Neurosurgical Forum
BROCA'S AREA

Time trend analysis of database and registry use in the neurosurgical literature: evidence for the advance of registry science

Anthony L. Asher, MD,[1] Sally El Sammak, MD,[2,3]
Giorgos D. Michalopoulos, MD,[2,3] Yagiz U. Yolcu, MD,[2,3]
A. Yohan Alexander, BA,[2,3] John J. Knightly, MD,[4]
Kevin T. Foley, MD,[5] Christopher I. Shaffrey, MD,[6]
Robert E. Harbaugh, MD,[7] Geoffrey A. Rose, MD,[8]
Domagoj Coric, MD,[1] Erica F. Bisson, MD, MPH,[9]
Steven D. Glassman, MD,[10]
Praveen V. Mummaneni, MD, MBA,[11] and
Mohamad Bydon, MD[2,3]

[1]Neuroscience Institute, Carolina Neurosurgery & Spine Associates, Carolinas Healthcare System, Charlotte, North Carolina; [2]Mayo Clinic Neuro-Informatics Laboratory, Department of Neurologic Surgery, Mayo Clinic, Rochester, Minnesota; [3]Department of Neurologic Surgery, Mayo Clinic, Rochester, Minnesota; [4]Atlantic Neurosurgical Specialists, Morristown, New Jersey; [5]Department of Neurosurgery, University of Tennessee, Memphis, Tennessee; [6]Duke Neurosurgery and Orthopaedic Surgery, Duke University Medical Center, Durham, North Carolina; [7]Department of Neurosurgery, College of Medicine, Pennsylvania State University, Hershey, Pennsylvania; [8]Sanger Heart & Vascular Institute, Atrium Health, Charlotte, North Carolina; [9]Department of Neurological Surgery, University of Utah, Salt Lake City, Utah; [10]Norton Leatherman Spine Center, Louisville, Kentucky; and [11]Department of Neurological Surgery, University of California, San Francisco, California

THE establishment of databases and registries has enhanced evidence-based care delivery and facilitated quality-of-care monitoring across medicine.[1,2] In the field of neurosurgery, these data sets are commonly used to facilitate the unification of institutional efforts to improve surgical outcomes, quality of patient care delivery, and cost containment.[3]

Large sample size, heterogeneity, and broad research potential are three main advantages of databases and registries. Simultaneous data entry from multiple institutions results in a volume of information that clinical studies cannot reach.[4,5] These institutions vary in population and academic level, providing a broad representation of the average clinical practice.[6]

Before the availability of specialty-specific registry programs, researchers in the field of neurosurgery had to rely on the administrative, hospital, and insurance claims data.[7,8] However, these data sets were limited by the lack of comprehensive operative variables, incomplete patient characteristics, and missing longitudinal patient-reported outcomes (PROs).[8] The introduction of the Quality Outcomes Database (QOD) by the NeuroPoint Alliance was the principal stepping stone toward utilizing specialty-specific data to promote individualized patient care in the field of neurosurgery; progress toward this goal has been reflected through the creation of predictive models that guide clinical decisions, the identification of risk factors for complications in neurosurgical procedures, and the revelation of socioeconomic disparities in quality of care delivery.[3,7]

In this article, we attempt to 1) describe the trends in the use of databases and registries from 2005 to 2020 across 6 top-ranked neurosurgical journals, 2) update the literature on trends in the most recent 5 years,[7] and 3) review the strengths and limitations of these data sets and their impact on neurosurgical research.

A literature search was performed for all articles published between January 1, 2005, and January 1, 2021, in the following journals: *Journal of Neurosurgery*, *Neurosurgery*, *Journal of Neurosurgery: Spine*, *Journal of Neurosurgery: Pediatrics*, *Neurosurgical Focus*, and *World Neurosurgery/Surgical Neurology*. Linear regression was used to estimate yearly trends in the utilization of registries and databases among the included journals.

