MINI-SENTINEL METHODS
16 HEALTH OUTCOMES OF INTEREST FOR SURVEILLANCE PREPAREDNESS

Prepared by: Cristin Freeman, MPH, MBE,[1,2] Charles Leonard, PharmD, MSCE,[1,2] Patrick Archdeacon, MD,[3] Ryan Carnahan, PharmD, MS, BCPP,[4] Elizabeth Chrischilles, MS, PhD,[4] Sean Hennessy, PharmD, PhD,[1,2]

Author Affiliations: 1. Center for Clinical Epidemiology and Biostatistics, Department of Biostatistics and Epidemiology, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA; 2. Center for Pharmacoepidemiology Research and Training, Department of Biostatistics and Epidemiology, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA; 3. Office of Medical Policy, Center for Drug Evaluation and Research, FDA, Silver Spring, MD; 4. Department of Epidemiology, College of Public Health, The University of Iowa, Iowa City, IA

Acknowledgements: The authors wish to thank the following staff members from the Harvard Pilgrim Health Care Institute: Aarthi Iyer, MPH, and Sunali Goonesekera, SM, for their excellent project management; Candace Fuller, PhD, MPH, for contributing meaningful edits to this report; and Lisa Trebino, MS, and Tiffany Siu Woodworth, MPH, for their assistance with questions related to the Mini-Sentinel Common Data Model.

May 21, 2014

Mini-Sentinel is a pilot project sponsored by the U.S. Food and Drug Administration (FDA) to inform and facilitate development of a fully operational active surveillance system, the Sentinel System, for monitoring the safety of FDA-regulated medical products. Mini-Sentinel is one piece of the Sentinel Initiative, a multi-faceted effort by the FDA to develop a national electronic system that will complement existing methods of safety surveillance. Mini-Sentinel Collaborators include Data and Academic Partners that provide access to healthcare data and ongoing scientific, technical, methodological, and organizational expertise. The Mini-Sentinel Coordinating Center is funded by the FDA through the Department of Health and Human Services (HHS) Contract number HHSF223200910006I.

Mini-Sentinel Methods
16 Health Outcomes of Interest for Surveillance Preparedness

Table of Contents

I. BACKGROUND

Certain health outcomes may be causally associated with exposure to a medical product and/or medical countermeasures. As many of these health outcomes are relatively rare events, detecting a corresponding signal in the treatment exposed population poses a challenge to post-marketing active surveillance systems. Electronic healthcare data environments with large population coverage, such as the Mini-Sentinel distributed database (MSDD), provide opportunities to conduct active surveillance for such health outcomes of interest (HOIs) after exposure to medical products and medical countermeasures. Thus, development of HOI identification algorithms for use in the MSDD is needed.

FDA identified 16 HOIs as priorities for algorithm development. The HOIs were selected based on several criteria: 1) likelihood of association with new medical products and/or cause of withdrawal of past medical products, 2) clinical and public health importance, 3) suitability for detection with active surveillance (e.g., short latency period, high hazard ratio), and 4) appropriateness of data within MSDD for detection of the HOI (i.e., does the MSDD capture the necessary data elements for the right population in the appropriate setting). Additional algorithms for the detection of other HOIs will be developed as new needs and capabilities emerge.

II. PURPOSE

The workgroup set forth to recommend or propose algorithms for the identification of 16 HOIs to FDA. The 16 HOIs of interest included: Achilles tendon rupture, erythema multiforme major (including Stevens-Johnson syndrome and toxic epidermal necrolysis), febrile seizure, Guillain-Barre syndrome, Henoch-Schönlein purpura, hip fracture, idiopathic thrombocytopenia purpura, peripheral neuropathy, pulmonary hypertension, rhabdomyolysis, severe acute liver injury, sudden cardiac death and ventricular arrhythmia, suicide (including attempted suicide), thrombocytopenia, type I diabetes, and valvulopathy. This report was the workgroup's main deliverable.

The focus of this activity was to identify algorithms, validated when possible, to identify selected HOIs relevant to medical products/countermeasures in the MSDD. Algorithms capable of identifying these HOIs within the MSDD will enable assessments of adverse outcomes after exposure to medical products/countermeasures. The workgroup identified or derived algorithms for the HOIs based on extensive review of existing literature and other data. This activity was not a systematic review and results presented within should be interpreted accordingly. Additionally, performing validation studies of these algorithms was beyond the scope of the workgroup's charge. Mini-Sentinel has published a report related to the current report, entitled Alternative Methods for Health Outcomes of Interest Validation that considered whether less resource-intensive approaches to validation might be feasible.[1] Readers interested in alternative methods to validate the subset of the outcomes included in this report (erythema multiforme/Stevens-Johnson syndrome, febrile seizure, Guillain-Barre syndrome, Henoch-Schönlein purpura, hip fracture, idiopathic thrombocytopenia purpura, pulmonary hypertension, rhabdomyolysis, sudden death, suicide, tendinopathies, type I diabetes, and valvulopathy), should see the above-referenced publication available on the Mini-Sentinel website.

III. LIMITATIONS

Many of the algorithms discussed in this report have limitations. Additionally, just because an algorithm is listed in this report does not necessarily mean that it will function well. Some are applicable only to specific populations, while others require a distinction between primary and secondary International Classification of Diseases (ICD-9) coding positions. Determining whether a diagnosis is primary (also denoted as principal or first-listed) or secondary is essential to the function of some HOI identification algorithms included in this report. Primary diagnosis can be difficult to identify consistently for all Data Partners in some electronic healthcare data environments included in the MSDD. Thus, the ability to apply some of the algorithms included in this report may be influenced by the data available in a particular electronic healthcare data environment.

IV. WORKGROUP MEMBERS

The findings of this workgroup resulted from a collaboration between the Center for Pharmacoepidemiology Research and Training at the Perelman School of Medicine of the University at Pennsylvania (as lead site), the University of Iowa, the Mini-Sentinel Operations Center (MSOC) and the FDA. **Table 1** lists workgroup members.

Table 1. Workgroup constituents and roles

Participating Site	Participant	Workgroup Role
University of Pennsylvania	Sean Hennessy, PharmD, PhD	Leader
	Cristin Freeman, MPH, MBE	Lead Evidence Evaluator
	Charles Leonard, PharmD, MSCE	Member
University of Iowa	Ryan Carnahan, PharmD, MS, BCPP	Member
	Elizabeth Chrischilles, MS, PhD	Member
Harvard Pilgrim Health Care Institute	Aarthi Iyer, MPH	MSOC support
	Sunali Goonesekera, SM	MSOC support
FDA	Patrick Archdeacon, MD	Oversight and guidance

V. WORKGROUP METHODS

Major workgroup activities included: a) survey and inventory of Mini-Sentinel Investigators and Data Partners regarding their prior experience in developing algorithms identifying the above-listed outcomes of interest; b) review of existing (e.g., literature-reported) outcome definitions, focusing on studies using U.S. and Canadian data using ICD-9 diagnosis codes; and c) developing recommendations for outcome-defining algorithms while documenting selection rationale and technical specifications.

The survey of Mini-Sentinel Investigators and Data Partners was intended to supplement our planned literature reviews for identifying existing algorithms for HOIs. The workgroup developed a personalized solicitation that briefly described the purpose of our workgroup and requested voluntary provision of published and unpublished algorithms from recipients. This solicitation was emailed by the MSOC. Responses were received by the MSOC and the workgroup Lead Evidence Evaluator and inventoried on the Sentinel Initiative's WebEx WebOffice site.

The Lead Evidence Evaluator then conducted an in-depth review of the literature on the 16 HOIs named in the workgroup opportunity to identify existing algorithms and characterize validation metrics. Principal search strategies included the use of PubMed, Google Scholar, and the review of the reference sections of manuscripts identified via these methods. In several instances, when detail on an algorithm(s) was not fully presented in an identified manuscript, its authors were contacted for supplemental information. This activity was not intended to be a systematic review. Results of the in-depth literature review and synthesis were cataloged and discussions of these findings were presented to the FDA during twice-monthly workgroup teleconferences. In some cases, a workgroup-developed algorithm, derived from an algorithm reported in the literature, was listed as a primary or secondary algorithm.

VI. WORKGROUP FINDINGS

A. SURVEY/INVENTORY OF MINI-SENTINEL INVESTIGATORS AND DATA PARTNERS

Mini-Sentinel Investigators and Data Partners were contacted via email during February 2013. The MSOC received responses from fifteen Mini-Sentinel collaborators. Three investigators/Data Partners responded that they had validation work underway that may or may not be finished in time to be included in this report. Most responses included the provision of published manuscripts describing the validity of specific algorithms and/or compiled code sets (i.e., operational definitions for given outcomes) without an evaluation of their validity.

B. DEVELOPMENT OF ALGORITHMS AND DOCUMENTATION OF SPECIFICATIONS

A literature review was conducted for the following HOIs. For each, one or more observed or derived algorithms are presented along with a rationale. In some instances, the algorithm presented has not been validated, yet the workgroup offers the algorithm as a possible approach to identify the HOI. The presence of an algorithm in this report does not mean that the workgroup knows how this algorithm will perform. Text on each HOI describes the process of review and selection for each primary/secondary observed or derived algorithm. If an algorithm has not been validated or has a PPV less than 75%, this was denoted in the **Appendix Table 1**. A high-level summary of estimated PPVs and sensitivities for primary observed or derived algorithms are presented in **Table 2** below.

Table 2. Estimated positive predictive values (PPVs) and sensitivities of primary observed or derived algorithms for identifying specific HOIs

HOIs	Estimated PPV (%)	Estimated sensitivity (%)
1. Achilles tendon rupture (ATR)	86	unknown[†]
2. Erythema multiforme major (EMM), including Stevens Johnson syndrome (SJS) and toxic epidermal necrolysis (TEN)	unknown	unknown
3. Febrile seizure[*]	unknown	unknown
4. Guillain-Barre syndrome (GBS)	70-88	unknown
5. Henoch-Schönlein purpura (HSP)[††]	unknown	unknown
6. Hip fracture	94	unknown
7. Idiopathic thrombocytopenia purpura (ITP)[††]	unknown	unknown
8. Peripheral neuropathy (PN)[††]	unknown	unknown

HOIs	Estimated PPV (%)	Estimated sensitivity (%)
9. Pulmonary hypertension (PH) [††]	unknown	unknown
10. Rhabdomyolysis[††]	8	76
11. Severe acute liver injury (SALI)	63-100	unknown
12. Sudden cardiac death and ventricular arrhythmia (SCD/VA)	85	unknown
13. Suicide/Attempted suicide	100	unknown
14. Thrombocytopenia	83	unknown
15. Type I diabetes mellitus	97	unknown
16. Valvulopathy	93	41

[†] Unknown can mean: 1) the metric was not included in the published algorithm, 2) the workgroup could not find a validated algorithm, 3) the workgroup chose to recommend an unvalidated- algorithm due to expert input and/or coding changes since the last validation study was performed
* Validation statistics to be released with the PRISM Febrile Seizure report, Fall/Winter 2013
[††] Lack of published validation studies and/or manuscripts using ICD-9 diagnosis codes to describe the HOI

1. Achilles tendon rupture (ATR)

Primary Observed or Derived Algorithm

Any claim type, any position
>= 1 ICD-9 code: 727.67
AND
>= 1 CPT code: 27605, 27606, 27650, 27652, 27654, or 01472

A literature review produced 7 publications regarding algorithms or codes to identify tendon rupture.[2-8]

Initially, the intent of the workgroup was to research validated algorithms for all tendon rupture. After searching the literature, it became apparent that only Achilles tendon rupture (ATR) had been validated using administrative data. Given this, the workgroup focused its review on ATR as opposed to the initial FDA-requested HOI, tendon rupture.

