Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.JournalofSurgicalResearch.com

Association for Academic Surgery

Not Further Specified: Unclassified Orthopedic Injuries in Trauma Registries, Cause for Concern?

Bryant W. Oliphant, MD, MBA, MSc,[a,*] *Chelsea A. Harris, MD, MSc,*[b] *Anne H. Cain-Nielsen, MS,*[c] *James A. Goulet, MD,*[a] *and Mark R. Hemmila, MD*[c]

[a] *Department of Orthopaedic Surgery, University of Michigan, Ann Arbor, Michigan*
[b] *Department of Surgery, University of Maryland, Baltimore, Maryland*
[c] *Department of Surgery, University of Michigan, Ann Arbor, Michigan*

ARTICLE INFO

Article history:
Received 2 March 2019
Received in revised form
18 April 2019
Accepted 14 June 2019
Available online xxx

ABSTRACT

Background: Data accuracy is essential to obtaining correct results and making appropriate conclusions in outcomes research. Few have examined the quality of data that is used in studies involving orthopedic surgery. A nonspecific data entry has the potential to affect the results of a study or the ability to appropriately risk adjust for treatments and outcomes. This study evaluated the proportion of Not Further Specified (NFS) orthopedic injury codes found into two large trauma registries.

Materials: Data from the National Trauma Data Bank (NTDB) from 2011 to 2015 and from the Michigan Trauma Quality Improvement Program (MTQIP) 2011-2017 were used. We selected multiple orthopedic injuries classified via the Abbreviated Injury Scale, version 2005 (AIS2005) and calculated the percentage of NFS entries for each specific injury.

Results: There were a substantial proportion of fractures classified as NFS in each registry, 18.5% (range 2.4%-67.9%) in MTQIP and 27% (range 6.0%-68.5%) in the NTDB. There were significantly more NFS entries when the fractures were complex *versus* simple in both MTQIP (34.5% *versus* 9.6%, $P < 0.001$) and the NTDB (41.8% *versus* 15.7%, $P < 0.001$). The level of trauma center affected the proportion of NFS codes differently between the registries.

Conclusions: The proportion of nonspecific entries in these two large trauma registries is concerning. These data can affect the results and conclusions from research studies as well as impact our ability to truly risk adjust for treatments and outcomes. Further studies should explore the reasons for these findings.

This study was presented at the 14th annual Academic Surgical Congress, February 5-7, 2019, in Houston, TX.

* *Corresponding author*. Department of Orthopaedic Surgery, University of Michigan, North Campus Research Complex, 014-G014 2800 Plymouth Road, Ann Arbor, MI 48109-2800. Tel.: (734) 936-9839, (248) 761-6530; fax: +734 647 3277.

E-mail address: bryantol@med.umich.edu (B.W. Oliphant).

0022-4804/$ – see front matter © 2019 Elsevier Inc. All rights reserved.
https://doi.org/10.1016/j.jss.2019.06.065

Introduction

Data accuracy is paramount to obtaining correct results and making robust conclusions in outcomes research. Previous evaluations of trauma surgery registry data have questioned its accuracy[1] as well as its overall quality.[2] As outcomes research of large registries becomes more prevalent in orthopedic trauma surgery,[3-5] commensurate evaluation of data quality should follow. However, a major barrier to appropriate registry utilization is missing data. In particular, abstractors entering orthopedic trauma data have an option to classify the injury as "Not Further Specified (NFS)," meaning they could not identify an exact diagnosis. This coding option populates the registry with nonspecific data which could possibly affect the results or conclusions of these studies. While some have commented on NFS entries into these registries,[6] few have closely examined the use of this nonspecific diagnosis code in these registries.

Little is known about the prevalence of nonspecific orthopedic data found in common trauma registries. Furthermore, the institutional or injury factors associated with missingness have not been established. As a result, when researchers evaluate the epidemiology or treatment of a given condition using registry data, the degree to which they have captured the population of interest and therefore made accurate conclusions is unclear. Understanding the degree of missingness is also critical when comparing hospitals, as risk adjustment underpins any valid comparative analysis. For example, when examining the mortality of patients with femoral shaft fractures, Cantu *et al.* selected the exact diagnosis code (821.01—closed fracture of shaft of femur) and did not include the nonspecific diagnosis (821.0—fracture of shaft or unspecified part of femur closed) in their analysis.[3] In that study, the authors ignored the nonspecified or missing data and employed a complete case analysis method for their project. By not acknowledging these different diagnosis codes, we are unable to accurately risk adjust for outcomes and/or decide if the rendered treatment is appropriate.