After screening 40,039 articles, we identified a total of 832 articles that gleaned their data from databases and registries. An evaluation of the rate of registry- or database-related publications revealed an increase from 0.57% in 2005 to 3.26% by 2020 (Fig. 1A). The rate also increased across all journals, particularly *Journal of Neurosurgery*, *Journal of Neurosurgery: Spine*, *Neurosurgery*, and *World*
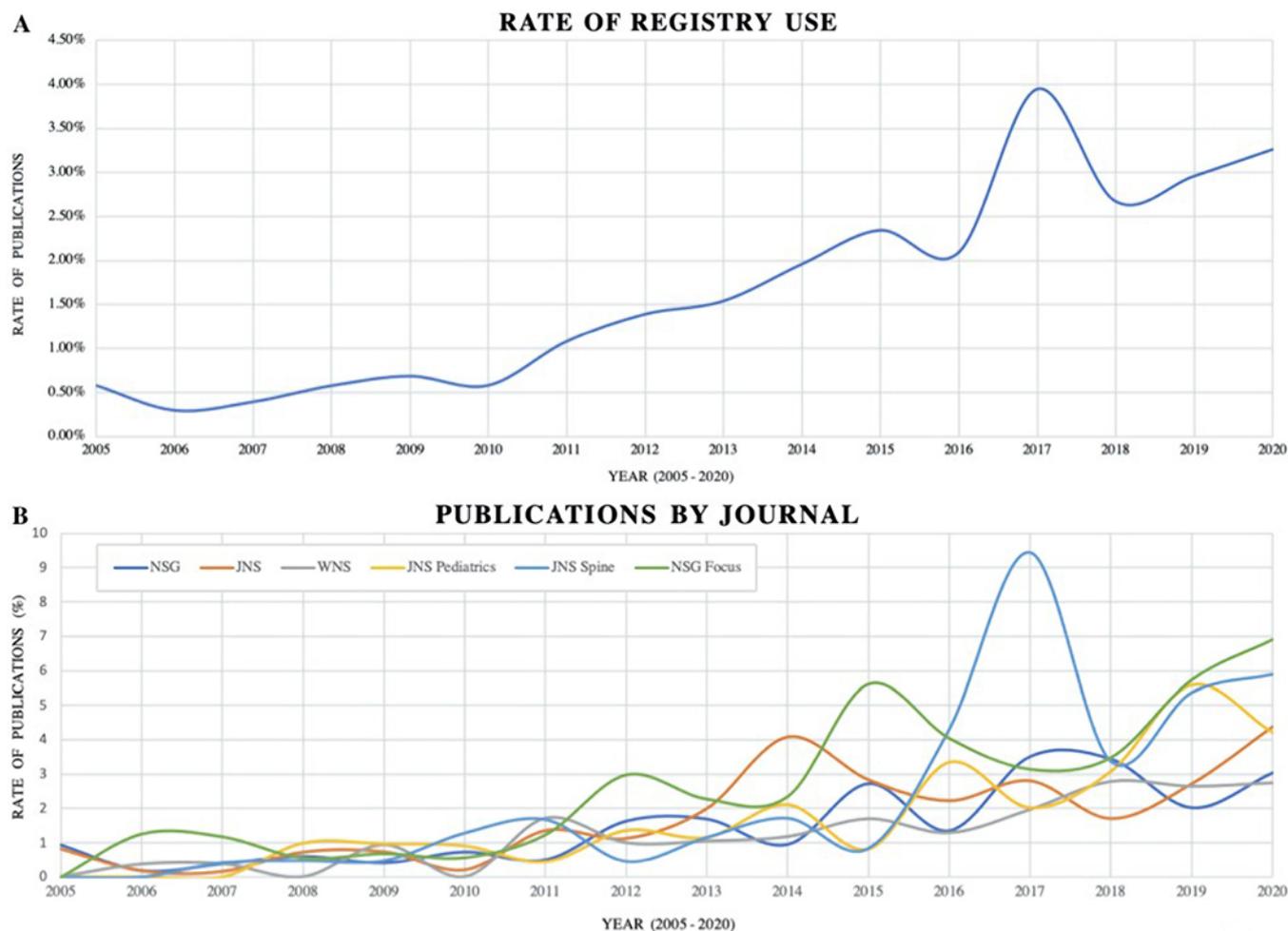
FIG. 1. Graphs showing the time series constructed with the rate of registry use (**A**) and the rate of registry publications by different journals (**B**) throughout the years 2005–2020. JNS = *Journal of Neurosurgery*; NSG = *Neurosurgery*; NSG Focus = *Neurosurgical Focus*; WNS = *World Neurosurgery/Surgical Neurology.* Figure is available in color online only.