Only one of the seven identified studies reported test characteristics for an ATR algorithm. Three of the studies were conducted using non-US/Canada data,[2-4] two did not use the ICD-9 codes to define their outcome diagnoses[5, 6] and one referred to a validated algorithm,[7] but was not a validation study. The remaining study was published by Seeger et al. in 2006.[8] This study reported a PPV of 91% using a combination of ICD-9 and Current Procedural Terminology (CPT) codes. Yet, this algorithm relied on provider specialty, a field unavailable within the Mini-Sentinel common data model (MSCDM). In order to gather more information on algorithm performance without the specialty information, the workgroup contacted the primary author. That author explained that after first identifying all patients with relevant ATR codes, the investigators made a distinction that required the patient to have both a diagnosis (inpatient, outpatient or emergency department (ED), ICD-9 code 727.67 [*nontraumatic rupture of Achilles tendon*], in any position) and a procedure code (inpatient or outpatient CPT code for Achilles tendon repair (27605 [*tenotomy, percutaneous, Achilles tendon (separate procedure); local anesthesia*], 27606 [*tenotomy, percutaneous, Achilles tendon (separate procedure); general anesthesia*], 27650 [*repair, primary, open or percutaneous, ruptured Achilles tendon*], 27652 [*repair, primary, open or percutaneous, ruptured Achilles tendon; with graft (includes obtaining graft)*], 27654 [*repair, secondary, Achilles tendon, with or without graft*], 01472 [a*nesthesia for procedures on nerves, muscles, tendons,*

and fascia of lower leg, ankle, and foot; repair of ruptured Achilles tendon, with or without graft])). If the answer was "no" (i.e., the patient had a diagnosis or a procedure, but not both), the PPV was 15%. If the answer was "yes" (i.e., the patient had both diagnosis and procedure), the PPV was 86%. The authors subsequently used provider specialty along with site of care to improve the latter PPV to 91%. In a personal communication, the primary author suggested that the small improvement in PPV may not be worth the resources needed to identify provider specialty and site of care, given that many existing databases do not include this information.

In the context of the discussion with the author, it was suggested that if a researcher were interested in using an algorithm that may be more sensitive, they could add CPT codes for medically treated ATR (i.e., walking boot). The author stated that medical treatment for ATR is more common in certain patient groups (i.e., unfavorable surgical candidates, such as older patients), so a broader algorithm that includes non-surgical codes should lead to a more complete capture of the outcome. Based on this suggestion, the workgroup searched for CPT codes for walking boot/ankle casting and were unable to find a well-defined list of codes that would clearly represent medical treatment for ATR. Given this, the workgroup decided not include codes for medically-treated ATR, even as a component of a secondary recommendation. The workgroup recognizes that addressing surgically-treated ATR only is a more narrow definition than if we were to include medically treated ATR; however the workgroup felt as though they could not recommend an un-validated algorithm to describe medically treated ATR in the absence of evidence.

Therefore, the workgroup recommends the validated Seeger et al. algorithm[8] for defining surgically-repaired ATR, which includes an ICD-9 code 727.67 in any claim type, any position and a CPT procedure code for ATR repair (27605, 27606, 27650, 27652, 27654, or 01472).

2. Erythema multiforme major (EMM), including Stevens-Johnson syndrome (SJS) and toxic epidermal necrolysis (TEN)

Primary Observed or Derived Algorithm
Any claim type, any position
>= 1 ICD-9 code: 695.1, 695.10, 695.11, 695.12, 695.13, 695.14, 695.15, or 695.19

A literature review produced 4 publications regarding algorithms or codes to identify erythema multiforme major, Stevens-Johnson syndrome, and/or toxic epidermal necrolysis.[9-12]

Validated algorithms to identify erythema multiforme (EM) exist, but are outdated. In 2009, Mini-Sentinel conducted an HOI evidence review on EM. All studies reported within consistently defined EM and related conditions via ICD-9-CM code 695.1 [*erythema multiforme*]. Prior to 2009, Erythema multiforme major (EMM), Stevens-Johnson syndrome (SJS), and toxic epidermal necrolysis (TEN) were defined by a single four-digit ICD-9 code 695.1, which also included staphylococcus scalded-skin syndrome (SSSS) and erythema multiforme minor. In October 2008, coding changes were made to add fifth-digit, specific ICD-9 codes to distinguish EM (major/minor) from SJS, TEN, and SJS–TEN overlap syndrome; SSSS was eliminated from ICD-9 695.1. In an attempt to make validation metrics interpretable post-coding changes, the Mini-Sentinel HOI evidence review reported the following PPV ranges with and without SSSS diagnoses: 61-66% and 54-60%, respectively.[9, 10]

In addition to the Mini-Sentinel review,[10] a few papers examined ICD-9 codes for EM/SJS/TEN; however, the data used pre-dates changes to the ICD-9 coding structure. One study by Eisenberg et al. reported that 695.1 poorly identified SJS in a large, commercially-insured U.S.population.[11] Knowing that the coding structure they were reporting on was already out of date, the authors performed an ad-hoc review in a more recent year of data (October 2009-November 2010), examining the five-digit ICD-9 code for SJS (695.13 [*Stevens-Johnson syndrome*]). This new code identified 565 patients, which was approximately 9% of all 695.1x codes found in that time. Further, Dreyfus et al. used the new five-digit ICD-9 codes to identify EM, SJS and TEN, but these outcomes were not validated as part of the study.[12]

While the newer five-digit ICD-9 codes have the potential to aid future examination of coding algorithms for EMM and related conditions, they have yet to be validated. Further, apart from skin biopsy, no specific laboratory technique exists to identify these conditions, so there are no additional metrics that the workgroup can propose in an effort to improve the primary algorithm.

For defining the outcome of EM, EMM, SJS and TEN, the workgroup proposes using the newer more specific ICD-9 codes together to define the general HOI and individually to define a specific condition: 695.1, 695.10 [*erythema multiforme, unspecified*], 695.11 [*erythema multiforme minor*], 695.12 [*erythema multiforme major*], 695.13, 695.14 [*Stevens-Johnson syndrome - toxic epidermal necrolysis overlap syndrome*], 695.15 [*toxic epidermal necrolysis*], or 695.19 [*other erythema multiforme*] in both inpatient and outpatient files (any position), acknowledging that these codes have not been validated. The workgroup suggests that researchers consider restricting their search to inpatient files when interested in looking for a subset of serious EMM (695.12)/SJS (695.13)/TEN (695.14 or 695.15).

3. Febrile seizure

Primary Observed or Derived Algorithm
ED or inpatient , any position
>= 1 ICD-9 code: 780.31 or 780.32

A literature review produced 10 publications regarding algorithms or codes to identify febrile seizure.[13-22]

Most of the published seizure literature has not focused on febrile seizure, and is limited to studies conducted in pediatric populations in post-immunization settings. One of the few studies which considered fever in the context of seizure was conducted by Tse et al. and focused on signal identification and risk evaluation of febrile seizures in the Vaccine Safety Datalink (VSD) project.[13] This study validated the algorithm for identifying febrile seizure (780.3 [*convulsions*], 780.31 [*febrile convulsions (simple), unspecified*]), (780.32 [*complex febrile convulsions*]), or 780.39 [*other convulsions*]) in the inpatient or ED setting) and found a PPV of 77%. Although 47 of the 61 charts reviewed had documentation of a seizure and concurrent fever in the specified risk and control windows, the ICD-9 codes used included both general convulsions codes and codes specific to febrile seizure. Performance of the individual codes used in this algorithm was not provided.

Other studies considered by the workgroup are as follows. Klein et al. used VSD data from years 2000-2008. This study examined adverse events in children receiving a combination measles-mumps-rubella-varicella (MMRV) vaccine versus separate MMR and varicella vaccines. Adverse events included febrile seizures identified using a combination of ED and inpatient epilepsy (345*[*epilepsy and recurrent*

seizures]) and seizure (780.3*) ICD-9 codes, which were later chart-reviewed. As a separate outcome, the authors examined ICD-9 codes for the outpatient diagnosis of fever (780.6 [*fever and other physiologic disturbances of temperature regulation*]).[14] The Klein algorithm also made exclusions based on history of seizure codes.[15]

The workgroup reviewed a variety of studies that looked at seizure and epilepsy codes in general, regardless of fever.[16-22] Shui et al.[17] provided additional detail on the frequency of ICD-9 codes 333.2 [*myoclonus*] and 779.0 [*convulsions in newborn*], which had low PPVs and accounted for very few of the overall epilepsy/seizure codes. These codes were not included in the Klein algorithm, so the workgroup was grateful for evidence showing the infrequency of their use. Additionally, MSOC confirmed that the MSCDM does not have a variable to capture body temperature data at this time.

The workgroup consulted with Alison Tse Kawai, ScD, SM, a Mini-Sentinel investigator studying febrile seizures post-influenza vaccination as a Post-Licensure Rapid Immunization Safety Monitoring (PRISM) activity. Dr. Kawai commented that to maximize PPV for this HOI, she would recommend restricting an algorithm to ICD-9 codes 780.31 or 780.32 found in an ED or inpatient setting, in any position. In a PRISM study identifying febrile seizure, the PPV for this algorithm was much greater than the PPV for another algorithm that used all seizure codes (excluding 780.33 [*post-traumatic seizures*]). Only about 8% of confirmed cases in the PRISM study lacked codes for febrile seizures, and the PPV for additional cases captured with more general ICD-9 code to define seizure (780.3 and 780.39 [*other convulsions*]) was extremely low, even if medically-attended fever on the same day in the ED or IP setting was required.

For defining the outcome of febrile seizure, the workgroup proposes the algorithm with the highest PPV from the to-be-published PRISM report on febrile seizure; specifically the definition requires an ED or inpatient ICD-9 code for 780.31 or 780.32 in any position. Validation statistics to support this algorithm will be published in the final report for the activity by spring 2014.

