Surg*. 2016;24(3):172-179.
122. Zhu VZ, Tuggle CT, Au AF. Promise and limitations of big data research in plastic surgery. *Ann Plast Surg*. 2016;76(4):453-458.
123. Zippin C, Lum D, Hankey BF. Completeness of hospital cancer case reporting from the SEER Program of the National Cancer Institute. *Cancer*. 1995;76(11):2343-2350.

124. Zippin C, Lum D. Study of completeness of the surveillance, epidemiology and end results (SEER) program case ascertainment by hospital size and casefinding source. *Health Rep.* 1993;5(1):87-90.
125. Kim HM, Goodman M, Kim BI, Ward KC. Frequency and determinants of missing data in clinical and prognostic variables recently added to SEER. *J Registry Manag.* 2011;38(3):120-131.
126. Coben JH, Steiner CA, Barrett M, Merrill CT, Adamson D. Completeness of cause of injury coding in healthcare administrative databases in the United States, 2001. *Inj Prev.* 2006;12(3):199-201.
127. Steiner C, Elixhauser A, Schnaier J. The healthcare cost and utilization project: an overview. *Eff Clin Pract.* 2002;5(3):143-151.
128. Riley GF. Administrative and claims records as sources of health care cost data. *Med Care* 2009;47(7 suppl 1):S51-S55.
129. Cleary R, Beard R, Coles J, et al. Comparative hospital databases: value for management and quality. *Qual Health Care.* 1994;3(1):3-10.
130. Grosse SD, Boulet SL, Amendah DD, Oyeku SO. Administrative data sets and health services research on hemoglobinopathies: a review of the literature. *Am J Prev Med.* 2010;38(4 suppl):S557-S567.
131. Johantgen M, Elixhauser A, Bali JK, Goldfarb M, Harris DR. Quality indicators using hospital discharge data: state and national applications. *Jt Comm J Qual Improv.* 1998;24(2):88-105.
132. Meyer GS, Battles J, Hart JC, Tang N. The US Agency for Healthcare Research and Quality's activities in patient safety research. *Int J Qual Health Care.* 2003;15( suppl 1):i25-i30.
133. Moore B, Karaca Z, Henke RM, Wong H. Characteristics and quality of inpatient stays at hospitals affiliated with health systems, 2009-2012: statistical brief #197. *Healthcare Cost and Utilization Project (HCUP) Statistical Briefs.* 2015.
134. Rice HE, Englum BR, Gulack BC, et al. Use of patient registries and administrative datasets for the study of pediatric cancer. *Pediatr Blood Cancers* 2015;62(9):1495-1500.
135. Witt WP, Weiss AJ, Elixhauser A. Overview of hospital stays for children in the United States, 2012: Statistical Brief #187. *Healthcare Cost and Utilization Project (HCUP) Statistical Briefs.* 2014.

## COMMENT

In the present study, the authors provide a description of many available databases with relevance to neurosurgery. They should be commended for carrying out the task of providing the strengths and limitations of each database, which can be a helpful resource to future investigators. Secondary analyses of large datasets provide a means for investigators to paint a broad disease portrait and address questions that would otherwise be difficult to study. Nevertheless, there are certain limitations to the utilization of such databases for research, including lack of granularity and coding errors. It is important to note that institutional series remain the only studies where coding accuracy can be verified by a direct view of the patient record. It is for this reason that complications data from administrative databases report far lower rates of adverse events than institutional series. Although studies on large databases will undoubtedly comprise a significant tool of future surgical outcomes and health services research, the value and role of institutional series to complement administrative and registry data should not be underestimated.

**Mohamad Bydon**  
Rochester, Minnesota