

# Translating coding lists in administrative claims-based research for cardiovascular procedures

Zachary J. Wanken, MD,<sup>a,b</sup> Peter B. Anderson, BA,<sup>c</sup> Sarah Y. Bessen, MPH,<sup>c</sup> John B. Rode, BA,<sup>c</sup> Jesse A. Columbo, MD, MS,<sup>a,b</sup> Spencer W. Trooboff, MD, MBA,<sup>a,b</sup> Kayla O. Moore, MPH,<sup>b</sup> and Philip P. Goodney, MD, MS,<sup>a,b</sup> Lebanon, NH

## ABSTRACT

**Background:** To effectively use administrative claims for healthcare research, clinical events must be inferred from coding data according to validated algorithms. In October 2015, the United States transitioned from the *International Classification of Diseases Ninth Revision (ICD-9)* to the Tenth Revision (ICD-10). We describe our method to derive new ICD-10 codes for outcomes after vascular procedures from our prior, validated ICD-9 codes.

**Methods:** We began with validated ICD-9 coding lists known to represent outcomes after lower extremity revascularization, thoracic aortic endograft placement, abdominal aortic aneurysm reintervention, and carotid revascularization. We used the publicly available general equivalence mapping tools to derive corresponding ICD-10 codes for each of the ICD-9 codes in our current lists. The resulting lists were then manually reviewed by multiple authors to ensure clinical relevance for appropriate event detection. Clinically nonrelevant and duplicated codes were removed.

**Results:** A total of 475 ICD-9 codes were translated to ICD-10 with a 98-fold increase ( $n = 46,630$ ) in the total number of codes. Overall, we found that 77% of codes ( $n = 35,833$ ) were either duplicated or not clinically relevant upon manual review. For example, for thoracic aortic endograft placement, 97 ICD-9 codes mapped to 14,661 ICD-10 codes in total. A total of 890 codes were removed as duplicates and 9035 codes were removed during manual clinical review. The resultant, reviewed list contained 4736 ICD-10 codes representing a 49-fold increase from the initial ICD-9 list. Findings were similar across the other procedures studied.

**Conclusions:** ICD-10 has expanded the number of codes necessary to describe outcomes after vascular procedures. More than 75% of the codes obtained using the general equivalence mapping database were either duplicated or not clinically relevant. Manual review of codes by researchers with clinical knowledge of the procedures is imperative. (*J Vasc Surg* 2019;■:1-7.)

**Keywords:** Outcomes research; Coding; Algorithm; Aortic aneurysm; Carotid

Accurate tracking of long-term outcomes after cardiovascular procedures is critical for evidence-based care. Long-term outcome assessment allows providers and companies to track the safety and efficacy of new treatments and new devices and provides evidence for improvement on existing therapies.

Identifying these outcomes in real-world practice is challenging. Patients are often treated at more than one hospital center, may not be entered into a procedural registry for outcome tracking, or may be lost to follow-up. To remedy these limitations, researchers seeking to measure outcomes after cardiovascular procedures have commonly used administrative claims data.<sup>1,2</sup> This method has several advantages. It allows for patients to be tracked across hospital systems and different diagnoses and, because it is tied to institutional reimbursement, patients are infrequently lost to follow-up.<sup>3</sup> However, the primary limitation of this type of data is that it is reliant on the use of billing codes as a surrogate for clinical events. Since the inception of the *International Classification of Diseases*, ninth revision (ICD-9) more than 30 years ago, researchers have put substantial effort into refining the individual ICD-9 codes that can be reliably used for clinical event detection, with great success.<sup>4</sup>

However, in October 2015 the United States transitioned from the ICD-9<sup>5</sup> to the tenth revision (ICD-10),<sup>6</sup> making the codes previously used for ICD-9 obsolete. Government entities including the Centers for Medicare and Medicaid Services and the National Bureau of

From the Section of Vascular Surgery, Dartmouth-Hitchcock Medical Center<sup>a</sup>; The Dartmouth Institute for Health Policy and Clinical Practice<sup>b</sup>; and The Geisel School of Medicine at Dartmouth College.<sup>c</sup>

This work was supported by the American Heart Association Grant # 18SFRN33900147. 2018. Population Health Project. Philip Goodney.

Author conflict of interest: none.

Presented at the American Heart Association, Vascular Discovery: From Genes to Medicine Meeting, Boston, Mass, May 14-16, 2019.

Additional material for this article may be found online at [www.jvascsurg.org](http://www.jvascsurg.org).

Correspondence: Zachary J. Wanken, MD, Section of Vascular Surgery, Dartmouth-Hitchcock Medical Center, One Medical Center Dr, Lebanon, NH 03766 (e-mail: [zwanken@gmail.com](mailto:zwanken@gmail.com)).