4. Guillain-Barre syndrome (GBS)

Primary Observed or Derived Algorithm
Inpatient, primary position
>= 1 ICD-9 code: 357.0
AND
Outpatient, any position
>= 1 ICD-9 code: 357.0
EXCLUDE
patients with the following in the 365 day baseline:
Any claim type, any position
>= 1 ICD-9 code: 357.0 or 357.81

A literature review produced 9 publications regarding algorithms or codes to identify Guillain-Barre syndrome.[23-31]

The challenge of identifying Guillain-Barre syndrome (GBS) in administrative data can be summarized by the following factors: a non-specific ICD-9 code, coding commonly referring to rule-out diagnoses or sometimes to history of GBS, and the miscoding of chronic inflammatory demyelinating polyneuropathy (CIDP) as GBS. The ICD-9 code commonly used when coding for GBS, 357.0 [a*cute infective polyneuritis*],

has inconsistent test characteristics when used alone. For example, while a PRISM study published by Yih et al. reported that 357.0 used in all care settings had a very low PPV of 12%-16%,[23] other studies examining the same code in all settings reported a PPV closer to 30%.[24, 25] This variation may be attributable to distinct types of data being used in different systems (e.g., claims vs. electronic medical record). Studies examining the code 357.0 in inpatient settings reported PPVs for any-position inpatient diagnoses of 357.0 in the range of 30%-45%.[26-28] When this definition is restricted to primary inpatient diagnosis codes only, PPVs increased to 68%-82%[28, 29] However, the higher PPVs reported by studies restricting their GBS definition to primary inpatient diagnosis codes should be interpreted with caution as they used several different case definitions which may be overly inclusive (confirmed/definite vs. probable/possible). In addition to varying case definitions, some studies had exclusions based on prior diagnoses of GBS, while others required the diagnosis to appear in a specific time frame post-vaccination (i.e., 126 days). A study by Burwen et al. further examined secondary position inpatient codes of 357.0, which validated poorly with a PPV of 20%,[29] supporting the use of this code in the primary position only.

The workgroup also identified GBS validation studies conducted in highly specific populations which may not generalize well to the MSDD population. For example, a study conducted in the active military population[30] required both a primary inpatient and outpatient diagnosis of 357.0 during a post vaccination time frame. Reported PPVs, sensitivities and specificities were high, ranging from 78%-88%, 92%-100%, and 78%-92% respectively. Another located study by Funch et al. used variables such as detailed provider specialty information for outpatient office visits, which is not currently available in the MSDD. A primary inpatient code of 357.0 in combination with an outpatient neurology visit with a 357.0 diagnosis had a PPV of 70%. While these algorithms performed well in specific populations, the unique populations under study, as well as use of variables not included in the MSCDM, suggest they will not generalize well in the MSDD environment.

After discussion of the available algorithms, the workgroup sought further input from a content expert, Dr. Katherine Yih. The workgroup described the literature outlined above and the possible GBS algorithm recommendations (requiring both an inpatient primary diagnosis of 357.0 and an outpatient diagnosis of 357.0; or requiring only a primary inpatient code). Dr. Yih agreed that these recommendations were reasonable, but suggested the workgroup consider expanding the GBS definition and provided an example definition proposed by PRISM for use during the 2012-2013 influenza season. This definition required a primary inpatient code along with some additional baseline exclusions for prior diagnoses of GBS and CIDP. Dr. Yih also described lessons learned from an influenza vaccine safety study conducted as part of the VSD, and stressed the importance of considering the relationship between GBS and CIDP, given that GBS can later be diagnosed as CIDP. Ideally, to identify people with GBS only, a person would be followed for some time after the initial GBS diagnosis to ensure that the condition is not in fact CIDP. Since looking ahead a year is not possible in a near-real-time surveillance activity, a baseline period of one year prior to GBS diagnosis should be examined for other GBS codes or an ICD-9 indicative of CIDP. A VSD study by Greene et al. chose to look back five years, rather than one year for prior GBS or CIDP diagnoses.[31] If codes were present, the (more recent) case was excluded. In the end, the PRISM definition ultimately did not require a principal diagnosis code, because it was not clear whether that data field was consistently or correctly filled out by the Mini-Sentinel/PRISM Data Partners. A chart review was not performed on the cases ascertained using the definition, so the proposed algorithm has not been validated by PRISM.

After synthesizing the evidence from the literature and expert opinion, for defining GBS the workgroup recommends the following as the primary recommendation: at least one principal inpatient ICD-9 code for 357.0 (if/when primary diagnosis is reliably available from Data Partners) followed by at least one outpatient ICD-9 code for 357.0 in any position, excluding persons with a diagnosis of GBS (357.0) or CIDP (357.81) in the 1-365 days prior to the hospitalization of interest, in any setting. As a secondary algorithm, the workgroup recommends one principal inpatient ICD-9 code for 357.0 for researchers who do not have access to outpatient records or who are looking for a simplified definition that does not involved excluding past diagnoses of GBS.

5. Henoch-Schönlein purpura (HSP)

Primary Observed or Derived Algorithm
Inpatient, any position
>= 1 ICD-9 code: 287.0
in patients <18 years

A literature review produced 2 publications regarding algorithms or codes to identify Henoch-Schönlein purpura.[32, 33]

Few studies were found in the literature pertaining to coding of Henoch-Schönlein purpura (HSP) in administrative data. The workgroup reviewed a study by Weiss et al. that included a cohort of children discharged with HSP between 2000 and 2007.[32] To be included in the study, the authors required a primary or secondary discharge diagnosis with ICD-9 code 287.0 [*allergic purpura*]. Authors did not present test characteristics with their algorithm. The workgroup also reviewed the PRISM Gap Report which designates HSP as only having 1-2 eligible studies and not fit for individual review.[33]

For defining the outcome of HSP, the workgroup proposes a primary algorithm including at least one any-position inpatient discharge diagnosis of 287.0 in patients <18 years of age, as used by Weiss et al.,[32] with the caveat that this algorithm has not been validated.

6. Hip fracture

Primary Observed or Derived Algorithm
Inpatient, primary position
>= 1 ICD-9 code: 820.00, 820.01, 820.02, 820.03, 820.09, 820.20, 820.21, or 820.8

A literature review produced 11 publications regarding algorithms or codes to identify hip fracture.[34-44] Published articles reviewed but not discussed in this report were ultimately excluded based relevance to the MSDD, coding source, country of origin, etc.[40-44]

There is considerable evidence available to support the use of ICD-9 codes to identify hip fracture in the inpatient administrative data.[34] Approximately 95% of hip fractures can be identified with inpatient data, and supplementing inpatient data with outpatient claims will allow for identification of an additional 5 %.[35] As there were numerous hip fracture studies available in the literature, the workgroup focused its review on the application of available hip fracture identification algorithms in the MSDD.

A 2013 systematic review by Hudson et al. examined the validity of diagnosis and procedure codes to identify hip fracture in administrative data.[34] This review reports PPVs and sensitivities of inpatient

claims for hip fracture ranging from 63%-96% and 69%-97% respectively, which increase to 86%-98%, and 83%-97% with the addition of physician claims diagnoses and procedure codes. Most studies in the review uniformly define hip fracture as an inpatient, primary or secondary, diagnosis of ICD-9 code 820.x [*fracture of neck of femur*], with some algorithms including 821.x [*fracture of other and unspecified parts of femur*] and others requiring CPT codes to identify procedures. Studies varied on whether or not they specified fourth and fifth digit ICD-9 codes.

Not included in the systematic review was a study by Narongroeknawin et al. which was published after the literature review was performed by Hudson et al. Narongroeknawin et al. performed a validation of ICD-9 codes from a large U.S. health system that are used to evaluate femoral fractures, in three groups: 1) subtrochanteric (820.22 [*fracture of subtrochanteric section of femur, closed*]), 2) diaphyseal (821.00 [*fracture of unspecified part of femur, closed*], 821.01 [*fracture of shaft of femur, closed*]), and 3) typical hip fractures (820.00 [*fracture of unspecified intracapsular section of neck of femur, closed*], 820.01[*fracture of epiphysis (separation) (upper) of neck of femur, closed*], 820.02[*fracture of midcervical section of femur, closed*], 820.03[*fracture of base of neck of femur, closed*], 820.09[*other transcervical fracture of femur, closed*], 820.20[*fracture of unspecified trochanteric section of femur, closed*], 820.21[*fracture of intertrochanteric section of femur, closed*], and 820.8 [*fracture of unspecified part of neck of femur, closed*]).[36] Authors intentionally excluded hospitalizations where codes were present for open and distal end femoral fractures (820.1x [*transcervical fracture, open*], 820.3x [*pertrochanteric fracture of femur, open*], 820.9 [*fracture of unspecified part of neck of femur, open*], 821.1x [*fracture of shaft or unspecified part of femur, open*], 821.2x [*fracture of lower end of femur, closed*], and 821.3x [*fracture of lower end of femur, open*]), which are often attributable to major trauma. The authors also included a large table of algorithms with varying definitions; however, given that a provider specialty variable is not currently included in the MSCDM, algorithms that required a surgeon's diagnosis of hip fracture were deemed irrelevant. For subtrochanteric, diaphyseal, and typical hip fractures algorithms that required one inpatient discharge diagnosis (any of the ICD-9 codes listed above) in any position, the PPVs reported were 69%, 89% and 85% respectively. When the definitions were restricted to include only primary listed inpatient discharge diagnoses, the PPVs for all three types of femoral fracture increased to 74%, 94% and 94%. Authors also reported algorithms that required a primary discharge diagnosis code and the same diagnosis code on an outpatient physician claim within ninety days of discharge, which resulted in a higher PPV for subtrochanteric fractures (86% vs. 74% for primary inpatient discharge diagnosis alone), but did not significantly change the PPV for diaphyseal (96% vs. 94%) and typical hip fractures (95% vs. 94%). Given the programmatic efforts needed to linked inpatient and outpatient claims, the inclusion of this second algorithm will be left to the discretion of the researcher. Hospitalizations with discharge diagnoses containing ICD-9 E-codes for major trauma (E800-E848, E881-E884, E908-E909, and E916-E928) were excluded in a sub-analysis, with PPVs not substantially different than those reported in the primary analyses. Administrative data was not successful in identifying atypical femoral fractures.

In a fracture validation study by Ray et al., authors demonstrated a complicated algorithm including ICD-9 diagnosis and procedure codes, CPT codes for fracture identification and confirmation, as well as CPT codes for imaging.[35] This algorithm relied on the manual assignment of a fracture site, based on claims found in a seven day window from the index date (first fracture claim). Although the authors provided a detailed list of codes, their methods are difficult to reproduce. In the discussion, authors offered an alternate simplified definition where each inpatient and ED ICD-9 diagnosis was counted as a fracture.

The PPV reported for the simplified definition hip fracture (820.x) is 90%, which is comparable to the any-listed inpatient algorithm described by Narongroeknawin et al. (PPV of 85%).

