With Great Power Comes Great Responsibility
Big Data Research From the National Inpatient Sample

Rohan Khera, MD; Harlan M. Krumholz, MD, SM

The use of large administrative databases is transforming clinical cardiovascular research. These sources of big data allow the study of practices and outcomes across a spectrum of health systems, providing real-world evidence. However, these databases have peculiarities to their design that require specialized expertise and distinct analytic practices for their appropriate interpretation. We discuss these issues in the context of the National Inpatient Sample (NIS) that is one such data set used in healthcare research. Compiled by the Agency for Healthcare Research and Quality annually since 1988, it comprises a large number of inpatient discharges from US community hospitals regardless of the payer (≈8 million/y), with each observation representing a unique hospitalization.[1] It has some features to its design and the content of its data that are essential to consider in the pursuit of studies with it.

The NIS includes information on patient demographics, administrative codes for primary diagnosis and secondary diagnoses, procedures, survival to discharge, disposition, hospital charges, and length of stay.[1] The NIS can be used to examine the use of hospital health services, practice variation, cost, and the impact of health policy interventions in the inpatient setting.[1] The data are easily accessible, inexpensive, and can be analyzed using ubiquitous statistical programs. Consequently, research publications from the NIS data have grown rapidly in recent years (Figure 1). Nevertheless, researchers, as well as scientific journals and their readers, may not yet be familiar with the nuances of this complex data set and therefore be challenged to determine if the data are interpreted correctly.

Although not an exhaustive list, we discuss 4 instances highlighting issues related to this widely used database that should be considered when using it as a scientist, evaluating it as a reviewer, or understanding it as a consumer of scientific studies. We think that these issues are pervasive in the literature and have identified several studies with similar problems. We have used a few representative examples to illustrate these issues but do not think it is appropriate to call out particular authors or papers. Nevertheless, we shared the specific studies discussed here with the Editors to have our conclusions verified.

1. Dynamic sample design: The NIS is constructed using a complex sampling design, and obtaining national estimates requires accounting for clustering at hospitals and stratification of sampled data and for changes in sampling over time.[2,3] During 1988 to 2011, the NIS was constructed annually by including 100% of the discharges from 20% of US hospitals and was redesigned in 2012 as a 20% national patient-level sample, with nonrepresentative sampling across hospitals.[2] Accounting for these changes is essential for an accurate study design. For example, a study using NIS 2003 to 2012 compared calendar year trends in rates of an invasive cardiovascular procedure between hospitals with, and without a second, more complex, operative procedure. Although appropriate within the 2003 to 2011 data, this was a problem with the 2012 data.[2] Because the NIS only captures a nonrepresentative fraction of hospital discharges after 2011, volumes of either procedure cannot be determined for this period.
2. Inpatient hospitalization record: The NIS does not identify individual patients, and recurrent hospitalizations appear as distinct observations.[3] Furthermore, it does not capture outpatient encounters or observation-only stays, and conditions and procedures occurring across multiple healthcare settings may be under-represented.[1,3] This may be an important consideration in interpreting a study performed in NIS 2001 to 2011 that reported a low use of a routine diagnostic imaging modality that is performed in both inpatient and outpatient settings and found that compared with hospitalizations where this study was performed, those without this procedure had higher mortality rates. The latter analysis incorrectly assumes that NIS captures all healthcare records of individual patients and does not account for other settings where the diagnostic test may have been performed during the same illness episode—either in an outpatient encounter directly preceding the hospitalization or in a recent prior hospitalization. Furthermore, the analysis may also be confounded by illness severity, and patients undergoing multiple procedures may not have the procedure code for this simple diagnostic test included in the record because of either limited additional reimbursement value or limited space on a claim record.
3. Volume assessments: Similar to the limited ability to perform hospital-level volume assessment since 2012, the data structure does not allow volume estimates for

From the Division of Cardiology, University of Texas Southwestern Medical Center, Dallas, TX (R.K.); Section of Cardiovascular Medicine, Department of Internal Medicine, School of Medicine, Department of Health Policy and Management, School of Public Health, and Robert Wood Johnson Clinical Scholars Program, Department of Internal Medicine, School of Medicine, Yale University, New Haven, CT; and Center for Outcomes Research and Evaluation, Yale New Haven Health, CT (H.M.K.).