The editors and reviewers of this article have no relevant financial relationships to disclose per the JVS policy that requires reviewers to decline review of any manuscript for which they may have a conflict of interest.

0741-5214

Published by Elsevier Inc. on behalf of the Society for Vascular Surgery.

<https://doi.org/10.1016/j.jvs.2019.09.040>

Economics Research (NBER) have provided translation tools known as general equivalence mappings (GEMs), but fewer than 3% of translations yield a direct match between ICD-9 and ICD-10 codes.<sup>7</sup> The objective of this study was to use our previously developed and validated ICD-9 codes and translate them to ICD-10 in a systematic way, maintaining clinical relevance for desired event detection. We applied this process across four categories of cardiovascular procedures, namely, lower extremity revascularization, thoracic aortic endograft placement, reintervention after abdominal aortic aneurysm repair, and carotid revascularization.

## METHODS

**Methodologic overview.** Administrative claims-based research requires the identification of clinical events using claims codes associated with hospitalizations, procedures, or office visits. We have previously developed and validated lists of codes (coding algorithms) for each clinical event that we want to capture. Algorithms have been categorized based on the type of procedure and event detection desired. For example, our algorithm for carotid revascularization focuses on codes related to interventions of head and neck vessels, whereas our algorithm for lower extremity revascularization focuses on additional revascularization procedures or amputation. Our current lists consist of ICD-9 codes that have been used to identify clinical events of interest. Our group's prior work has focused on the refinement of these algorithms with validation against registry data as well as manual chart review to ensure accurate event capture.<sup>4</sup> Prior work has also outlined our initial translation work between ICD-9 and ICD-10.<sup>8</sup> This study builds on our previous work and describes further refinement of our translation process. An iterative six-step process was developed to translate our current ICD-9 based algorithms to ICD-10 (Fig 1).

**Step 1: Refining the initial list of ICD-9 codes.** In our first step, we labelled the ICD-9 codes in our existing coding algorithms. We used the `VLOOKUP` command in Microsoft Excel (Redmond, Wash) to automate the labelling of all included ICD-9 codes from a bank of all ICD-9 codes and associated labels. We then used these labels in latter steps of the clinical review process to ensure carryover of the events of interest.

**Steps 2 and 3: Translation of ICD-9 codes.** In step 2, we performed a direct translation from ICD-9 to ICD-10 using publicly available GEMs from NBER.<sup>9</sup> Procedure codes and diagnosis codes were mapped separately. First, we performed this step using Stata/IC v15.1 (StataCorp LLC, College Station, Tex). For a given set of ICD-9 codes, we used three data files to generate the corresponding set of ICD-10 codes: (1) previously validated ICD-9 code list, (2) GEM-based ICD-9 to ICD-10 crosswalk, which contains one entry for each ICD-10 code