Also considered was a study by Jean et al. which reported a high PPV and sensitivity (93% and 95% respectively), for an algorithm to identify incident hip fracture in a Quebec physician claims database.[37] This algorithm was ultimately deemed not relevant to the MSDD due to the lack of provider specialty information. A study by Tamblyn et al. reported sensitivities of outpatient physician claims by diagnostic, procedure or diagnostic and procedure coding algorithms (83%, 94% and 97%).[38] Authors noted that the sensitivities varied depending on whether the exact procedure date or date window was used. Reported sensitivities were much lower when an exact procedure date was examined. Virnig et al. examined inpatient diagnosis of ICD-9 code 820.x in women aged 55-69 years of age receiving Medicare.[39]

For defining hip fracture, the workgroup's primary recommendation is derived from the definition of typical hip fracture described by Narongroeknawin et al.: a primary inpatient ICD-9 code for typical hip fracture (820.00, 820.01, 820.02, 820.03, 820.09, 820.20, 820.21, or 820.8). This algorithm also performs well when the definition is expanded to include both primary and secondary (any position) inpatient diagnosis codes; therefore, the workgroup's secondary algorithm includes: at least one any-position inpatient ICD-9 code for typical hip fracture (820.00, 820.01, 820.02, 820.03, 820.09, 820.20, 820.21, or 820.8), for use when information distinguishing primary diagnosis is not available.

7. Idiopathic thrombocytopenia purpura (ITP)

Primary Observed or Derived Algorithm
Any claim type, any position
>= 1 ICD-9 code: 287.31

A literature review produced 6 publications regarding algorithms or codes to identify idiopathic thrombocytopenia purpura.[45-50]

Up until 2006, ICD-9 coding for idiopathic thrombocytopenia purpura (ITP) was limited to a general, four-digit code for thrombocytopenia (ICD-9 287.3 [*primary thrombocytopenia*]). Most of the published validation studies for ITP evaluated this ICD-9 code in data derived from different patient care settings.[45-48] These settings included a combination of adult and/or pediatric patients and inpatient and/or outpatient records and studies reported PPVs ranging from 65%-83%, with use of both pediatric and inpatient claims achieving the highest performance. The expansion of 287.3 in 2006 included the addition of a fifth digit to designate specific conditions. As a result of this update, 287.31 [*immune thrombocytopenic purpura*] -- a five-digit code specific for immune thrombocytopenic purpura, now exists. A study by Chiao et al. used 287.31 to define a population of ITP patients but did not validate their algorithm.[49]

The literature review identified one study validating this five-digit code, yet it did not use U.S. or Canadian data.[50] Galdarossa et al. examined ICD-9 287.31 in the inpatient setting and reported a high PPV, NPV, sensitivity and specificity of 80%, 88%, 80%, and 88% respectively.[50] This is the only published study validating the new ITP-specific code. After much discussion, the workgroup decided to expand Galdarossa's definition to the outpatient setting as part of a proposed composite algorithm, given the 287.3 code had validated similarly to the inpatient codes, prior to the fifth digit expansion.

In additional to the published literature, an algorithm developed by FDA and HealthCore to identify ITP was provided to the workgroup through a personal communication. At the time of this communication, this algorithm had not yet been applied or validated, but will be used to by FDA to examine the association between seasonal flu vaccine and ITP.

Based on extensive workgroup discussion, for defining the outcome of ITP, the primary proposed algorithm is as follows: an inpatient or outpatient ICD-9 code 287.31, in any position. The workgroup also proposes two secondary algorithms: 1) ICD-9 287.30 [*primary thrombocytopenia, unspecified*], 287.31, 287.32 [*Evans' syndrome*], or 287.39 [*other primary thrombocytopenia*] in any position on an inpatient or outpatient claim, to be used when looking to describe a broader HOI; 2) ICD-9 code 287.31 and one of the following complete blood count (CBC) or blood smear CPT codes on the same day or one day prior: 85025 [b*lood count; complete (CBC), automated and automated differential WBC count*], 85027 [b*lood count; complete (CBC)*], 85032 [*blood count; manual cell count*], 85049 [*blood count; platelet, automated*], 85060 [*blood smear, peripheral, interpretation by physician with written report*], or 85097 [*bone marrow, smear interpretation*]—a non-validated algorithm provided as the result of a HealthCore-FDA collaboration, to be used when laboratory data is available. In general, the workgroup did not typically consider performance metrics reported for algorithms developed in non-U.S./Canadian databases. Given that the paper by Galdarossa et al. is the only validation study of the new codes, the workgroup considered the algorithm in developing the proposed composite algorithm; however one should use this algorithm with caution, as performance metrics reported in non-US databases may not reliably predict performance in the MSDD.

8. Peripheral neuropathy (PN)

Proposed Composite Algorithm

Definition #1:
Any claim type, any position
>= 1 ICD-9 code (broad): 250.6, 250.60, 250.61, 250.62, 250.63, 337.0, 337.00, 337.01, 337.09, 337.1, 337.9, 349.9, 356.4, 356.8, 356.9, 357, 357.0, 357.1, 357.2, 357.3, 357.4, 357.5, 357.6, 357.7, 357.8, 357.81, 357.82, 357.89, 357.9, or 729.2
AND
>= PN 1 drug

Definition #2:
Any claim type, any position
>= 1 ICD-9 code (broad): 250.6, 250.60, 250.61, 250.62, 250.63, 337.0, 337.00, 337.01, 337.09, 337.1, 337.9, 349.9, 356.4, 356.8, 356.9, 357, 357.0, 357.1, 357.2, 357.3, 357.4, 357.5, 357.6, 357.7, 357.8, 357.81, 357.82, 357.89, 357.9, or 729.2
without mention of a PN drug

Definition #3:
Any claim type, any position
>= 1 ICD-9 code (narrow): 356.4, 356.8, 356.9, 357, 357.0, 357.1, 357.6, 357.7, 357.8, 357.81, 357.89, 357.9, or 729.2
AND
>= 1 PN drug

Definition #4:
Any claim type, any position

>= 1 ICD-9 code (narrow): 356.4, 356.8, 356.9, 357, 357.0, 357.1, 357.6, 357.7, 357.8, 357.81, 357.89, 357.9, or 729.2
without mention of a PN drug

A literature review produced 2 publications regarding algorithms or codes to identify peripheral neuropathy.[51, 52]

Since peripheral neuropathy (PN) can be defined in many ways, there is no one study that perfectly describes this outcome and no study in which all possible codes have been validated. The clinical complexity PN of is important to note, and very relevant when attempting to define this condition in administrative data. PN can be clinically defined as either be inherited or acquired, uni- or bi-lateral and focused or extensive. Drug-induced PN would likely present as symmetrical and distal in its features.[53] Hartfield et al. examined both a long-detailed list of ICD-9 diagnosis codes and pharmacy data to validate codes that describe painful diabetic neuropathy.[51] Their neuropathy identification algorithm also required exclusions of nondiabetic etiologies. In an effort to identify a full list of specific neuropathic disorder codes, the workgroup reviewed a paper by Berger et al. which does not validate ICD-9 codes, but provides a comprehensive list of painful neuropathic disorders.[52]

Because of the lack of validation studies on PN, the workgroup developed four different approaches to identifying PN, which used a combination of diagnosis codes and drugs prescribed to treat the symptoms of PN. The broad categories were developed to include codes for widespread manifestation of PN symptoms, regardless of the suggested source or cause. The narrow list excluded codes that pertain to diabetic PN, autonomic disease and any codes that are suggestive of a specific source not considered to be used for drug-induced PN. Any ICD-9 codes that had a definition which indicated a mononeuropathy or a specific site were removed. ICD-9 codes for cancer and HIV related neuropathies were reviewed and excluded. A list of prescriptions drugs used to treat PN was developed by combining treatment recommendations[54] and workgroup expertise.

The workgroup proposes the following four definitions of PN (in no particular order), intended to be used at the discretion of the researcher (i.e., each individual study will need to determine which file types and coding positions to include): 1) broad diagnosis code for PN (see **Table 3**) + PN treatment drug (see **Table 4**), 2) broad diagnosis code for PN without a PN treatment drug, 3) narrow diagnosis code for PN (see **Table 5**) + PN treatment drug, or 4) narrow diagnosis code for PN without a PN treatment drug. Where broad is defined as codes inclusive of all PN codes and narrow describes a list of codes that excludes certain conditions, including diabetic PN. Researchers who chose to use an algorithm requiring the presence of a PN drug will need to develop their own code set of relevant National Drug Codes.

Table 3. BROAD list of ICD-9 codes to define PN (wide-spread disseminated, regardless of suggested source/cause)

ICD-9 Code	Definition
250.6	*diabetes with neurological manifestations*
250.60	*diabetes mellitus with neurological manifestations, type two or unspecified type, not stated as uncontrolled*
250.61	*diabetes mellitus with neurological manifestations, type one [juvenile type] not stated as uncontrolled*
250.62	*diabetes mellitus with neurological manifestations, type two or unspecified type, uncontrolled*
250.63	*diabetes mellitus with neurological manifestations, type one [juvenile type] uncontrolled*
337.0	*idiopathic peripheral autonomic neuropathy*
337.00	*idiopathic peripheral autonomic neuropathy, unspecified*
337.01	*carotid sinus syndrome*
337.09	*other idiopathic peripheral autonomic neuropathy*
337.1	*peripheral autonomic neuropathy in disorders classified elsewhere*
337.9	*unspecified disorder of autonomic nervous system*
349.9	*unspecified disorders of nervous system*
356.4	*idiopathic progressive polyneuropathy*
356.8	*other specified idiopathic peripheral neuropathy*
356.9	*unspecified idiopathic peripheral neuropathy*
357	*inflammatory and toxic neuropathy*
357.0	*acute infective polyneuritis*
357.1	*polyneuropathy in collagen vascular disease*
357.2	*polyneuropathy in diabetes*
357.3	*polyneuropathy in malignant disease*
357.4	*polyneuropathy in other diseases classified elsewhere*
357.5	*alcoholic polyneuropathy*
357.6	*polyneuropathy due to drugs*
357.7	*polyneuropathy due to other toxic agents*
357.8	*other inflammatory and toxic neuropathies*
357.81	*chronic inflammatory demyelinating polyneuritis*
357.82	*critical illness polyneuropathy acute motor neuropathy*
357.89	*other inflammatory and toxic neuropathy*
357.9	*unspecified inflammatory and toxic neuropathies*
729.2	*neuralgia, neuritis, and radiculitis, unspecified*

Table 4. Drugs used to treat symptoms and prevent complications of PN.