Our goal was to better quantify the number of orthopedic NFS data in these trauma registries. We examined two large registries, the National Trauma Data Bank (NTDB) and the Michigan Trauma Quality Improvement Program (MTQIP) and focused on multiple orthopedic injuries. Our primary outcome was to calculate the prevalence of NFS entries in these databases. Our secondary outcomes were to examine other factors that influenced the percentage of NFS entries in these registries.

Material and methods

Data sources

MTQIP data from 2011 to 2017 and from the NTDB data from 2011 to 2015 were utilized. MTQIP is a collaborative quality initiative comprised of 29 American College of Surgeons - Committee on Trauma verified Level 1 and 2 trauma centers in the state of Michigan.[7] MTQIP utilizes a data definitions dictionary, based upon the National Trauma Data Standard, which is published online and updated annually.[8] Trauma data abstractors from participating hospitals undergo training in MTQIP and National Trauma Data Standard data definitions.[9] Data are transmitted from the trauma registry at participating hospitals to the coordinating center at 2-mo intervals. The inclusion criteria applied to form the MTQIP patient cohort are as follows:

- Age ≥16 y.
- At least one valid trauma International Classification of Diseases, ninth Revision, Clinical Modification (ICD-9-CM) code in the range of 800-959.9. Excluding the following isolated injuries: late effects (905-909.9), superficial injuries (910-924.9), and foreign bodies (930-930.9) or at least one valid trauma International Classification of Diseases, 10th Revision, Clinical Modification (ICD-9-CM) code S00-S99 with seventh character modifiers of A, B, or C only *(injuries to specific body parts—initial encounter)*, T07 *(unspecified multiple injuries)*, T14 *(injury of unspecified body region)*, *T79.A1-T79.A9 with seventh character modifier of A only (Traumatic Compartment Syndrome—initial encounter)*. Excluding the following isolated injuries: SXX (superficial injuries of a body region), T20-T28 with seventh character modifier of A only *(burns by specific body parts—initial encounter)*, T30-T32 *(burn by TBSA percentages)*. Late effect codes, which are represented using the same range of injury diagnosis codes but with the seventh digit modifier code of D through S, are also excluded.
- Primary mechanism of injury classified as either blunt or penetrating:
  - Blunt is defined as an injury where the primary E-code is mapped to the following categories: fall, machinery, motor vehicle traffic, pedestrian, cyclist, and struck by against.
  - Penetrating is defined as an injury where the primary E-code is mapped to the following categories: cut/pierce, and firearm.
- Calculated injury severity score ≥ 5.
- Emergency department discharge disposition and/or hospital discharge disposition must be known.

All injury severity score values were derived from registrar abstracted and recorded Abbreviated Injury Scale 2005 codes with 2008 updates (AIS 2005).

In general, orthopedic surgery utilizes unique classification systems that are specific to a certain anatomical fracture or injury (i.e., Neer—proximal humerus fracture, Garden—femoral neck fracture, Tile or Young & Burgess—pelvic ring injuries, etc.). Furthermore, the Arbeitsgemeinschaft für Osteosynthesefragen/Orthopedic Trauma Association (AO/OTA) classification system[10] is a comprehensive mechanism that is able to describe all musculoskeletal injuries within one framework. However, no large-scale trauma registries (NTDB, MTQIP, etc.) utilize any of these orthopedic specific classification systems. Instead, they rely on ICD-9 or ICD-10 or the Abbreviated Injury Scale, version 2005 (AIS2005). We chose to study the AIS2005 because it actually incorporates components of some orthopedic classification systems into its overall structure. For example, a partially stable pelvic ring injury (Tile B type pattern) has a designated AIS2005 code (856161.3).[11] We felt that these similarities between the systems would allow for the best opportunity to evaluate the ability of the trained registrars to accurately extract this information. Injury diagnoses are made from data that are extracted from various sources (Fig. 1) by the registrars. When there is not enough information available for the registrar to make an accurate diagnosis, NFS is chosen. NFS is also the default entry in some of the registry software systems that are used.