## ARTICLE HIGHLIGHTS

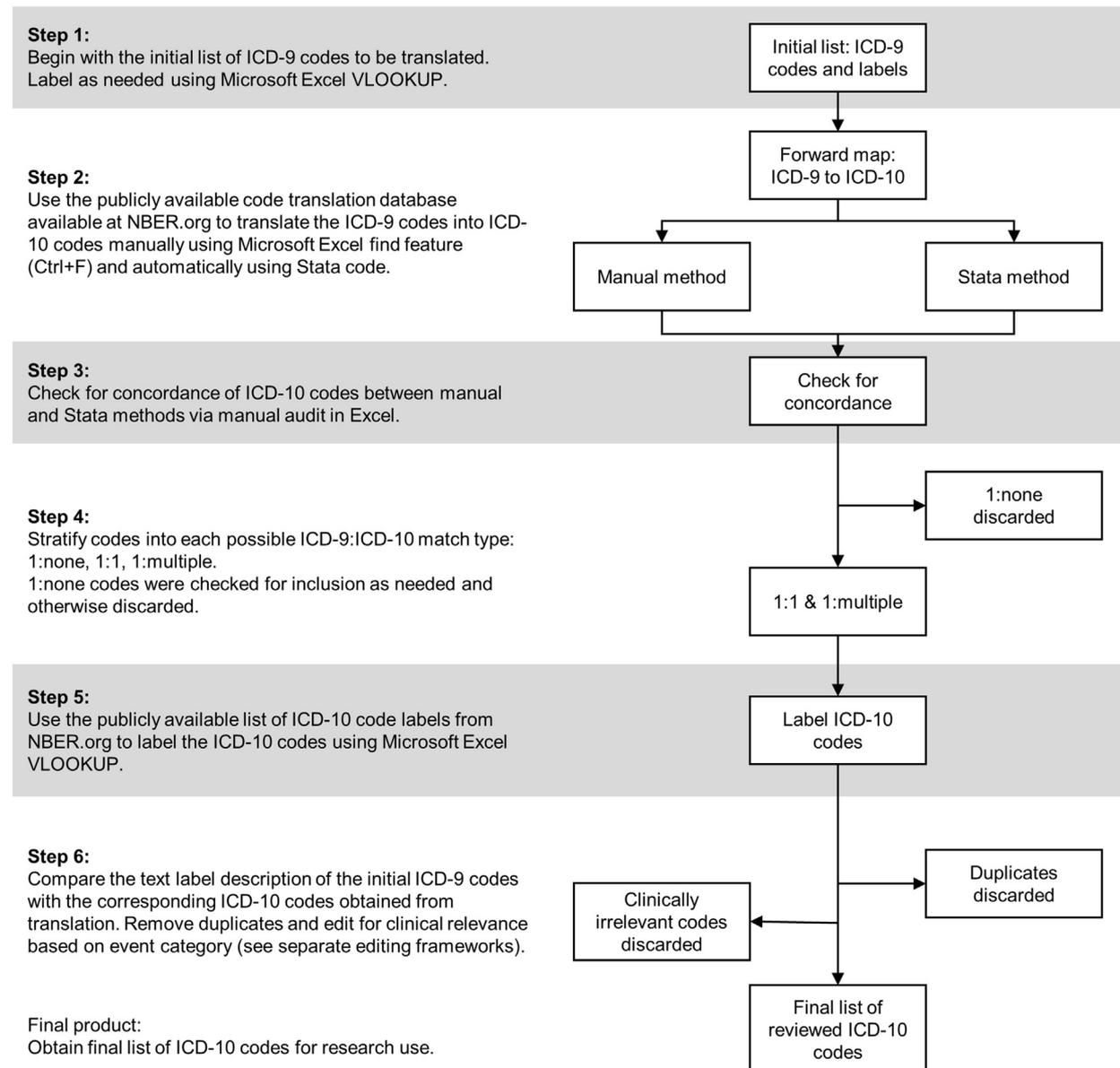
- **Type of Research:** Translation of coding algorithm between the *International Classification of Diseases, Ninth Revision* and *Tenth Revision* to identify clinical diagnoses and procedures for claims-based research
- **Key Findings:** Direct translation of coding algorithms using publicly available crosswalks resulted in a 70-fold increase in the number of codes. Upon review, nearly 80% of these codes were either duplicate or were not clinically relevant for event detection.
- **Take Home Message:** A manual review of codes from researchers with clinical knowledge of the procedures is imperative.

with its corresponding ICD-9 code, and (3) GEM-based ICD-10 label file. The GEM crosswalk file was merged many:1 with the validated ICD-9 code list, followed by the ICD-10 label file. Only matched entries were retained, reflecting the raw list ICD-10 codes forward mapped from ICD-9.

To ensure the accuracy of this method, each list was also manually mapped using Microsoft Excel. Each ICD-9 code was identified in the GEM ICD-9 to ICD-10 forward mapping crosswalk using the find function (Ctrl+F). All associated ICD-10 codes were then copied (Ctrl+C) and pasted (Ctrl+V) into an associated column. This process was repeated for every ICD-9 code in each of our algorithms. The resultant lists were compared to the lists produced using the Stata method and noted to be identical (step 3).

**Step 4: Review of nontranslated codes.** In step 4, we reviewed all ICD-9 codes that did not translate to an ICD-10 code using the GEM crosswalk (1:none). Upon manual review, we noted that the ICD-9 codes that mapped 1:none were broad in scope and the events that these codes were meant to detect were captured by ICD-10 codes mapped from other ICD-9 codes in our list. For example, ICD-9 code 0063, percutaneous insertion of carotid artery stent, did not forward map to any ICD-10 codes using the GEM. However, review of the available ICD-10 codes demonstrated that all carotid stent codes had been appropriately included by mapping another ICD-9 code in this category (code 0061: percutaneous angioplasty of extracranial vessel). Manual review demonstrated a similar pattern across all categories and no additional ICD-10 codes were included for the ICD-9 codes that mapped 1:none.