Drug Class	Active Ingredient	Neuropathy Related Indication[55]
Anticonvulsants	gabapentin	Postherpetic neuralgia, off-label diabetic neuropathy
	pregabalin	Off-label neuralgia/neuropathy/chronic pain
	valproate	Off-label postherpetic neuralgia
	carbamazepine	Trigeminal neuralgia
	phenytoin	No neuropathy related diagnoses listed
	topiramate	Off-label, neuralgias/neuropathy/pain
Antidepressant drugs		
Tricyclic Antidepressants	amitriptyline	Off-label neuropathic pain
	desipramine	Off-label diabetic neuropathy
Serotonin and Norepinephrine Reuptake Inhibitors	nortriptyline	Off-label postherpetic neuralgia
	duloxetine	Diabetic peripheral neuropathic pain
	venlafaxine	Off-label diabetic neuropathy
Skeletal Muscle Relaxants	baclofen	Other possible off-label use neuropathic pain
Antiarrhythmic Agents	mexiletine	Off-label diabetic neuropathy, neuropathy (nondiabetic)

Table 5. NARROW list of ICD-9 codes to define PN (excluding diabetic PN, autonomic disease and any codes that are suggestive of a specific source not considered to be used for drug-induced PN)

ICD-9 Code	Definition
356.4	*idiopathic progressive polyneuropathy*
356.8	*other specified idiopathic peripheral neuropathy*
356.9	*unspecified idiopathic peripheral neuropathy*
357	*inflammatory and toxic neuropathy*
357.0	*acute infective polyneuritis*
357.1	*polyneuropathy in collagen vascular disease*
357.6	*polyneuropathy due to drugs*
357.7	*polyneuropathy due to other toxic agents*
357.8	*other inflammatory and toxic neuropathies*
357.81	*chronic inflammatory demyelinating polyneuritis*
357.89	*other inflammatory and toxic neuropathy*
357.9	*unspecified inflammatory and toxic neuropathies*
729.2	*neuralgia, neuritis, and radiculitis, unspecified*

9. Pulmonary hypertension (PH)

Primary Observed or Derived Algorithm
Any claim type, any position
>= 1 ICD-9 code: 416.0, 416.8, or 416.9

A literature review produced 9 publications regarding algorithms or codes to identify pulmonary hypertension.[56-64]

While the workgroup's literature review was unable to locate any studies which validated ICD-9 codes for pulmonary hypertension (PH), it located many studies which proposed ICD-based PH identification algorithms. Studies varied whether they used the 416.0 [*primary pulmonary hypertension*], 416.8,

and/or 416.9 [*Chronic pulmonary heart disease, unspecified*] codes to define patients with PH,[56-61] with a study by Lowe et al. using all three four-digit codes.[62] In one instance, the ICD-9 code 416.8 [*other chronic pulmonary heart diseases*] was used as part of a larger algorithm to define patients at low risk for a pulmonary embolism, but the codes were not individually validated, so authors did not report test characteristics relevant to this HOI.[63] Additionally, the Centers for Disease Control and Prevention (CDC) published a surveillance report which used all three four-digit codes to assess trends over time in deaths and hospitalizations attributable to PH.[64]
After discussion with the workgroup, members from the University of Pennsylvania consulted with internal experts to learn to how these codes are used in practice for patients with suspected drug-induced PH. Experts agreed, that in the absence of evidence, all three four-digit ICD-9 codes should be included in the algorithm.

For the defining the outcome of PH, the workgroup proposes an inpatient or outpatient diagnosis of ICD-9 code 416.0, 416.8, 416.9 in any position, with the caveat that these codes have not been validated.

10. Rhabdomyolysis

Primary Observed or Derived Algorithm
Any claim type, any position
>= 1 ICD-9 code: 728.88

A literature review produced 4 publications regarding algorithms or codes to identify rhabdomyolysis.[65-68]

Rhabdomyolysis had a general, non-specific four-digit ICD-9 code which was updated to a more specific a five-digit ICD-9 code (728.88 [*rhabdomyolysis*]) in October 2003. A HealthCore study by Cziraky et al. used this new code to identify rhabdomyolysis, but authors did not validate their algorithm. In the paper the authors cited a paper by Andrade et al. (described below) which published PPVs for the outdated ICD-9 codes.[65] Another study by Setoguchi et al. used Andrade's algorithm for rhabdomyolysis up until October 2003, and after, defined rhabdomyolysis as a hospitalization with a primary or secondary diagnosis of ICD-9 728.88. Authors used Medicare data (with some Medicaid for dual-eligibles) from 1999-2005. The authors did not validate their algorithm.[66]

Floyd et al. performed a validation of the new code (728.88) for rhabdomyolysis in a cohort of statin users, but very poor performance of this code was observed.[67] PPVs reported for this study were 7.5%, 7.2% and 9.5% for all settings, inpatient, and outpatient claims respectively. Reported sensitivity was 76%. The authors concluded that the use of the new administrative diagnosis code for rhabdomyolysis (728.88) was highly nonspecific for statin-related rhabdomyolysis. It is also important to note that the validation studies by Andrade et al. and Floyd et al. were conducted in patients exposed to at least one statin.

Given the low PPVs found in the one validation study relevant to the new 728.88 code, the workgroup revisited validation work performed on the older, more general codes used to describe rhabdomyolysis. In the only study published that presented PPVs, Andrade et al. examined cohorts of new statin or fibrate users in U.S. managed care organizations for the purposes of screening.[68] As such, criteria listed were not fully fleshed-out algorithms. The authors of this study offer five proposed criteria to identify

serious cases of myopathy and rhabdomyolysis, including hospital discharge diagnoses for muscle disorders and laboratory value claims within a week of the hospitalization. Individual components included in each of the five criteria were validated. Three individual ICD-9 codes (791.3 [*myoglobinuria*], 728.89 [*other disorders of muscle, ligament, and fascia*], 729.1 [*myalgia and myositis, unspecified*]) included had high PPVs (100%, 80%, 100%); however, authors reported very few cases available for review (n=1, n=4, n=1 respectively). Although these PPVs seem promising, when the workgroup examined the proposed criteria overall, PPVs were not reported. Thus, the workgroup calculated a PPV for the criterion deemed most applicable to the MSDD, which resulted in a PPV of 32% (based on individual PPVs provided by the authors). Ultimately, Andrade et al. concluded that, while their algorithm may be useful for signal detection, generalizability may be limited by the included study population (i.e., all study participants were users of lipid-lowering drugs).

In light of the evidence reviewed for defining the outcome of rhabdomyolysis, the workgroup proposes the following observed algorithm: an inpatient or outpatient diagnosis of ICD-9 code 728.88 in any position, with the caveat that this code is not well studied and validated poorly in the single validation study. The workgroup is anticipating a potential for a shift in how rhabdomyolysis is coded, due to the addition of the five-digit code specific to the condition and holds the expectation that the new, more-specific code will better define rhabdomyolysis in the ICD-9 coding structure. Additionally, the workgroup would recommend the use of laboratory data when specific test result variables become available in the MSCDM.

11. Severe acute liver injury (SALI)

Primary Observed or Derived Algorithm
Inpatient, any position
>= 2 ICD-9 codes: 573.3, 573.8, 570, 572.2, 572.4, 572.8, or V42.7
OR
Inpatient, any position
>= 1 ICD-9 code: 573.3, 573.8, 570, 572.2, 572.4, 572.8, or V42.7
AND
>= 1 ICD-9 procedure code: 50.1x, 50.9x OR
>= 1 CPT code: 47000, 47001, or 47100

A literature review produced 3 publications regarding algorithms or codes to identify severe acute liver injury.[69-71]

In 2011, the Observational Medical Outcomes Partnership (OMOP) published a presentation focusing on algorithms to identify acute liver injury (ALI).[69] OMOP concluded that their systematic review for ALI did not identify any algorithms with a high PPV. Based on codes identified in ALI validation studies, OMOP developed a series of composite ALI definitions. As a result of this work, a validation study was conducted to examine four possible ALI definitions, three of which were found to have PPVs <10%. The remaining definition, based solely on laboratory data, had a PPV = 51%.[70]

In addition to OMOP efforts, a Mini-Sentinel workgroup was assembled to examine ICD-9 diagnosis codes for identifying severe acute liver injury (SALI) in automated healthcare claims data. In 2013, this workgroup published a study examining the validity of diagnosis codes to identify severe acute liver injury (SALI) in the MSDD.[71] ICD-9 codes for hospitalized SALI generally yielded low PPVs (6.5%-54.3%) in the MSDD. Mini-Sentinel investigator and lead author of the SALI validation work above joined in the

workgroup discussion on validating this HOI; Dr. Lo Re explained that select combinations of codes indicative of SALI had high PPVs (60.0%-100.0%) for confirmed outcomes, although researchers must acknowledge that these algorithms will miss events. Additionally, the Mini-Sentinel workgroup on SALI decided there may be a difference in the accuracy of codes based on the prior medical history involving chronic liver/biliary disease. Therefore, the validation project was split into two cohorts with and without previously diagnosed chronic liver disease. The coding for each of these cohorts is included in the master algorithm database.

After discussing the multiple SALI code combinations presented in the validation paper, the workgroup proposes the following algorithm: any 2 inpatient SALI ICD-9 codes (573.3 [*hepatitis, unspecified. toxic hepatitis*], 573.8 [*other specified disorders of liver*], 570 [*acute and subacute necrosis of liver*], 572.2 [*hepatic encephalopathy*], 572.4 [*hepatorenal syndrome*], 572.8 [*other sequelae of chronic liver disease*], or V42.7 [*liver replaced by transplant*] in any position (PPV = 63.2%) or an inpatient SALI code (573.3, 573.8, 570, 572.2, 572.4, 572.8, or V42.7 with a procedure code indicative of a liver biopsy (ICD-9 procedure codes: 50.1x [*diagnostic procedures on liver*], or 50.9x [*other operations on liver*]; or CPT codes: 47000 [*biopsy of liver, needle; percutaneous*], 47001 [*biopsy of liver, needle; when done for indicated purpose at time of other major procedure*], 47100 [*biopsy of liver, wedge*]) (PPV = 75%-100%), for defining the outcome of SALI. In addition to these coding requirements, any patient with an inpatient or outpatient ICD-9 diagnosis code of a liver/biliary disease within the 12 months prior to the SALI diagnosis should be excluded. This is in part a balance of the highest PPVs and an effort to reduce the programming burden. Future efforts should include the incorporation of laboratory values. See the Mini-Sentinel publication[71] for more discussion on strengths and limitations, as well as recommended laboratory values to be used should they become available in the MSDD.

12. Sudden cardiac death (SCD) and ventricular arrhythmia (VA)

Primary Observed or Derived Algorithm
ED or Inpatient, primary/first-listed position
>= 1 ICD-9 code: 427.5, 798, 798.1, 798.2, 427.1, 427.4, 427.41, or 427.42
EXCLUDE
patients with a prior SCD/VA diagnosis:
Any claim type, any position
>= 1 ICD-9 code: 427.5, 798, 798.1, 798.2, 427.1, 427.4, 427.41, or 427.42

A literature review produced 13 publications regarding algorithms or codes to identify sudden cardiac death.[72-84]

Sudden cardiac death (SCD) is difficult to define in administrative data due to many factors attributable to the nature of the condition (i.e., rapid onset, prior symptoms often absent, person often dies in an outpatient setting). Given these difficulties, many researchers have examined the validity of death certificates for classification of SCD.[72-74] However, with the exception of one report by Chung et al.,[72] most of studies examining validity of death certificates for classification of SCD reported low PPVs (PPV = 19-32%).[75]

Chung et al. performed a validation study with death record data for Tennessee Medicaid enrollees, examining SCD occurring between 1990 and 1993.[72] The study was limited to patients aged 15–84 years,

who were not in a long-term facility and did not have a life-threatening illness. Authors linked their Medicaid database with death certificate files and a state hospital discharge database.
Hennessy et al. developed a composite algorithm to exam outpatient-originating SCD/ventricular arrhythmia (VA),[75] using data from 1999-2002 for 5 large state Medicaid programs. Authors chose to include VA in their definition, as SCD is often attributable to VA, and this SCD/VA event did not necessarily have to result in death.[76] A validation study specifically examining VA found a sensitivity and specificity of 77% and 94% respectively for certain VA ICD-9 codes, giving the inclusion of these codes in the Hennessy algorithm some validity.[76] Hennessy et al. presented PPVs for the overall algorithm, and for the individual component codes. The PPV reported for the overall composite SCD/VA algorithm was 85%. Authors included inpatient and ED claims, as many cases of SCD or VA are treated in the ED and may not result in a hospital admission (e.g., if the patient dies before admission). In addition, documentation of an outpatient witnessed sudden collapse, or a person found dead or unconscious in the field with evidence that the individual had been alive in the prior 24 hours, met the validation criterion. VA diagnoses predominated in hospitalization claims, while SCD diagnoses predominated within ED claims.

