New Statistical Editor—February 2018

Melina R. Kibbe, MD

JAMA Surgery sincerely thanks Terry Hyslop, PhD, for her service as Statistical Editor from 2015 through 2017. Terry was an asset to *JAMA Surgery*, and we will miss her greatly.

JAMA Surgery is pleased to announce the appointment of Jason S. Haukoos, MD, MSc, as Statistical Editor effective February 1, 2018. Dr Haukoos earned his medical degree from the UCLA (University of California, Los Angeles) School of Medicine. He completed his residency in emergency medicine and a clinical research fellowship at Harbor-UCLA Medical Center in Torrance, California. During his fellowship, he earned his Master of Science degree in epidemiology from the UCLA School of Public Health.

Dr Haukoos serves as professor of emergency medicine at the University of Colorado School of Medicine, Denver, with a secondary appointment as professor of epidemiology at the Colorado School of Public Health. He also serves as Director of Research and Director of the Clinical Scientist and Health Services Research Fellowships for the Department of Emergency Medicine at Denver Health, Denver, Colorado.

Dr Haukoos is a past recipient of an Individual National Research Service Award and an Independent Scientist Award from the Agency for Healthcare Research and Quality and has received investigator-initiated funding from the Centers for Disease Control and Prevention, the Agency for Healthcare Research and Quality, and the National Institutes of Health. He has published extensively, with over 150 peer-reviewed original articles to date, and currently serves as a Research Methodology Editor for *Annals of Emergency Medicine.*

Jason Haukoos, MD

A Checklist to Elevate the Science of Surgical Database Research

Adil H. Haider, MD, MPH; Karl Y. Bilimoria, MD, MS; Melina R. Kibbe, MD

Each year, JAMA Surgery receives hundreds of submissions that retrospectively analyze large surgical databases. Although many of these attempt to shed light on new and important questions, most do not get published. A majority of submissions are not even sent out for peer review because they have clear flaws in the data analytic techniques or they attempt to address a research

question that cannot be adequately answered with the proposed data set. Of those that are sent out for peer review, many are recommended to be rejected by expert peer reviewers as they find major methodological flaws in the use of these otherwise powerful data sets. Articles that are published frequently come from a select group of investigators who have developed a mastery of specific data sets and the analytic techniques required to truly harness their potential.

To help more and more investigators develop the skills needed to appropriately use the increasing number of large surgical data sets available, the editors of *JAMA Surgery* have commissioned this current series of statistical methodology articles. The series is aimed at providing a short, practical guide for academic surgeons and researchers in the use of the most widely available surgical data sets that can be used across the research continuum, from conceptualization to peer-reviewed publication. To achieve this, *JAMA Surgery* is pleased to partner with the Surgical Outcomes Club (http://www.surgicaloutcomesclub.com) to publish a series that will be instrumental in elevating the science used in surgical outcomes research.

This 13-part series provides a succinct overview of the 11 most widely used data sets[1-11] (**Box 1**), their specific features, strengths, limitations, and some important statistical considerations. In addition, we present a 10-item checklist (**Box 2**) in this Editorial that authors can use to ensure that

Box 1. Databases Covered in This Series

Agency for Healthcare Research and Quality Healthcare Cost and Utilization Project databases: National Inpatient Sample, State Inpatient Databases, and Kids' Inpatient Database[1]

Surveillance, Epidemiology, and End Results Program[2]

Medicare Claims Data[3]

Military Health System Tricare Encounter Data[4]

Veterans Affairs Surgical Quality Improvement Program[5]

National Surgical Quality Improvement Program[6]

Metabolic and Bariatric Surgery Accreditation and Quality Improvement Program[7]

National Cancer Database[8]

National Trauma Data Bank[9]

Society for Vascular Surgery Vascular Quality Initiative[10]

The Society of Thoracic Surgeons National Database[11]