**Steps 5 and 6: Review of ICD-10 codes and compiling of final list for clinical event detection.** In step 5, we applied labels to the resultant ICD-10 code lists using the same process that was described in step 1. The `VLOOKUP` command was used in Microsoft Excel to



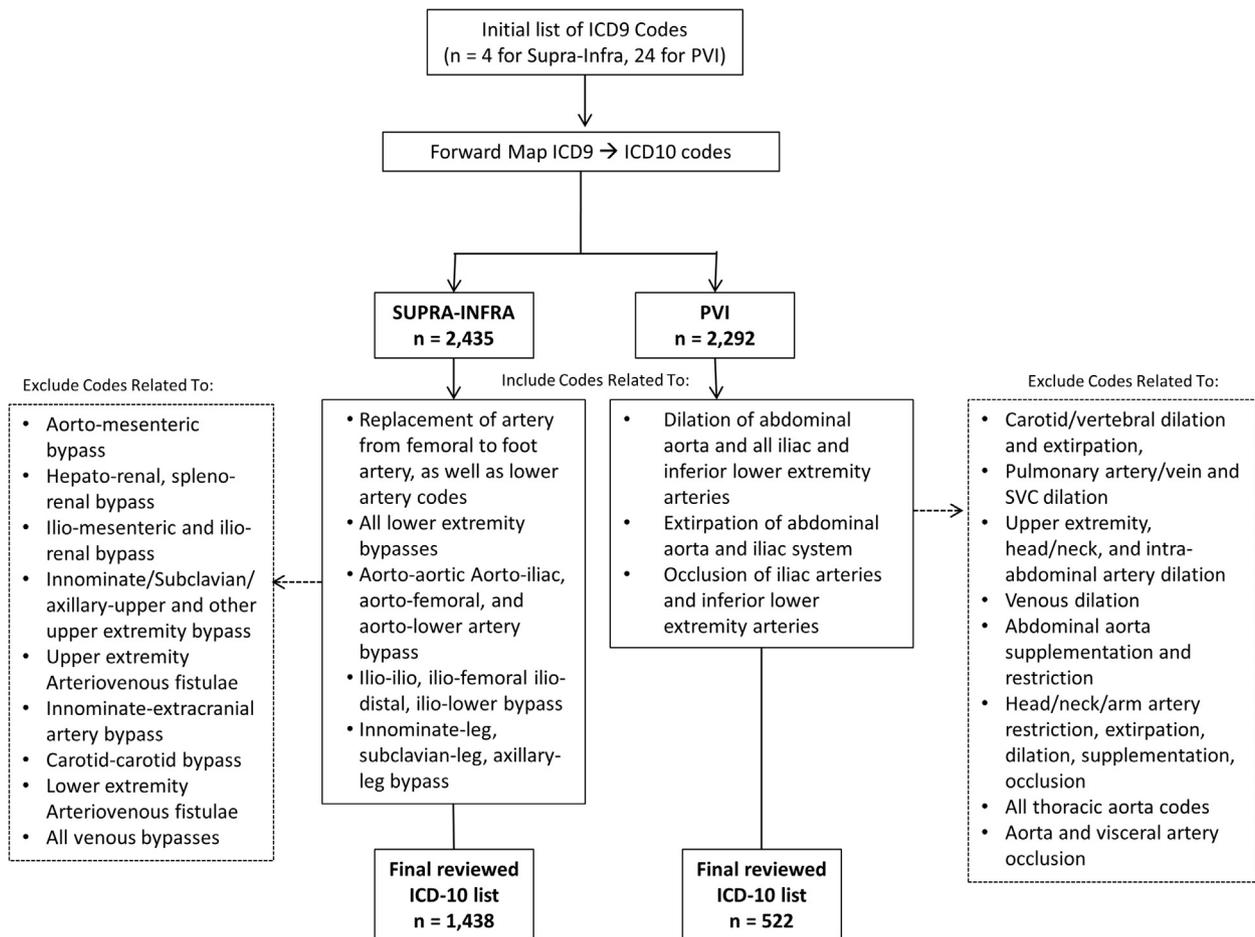
**Fig 1.** Process map for translation from *International Classification of Diseases (ICD) Ninth Revision* to *Tenth Revision* and validation of clinically relevant codes.

automate labelling of the codes using code labels from NBER.<sup>9</sup>

We then completed manual review of all translated ICD-10 codes and labels in step 6. We first removed all duplicated codes using the remove duplicates command in Excel. Duplicated codes occurred because many ICD-9 codes mapped to a similar set of ICD-10 codes. Therefore, many ICD-10 codes were present multiple times. The remove duplicates command eliminated the duplicate codes leaving just one instance of each ICD-10 code included. Next, a team of at least two clinicians separately reviewed each ICD-10 code and associated label to determine the clinical relevance of

inclusion. If there was uncertainty during the initial review, the code was flagged for full research team review. This process was facilitated via the use of inclusion criteria frameworks developed by subject experts familiar with the procedures being studied (Figs 2-4).