Analysis

Inclusion and exclusion criteria were designed to capture patients who would be most likely to have enough clinical information (i.e., radiographic studies, consult notes, etc.) to make an accurate diagnosis. We subsequently excluded patients who were dead on arrival, died in the emergency department or during their hospital stay, were transferred out, spent less than 12 h in the hospital, or left against medical advice. The remaining patients were included in the analysis as they represented a cohort that should have the necessary information in their medical record to make an accurate injury diagnosis.

We aimed to study injuries common to the field of orthopedic trauma, and so we included pelvic, acetabular, long bone, periarticular fractures and some tarsal bones in our analysis. We did not include hand and spine (cervical, thoracic, and lumbar) fractures as these injuries are usually treated by other subspecialists. We excluded the fibula in our analysis as the AIS2005 scale does not correlate well with the Danis-Weber classification system on which this is section is based. Injuries were also grouped into “simple” (i.e., tibial diaphysis, proximal femur, etc.) or “complex” (i.e., pelvic ring, acetabulum, periarticular, etc.) (Table 1). Senior orthopedic surgeons determined complexity based on the ease of classifying the facture as well as the level of expertise required to treat them. The percentage of NFS injuries in each injury group was calculated by dividing the number of NFS injuries by the total number of injuries in that group. The percentage of NFS injuries was also evaluated based on trauma center level and complexity of the injury.

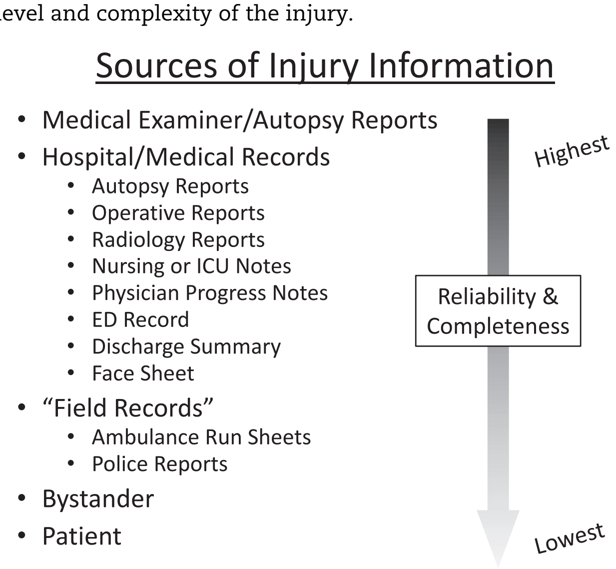

Fig. 1 – Sources of information used by trauma registrars to populate the registry. ICU = intensive care unit; ED = emergency department.

Table 1 – Fracture groups.

Simple	Complex
Proximal femur	Pelvic ring
Femoral shaft	Acetabulum
Tibial shaft	Distal femur
Clavicle	Proximal tibia
Humerus	Distal tibia
Radius	Talus
Ulna	Calcaneus
	Navicular
	Scapula
	Proximal humerus

Statistical methods

Statistical analyses were performed using Stata MP, version 14 (StataCorp, College Station, TX). Statistical significance was defined as a *P* value < 0.05. We used a proportion test (prtesti) to determine a difference in proportions between groups (level 1 *versus* level 2, simple *versus* complex). Approval for this study was obtained from the Michigan Medicine Institutional Review Board.

Results

We identified 70,918 total fractures in MTQIP, and 13,116 (18.5%) were classified as NFS. Percent NFS varied based on specific fracture type from 2.4% (radius) to 67.9% (navicular) (Table 2 and Fig. 2). There was a significantly higher proportion of complex NFS fractures than simple NFS fractures (34.5% *versus* 9.6%, $P < 0.001$) (Table 3). Level 1 trauma centers saw fewer complicated fractures (11,796 *versus* 13,593) but classified a higher proportion of them as NFS than level 2 trauma centers (21.2% *versus* 16.6%, $P < 0.001$) (Table 4).