Box 2. Checklist to Elevate the Science of Surgical Database Research

1. Have a solid research question and clear hypothesis. Consider using the FINER (Feasible, Interesting, Novel, Ethical, Relevant) or PICO (Patient, Population, or Problem; Intervention, Prognostic Factor, or Exposure; Comparison or Intervention; Outcome) criteria to develop these.
2. Ensure compliance with the institutional review board and data use agreements.
3. Conduct a thorough literature review. Use a reference management program for ease in manuscript development.
4. Make sure this is the best data set available and that it has the appropriate variables to answer your research question.
5. Clearly define the inclusion criteria, exclusion criteria, and outcome variables. Use a flow diagram to describe final patient selection.
6. Identify potential confounders and use risk adjustment to minimize bias. Consider using a directed acyclic graph to represent potential associations. Avoid use of causal language in reporting results of these observational studies.
7. Ensure that the data variables have not changed over time. If so, account for this.
8. Ensure that competing risks are identified and addressed.
9. Ensure that data issues, such as missing data, are discussed and that any sensitivity analyses or imputations performed are reported in a clear and cohesive way.
10. Ensure that your article has a clear take-home message that addresses how your research advances current knowledge and has important policy or clinical implications.

they have covered what is "at minimum" expected from a manuscript that uses 1 of these databases. Finally, we support this series with an Editorial[12] by our biostatistician colleagues, who provide more in-depth information on statistical methodologies mentioned in the practical guides as well as potential pitfalls that need to be avoided. To ensure that these guides are truly practical and relevant, we have leveraged our partnership as the official journal of the Surgical Outcomes Club to develop a 3-person authorship team that includes (1) a surgeon investigator who is a senior member of the Surgical Outcomes Club with extensive experience using that particular data set; (2) a member of the *JAMA Surgery* Editorial Board who commonly reviews such manuscripts; and (3) a *JAMA Surgery* biostatistician who is routinely consulted to knowledgeably evaluate the methods for these types of papers (in some cases, the *JAMA Surgery* board member is also an expert methodologist, obviating the need for a biostatistician). This authorship strategy has ensured that each guide is presented in terms that are relevant to surgeons, even if they do not have previous experience with the biostatistics or the data set involved and includes basic information required to prepare a manuscript for the rigorous *JAMA Surgery* peer review process.

To help authors improve the quality of their submissions, we have developed a 10-item checklist (Box 2). The first item in our checklist encourages authors to pursue hypothesis-driven science. Defining a solid research question is key to translating a problem into an operational hypothesis. The FINER (Feasible, Interesting, Novel, Ethical, Relevant) criteria or the PICO (Patient, Population, or Problem; Intervention, Prognostic Factor, or Exposure; Comparison or Intervention; Outcome) format can help develop a meaningful research question.[13,14] Adequately defining the population of interest lays a solid groundwork for the interpretation, applicability, and generalizability of the research findings. We understand that in many cases, authors may be using these large databases for "hypothesis-generating" research. That is of course acceptable, but one must start with a solid research question to conduct a

meaningful research project that will generate important hypotheses from the large data sets that can then be further studied with translational or prospective approaches. Some authors ask if it is acceptable to try and see what they can find in a data set that they may have access to without a real research question. This is never acceptable.

Second, we remind authors to seek approval or an exemption from an institutional review board and to properly document and comply with applicable data use agreements. These are often overlooked, but compliance with applicable rules are necessary for patient privacy and a variety of important reasons. Third, a thorough literature review will assist in making sure the best database is selected to answer research questions and to make sure the research question has not been previously answered. Fourth, we encourage authors to invest enough time early on to get to know the database, confirm that it has the appropriate variables, and understand methodological considerations to make sure this is the best data set available for the study. Fifth, a clear definition of the inclusion and exclusion criteria, as well as outcome variables, is necessary for reviewers and readers to understand the population under study. This also helps facilitate data query and extraction of a complete and useful data set.

Another important aspect of working with databases is the need to identify potential confounders or covariates and use risk adjustment to minimize bias. Given the observational nature of data in these surgical registries, 1 approach to do this is to create a directed acyclic graph,[15] which will allow a visual depiction of the potential association being

explored along with the covariates and confounders that need to be kept in mind or accounted for while studying the association. Please refer to the Editorial by Kaji et al[12] for further details. Authors should also avoid use of causal language when describing the results of these observational studies. Seventh, authors must account for any updates or significant changes to the variables of interest over time as this might jeopardize comparison between and across years (for example, in the National Cancer Database, the definition of sentinel lymph node biopsy for breast and melanoma has changed during the last 10 years, and this must be accounted for). Eighth, authors are encouraged to identify if competing risks exist in outcomes.[16] For example, if authors are studying complication rates 30 days after surgery, one must account for patients who may have already died and are not at risk for developing these complications. Ninth, authors must ensure that any data issues, such as missing data, are openly discussed in a clear, cohesive, and replicable way. Authors must lay out any data limitations, how they were addressed, and measures taken to reduce their impact (eg, sensitivity analyses, multiple imputation[17] for missing data). Finally, as our last item in the checklist, we encourage authors to clearly state a take-home message. It is best to communicate how the study advances the science, addresses gaps in knowledge, highlights further research opportunities, and discusses important policy or clinical implications of the work.