ICD-9 code 38.12 can be used as an example for our process overall. The associated ICD-9 label for code 38.12 is "Endarterectomy, other vessels of head and neck." Using the GEM translation tool, code 38.12 maps to 31 ICD-10 codes. After clinical review, we included 16 of these codes in our final ICD-10 list. For example, code O3CH0ZZ with associated label Extirpation of matter from right common carotid artery, open approach" was included as a



**Fig 2.** Lower extremity revascularization framework. *ICD*, International Classification of Diseases; *PVI*, endovascular peripheral vascular intervention; *Supra-Infra*, suprainguinal and infrainguinal open revascularization; *SVC*, superior vena cava.

clinically relevant code. Alternatively, 15 codes were not clinically relevant and eliminated from our final ICD-10 list. For example, code O3CS4ZZ with associated label “Extirpation of matter from right temporal artery, percutaneous endoscopic approach” was excluded.

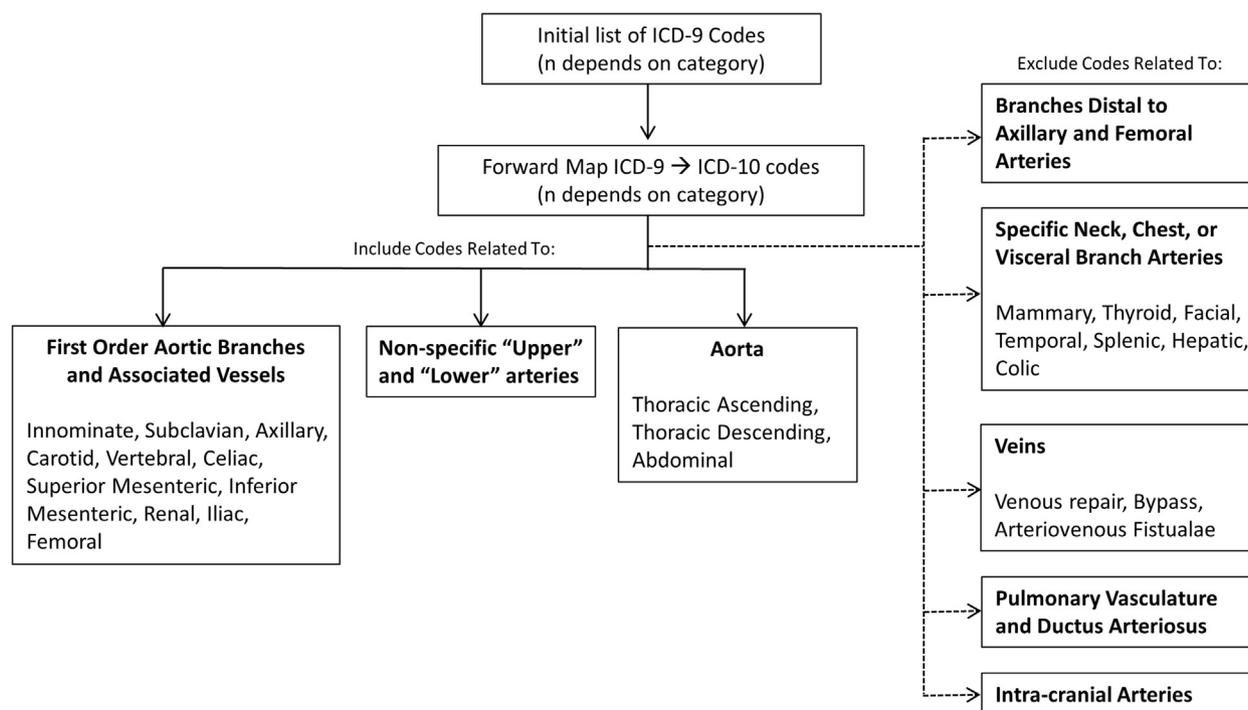
This process was repeated across four categories of cardiovascular procedures: lower extremity revascularization, thoracic aortic endograft placement, reintervention after abdominal aortic aneurysm repair, and carotid revascularization including endarterectomy and stenting. In completing the process, we derived four new ICD-10-based coding algorithms. This work used publicly available coding lists and crosswalks and did not involve human patients. Therefore, informed consent and institutional review board evaluation were not required.