Correspondence to Harlan M. Krumholz, MD, SM, Department of Internal Medicine, Yale School of Medicine, 1 Church St, Suite 200, New Haven, CT 06510. E-mail harlan.krumholz@yale.edu

(*Circ Cardiovasc Qual Outcomes.* 2017;10:e003846. DOI: 10.1161/CIRCOUTCOMES.117.003846.)

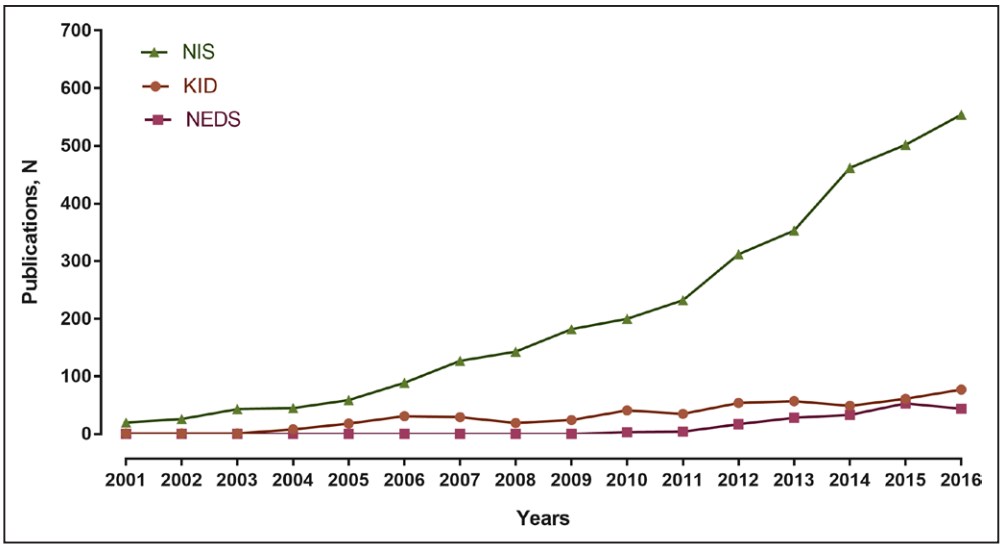

Figure 1. Calendar year trends in publications from the National Inpatient Sample (NIS). Number of peer-reviewed publications from the NIS have increased rapidly in recent years. Data from other Healthcare Cost and Utilization Project (HCUP) data sets are presented for comparison—Kids' Inpatient Database (KID) and Nationwide Emergency Department Sample (NEDS). Data taken from: HCUP Publications. Agency for Healthcare Research and Quality, Rockville, MD. www.hcup-us.ahrq.gov/reports/pubsearch/pubsearch.jsp.

certain subgroups. First, US states are not a part of the sampling framework of the NIS, and therefore, sampled discharges from a given state are not representative of all discharges from that state.[4] States contribute hospitalizations based on how representative their hospitals and patient population are to the national landscape. Hence, unless a state's hospital characteristics (ownership, urban/rural location, teaching status, and bed size) and patient features (diagnosis-related groups), which are components of NIS sampling methodology, are nationally representative, state-level samples are not representative of the state's discharges. Hence, in a study that assesses state-level rates of a specific procedure performed for an acute cardiovascular condition before and after changes in public-reporting regulations in that state, as compared with other states in the NIS, may be biased by the sampling in the respective states. State-to-state comparisons assume representative samples and are better conducted using databases that have this property. Second, analysis of provider-level volumes is particularly challenging. A study evaluating volume–outcomes associations for procedures performed by individual providers is also not appropriate because the provider code-field in NIS does not link to a specific procedure and is not reported uniformly across hospitals and states, referring to individual physicians at some hospitals, and physician groups at others.[5]