We recommend that authors use this checklist, the practical guide for their chosen data set, and the statistical tips for analyzing data sets as a 3-part series to consult before submission of their manuscript. We hope that by following these simple guides, authors can benefit from the collective wisdom of so many colleagues who have successfully completed similar analyses in the past. We look forward to the opportunity to publish analytically advanced studies and hope that these guides will help elevate the science of surgical database research.

REFERENCES

1. Stulberg JJ, Haut ER. AHRQ Healthcare Cost and Utilization Project Databases: National Inpatient Sample (NIS) [published online April 4, 2018]. *JAMA Surg*. doi:10.1001/jamasurg.2018.0542

2. Doll KM, Rademaker A, Sosa JA. Longitudinal outcomes reporting using the Surveillance, Epidemiology, and End Results (SEER) Database [published online April 4, 2018]. *JAMA Surg*. doi:10.1001/jamasurg.2018.0501

3. Ghaferi AA, Dimick JB. Longitudinal outcomes reporting using Medicare claims [published online April 4, 2018]. *JAMA Surg*. doi:10.1001/jamasurg.2018.0489

4. Schoenfeld AJ, Kaji AH, Haider AH. Outcomes reporting using Tricare claims [published online April 4, 2018]. *JAMA Surg*. doi:10.1001/jamasurg.2018.0480

5. Massarweh NM, Kaji AH, Itani KMF. Veterans Affairs Surgical Quality Improvement Program [published online April 4, 2018]. *JAMA Surg*. doi:10.1001/jamasurg.2018.0504

6. Raval MV, Pawlik TM. National Surgical Quality Improvement Program (NSQIP) and pediatric NSQIP [published online April 4, 2018]. *JAMA Surg*. doi:10.1001/jamasurg.2018.0486

7. Telem DA, Dimick JB. Metabolic and Bariatric Surgery Accreditation and Quality Program (MBSAQIP) [published online April 4, 2018]. *JAMA Surg*. doi:10.1001/jamasurg.2018.0495

8. Merkow RP, Rademaker AW, Bilimoria KY. National Cancer Database [published online April 4, 2018]. *JAMA Surg*. doi:10.1001/jamasurg.2018.0492

9. Hashmi ZG, Kaji AH, Nathens AB. National Trauma Data Bank [published online April 4, 2018]. *JAMA Surg*. doi:10.1001/jamasurg.2018.0483

10. Desai SS, Kaji AH, Upchurch G. Society for Vascular Surgery Vascular Quality Improvement Program [published online April 4, 2018]. *JAMA Surg*. doi:10.1001/jamasurg.2018.0498

11. Farjah F, Kaji AH, Chu D. Society of Thoracic Surgery (STS) Dataset [published online April 4, 2018]. *JAMA Surg*. doi:10.1001/jamasurg.2018.0545

12. Kaji AH, Rademaker AW, Hyslop T. Tips for analyzing large data sets from the *JAMA Surgery* statistical editors [published online April 4, 2018]. *JAMA Surg*. doi:10.1001/jamasurg.2018.0647

13. Cummings SR, Browners WS, Hulley SB. Conceiving the research question and developing the study plan. In: Hulley SB, Cummings SR, Browner WS, Grady DG, Newman TB, eds. *Designing Clinical Research*. 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2007:19-22.

14. Brian Haynes R. Forming research questions. *J Clin Epidemiol*. 2006;59(9):881-886.

15. Shrier I, Platt RW. Reducing bias through directed acyclic graphs. *BMC Med Res Methodol*. 2008;8:70.

16. Sun M, Choueiri TK, Hamnvik OP, et al. Comparison of gonadotropin-releasing hormone agonists and orchiectomy: effects of androgen-deprivation therapy. *JAMA Oncol*. 2016;2(4):500-507.

17. Oyetunji TA, Crompton JG, Ehanire ID, et al. Multiple imputation in trauma disparity research. *J Surg Res*. 2011;165(1):e37-e41.