## RESULTS

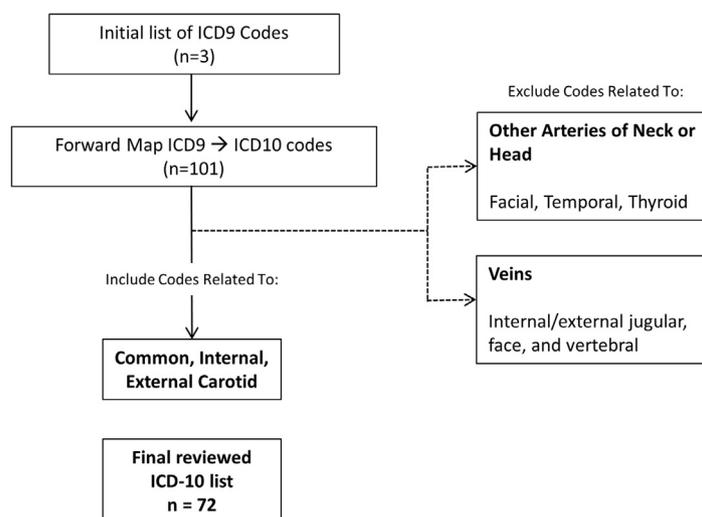
**Overview of codes for all procedures.** We started with a total of 475 ICD-9 codes across four categories: 28 related to lower extremity revascularization, 97 related to thoracic aortic endograft placement, 347 related to

abdominal aortic aneurysm repair, and 3 related to carotid revascularization (Fig 5). These 475 initial codes were directly translated to 46,630 total ICD-10 codes using the GEM translation tools, a 98-fold increase. Of these, 16,204 were duplicates and were removed resulting in 30,426 ICD-10 codes. An additional 19,629 codes were removed through manual clinical review, resulting in 10,797 ICD-10 codes across all four categories and representing a 23-fold increase from the original ICD-9 lists. The comprehensive lists of codes for initial ICD-9, final ICD-10, and eliminated ICD-10 codes can be found in [Appendices A to C](#) (online only), respectively.

**Lower extremity revascularization.** Our initial ICD-9 codes for lower extremity revascularization were separated into two additional categories: peripheral endovascular intervention (endovascular peripheral vascular intervention [PVI], 24 ICD-9 codes) and open surgical bypass (4 ICD-9 codes; Fig 2). All of these were procedure codes, because there were no diagnosis codes for this category. In total, this collection of 28 ICD-9 codes mapped to 4727 ICD-10 codes (2292 PVI, 2435 surgical



**Fig 3.** Thoracic and abdominal aortic repair framework. *ICD, International Classification of Diseases.*



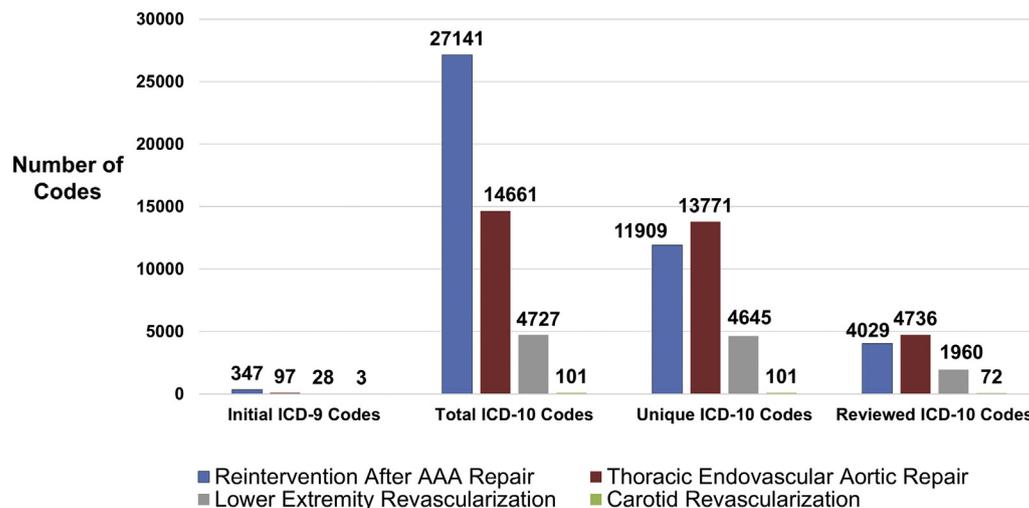
**Fig 4.** Carotid revascularization framework. *ICD, International Classification of Diseases.*

bypass). A total of 82 codes were removed as duplicates and 2685 codes were removed because they were not clinically relevant. The resultant, reviewed list contained 1960 ICD-10 codes (522 PVI, 1438 surgical bypass) representing a 70-fold increase from the initial ICD-9 list.

**Thoracic aortic endograft placement.** Our initial ICD-9 code list for thoracic aortic endograft placement included 97 codes and were divided into four categories: 37 index procedure codes, 3 codes pertaining to aortic rupture, 37 procedure codes for reintervention, and 20 diagnosis codes for reintervention (Fig 3). In total, these 97 codes mapped to 14,661 ICD-10 codes. A total of 890

codes were removed as duplicates and 9035 codes were removed because they were not clinically relevant. The resultant, reviewed list contained 4736 ICD-10 codes (2349 index procedure, 4 rupture, 2332 procedure codes for reintervention, 51 diagnosis codes for reintervention) representing a 49-fold increase from the initial ICD-9 list.