4. Administrative codes: A final consideration is the identification of disease conditions or procedures based on their descriptive connotations without formal validation. The claim codes that do not affect reimbursement directly may be prone to variation in coding practices. As an example, a study conducted using NIS 1993 to 2007 found that rates of pulmonary artery hypertension hospitalizations declined abruptly during the study period.[6] The authors, however, appropriately investigated this trend in other data sets and inferred that this did not represent a true demographic trend but was likely because of a recommendation to limit the use of the pulmonary artery hypertension –specific claim code as a default for all pulmonary hypertension–related hospitalizations during this period. Similarly, using codes to identify specific diagnostic subgroups, such as the ST-segment–elevation myocardial infarction among all acute myocardial infarction, heart failure with preserved ejection fraction among all heart failure, and in-hospital cardiac arrest, without a subgroup-specific reimbursement value, may also be inaccurate, with noise or bias introduced. In addition to the primary diagnosis code, secondary diagnoses should be interpreted with caution, particularly for identifying events that may have occurred during a hospitalization. Because the NIS does not have present-on-admission flags accompanying its secondary diagnosis codes, or allow longitudinal assessment of patients, most secondary codes may not be sufficiently reliable in distinguishing complications from comorbid conditions. A rigorous literature review for prior validation studies before conducting such an investigation is warranted.

Given its complexity and ever-evolving data structure, the Agency for Healthcare Research and Quality recommends a careful review of NIS' publicly available documentation.[7] This includes details on year-specific data structure,[7] statistical best-practices,[3] and analytic tools.[8] In addition, it offers HCUPnet,[9] a publicly accessible, web-based portal that provides national estimates for individual administrative diagnosis/procedure codes, which can help with appropriately vetting proposed methodological strategies. Furthermore, it may be prudent for investigators to clarify additional questions directly with the Agency for Healthcare Research and Quality rather than solely relying on the methodology of published studies in the literature.

Finally, we think that a simple checklist, like the one we propose in Figure 2, may help prevent common errors early in the study-design phase and improve the validity and generalizability of studies using the NIS. Furthermore, to communicate that best-practices are followed, there is specific information that should be specifically highlighted in manuscripts. (1)

Section A: Research Design

- ☐ Does the study consider that it can only detect disease conditions, procedures, and diagnostic tests in hospital settings?*
- ☐ Does the study acknowledge that it includes encounters, not individual patients?*
- ☐ Does the study avoid diagnosis/procedure-specific volume assessments for units that are not a part of the sampling frame of the NIS, and are therefore not representatively sampled, including
  - a) geographic units, like U.S. states
  - b) healthcare facilities (after 2011)
  - c) individual healthcare providers?

Section B: Data Interpretation

- ☐ Does the study attempt to identify disease conditions or procedures of interest using administrative codes or their combinations that have been previously validated?*
- ☐ Does the study limit its assessment to only in-hospital outcomes, rather than those occurring after discharge?*
- ☐ Does the study distinguish complications from comorbidities or clearly note where it cannot?*

Section C: Data Analysis

- ☐ Does the study clearly account for the survey design of the NIS and its components -clustering, stratification, and weighting?*
- ☐ Does the study adequately address changes in data structure over time (for trend analyses)?*

*Fields marked with asterisk may specifically be included as a checklist in published studies

Figure 2. Proposed study design checklist for studies published using the National Inpatient Sample (NIS). The fields marked with an asterisk (*) may be included as a checklist in published studies.

Data source: (i) The years of NIS data included and (ii) if the NIS data structure changed during the study period, how these changes are germane to the study question and addressed. (2) Research design: It is clearly stated that (i) captured encounters represented hospitalization records and not distinct patients; (ii) validated administrative codes are used to identify diseases/procedures, or the lack of validation is acknowledged as a study limitation; (iii) outcome assessment is limited to the in-hospital setting, and post-discharge outcomes are not inferred; and (iv) secondary diagnosis codes are not used to infer complications because these may represent comorbid conditions, unless they are specific for in-hospital events or present-on-admission codes are used. (3) Data analysis: The study clearly (i) accounts for the survey design of the NIS and its components—clustering, stratification, and weighting; (ii) reports the software program, as well as the survey-specific commands used to generate national estimates; and (iii) states how trend analyses are modified to account for changes in data structure. (4) Data interpretation: The study clearly states that (i) the estimates for disease conditions and procedures from the NIS only represent their occurrence in an inpatient setting and do not account for outpatient occurrences, (ii) an assessment of possible confounding through appropriate statistical models and necessary sensitivity/subgroup analyses was performed, and (iii) the findings of the study were not sensitive to interpreting complications as comorbidities, or vice versa, given the challenges in differentiating the 2 in administrative data. In the future, studies will also need to be clear about how they handled the transition from *International Classification of Diseases, Ninth Revision* to *International Classification of Diseases, Tenth Revision*.