Neurosurgery/Surgical Neurology (Fig. 1B). Research topics related to spine and neurooncology took precedence over the years, with a substantial climb in 2015; by 2020, spine topics represented 7.2% of all research output, and neurooncology topics represented 4.6%. Vascular neurosurgery also increased throughout the years, reaching 2.5% in 2020. Meanwhile, topics pertaining to pediatrics remained stable at roughly 0.3% (Fig. 2).

We observed that administrative (e.g., National [Nationwide] Inpatient Sample [NIS]) and nonsurgical (National Cancer Database [NCDB] and Surveillance, Epidemiology, and End Results [SEER]) databases slowly lost ground to surgical (American College of Surgeons National Surgical Quality Improvement Program [ACS NSQIP]) and neurosurgical (QOD) registries. The rate of ACS NSQIP–related publications was initially steady and eventually climbed to 0.53% in 2016 and a peak at 0.74% in 2019 (coefficient = 1.32, 95% CI 0.82–1.81). A total of 51 articles that had used the QOD were identified. Evident increase in QOD-related publications was seen across the literature, with an increase from 0.043% in 2015 to 0.33% in 2020 (coefficient = 2.19, 95% CI 1.05–3.32).

Real-World and Big Data Applications

Undeniably, randomized controlled trials (RCTs) provide the highest level of evidence in the medical literature; however, they bear some disadvantages.[9] The application of clinical trials may often be limited by sample size, typically due to low-caseload diseases, cost containment, and feasibility.[10] Moreover, the inherent and necessary inflexible nature of RCTs limits their external validity; narrow criteria complicate extrapolating RCT findings to the real world of clinical practice (i.e., establishing *external* validity).[11] For these reasons, complementing real-world evidence with RCTs is thought to provide the most powerful evidence-based research method.[12] One way to do this is via observational studies; community equipoise (variation in practice), clinical equipoise (clinician perception), and patient equipoise (patient preference) are concepts that may have varying implications in clinical design and conduct but are readily available in real-world data.[6] The trends in mass digitization of the healthcare system have revolutionized the accessibility of real-world and big data and advanced the applicability of these data in clinical practice.[5,13] This has allowed the application of advanced
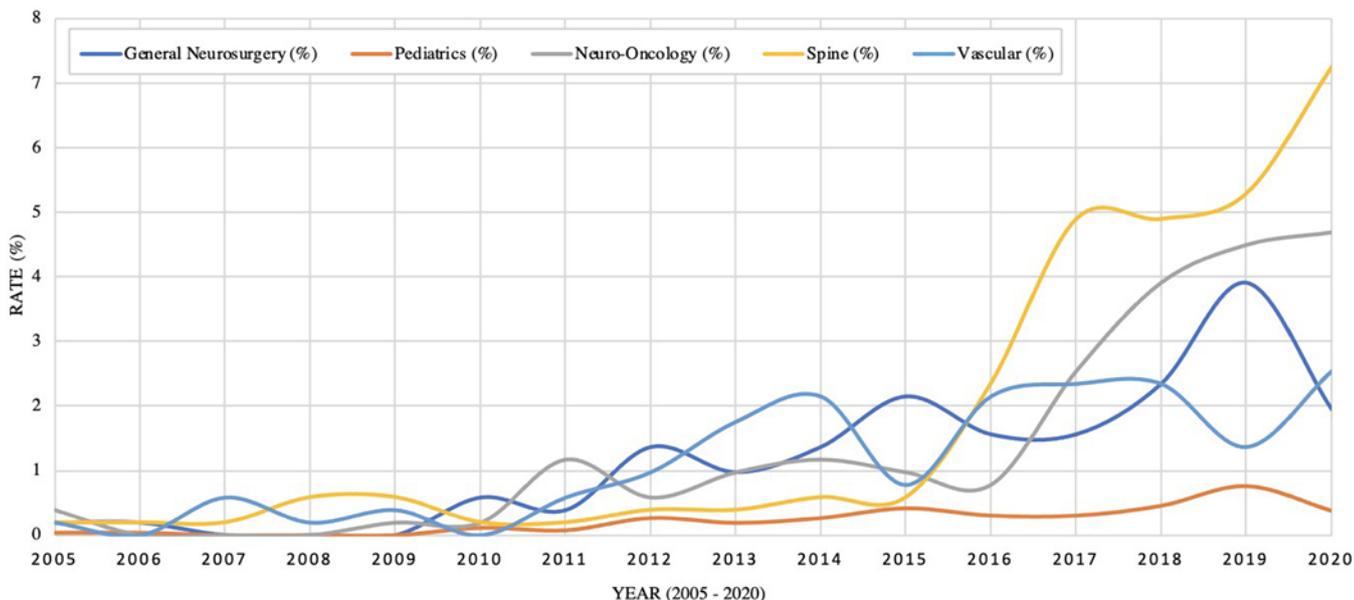
FIG. 2. Graphs showing the trend of research topics for spine, neuro-oncology, pediatrics, general neurosurgery, and vascular topics throughout the years 2005–2020. Figure is available in color online only.