In the NTDB, we identified 1,269,278 total fractures of which 342,472 (27%) were classified as NFS. Broken down by fracture type, this ranged from 6.0% (proximal femur) to 68.5% (navicular) (Table 2 and Fig. 3). There were also a higher proportion of complex fractures that were classified as NFS than simple fractures (41.8% *versus* 15.7%, $P < 0.001$). In contrast to our results from MTQIP, level 2 trauma centers in NTDB had a higher proportion of fractures classified as NFS than level 1 trauma centers, 26.6% *versus* 24.4% ($P < 0.001$). Also, level 2 trauma centers saw fewer complex fractures (92,134 *versus* 196,290) than level 1 trauma centers.

Discussion

In this study, we found a substantial number of unclassified orthopedic injuries in two large trauma registries. There was also a tremendous range of NFS entries in the specific fracture groups, from less than 10% to some with greater than 50%. Also, more complicated injuries patterns were associated with a higher proportion of NFS entries. Finally, the effect of level of trauma center on the proportion of NFS codes varied between the registries.

Data accuracy is paramount to obtaining accurate results and making robust conclusions in comparative effectiveness research. Missing data, and the methods used to compensate

Table 2 – Fracture type distribution.

	MTQIP		NTDB*	
	All fractures	NFS fractures (%)	All fractures	NFS fractures (%)
All fractures	70,918	13,116 (18.5)	1,269,278	342,472 (27)
Pelvic ring	7009	1639 (23)	146,938	52,168 (36)
Acetabulum	3083	1198 (39)	69,310	30,705 (44)
Proximal femur	23,571	825 (3.5%)	203,886	12,170 (6)
Femoral shaft	4334	1483 (34)	78,590	30,882 (39)
Distal femur	3038	1059 (35)	44,375	18,947 (43)
Proximal tibia	2518	530 (21)	64,648	17,774 (27)
Tibial shaft	2322	624 (27)	57,110	19,378 (34)
Distal tibia	2252	1087 (48)	64,172	33,443 (52)
Talus	884	507 (57)	19,562	11,192 (57)
Calcaneus	1084	625 (58)	24,769	14,245 (58)
Navicular	308	209 (68)	5545	3796 (68)
Clavicle	4065	1008 (25)	78,208	26,849 (34)
Scapula	2633	1209 (46)	53,378	27,132 (51)
Proximal humerus	2580	703 (27)	54,369	19,544 (36)
Humerus	1956	135 (6.9)	88,517	9219 (10)
Radius	5357	128 (2.4)	128,154	7715 (6)
Ulna	3924	147 (3.8)	87,747	7313 (8.3)

*Includes level 1, 2, 3, and 4 trauma centers.

for this, have been shown to affect the final results.[12] An NFS entry is essentially missing data as it creates a nonspecific entry in that variable field. Methodologists have stated that up to 10% of missing data could be tolerated[13,14]; however, more importantly, the reasons for this should be explored.[15] Further, while there are multiple ways to account for missing data,[16-18] it is unclear if these methods can be applied to categorical data in these registries. We have found that many commonly studied injuries in trauma care (pelvis, femoral diaphysis, etc.) have far greater than 10% NFS entries. With this number of essentially missing data, there should be caution in interpreting the conclusions of these studies.

Data accuracy is also vital when collaboratives, such as the MTQIP, use registries to evaluate performance (e.g., was the appropriate treatment rendered). For example, entering an NFS humerus fracture would not allow auditors to discern if the injury was a diaphyseal fracture that could be treated with a brace *versus* a distal intraarticular fracture which would require a surgical open reduction and internal fixation. Further, because collaboratives often include highly disparate institutions with variations in resources and systems, our ability to adequately risk adjust outcomes is impaired with a nonspecific diagnosis. The development of posttraumatic arthritis is not a rare outcome when treating intraarticular fractures (involving the joint surface); however, it is much more unlikely to occur in fractures that are extraarticular in nature. While this clinical distinction might seem obvious, we are unable to discern this key difference if NFS is entered into the registry. This could become an issue if institutions or providers are erroneously penalized or rewarded because of the inaccurate coding of these injuries.