**Reintervention after abdominal aortic aneurysm repair.** Our initial ICD-9 code list for reintervention after abdominal aortic aneurysm repair included 347 codes (328 procedure codes, 19 diagnosis codes) which mapped to 27,141 ICD-10 codes in total (26,992 procedure



**Fig 5.** International Classification of Diseases (ICD) Ninth and Tenth Revisions code counts across the translation process.

codes, 149 diagnosis codes; Fig 3). A total of 15,232 codes were removed as duplicates and 7880 codes were removed during manual clinical review. The resultant, reviewed list contained 4029 ICD-10 codes (3979 procedure codes, 50 diagnosis codes) representing a 12-fold increase from the initial ICD-9 list.

**Carotid revascularization.** Our initial ICD-9 code list for carotid revascularization included three codes which mapped to 101 ICD-10 codes in total (Fig 4). There were no duplicated codes on this list and 29 codes were removed during manual clinical review. The resultant, reviewed list contained 72 ICD-10 codes representing a 24-fold increase from the initial ICD-9 list.

## DISCUSSION

In this study, we describe a six-step process for translating validated lists of ICD-9 codes to ICD-10 for event detection in claims-based research. Overall, we found a 23-fold increase across four categories of cardiovascular procedures. This ranged from a 12-fold increase for reintervention after abdominal aortic aneurysm repair to a 70-fold increase for codes pertaining to lower extremity revascularization. Duplicated codes accounted for more than one-third of the ICD-10 lists when using the GEM translation tool alone. Another third of the translated codes were eliminated after clinical review revealed that they were not relevant for clinical events of interest. Despite eliminating more than one-half of the translated codes, our final lists still represented a 23-fold increase with ICD-10 compared with ICD-9.

More broadly, the transition from ICD-9 to ICD-10 resulted in more than an eight-fold increase in the total number of codes from about 17,000 to about 140,000 codes. For example, ICD-9 includes approximately 13,000 diagnosis codes and 3800 procedure codes compared with 69,000 diagnosis and 72,000 procedure

codes in ICD-10.<sup>10,11</sup> The main reason for the code increase is the added level of specificity in ICD-10, and this is particularly true for procedure codes. Health services researchers across all specialties must translate algorithms for ongoing research beyond 2015. However, translation of cardiovascular procedures has been shown to be particularly complex. Boyd et al found<sup>7</sup> that cardiovascular procedures were the sixth most complex specialty category for translation between ICD-9 and ICD-10. Only 1% of codes were found to map identically with one new code and 55% of translations were categorized as convoluted. We similarly found that very few ICD-9 codes mapped 1:1 with ICD-10 resulting in a massive increase in the number of codes in each of our algorithms.

When used alone, publicly available GEM crosswalks result in translations with many duplicated ICD-10 codes. In our experience, the resultant lists are also excessively inclusive for nonrelevant and unwanted clinical events. Therefore, GEMs are inadequate when used alone to translate coding algorithms for claims-based research. For this reason, close manual review by researchers with clinical knowledge of the procedures is imperative. For example, our ICD-9 lists include several codes that are quite broad in scope such as code 3929 "Other (peripheral) vascular shunt or bypass." This single ICD-9 code mapped to 1287 ICD-10 codes which specify the location, conduit, and target of each shunt or bypass. Another example is code 3950 "Angioplasty of other non-coronary vessel(s)," which maps to 1094 ICD-10 codes with similar added specificity. There were even some more specific ICD-9 codes that mapped to a surprising number of ICD-10 codes. For example, code 3925 "Aorta-iliac-femoral bypass" maps to 320 ICD-10 codes owing to the myriad combinations of inflow and outflow targets as well as conduit options. The added level of

specificity inherent to ICD-10 allowed us to narrow the scope of our algorithms specifically to our clinical events of interest. However, manual clinical review was a time-intensive process in our experience. Whereas direct translation and removal of duplicated codes was largely automated, our clinical review process required our research team to render a clinical decision on each individual code.

The transition to ICD-10 represents a paradigm shift from ICD-9 where nonspecific codes have been largely eliminated in favor of codes with a considerably greater level of specificity. Hopefully, claims-based research will benefit from the added level of detail. Each added level of specification could be used to track the individual procedures, complications, and reinterventions associated with specific device types or anatomic location. However, it remains to be seen how these codes are used in real-world clinical practice. It may be very difficult for coders and billers to translate the medical documentation to the level of specificity possible in ICD-10. Additionally, billing practices are likely to vary from institution to institution, potentially limiting generalizability. Variation will also likely be seen over time as coders become more familiar with ICD-10 and adapt their practices. Interestingly, many potential operations have been included in ICD-10 that have not been described and are unlikely to be performed. However, the codes have clearly been designed to be all encompassing so that specificity can be maximized.