analytics and thus the assessment of interactions between covariates and highly interactive outcomes. This can be useful in detecting higher-order correlations and patterns that may otherwise be missed in parametric statistics.[14] These analyses thus aid in detecting higher-order correlations and patterns that may otherwise be missed in parametric statistics, and eventual translation to personalized and precision medicine in neurosurgery, medical device design, decision-making, and preventive care.[5]

Socio-Political Background

Registries have been present in the United States since the 1960s. However, their robust employment was initially primarily influenced by political decisions. The Tax Relief and Health Care Act of 2006 (TRHCA) established the Physician Quality Reporting System (PQRS), which enclosed a 1.5% incentive payment for physicians who reported data on quality measures.[15] In 2010, the Affordable Care Act changed this positive reinforcement to negative by issuing financial penalties to the physicians who failed to report quality outcomes data.[16] This new paradigm for public and societal accountability led to a surge in the establishment and the subsequent use of databases and registries.[17] This is reflected in our results, with a doubling in the overall rate between 2010 and 2012.

Registries and Databases

The terms "registry" and "database" are often used interchangeably in the medical literature. Administrative healthcare databases are typically created from records of service utilization and payments for hospital billing purposes.[17,18] However, we suggest the term "registry" be reserved for prospectively collected information on data elements with specified clinical and research intentions. The information found within a clinical registry is used to evaluate patient outcomes for a particular population

with a specified condition, disease, or exposure.[19] Opportunities for registries have been well recognized across different disciplines. For example, registries that offer quality metrics have been frequently employed in cardiovascular disease, allowing for performance measurement and quality care delivery.[20] Nonetheless, understanding the strengths and limitations of the available data sets is important to guide meaningful clinical and research applications. Claus et al. described the paradigm as a "data mining expedition."[4] However, the use of large data sets and the generation of multiple hypotheses would often achieve statistical significance; conclusions should always be tailored to a clinical and scientific knowledge platform.

Administrative Databases

As illustrated by our results, early initiatives centered on the use of administrative databases. NIS is one example of an administrative database developed for the Healthcare Cost and Utilization Project (HCUP).[17] Although NIS does provide substantial information, limitations have been previously highlighted, specifically, coding errors and diagnosis.[21,22] Furthermore, the lack of operative details is another limitation.[21] The Nationwide Readmissions Database (NRD) is another example of an administrative database. This database allows for the evaluation of readmission analyses across hospital types for all payers and the uninsured. Underestimations of crucial factors that affected patient outcomes were also seen in the NRD.[23] While these findings were limited to readmission claims, several similar findings have highlighted the limitations in using administrative databases for critical research.[21]

Discipline-Specific Databases

NCDB, sponsored by the American College of Surgeons and the American Cancer Society, and SEER, established by the National Cancer Institute (NCI), are

oncological databases. NCDB captures cancer variables such as stage, interventions, histological information, while SEER captures variables of demographics and epidemiology.[24,25] The National Trauma Data Bank (NTDB) captures a compilation of all information about traumatic injuries and outcomes in the United States.[24,26] The ACS NSQIP database was initially built in the early 1990s to monitor the quality of surgical care in Veterans Affairs hospitals.[27] This registry contains variables on baseline patient characteristics, operative details, and 30-day postoperative complications.[27] One particular limitation of these data sets is their cross-sectionality. For example, the focus of ACS NSQIP is considered perioperative and hence only captures 30-day outcomes, which can hinder procedural analyses.[27] While these data sets and several others available via claims data, financial data, and private databases (e.g., MarketScan Research Database and PearlDriver) were not specific for the field of neurosurgery, the need to oblige to the momentum of "big data" research took priority. However, the information within these databases might not always have been sufficient to analyze neurosurgical trends and outcomes of operations.