It is somewhat intuitive that more complicated injuries would have a higher proportion of NFS entries. Simple fractures, such as the proximal femur group (hip fracture), is a common injury in both registries, and the classification system used by orthopedic surgeons has terms that are similar to those found in the AIS2005 descriptions. In contrast, for pelvic ring injuries, the Modified Tile classification[11] that AIS2005 employs is only one of multiple classification systems available. In fact, may orthopedic surgeons prefer the Young and Burgess system[19] to classify the injury, which may result in translation difficulties for data coders. These nuanced differences in terminology might explain why there is up 36% NFS entries for the pelvic ring. With the substantial number of complex patterns that were classified as NFS, we should examine the possible reasons for this problem such as the terminology used by surgeons, the training of registrars, and the means for a diagnosis adjudication.

It should be noted that in theory, all fractures can be appropriately classified in some fashion. Orthopedic surgeons have developed multiple classification systems for almost all fractures encountered in practice.[20] While there might be a complex fracture pattern or multiple distinct injuries in the bone, the fracture can still be classified. Thus, there are likely a variety of factors or barriers at play that might explain our findings. As mentioned before, the classification system and subsequent terminology used by the orthopedic surgeon might not be the same as the terminology known by the registrar, possibly leading to confusion. It is unknown if there are open communication channels between the orthopedic surgeons and the registrars to ask any questions or to gain further clarification before entering an NFS diagnosis. It is also unknown how much the current medical record lacks in needed information that would help discern a specified diagnosis. These barriers should be explored further so that we can lessen the proportion of NFS entries in these registries.