Both Hennessy and Chung focused their validation on outpatient-originating SCD (/VA). Hennessy et al. accomplished this by restricting ICD-9 codes to a primary position, as inclusion of secondary codes often identified persons with inpatient episodes of non-sustained ventricular tachycardia.[77] Chung et al. did so by excluding cases in which the patient died in a hospital or nursing home or was successfully resuscitated. There are three main differences between the validation studies presented above. The study by Chung et al. used death record cause of death, which may not be available to all researchers, and is not currently included in the MSCDM. In contrast, Hennessy et al. used an ICD-9 based SCD algorithm in inpatient and ED claims data. Additionally, Hennessy et al. used a broader population (only making exclusions based on prior SCD diagnoses) and encounter type (allowing ED visits). The Chung study used a restricted population which excluded: patients with serious non-cardiac illness, nursing home residents, and those older than age 75; thus their algorithm may not be applicable to individuals in the excluded groups. Lastly, Chung et al. used a more restrictive case definition 'a sudden pulseless condition (arrest) that was immediately fatal, where Hennessy et al. allowed for patients to survive their SCD event. Since the Hennessy paper did not exclude events in patients with serious non-cardiac illness, a commentary reviewing these two papers suggested the PPV in ED records may be inflated by those with terminal illness arriving at the ED in cardiac arrest.[78] Authors from the Hennessy paper conducted a *post hoc* sensitivity analysis which excluded seriously ill patients, and PPVs were similar to their original study results.[79]

The validity of the algorithm developed by Chung et al. was duplicated in a study by the same research group (Kawai et al.), which reported a PPV of 88%.[80] Of the 140 sampled deaths, 81 were adjudicated; 73 (90%) were sudden cardiac deaths. After removing two deaths possibly attributable to opioid overdose, the PPV was 88%.

It is important to note that validation results from the study by Hennessy et al. were unable to be duplicated when the algorithm was applied to a pediatric population, in a study done by the same research group. Schelleman et al. reported a PPV of 41% for the above listed algorithm by Hennessy et al., in subjects aged 3 to 17 years,[81] presumably because of the low frequency of this event in children. Given this, the algorithm's validity may be limited to an adult population.

The workgroup considered other papers by Ray et al. and Staffa et al., which did not report specific codes for their algorithms and had variable performance, reporting PPVs of 47% and 73% respectively.[82,83] Mini-Sentinel performed a HOI review on cardiac arrhythmias which concluded algorithms that included VA and SCD codes performed better than VA codes alone.[84]

For defining the outcome of SCD, the workgroup proposes the SCD and VA composite algorithm by Hennessy et al., which requires a primary inpatient or first-listed ED ICD-9 code of 427.5 [*cardiac arrest*], 798 [*Sudden death, cause unknown*], 798.1 [*instantaneous death*], 798.2 [*Death occurring in less than 24 hours from onset of symptoms, not otherwise explained*], 427.1 [*paroxysmal ventricular tachycardia*], 427.4 [*ventricular fibrillation and flutter*], 427.41 [*ventricular fibrillation*], or 427.42 [*ventricular flutter*], in patients >17 years old. Given the limited access to cause of death information at this time, the workgroup does not propose a secondary algorithm.

13. Suicide, including attempted suicide

Primary Observed or Derived Algorithm
Inpatient or outpatient, any position
>= 1 ICD-9 E-code: E950, E951, E952, E953, E954, E955, E956, E957, E958
90 days before or 180 days after initial antidepressant prescription or psychotherapy visit related to a depression code

A literature review produced 8 publications regarding algorithms or codes to identify suicide, including attempted suicide.[85-92]

In 2012, Mini-Sentinel investigators published a review paper of the suicide literature.[85] The results of the review revealed that E-codes[a] for intentional self-injury are most often used in the suicide literature and authors report a widely variant PPV range, 4%-100%.

Validation manuscripts related to suicide can generally be split into one of two subject categories, 1) completed or 2) attempted suicide. The literature review resulted in two studies that attempted to validate completed suicide in U.S. databases. Both available studies were validated using death registry information. The first, a study by Shevchenko et al., examined discharges from thirty-five Connecticut acute care hospitals.[86] The authors included all hospital stays that ended in death with a discharge diagnosis of E950 [*suicide and self-inflicted poisoning by solid or liquid substances*], E951 [*suicide and self-inflicted poisoning by gases in domestic use*], E952 [*suicide and self-inflicted poisoning by other gases and vapors*], E953 [*suicide and self-inflicted injury by hanging, strangulation, and suffocation*], E954 [*suicide and self-inflicted injury by submersion (drowning)*], E955 [*suicide and self-inflicted injury by firearms, air guns, and explosives*], E956 [*suicide and self-inflicted injury by cutting and piercing instrument*], E957 [*suicide and self-inflicted injuries by jumping from high place*], E958 [*suicide and self-inflicted injury by other and unspecified means*], E959 [*late effects of self-inflicted injury*] from uniform hospital discharge data sets. The PPV and sensitivity reported were moderate, 60% and 65% respectively. Charts were validated by linking a hospital discharge database with the Connecticut death registry. A second validation study conducted in the South Carolina Violent Death Reporting System,

[a] As part of the ICD-9 coding hierarchy, E-codes provide information on supplementary classification of external causes of injury and poisoning.

looked for either ED or inpatient visits dated on the day of the confirmed suicide. This study, by Weis et al. used suicide-related claims (E950-E959) from statewide hospital and ED billing records during one year to identify suicide captured in a claims database. The authors reported a sensitivity of 13.8% for confirmed suicides that have a hospitalization or ED visit within one day of death. When authors included suicide-related events for the year prior to death, the sensitivity only marginally improved (14.3%).[87] The workgroup determined that death information and access to death certificates are needed to assess each death reported as a completed suicide. This is beyond the current capability of Mini-Sentinel.

Inpatient discharge E-codes for self-harm (E950-E959) excluding E-codes for assault or undetermined intention were well validated (PPVs ranged 70%-100%). If codes for undetermined intent (E980-E987) were added to the algorithm, it becomes more difficult to differentiate between self-harm and suicidal intent. Simon et al. studied a cohort of health maintenance organization outpatients starting an antidepressant for treatment of their depression.[88] PPVs reported in this study are high, ranging from 70%-100% depending on which codes are included. The PPV for medical record-verified intentional self-harm and suicidal intent among patients with intentional self-harm codes (E950-E958) were both 100%. PPVs for intentional self-harm and suicidal intent dropped to 80% and 70% respectively when codes for undetermined intent were added to the algorithm (E980 [*poisoning by solid or liquid substances, undetermined whether accidentally or purposely inflicted*], E981 [*poisoning by gases in domestic use, undetermined whether accidentally or purposely inflicted*], E982 [*poisoning by other gases, undetermined whether accidentally or purposely inflicted*], E983 [*hanging, strangulation, or suffocation, undetermined whether accidentally or purposely inflicted*], E984 [*submersion (drowning), undetermined whether accidentally or purposely inflicted*], E985 [*injury by firearms, air guns, and explosives, undetermined whether accidentally or purposefully inflicted*], E986 [*injury by cutting and piercing instruments, undetermined whether accidentally or purposely inflicted*], E987[*falling from high place, undetermined whether accidentally or purposely inflicted*]). In a validation study examining a slightly different set of suicide codes, Iribarren et al. used Kaiser Permanente data from Northern California (1979-1993) to identify first hospitalizations for ICD-9 codes E950-E959.[89] This algorithm included a code used for late effects of self-inflicted injury (E959). The PPV, sensitivity and specificity reported for this algorithm were 86%, 95% and 87% respectively. The primary measure in this study was the percentage of E-code discharge diagnoses with suicide attempt confirmed by medical chart review. In an attempt to identify missed cases, authors extended their algorithm to include selected injuries and a depression diagnosis, but not an E-code. The PPV for this alternative algorithm was low, 26%, and dropped to 4% when the requirement of a depression diagnosis was removed. Two additional studies examining similar E-codes reported some interesting information on intention category and self-poisoning, but did not validate the codes included in their studies. A study by Blanc et al. showed that only 32.1% of E-codes and medical record review are in agreement when examined by intention category (unintentional, suicide/intentional, assault by poisoning, or undetermined intent).[90] Rhodes et al. reported 36.5% of discharges indicating self-poisoning were coded as intentional in the medical record, but 59.5% of self-poisoning classified as intentional based on expert review.[91] This suggested that codes used for unintentional (E800-E869, E880-E929) or undetermined (E980–E989) intent may be used in place of the intentional self-harm E-codes.

For defining the outcome of attempted suicide, the workgroup proposes an algorithm presented by Simon et al. which requires, in any position, an inpatient or outpatient ICD-9 E-code for intentional self-harm (E950–E958) in 90 days before or 180 days after initial antidepressant prescription or

psychotherapy visit related to a depression code. The workgroup recognizes the lack of validation studies performed in a broader population, thus have no evidence as to how this algorithm will perform outside of those with depression or who are on an antidepressant. Despite the age of the study, the workgroup proposes the following secondary algorithm for researchers who want to define a broader HOI, based on work by Iribarren et al.: an inpatient hospital discharge E-code for deliberate self-harm (E950–E959) in the primary or secondary position.

One final note of caution, E-codes may be inconsistently coded and coding may depend on local practices and regulations. Therefore it is recommended that E-code use be evaluated across Data Partners, and potentially evaluated within Data Partners stratified by geographic location. Such stratification will assist in the understanding of the completeness of E-code data in the MSCDM prior to using this algorithm.[92]

14. Thrombocytopenia

Primary Observed or Derived Algorithm
Inpatient, any position
>= 1 ICD-9 code: 287.1, 287.30, 287.31, 287.32, 287.33, 287.39, 287.4, 287.5, or 289.84

A literature review produced 9 articles regarding algorithms or codes to identify thrombocytopenia.[45-47, 50, 93-97]

The literature review resulted in the identification of studies using non-US/Canadian databases, which may not be generalizable to the MSDD. There was little information on codes other than 287.3 (including secondary or unspecified thrombocytopenia (287.4 [*secondary thrombocytopenia*], 287.5 [*thrombocytopenia, unspecified*]) which would likely be of more interest for this outcome. Heparin induced thrombocytopenia (HIT) had its own code (289.84 [*heparin-induced thrombocytopenia (HIT)*]), which was implemented in late 2008. It is possible that diagnostic changes occurred since this new code has been implemented, but the HIT specific code was not validated at the time of this review. Prior to 2008, some investigators used an E-code for anticoagulant related adverse event in combination with 287.4 to determine whether heparin was the cause of the event.