Neurosurgical Databases: QOD

The allure of large, prospectively collected real-world evidence and big data applications led to the development of registries that have more meticulous and targeted methodologies. The QOD is a multiinstitutional registry focusing solely on patients undergoing neurosurgical procedures. It encompasses prospectively collected operative and patient-reported outcomes (PROs).[28,29] The QOD is the first database explicitly designed to reflect a unified national quality reporting mechanism in neurosurgery; this paved the way for comparative effectiveness analysis and evidence-based practice amendments.[29,30] One value of the QOD compared with other databases is its specificity to subsections of neurosurgery; examples of registries currently available are the Spine Registry QOD, including Lumbar Spine, Spinal Deformity, and Cervical Spine Registries, the Neurovascular Quality Initiative QOD, the Tumor QOD, and the Stereotactic Surgery Registry.[31] Additionally, these registries contain a multitude of prospectively collected outcomes, including surgical complications, reoperations, and PROs, at multiple follow-up time points.[32] The availability of long-term PROs and the granularity of patient and surgical characteristics are distinguishing features of the QOD. While short-term data typically measure safety, long-term data offer better insight into the quality of care delivered and the sustainability of an intervention.[32] The QOD also paved the way for the routine collection of longitudinal data as part of the standard healthcare practice and facilitated the regulatory application of these resources; this led to a subsequent increase in data accumulation and research generation.[32] Additionally, establishing a predetermined purpose for these registries ensured the availability of the abundant information needed to address a myriad of clinical questions. Another practical importance is its reliance on clinical data entered only by practice groups with clinical expertise in neurosurgery. Moreover enrollment in QOD is systematically standardized across all sites.[29]

Pitfalls in Database and Registry Use

The pitfalls in the utility of data sets are granularity, accuracy, and observational nature. The cost of maintaining databases or registries with a variety of information is high. Furthermore, a granular database requires a dedicated staff with educational training to ensure accurate and consistent information. Moreover, the inclusion of big data in a meta-analysis can be difficult. First, we would expect the effect size of a database study to cause asymmetrical distribution. Second, data included within an analysis must be mutually exclusive; institutional studies using patient cohorts included within a national registry must be identified and excluded. Additionally, large registries emphasizing prospectively collected data are at an increased risk of significant sampling bias with loss to follow-up.[4] The phenomenon referred to as "voltage drop" depicts the concept that effects of interventions are expected to decrease as populations become more heterogeneous and complex.[33] Finally, clinical practice is multiparametric, and decisions are typically influenced by several variables, including clinician and patient preferences. The observational nature of registries necessitates the presence of confounders and concealed variables that cannot otherwise be adjusted.

Future Directions

Our findings indicate that specialty-specific registries, during their short history, have impacted neurosurgical research. One interesting direction for future efforts would include annotated or "raw" imaging in these registries. Attempts so far have not been systematic enough, and computer vision is underutilized in clinical neurosurgical research. Another promising endeavor would be to pursue cross-talking between registries; patients enrolled into more than one registry would be identified, and data from both would be available to the researchers. This would enhance long-term, holistic follow-up and would prevent the duplication of efforts. Registry-based science methods, along with the information within these data sets, may actually constitute a new "academic currency," a novel medium to promote collaboration, exchange, and synergies across institutions. The proliferation of clinical registries and the concurrent increase in the associated academic output, as evidenced by the significant increase in registry-related articles, suggest that quality science has emerged as a new and productive avenue for academic activity. Unlike traditional academic pursuits, it uniquely integrates with clinical workflows, advances collective knowledge, and has immediate implications for assessing therapeutic and economic efficacies of care. Therefore, resource allocation toward quality science represents the potential for excellent returns on healthcare dollar investment, particularly for academic healthcare systems. The trend will likely further encourage the development of legions of healthcare workers and hospitals rooted in value-based care delivery, systemic quality improvement, scientific research production, and data science.

Conclusions

Our findings summarize the trend in database and registry use in neurosurgical research over the past 15 years.

Data use has shifted from administrative databases to surgical and neurosurgical registries. As the application of big data in neurosurgery is expanding, the possibility of performing advanced analytics, establishing clinical guidelines, and evaluating surgeon-generated performance standards is also evolving. We have likely entered into a new era in which research related to advanced uses of registry data will continue to expand. The product of registry-based science may be emerging as the new "academic currency," through which our clinical communities can simultaneously facilitate collaboration, advance science, enhance safety measures, and improve the economic and clinical outcomes of care.