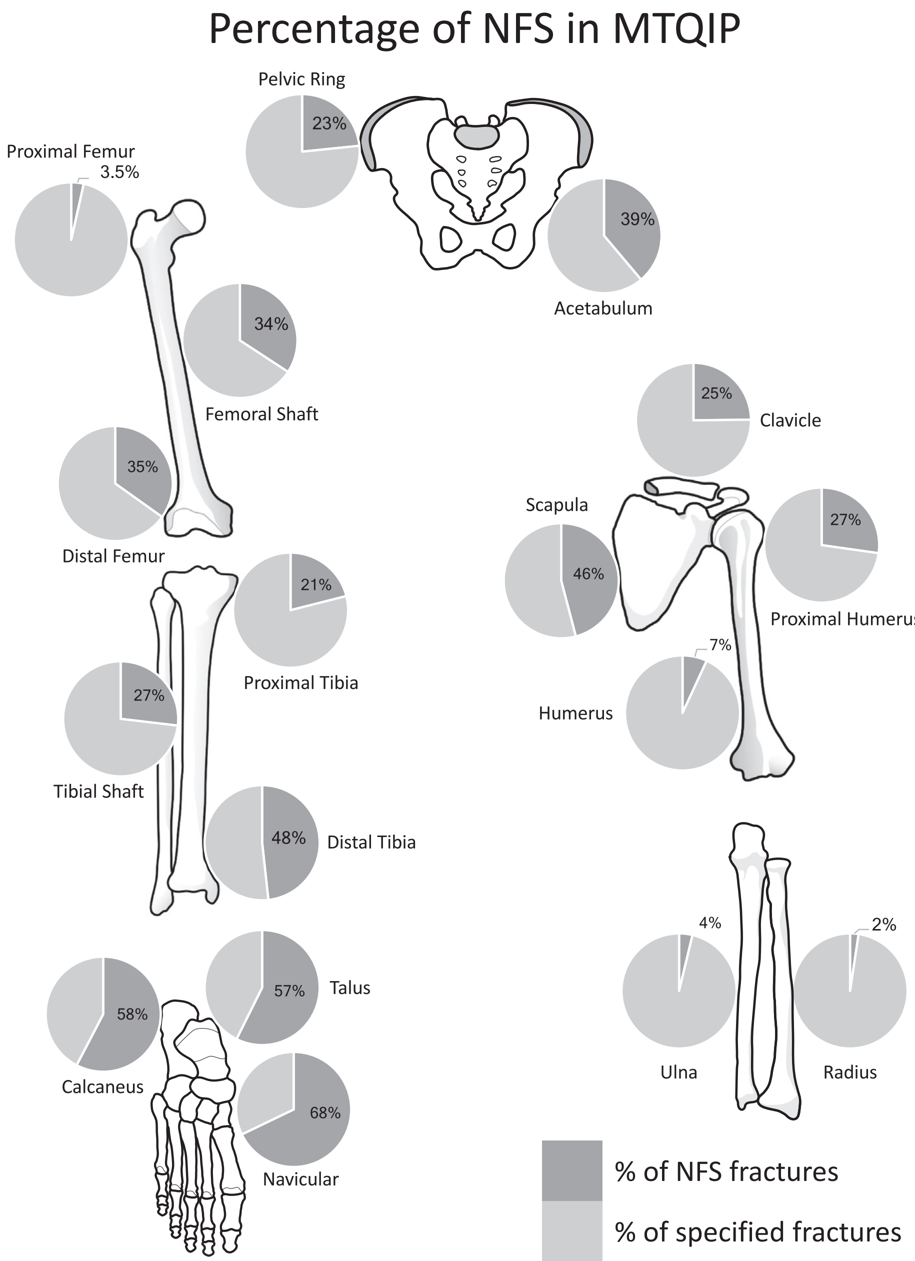

Fig. 2 – Percentage of Not Further Specified fractures found in the MTQIP Registry.

The correlation of trauma center level and proportion of NFS entries differed between MTQIP and NTDB. However, it should be noted that the centers that saw fewer complex injuries (level 1 centers in MTQIP and level 2 centers in NTDB) had a higher proportion of NFS entries. It has been shown with outcomes that trauma volume does make a difference.[21-23] It

Table 3 – NFS fracture based on fracture complexity.

	MTQIP			NTDB*		
	Simple	Complex	P value	Simple	Complex	P value
All fractures	45,529	25,389		722,212	547,066	
NFS fractures, *n* (%)	4350 (9.6)	8766 (34.5)	**<0.001**	113,526 (15.7)	228,946 (41.8)	**<0.001**

The *P* values were bolded for emphasis as being significant.
* Includes level 1, 2, 3, and 4 trauma centers.

Table 4 – NFS fracture based on trauma center level.

	MTQIP			NTDB		
	Level 1	Level 2	*P* value	Level 1	Level 2	*P* value
All fractures	29,122	41,796		427,149	220,077	
NFS fractures, *n* (%)	6187 (21.2)	6929 (16.6)	**<0.001**	104,358 (24.4)	58,527 (26.6)	**<0.001**

The *P* values were bolded for emphasis as being significant.

is possible that the registrars at these centers might have gained more experience with coding these more complicated injuries because of the higher volume. It should also be noted that our institution was not coded as a level 1 trauma center in NTDB. We are unsure if this misclassification might also affect other centers in that database which might then explain these counterintuitive results.

There are several limitations of this study. We included years 2011-2017, and there might have been varied adoption of AIS2005 from prior versions at different institutions that affected how comfortable the registrars were with this update. We are also unable to account for the variation in the quality of medical records or the experience of the registrars across the different institutions. That our institution was