A study by Poordad et al. reviewed the ICD-9 diagnoses (287.3-287.5) for 2,500 patients using laboratory data as the gold standard.[93] These were Hepatitis C virus patients with diagnosed thrombocytopenia and all patients included had complete laboratory results available. The authors reported 65% concordance between diagnoses and platelet count <100,000/µL. Another study by Poordad et al. performed a chart validation in patients with chronic liver disease, from a large integrated U.S. commercial health plan.[94] ICD-9 codes examined were 287.3, 287.4, 287.5, using blood platelet count as the gold standard. PPV and sensitivity were low (31%, 40%) while NPV and specificity were high (99%, 98%). Galdarossa et al. (discussed in section on ITP (#7)) reported a PPV of 83%, but did not provide a definition for thrombocytopenia in their validation study.[50]

A few other studies were identified but deemed not relevant to this particular review.[45-47, 95, 96] One study conducted in the Netherlands by ten Berg et al. examined inpatient codes for 287.3,287.4,287.5 (authors did not specify primary or secondary diagnosis).[97] PPV, NPV, sensitivity, and specificity for this algorithm were reported as 86%, 99%, 12% and 99% respectively. This study also used "at least one platelet count <100,000/µL" as the gold standard.

For defining the outcome of thrombocytopenia, the workgroup principally proposes the following composite algorithm based on published literature and expert advice: at least one inpatient, any position, diagnosis of the following ICD-9 codes: 287.1 [*qualitative platelet defects*], 287.30, 287.31, 287.32, 287.33 [*congenital and hereditary thrombocytopenic purpura*], 287.39, 287.4, 287.5, 289.84. The workgroup decided to add a code specific for HIT (289.84) to Galdarossa's algorithm. This code has not yet been validated but holds great potential to identify HIT specifically in the absence of laboratory values. The workgroup does not prefer to recommend studies using non-US data based on concerns with generalizability; however, the paper by Galdarossa et al. is the only published validation study evaluating the new codes for thrombocytopenia. Therefore, as a secondary algorithm, the workgroup proposes the algorithm from Galdarossa et al.: at least one inpatient diagnosis of the following ICD-9 codes, in any position: 287.1, 287.30, 287.31, 287.32, 287.33, 287.39, 287.4, 287.5. This secondary algorithm may be preferred if HIT is not of interest.

15. Type I diabetes

Primary Observed or Derived Algorithm
Inpatient or outpatient, any position
>= 1 ICD-9 code: 250.01, 250.03, 250.11, 250.13, 250.21, 250.23, 250.31, 250.33, 250.41, 250.43, 250.51, 250.53, 250.61, 250.63, 250.71, 250.73, 250.81, 250.83, 250.91, or 250.93

A literature review produced 14 articles regarding algorithms or codes to identify type I diabetes.[98-111]

Type I diabetes is explicably intertwined with type II diabetes in the validation literature. Many validation studies exist that examine codes for both type I and type II diabetes, report PPVs and sensitivities ranging from 64%-98% and 73%-97%, respectively.[98-107] Much less information is available when considering the distinct ICD-9 codes to distinguish between Type I and Type II diabetes, but validation studies do exist.

In a study by Bobo et al., authors validated an algorithm for identifying persons with type I diabetes within Tennessee Medicaid data, finding a PPV of 80%, sensitivity of 65% and specificity being > 99%.[108] Operationally, since this study included codes for both Type I and II diabetes, the algorithm required: 1) a primary discharge diagnosis of 250 [*diabetes mellitus*], 250.0x [*diabetes mellitus without mention of complication*], 250.1x [*diabetes with ketoacidosis*], 250.2x [*diabetes with hyperosmolarity*], 250.3x [*diabetes with other coma*], or 250.9x [*diabetes with unspecified complication*]; 2) an inpatient stay with a secondary discharge diagnosis for one of these same ICD-9 codes + no diagnosis for 256.4 [*polycystic ovaries*] within 120 days of the diabetes diagnosis + a confirmatory antidiabetic prescription or an additional any-setting any-position ICD-9 code for diabetes within 120 days; or 3) an outpatient visit with a primary diagnosis for one of these same ICD-9 codes + a confirmatory antidiabetic prescription or an any-position inpatient ICD-9 code for diabetes within 120 days. In any of these three scenarios, >=1 prescription for insulin was also required within 120 days of the diabetes diagnosis, with no more than a single prescription for an oral antidiabetic drug during that interval. A single prescription for an oral agent did not serve as an exclusion criterion, because such drugs may be occasionally prescribed while awaiting the results of confirmatory testing for type I diabetes. If the aforementioned definition was not met, the individual was classified as a type II diabetic. Of note, the study population consisted solely of a small number of pediatric, adolescent, and young adult atypical antipsychotic users aged 6-24 years.

Rhodes et al. also examined a pediatric, adolescent, and young adult population, yet within the Endocrine/Diabetes or Obesity Programs at Children's Hospital in Boston, finding a PPV of 97% for Type I diabetes codes.[109] To identify Type I diabetes, the algorithm required an inpatient or outpatient ICD-9 code for 250.x1 or 250.x3 which includes: 250.01 [*diabetes mellitus without mention of complication, type one [juvenile type], not stated as uncontrolled*], 250.03 [*diabetes mellitus without mention of complication, type one [juvenile type], uncontrolled*], 250.11 [*diabetes mellitus with ketoacidosis, type one [juvenile type] not stated as uncontrolled*], 250.13 [*diabetes mellitus with ketoacidosis, type one [juvenile type], uncontrolled*], 250.21 [*diabetes mellitus with hyperosmolar coma, type one [juvenile type], not stated as uncontrolled*], 250.23 [*diabetes mellitus with hyperosmolar coma, type one [juvenile type], uncontrolled*], 250.31 [*diabetes mellitus with other coma, type one [juvenile], not stated as uncontrolled*], 250.33 [*diabetes mellitus with other coma, type one [juvenile], uncontrolled*], 250.41 [*diabetes mellitus with renal manifestations, type one [juvenile type], not stated as uncontrolled*], 250.43 [*diabetes mellitus with renal manifestations, type one [juvenile type], uncontrolled*], 250.51 [*diabetes mellitus with ophthalmic manifestations, type one [juvenile type], not stated as uncontrolled*], 250.53 [*diabetes mellitus with ophthalmic manifestations, type one [juvenile type], uncontrolled*], 250.61 [*diabetes mellitus with neurological manifestations, type one [juvenile type] not stated as uncontrolled*], 250.63 [*diabetes mellitus with neurological manifestations, type one [juvenile type] uncontrolled*], 250.71 [*diabetes mellitus with peripheral circulatory disorders, type one [juvenile type], not stated as uncontrolled*], 250.73 [*diabetes mellitus with peripheral circulatory disorders, type one [juvenile type], uncontrolled*], 250.81 [*diabetes mellitus with other specified manifestations, type one [juvenile type], not stated as uncontrolled*], 250.83 [*diabetes mellitus with other specified manifestations, type one [juvenile type], uncontrolled*], 250.91 [*diabetes mellitus with unspecified complication, type one [juvenile type], not stated as uncontrolled*], or 250.93 [*diabetes mellitus with unspecified complication, type one [juvenile type], uncontrolled*]. This algorithm is much less complicated than the algorithm proposed by Bobo et al. and has a higher PPV, but authors did not report sensitivity or specificity.

Additionally, Klompas et al. developed algorithms for identifying individuals with type I diabetes, without regard to patient age, within Atrius Health electronic medical record data.[110] The algorithm required >=2 ICD-9 diagnoses for 250.x1 or 250.x3 (as described above), a current prescription for insulin and no prescription for an oral antidiabetic agent at any time (excluding metformin). This yielded a PPV of 81% and sensitivity of 32%. Twenty-one other candidate algorithms were presented by the authors in the manuscript. Of note, among persons meeting screening criteria for potential diabetes, algorithms that maximized sensitivity (often at the cost of PPV) included individual components such as: a prescription for insulin; no record of any oral antidiabetic drug; and no record of any oral antidiabetic drug (excluding metformin). Algorithms that maximized PPV (often at the cost of sensitivity) included individual components such as: a ratio of type I to type II ICD-9 codes >0.5; a ratio of type I to type II ICD-9 codes >0.5 + prescription for insulin; a ratio of type I to type II ICD-9 codes >0.5 + prescription for glucagon; C-peptide <0.8; and a prescription for urine acetone test strips. An algorithm that maximized PPV (96%) while maintaining an acceptable level of sensitivity (61%) included a requirement for a ratio of type I to type II ICD-9 codes >0.5 and no prescription for an oral antidiabetic drug (excluding metformin. The authors also developed an "optimized" algorithm, achieving a high PPV (96%) and perfect sensitivity (100%), yet this definition requires laboratory components that are not currently supported by the MSCDM (e.g., C-peptide, diabetes autoantibodies results).

In addition to the studies validating ICD-9 codes, a study by Vanderloo et al. conducted in British Columbia, reviewed both ICD-9 and ICD-10 codes.[111] The PPV, sensitivity and specificity reported were

98%, 99% and 78%, respectively for type I diabetes. The workgroup did not recommend this algorithm, based its lumping of ICD-9 and ICD-10 codes.

For defining the outcome of type I diabetes, the workgroup favors the algorithm proposed by Rhodes et al., which includes inpatient or outpatient ICD-9 codes 250.x1 or 250.x3, in any position. This algorithm will best identify pediatric, adolescent, and young adults with type I diabetes, and although similar to Bobo et al., the algorithm was evaluated in more persons. As a secondary algorithm, the workgroup proposes the algorithm by Klompas et al. which includes: a ratio of type I to type II ICD-9 codes >0.5 + no prescription for an oral antidiabetic drug (excluding metformin), when researchers are interested in an adult patient population. Should requisite laboratory data be made available within the distributed database, the Klompas et al. "optimized" algorithm should be used.