References

1. Hoque DME, Kumari V, Hoque M, Ruseckaite R, Romero L, Evans SM. Impact of clinical registries on quality of patient care and clinical outcomes: a systematic review. *PLoS One*. 2017;12(9):e0183667.
2. Williams WG. Uses and limitations of registry and academic databases. *Semin Thorac Cardiovasc Surg Pediatr Card Surg Annu*. 2010;13(1):66-70.
3. The impact of big data in neurosurgery. *AANS Neurosurgeon*. April 10, 2019. Accessed October 14, 2021. https://aansneurosurgeon.org/aansstudent/the-impact-of-big-data-in-neurosurgery/
4. Claus CF, Lytle E, Carr DA, Tong D. Big data registries in spine surgery research: the lurking dangers. *BMJ Evid Based Med*. 2021;26(3):103-105.
5. Raju B, Jumah F, Ashraf O, Narayan V, Gupta G, Sun H, et al. Big data, machine learning, and artificial intelligence: a field guide for neurosurgeons. *J Neurosurg*. 2021;135(2):373-383.
6. Walicke P, Abosch A, Asher A, Barker FG II, Ghogawala Z, Harbaugh R, et al. Launching effectiveness research to guide practice in neurosurgery: a National Institute Neurological Disorders and Stroke workshop report. *Neurosurgery*. 2017; 80(4):505-514.
7. Oravec CS, Motiwala M, Reed K, Kondziolka D, Barker FG II, Michael LM II, Klimo P Jr. Big data research in neurosurgery: a critical look at this popular new study design. *Neurosurgery*. 2018;82(5):728-746.
8. Karhade AV, Larsen AMG, Cote DJ, Dubois HM, Smith TR. National databases for neurosurgical outcomes research: options, strengths, and limitations. *Neurosurgery*. 2018;83(3): 333-344.
9. Davis K, Minckas N, Bond V, Clark CJ, Colbourn T, Drabble SJ, et al. Beyond interviews and focus groups: a framework for integrating innovative qualitative methods into randomised controlled trials of complex public health interventions. *Trials*. 2019;20(1):329.
10. Jamjoom AAB, Gane AB, Demetriades AK. Randomized controlled trials in neurosurgery: an observational analysis of trial discontinuation and publication outcome. *J Neurosurg*. 2017;127(4):857-866.
11. Spiotta AM. Twinkle, twinkle little STAR, how I wonder what you are: the case for high-quality, large-scale, "real-world" databases. Letter. *Neurosurgery*. 2020;87(2):E271-E272.
12. Kim HS, Lee S, Kim JH. Real-world evidence versus randomized controlled trial: clinical research based on electronic medical records. *J Korean Med Sci*. 2018;33(34):e213.
13. Bydon M, Schirmer CM, Oermann EK, Kitagawa RS, Pourratian N, Davies J, et al. Big data defined: a practical review for neurosurgeons. *World Neurosurg*. 2020;133:e842-e849.
14. Ryo M, Rillig MC. Statistically reinforced machine learning for nonlinear patterns and variable interactions. *Ecosphere*. 2017;8(11):e01976.
15. Anumula N, Sanelli PC. Physician Quality Reporting System. *AJNR Am J Neuroradiol*. 2011;32(11):2000-2001.
16. Gillam S, Siriwardena N. *The Quality and Outcomes Framework: QOF—Transforming General Practice*. CRC Press; 2018.
17. Harbaugh CM, Cooper JN. Administrative databases. *Semin Pediatr Surg*. 2018;27(6):353-360.
18. Groth SS, Habermann EB, Massarweh NN. United States administrative databases and cancer registries for thoracic surgery health services research. *Ann Thorac Surg*. 2020; 109(3):636-644.
19. Agency for Healthcare Research and Quality. *Registries for Evaluating Patient Outcomes: A User's Guide*. 4th ed. AHRQ; 2020. Accessed October 14, 2021. https://effectivehealthcare.ahrq.gov/sites/default/files/pdf/registries-evaluating-patient-outcomes-4th-edition.pdf
20. Bhatt DL, Drozda JP Jr, Shahian DM, Chan PS, Fonarow GC, Heidenreich PA, et al. ACC/AHA/STS statement on the future of registries and the performance measurement enterprise: a report of the American College of Cardiology/American Heart Association Task Force on Performance Measures and the Society of Thoracic Surgeons. *J Am Coll Cardiol*. 2015;66(20):2230-2245.
21. Gologorsky Y, Knightly JJ, Chi JH, Groff MW. The Nationwide Inpatient Sample database does not accurately reflect surgical indications for fusion. *J Neurosurg Spine*. 2014; 21(6):984-993.
22. Kurian SJ, Yolcu YU, Zreik J, Alvi MA, Freedman BA, Bydon M. Institutional databases may underestimate the risk factors for 30-day unplanned readmissions compared to national databases. *J Neurosurg Spine*. 2020;33(6):845-853.
23. Murphy M, Alavi K, Maykel J. Working with existing databases. *Clin Colon Rectal Surg*. 2013;26(1):5-11.
24. Bilimoria KY, Stewart AK, Winchester DP, Ko CY. The National Cancer Data Base: a powerful initiative to improve cancer care in the United States. *Ann Surg Oncol*. 2008; 15(3):683-690.
25. National Cancer Institute NC. SEER Cancer Statistics Review 1975-2013. Accessed October 14, 2021. https://seer.cancer.gov/csr/1975_2013/browse_csr.php
26. Tsiklidis EJ, Sims C, Sinno T, Diamond SL. Using the National Trauma Data Bank (NTDB) and machine learning to predict trauma patient mortality at admission. *PLoS One*. 2020;15(11):e0242166.
27. Khuri SF, Daley J, Henderson WG. The comparative assessment and improvement of quality of surgical care in the Department of Veterans Affairs. *Arch Surg*. 2002;137(1): 20-27.
28. Asher AL, Speroff T, Dittus RS, Parker SL, Davies JM, Selden N, et al. The National Neurosurgery Quality and Outcomes Database (N2QOD): a collaborative North American outcomes registry to advance value-based spine care. *Spine (Phila Pa 1976)*. 2014;39(22)(suppl 1):S106-S116.
29. 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.
30. NeuroPoint Alliance. Participate in the NeuroPoint's QOD Spine Surgery Registry. Accessed October 14, 2021. https://www.neuropoint.org/registries/qod-spine/#1519913083389-a5b639c4-145e
31. NeuroPoint Alliance. Current QOD Registries. Accessed October 14, 2021. https://www.neuropoint.org/registries/qod/
32. Asher AL, McGirt MJ, Glassman SD, Groman R, Resnick DK, Mehrlich M, et al. Regulatory considerations for prospective patient care registries: lessons learned from the National Neurosurgery Quality and Outcomes Database. *Neurosurg Focus*. 2013;34(1):E5.
33. Kilbourne AM, Neumann MS, Pincus HA, Bauer MS, Stall