We do acknowledge limitations of our study. Although we completed manual clinical review of each list, true event detection from real world claims data remains unknown. In our future work, we plan to compare event rates in procedures after 2015 (ICD-10 based) with those same procedures before 2015 (ICD-9 based) for comparison of event rates in real-world practice. Additionally, we plan to compare events identified in ICD-10 with events noted in the Vascular Quality Initiative registry<sup>12</sup> for comparison across platforms. We also did not define discrete clinical outcomes in the resultant lists. We organized our lists into general categories for use in identifying reinterventions or open revascularization compared with a percutaneous procedure. However, we did not define code groupings for specific outcomes, and we have not included an exhaustive list to identify all important outcomes after vascular interventions.

## CONCLUSIONS

Research using administrative claims data requires ongoing adaptation of algorithms for accurate event detection. The transition from ICD-9 to ICD-10 broadly expanded the number and specificity of codes used to describe cardiovascular diagnoses and procedures. When GEM translation tools were applied to existing ICD-9 algorithms, more than 75% of the resulting

ICD-10 codes were either duplicated or not clinically relevant. Therefore, a manual review of codes by researchers with clinical knowledge of the procedures is imperative.

## AUTHOR CONTRIBUTIONS

Conception and design: ZW, PA, SB, JR, JC, ST, KM, PG  
Analysis and interpretation: ZW, PA, SB, JR, ST, KM, PG  
Data collection: ZW, PA, SB, JR, ST  
Writing the article: ZW, PA, SB, JR, ST  
Critical revision of the article: ZW, PA, SB, JR, JC, ST, KM, PG  
Final approval of the article: ZW, PA, SB, JR, JC, ST, KM, PG  
Statistical analysis: Not applicable  
Obtained funding: Not applicable  
Overall responsibility: ZW

## REFERENCES

- Alluri RK, Leland H, Heckmann N. Surgical research using national databases. *Ann Transl Med* 2016;4:393.
- Bensley RP, Yoshida S, Lo RC, Fokkema M, Hamdan AD, Wyers MC, et al. Accuracy of administrative data versus clinical data to evaluate carotid endarterectomy and carotid stenting. *J Vasc Surg* 2013;58:412-9.
- Gavriellov-Yusim N, Friger M. Use of administrative medical databases in population-based research. *J Epidemiol Community Health* 2014;68:283-7.
- Columbo JA, Kang R, Hoel AW, Kang J, Leinweber KA, Tauber KS, et al. A comparison of reintervention rates after endovascular aneurysm repair between the Vascular Quality Initiative registry, Medicare claims, and chart review. *J Vasc Surg* 2019;69:74-9.
- International Classification of Diseases, Ninth Revision (ICD-9). Centers for Disease Control and Prevention. Available at: [www.cdc.gov/nchs/icd/icd9.htm](http://www.cdc.gov/nchs/icd/icd9.htm). Accessed September 10, 2018.
- International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM). Centers for Disease Control and Prevention. Available at: [www.cdc.gov/nchs/icd/icd10cm.htm](http://www.cdc.gov/nchs/icd/icd10cm.htm). Accessed September 10, 2018.
- Boyd AD, Li J, Kenost C, Zaim SR, Krive J, Mittal M, et al. ICD-10 procedure codes produce transition challenges 2017. *AMIA Jt Summits Transl Sci Proc*; 2018. p. 35-44.
- Columbo JA, Kang R, Trooboff SW, Jahn KS, Martinez CJ, Moore KO, et al. Validating publicly available crosswalks for translating ICD-9 to ICD-10 diagnosis codes for cardiovascular outcomes research. *Circ Cardiovasc Qual Outcomes* 2018;11:e004782.
- The National Bureau of Economic Research. Available at: <http://nber.org/>. Accessed September 10, 2018.
- Jones LMN. Use caution when entering the crosswalk: a warning about relying on GEMs as your ICD-10 solution. ICD LOGIC. Available at: [www.cms.org/uploads/ICDLogicGEMSWHITEPAPER.pdf](http://www.cms.org/uploads/ICDLogicGEMSWHITEPAPER.pdf). Accessed September 10, 2018.
- Centers for Medicare & Medicaid Services. Available at: [www.cms.gov/](http://www.cms.gov/). Accessed September 10, 2018.
- Vascular Quality Initiative. 2018. Available at: [www.vqi.org/](http://www.vqi.org/). Accessed September 10, 2018.

Submitted Jun 5, 2019; accepted Sep 13, 2019.

Additional material for this article may be found online at [www.jvascsurg.org](http://www.jvascsurg.org).