In summary, with the increasing access and use of NIS in clinical investigations, there is a potential for errors based on an inadequate understanding of the database design and how it has changed over time. The clinical research community and scientific journals have a responsibility of vetting

research ideas and ensuring appropriate interpretation of study results that ensure consistency with the design of this otherwise powerful data set. Because more, large, complex existing data sets are becoming available, the importance of understanding their particular features and their limitations will be increasingly important.

Sources of Funding

Dr Khera is supported by the National Heart, Lung, and Blood Institute (5T32HL125247-02) and the National Center for Advancing Translational Sciences (UL1TR001105) of the National Institutes of Health.

Disclosures

None.

References

1. HCUP Databases. *Healthcare Cost and Utilization Project—Overview of the National (Nationwide) Inpatient Sample (NIS).* Rockville, MD: Agency for Healthcare Research and Quality; 2016. www.hcup-us.ahrq.gov/nisoverview.jsp. Accessed December 15, 2016.
2. Houchens RL, Ross DN, Elixhauser A, Jiang J. Nationwide Inpatient Sample Redesign: Final Report. April 4, 2014. https://www.hcup-us.ahrq.gov/db/nation/nis/reports/NISRedesignFinalReport040914.pdf. Accessed July 20, 2014.
3. HCUP Methods Series. Healthcare Cost and Utilization Project (HCUP). Rockville, MD: Agency for Healthcare Research and Quality; 2016. www.hcup-us.ahrq.gov/reports/methods/methods.jsp. Accessed December 5, 2016.
4. Healthcare Cost and Utilization Project (HCUP). Why the NIS should not be used to make State-level estimates. Rockville, MD: Agency for Healthcare Research and Quality; 2016. www.hcup-us.ahrq.gov/db/nation/nis/nis_statelevelestimates.jsp. Accessed December 15, 2016.
5. HCUP NIS Description of Data Elements. Healthcare Cost and Utilization Project (HCUP). Rockville, MD: Agency for Healthcare Research and Quality; 2008. www.hcup-us.ahrq.gov/db/vars/mdnum2_r/nisnote.jsp. Accessed September 17, 2014.
6. Link J, Glazer C, Torres F, Chin K. International Classification of Diseases coding changes lead to profound declines in reported idiopathic pulmonary arterial hypertension mortality and hospitalizations: implications for database studies. *Chest*. 2011;139:497–504. doi: 10.1378/chest.10-0837.
7. NIS Database Documentation Archive. Healthcare Cost and Utilization Project (HCUP). Rockville, MD: Agency for Healthcare Research and Quality; 2016. www.hcup-us.ahrq.gov/db/nation/nis/nisarchive.jsp. Accessed March 15, 2017.
8. HCUP Frequently Asked Questions. Healthcare Cost and Utilization Project (HCUP). Rockville, MD: Agency for Healthcare Research and Quality; 2016. www.hcup-us.ahrq.gov/tech_assist/faq.jsp. Accessed December 15, 2016.
9. Healthcare Cost and Utilization Project. HCUPnet. 2017. https://hcupnet.ahrq.gov/#setup. Accessed January 1, 2017.

KEY WORDS: biostatistics ■ database ■ length of stay ■ patient discharge ■ public policy

Circulation
Cardiovascular Quality and Outcomes

With Great Power Comes Great Responsibility: Big Data Research From the National Inpatient Sample

Rohan Khera and Harlan M. Krumholz

Circ Cardiovasc Qual Outcomes. 2017;10:
doi: 10.1161/CIRCOUTCOMES.117.003846
Circulation: Cardiovascular Quality and Outcomes is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231

Print ISSN: 1941-7705. Online ISSN: 1941-7713