R. Implementing evidence-based interventions in health care: application of the replicating effective programs framework. *Implement Sci.* 2007;2:42.

Disclosures

Dr. Foley: consultant for Medtronic; direct stock ownership in Accelus, Companion Spine, Discgenics, DuraStat, Medtronic, NuVasive, Practical Navigation, RevBio, Spine Wave, Tissue Differentiation Intelligence, Triad Life Sciences, and True Digital Surgery; patent holder with Medtronic and NuVasive; royalties from Medtronic; and board of directors of Discgenics, DuraStat, RevBio, Tissue Differentiation Intelligence, Triad Life Sciences, and True Digital Surgery. Dr. Shaffrey: consultant for NuVasive, Medtronic, and SI Bone; direct stock ownership in NuVasive; patent holder with Medtronic, NuVasive, SI Bone, and Zimmer Biomet; and royalties from Medtronic and NuVasive. Dr. Coric: consultant for Globus Medical, Medtronic, and Spine Wave; royalties from Integrity Implants, Spine Wave, Stryker, Medtronic, RTI Surgical, and Globus Medical. Dr. Bisson: consultant for MiRus, Stryker, and Medtronic; and direct stock ownership in nView and MiRus. Dr. Glassman: employee of Norton Healthcare; consultant for Stryker and Medtronic; patent holder with Medtronic; clinical or research support for the study described from NuVasive; royalties from Medtronic; past president of the Scoliosis Research Society; and co-chair of the American Spine Registry. Dr. Mummaneni: consultant for DePuy Spine, Globus, and Stryker; royalties from DePuy Spine, Thieme Publishing, and Springer Publishing; direct stock ownership in Spinicity/ISD; support of non–study-related clinical or research effort from AO Spine and ISSG; and clinical or research support for the study described from NREF.

Correspondence

Mohamad Bydon: bydon.mohamad@mayo.edu.