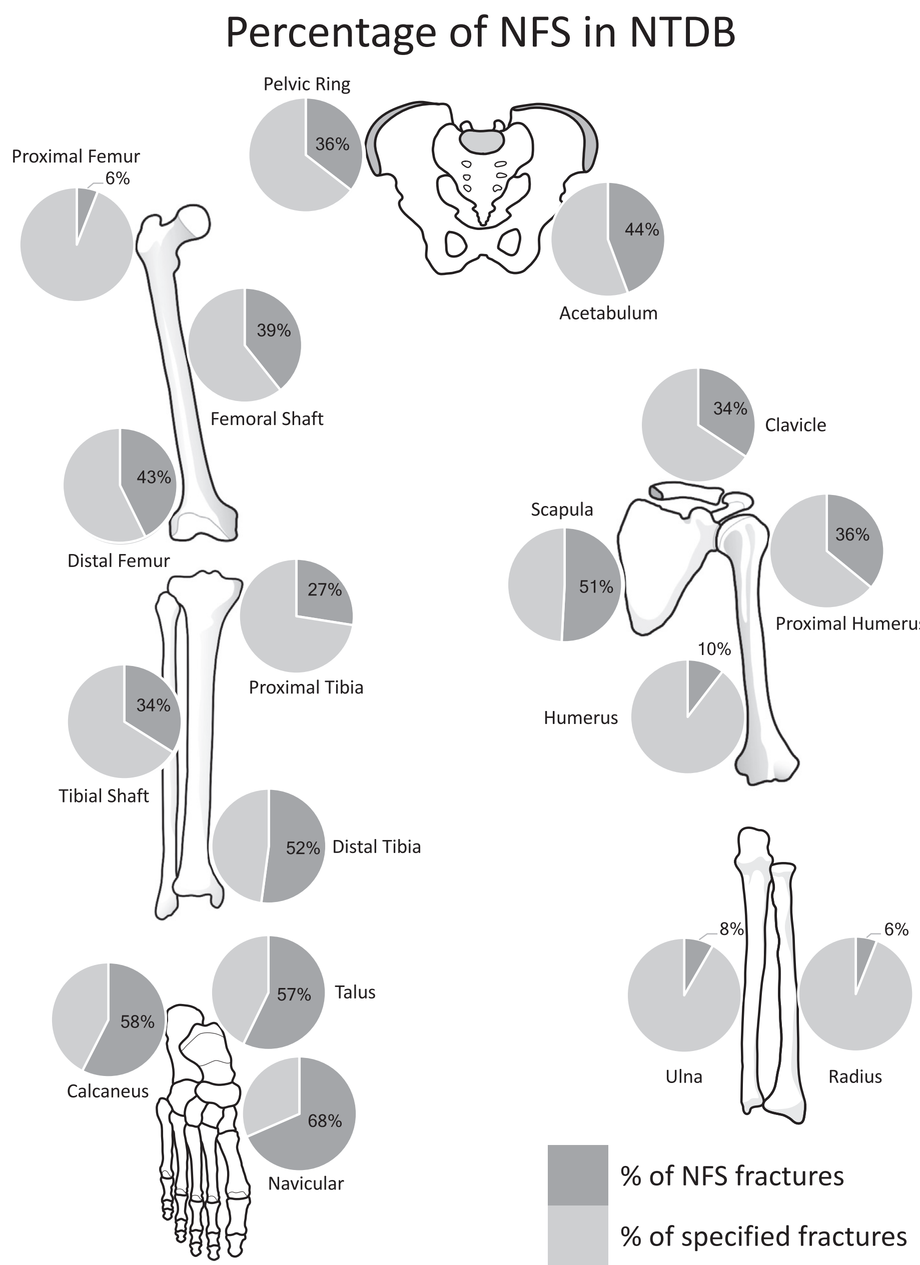

Fig. 3 – Percentage of Not Further Specified fractures found in the NTDB.

misclassified raises concerns that other institutions may have also been misclassified, which may in turn skew our results. We also did not include all potential orthopedic injuries such as spine or hand, and it is unknown if including these would lessen the proportion of NFS entries. Lastly, we did not examine the prevalence of nonspecific diagnosis codes found in the ICD-9 or ICD-10. These systems are also utilized in the registries to classify injuries; however, their ability to define a nonspecific diagnosis is not as clear as in AIS2005.

Conclusions

In conclusion, there is a significant number of nonspecific orthopedic injury data in two large trauma registries. Complicated fractures were more likely to be coded this way; however, the level of trauma center did not affect the proportion of NFS data. The lack of specificity in these registries is concerning, and this can affect the results and conclusions in research studies that use these sources. This also likely impacts our ability to truly risk adjust for treatments and outcomes in trauma centers. There are likely multiple reasons for these findings, and future studies should aim to explore these further.

Acknowledgment

The authors wish to acknowledge Mr. Adrian S. Wilkins for his help in creating the anatomical icons used in the figures.

Disclosure

Bryant W. Oliphant, Chelsea A. Harris, and James A. Goulet have no relevant financial interests or other conflicts of to disclose. Anne H. Cain-Nielsen receives salary support from Blue Cross Blue Shield of Michigan and Blue Care Network (a nonprofit mutual company) through their funding of the Michigan Trauma Quality Improvement Program. The company had no role in the study. She also receives grant support from the Michigan Department of Health and Human Services. There are no other conflicts to disclose. Mark R. Hemmila receives salary support from Blue Cross Blue Shield of Michigan and Blue Care Network (a nonprofit mutual company) through their funding of the Michigan Trauma Quality Improvement Program. The company had no role in the study. He also receives grant support from the Michigan Department of Health and Human Services. There are no other conflicts to disclose.

REFERENCES

1. Arabian SS, Marcus M, Captain K, et al. Variability in interhospital trauma data coding and scoring: a challenge to the accuracy of aggregated trauma registries. *J Trauma Acute Care Surg*. 2015;79:359–363.
2. Porgo TV, Moore L, Tardif PA. Evidence of data quality in trauma registries: a systematic review. *J Trauma Acute Care Surg*. 2016;80:648–658.
3. Cantu RV, Graves SC, Spratt KF. In-hospital mortality from femoral shaft fracture depends on the initial delay to fracture fixation and Injury Severity Score: a retrospective cohort study from the NTDB 2002-2006. *J Trauma Acute Care Surg*. 2014;76:1433–1440.
4. Morshed S, Miclau T, Bembom O, Cohen M, Knudson MM, Colford JM. Delayed internal fixation of femoral shaft fracture reduces mortality among patients with multisystem trauma. *J Bone Joint Surg Am*. 2009;91:3–13.
5. Oliphant BW, Tignanelli CJ, Napolitano LM, Goulet JA, Hemmila MR. American College of Surgeons Committee on Trauma verification level affects trauma center management of pelvic ring injuries and patient mortality. *J Trauma Acute Care Surg*. 2019;86:1–10.
6. Haws BE, Wuertzer S, Lenchik L, Miller AN. Misclassification of pelvic ring injuries in the National trauma Data Bank. *J Orthop Trauma*. 2015;29:460–462.
7. Hemmila MR, Jakubus JL, Cain-Nielsen AH, et al. The Michigan trauma quality improvement program: results from a collaborative quality initiative. *J Trauma Acute Care Surg*. 2017;82:867–876.
8. Michigan Trauma Quality Improvement Program. Available at: http://www.mtqip.org. Accessed March 23, 2018.
9. Hemmila MR, Jakubus JL. Trauma quality improvement. *Crit Care Clin*. 2017;33:193–212.
10. Kellam JF, Meinberg EG, Agel J, Karam MD, Roberts CS. Introduction: fracture and dislocation classification compendium-2018: international comprehensive classification of fractures and dislocations committee. *J Orthop Trauma*. 2018;32(Suppl 1):S1–S10.
11. Pennal GF, Tile M, Waddell JP, Garside H. Pelvic disruption: assessment and classification. *Clin Orthop Relat Res*. 1980;151:12–21.
12. Basques BA, McLynn RP, Fice MP, et al. Results of database studies in spine surgery can be influenced by missing data. *Clin Orthop Relat Res*. 2016;475:1–12.
13. Schafer JL. Multiple imputation: a primer. *Stat Methods Med Res*. 1999;8:3–15.
14. Bennett DA. How can I deal with missing data in my study? *Aust N Z J Public Health*. 2001;25:464–469.
15. Dong Y, Peng C-YJ. Principled missing data methods for researchers. *SpringerPlus*. 2013;2:222.
16. Van Der Palm DW, Van Der Ark LA, Vermunt JK. A comparison of incomplete-data methods for categorical data. *Stat Methods Med Res*. 2012;25:754–774.
17. Mallinckrodt C. Analyzing incomplete categorical data. In: *Preventing and Treating Missing Data in Longitudinal Clinical Trials. A Practical Guide*. Cambridge: Cambridge University Press; 2013:121–128.
18. Henry AJ, Hevelone ND, Lipsitz S, Nguyen LL. Comparative methods for handling missing data in large databases. *J Vasc Surg*. 2013;58:1353–1359.e1356.
19. Young JW, Burgess AR, Brumback RJ, Poka A. Pelvic fractures: value of plain radiography in early assessment and management. *Radiology*. 1986;160:445–451.
20. Marsh J, Slongo T, Agel J, et al. Fracture and Dislocation classification Compendium - 2007. *J Orthop Trauma*. 2007;21:S1–S6.
21. Freeman J, Nicholl J, Turner J. Does size matter? The relationship between volume and outcome in the care of major trauma. *J Health Serv Res Policy*. 2006;11:101–105.
22. Bennett KM, Vaslef S, Pappas TN, Scarborough JE. The volume-outcomes relationship for United States Level I trauma centers. *J Surg Res*. 2011;167:19–23.
23. Nathens AB, Jurkovich GJ, Maier RV, et al. Relationship between trauma center volume and outcomes. *JAMA*. 2001;285:1164–1171.