16. Valvulopathy

Primary Observed or Derived Algorithm
Inpatient, any position
>=1 ICD-9 code: V42.2 , V43.3, 415, or 428
AND
>= 1 ICD-9 code: 394.x-397.x , 398.9, or 424.x

A literature review produced 6 articles regarding algorithms or codes to identify valvulopathy.[104, 112-116]

Literature searches resulted in one validated algorithm for valvular heart disease, which also included valve replacement. Birman-Deych et al. performed a validation of valvular heart disease or valve replacement in Medicare patients, as a risk factor for stroke in patients hospitalized for atrial fibrillation.[104] This algorithm included codes for valve disease, other heart disease, valve disorders and endocarditis (394.x [*diseases of mitral valve*], 395.x [*diseases of aortic valve*], 396.x [*diseases of mitral and aortic valves*], 397.x [*diseases of other endocardial structures*], 398.9 [*other and unspecified rheumatic heart diseases*], or 424.x [other diseases of endocardium]), as well as V-codes[b] for valve replacement (V42.2 [heart valve replaced by transplant], or V43.3 [heart valve replaced by other means])). The PPV and specificity reported were high, at 93% and 97%, respectively. The NPV was more modest at 68%, while the sensitivity was sub-par at 41%. When the authors restricted this algorithm to require the diagnosis be in a primary position at the baseline hospitalization, the sensitivity dropped to 4%.

Other studies provided algorithms for defining valvular heart disease or valvulopathy in claims data, but these algorithms have not been validated. Abbott et al. defined hospitalizations for valvular heart disease using ICD-9 codes 394.x, 395.x, 396.x, 397.x, 424.0 [*mitral valve disorders*], 424.1 [*aortic valve disorders*] and authors exclude endocarditis (424.9 [*endocarditis, valve unspecified*], 424.90 [*endocarditis, valve unspecified, unspecified cause*], 424.91 [*endocarditis in diseases classified elsewhere*], 424.99[*other endocarditis, valve unspecified*]).[112] A study by DeBruin et al. used codes 424.0, 424.1, 424.2 and 424.3 to define valvulopathy as a potential confounder.[113] While Sundstrom et al. used ICD-9 codes 394-397 and 424, or ICD-10 codes I05 -I08 and I34 - I37 to exclude patients with valvular

[b] As part of the ICD-9 diagnosis code hierarchy, V-codes provide information on supplementary classification of factors influencing health status and contact with health services.

disease in their community based sample.[114] There was also a French study that provides ICD-10 codes for cardiac valvular insufficiency, but these are also not validated.[115]

After much discussion, the workgroup decided the primary algorithm should only include more severe valvulopathy. It was also suggested that the literature be searched for heart failure (HF/CHF) codes combined with codes for valvulopathy. One study performed in Ontario administrative data, by Zadikoff et al. defined their outcome of valvular heart disease, valvular repair/replacement by ICD-9 codes 394, 396, 397, 424.[116] This study excluded diseases of the aortic valve (395), valvular replacement/major cardiac surgery (V43.3) and all rheumatic heart disease (ICD-9: 391 [*rheumatic fever with heart involvement*], 392 [*rheumatic chorea*], 394, 395, 397, 398 [*other rheumatic heart disease*]) in a 3 year baseline period. Codes for rheumatic heart disease were not included in the outcome definition, which differs from the Birman-Deych algorithm listed above. Authors of the Zadikoff study discussed the issue that valvular heart disease has not been systematically studied in their database, yet they chose not to validate the algorithms proposed to identify valvular heart disease. Instead, authors used CHF as an outcome of interest, since it is the most likely clinically overt outcome of valvulopathy and codes to define CHF have been validated.

For defining the outcome of valvulopathy, the workgroup proposes a composite algorithm based restricting codes to the most severe cases of valvulopathy, which is essentially valve replacement. The primary algorithm consists of an inpatient or outpatient ICD-9 code V42.2 or V43.3, in the primary or secondary position. The workgroup would also like to recommend a secondary algorithm. As a secondary algorithm; the workgroup recommends the Birman-Deych algorithm, which includes any position, inpatient diagnosis of one of the following ICD-9 codes. 394.x, 395.x, 396.x , 397.x , 398.9, 424.x and V42.2, V43.3. This algorithm should be used when researchers are not limiting valvulopathy to the most severe cases.

VII. REFERENCES

1. Schumock GT, Lee TA, Pickard S, et al. Alternative methods for health outcomes of interest validation. 2013.
2. van der Linden PD, van de Lei J, Nab HW, Knol A, Stricker BH. Achilles tendinitis associated with fluoroquinolones. Br J Clin Pharmacol. 1999;48(3):433-437.
3. van der Linden PD, Nab HW, Simonian S, Stricker BH, Leufkens HG, Herings RM. Fluoroquinolone use and the change in incidence of tendon ruptures in the Netherlands. Pharm World Sci. 2001;23(3):89-92.
4. van der Linden PD, Sturkenboom MC, Herings RM, Leufkens HG, Stricker BH. Fluoroquinolones and risk of Achilles tendon disorders: Case-control study. BMJ. 2002;324(7349):1306-1307.
5. Sode J, Obel N, Hallas J, Lassen A. Use of fluroquinolone and risk of Achilles tendon rupture: A population-based cohort study. Eur J Clin Pharmacol. 2007;63(5):499-503. doi: 10.1007/s00228-007-0265-9.
6. Yee CL, Duffy C, Gerbino PG, Stryker S, Noel GJ. Tendon or joint disorders in children after treatment with fluoroquinolones or azithromycin. Pediatr Infect Dis J. 2002;21(6):525-529.
7. Wahl PM, Gagne JJ, Wasser TE, et al. Early steps in the development of a claims-based targeted healthcare safety monitoring system and application to three empirical examples. Drug Saf. 2012;35(5):407-416. doi: 10.2165/11594770-000000000-00000; 10.2165/11594770-000000000-00000.
8. Seeger JD, West WA, Fife D, Noel GJ, Johnson LN, Walker AM. Achilles tendon rupture and its association with fluoroquinolone antibiotics and other potential risk factors in a managed care population. Pharmacoepidemiol Drug Saf. 2006;15(11):784-792. doi: 10.1002/pds.1214.
9. Schneider G, Kachroo S, Jones N, Crean S, Avetisyan R, Reynolds MW. Mini-sentinel systematic evaluation of health outcome of interest definitions for studies using administrative data. erythema multiforme Major/Minor/Not otherwise specified, Stevens-Johnson syndrome, or toxic epidermal necrolysis report. 2010.
10. Schneider G, Kachroo S, Jones N, et al. A systematic review of validated methods for identifying erythema multiforme major/minor/not otherwise specified, Stevens-Johnson syndrome, or toxic epidermal necrolysis using administrative and claims data. Pharmacoepidemiol Drug Saf. 2012;21 Suppl 1:236-239. doi: 10.1002/pds.2331; 10.1002/pds.2331.
11. Eisenberg DF, Daniel GW, Jones JK, et al. Validation of a claims-based diagnostic code for Stevens-Johnson syndrome in a commercially insured population. Pharmacoepidemiol Drug Saf. 2012;21(7):760-764. doi: 10.1002/pds.3276; 10.1002/pds.3276.
12. Dreyfus B, Kawabata H, Gomez A. Selected adverse events in cancer patients treated with vascular endothelial growth factor inhibitors. Cancer Epidemiol. 2013;37(2):191-196. doi: 10.1016/j.canep.2012.11.001; 10.1016/j.canep.2012.11.001.
13. Tse A, Tseng HF, Greene SK, Vellozzi C, Lee GM, VSD Rapid Cycle Analysis Influenza Working Group. Signal identification and evaluation for risk of febrile seizures in children following trivalent inactivated influenza vaccine in the Vaccine Safety Datalink project, 2010-2011. Vaccine. 2012;30(11):2024-2031. doi: 10.1016/j.vaccine.2012.01.027; 10.1016/j.vaccine.2012.01.027.

14. Klein NP, Fireman B, Yih WK, et al. Measles-mumps-rubella-varicella combination vaccine and the risk of febrile seizures. Pediatrics. 2010;126(1):e1-8. doi: 10.1542/peds.2010-0665; 10.1542/peds.2010-0665.

15. Klein NP, Lewis E, Baxter R, et al. Measles-containing vaccines and febrile seizures in children age 4 to 6 years. Pediatrics. 2012;129(5):809-814. doi: 10.1542/peds.2011-3198; 10.1542/peds.2011-3198.

16. Kee VR, Gilchrist B, Granner MA, Sarrazin NR, Carnahan RM. A systematic review of validated methods for identifying seizures, convulsions, or epilepsy using administrative and claims data. Pharmacoepidemiol Drug Saf. 2012;21 Suppl 1:183-193. doi: 10.1002/pds.2329; 10.1002/pds.2329.

17. Shui IM, Shi P, Dutta-Linn MM, et al. Predictive value of seizure ICD-9 codes for vaccine safety research. Vaccine. 2009;27(39):5307-5312. doi: 10.1016/j.vaccine.2009.06.092; 10.1016/j.vaccine.2009.06.092.

18. Huang WT, Gargiullo PM, Broder KR, et al. Lack of association between acellular pertussis vaccine and seizures in early childhood. Pediatrics. 2010;126(2):263-269. doi: 10.1542/peds.2009-1496; 10.1542/peds.2009-1496.

19. Mullooly JP, Donahue JG, DeStefano F, Baggs J, Eriksen E, VSD Data Quality Working Group. Predictive value of ICD-9-CM codes used in vaccine safety research. Methods Inf Med. 2008;47(4):328-335.

20. Faught E, Richman J, Martin R, et al. Incidence and prevalence of epilepsy among older U.S. Medicare beneficiaries. Neurology. 2012;78(7):448-453. doi: 10.1212/WNL.0b013e3182477edc; 10.1212/WNL.0b013e3182477edc.

21. Barlow WE, Davis RL, Glasser JW, et al. The risk of seizures after receipt of whole-cell pertussis or measles, mumps, and rubella vaccine. N Engl J Med. 2001;345(9):656-661. doi: 10.1056/NEJMoa003077.

22. Zangwill KM, Eriksen E, Lee M, et al. A population-based, post licensure evaluation of the safety of a combination diphtheria, tetanus, acellular pertussis, hepatitis B, and inactivated poliovirus vaccine in a large managed care organization. Pediatrics. 2008;122(6):e1179-85. doi: 10.1542/peds.2008-1977; 10.1542/peds.2008-1977.

23. Yih WK, Lee GM, Lieu TA, et al. Surveillance for adverse events following receipt of pandemic 2009 H1N1 vaccine in the post-licensure rapid immunization safety monitoring (PRISM) system, 2009-2010. Am J Epidemiol. 2012;175(11):1120-1128. doi: 10.1093/aje/kws197; 10.1093/aje/kws197.

24. Funch D, Holick C, Velentgas P, et al. Algorithms for identification of Guillain-Barre syndrome among adolescents in claims databases. Vaccine. 2013;31(16):2075-2079. doi: 10.1016/j.vaccine.2013.02.009; 10.1016/j.vaccine.2013.02.009.

25. Shui IM, Rett MD, Weintraub E, et al. Guillain-Barre syndrome incidence in a large united states cohort (2000-2009). Neuroepidemiology. 2012;39(2):109-115. doi: 10.1159/000339248; 10.1159/000339248.

26. Jones TF, McMillian M, Boothe E, Hanna S, Ingram LA. Hospital discharge data for Guillain-Barre syndrome and influenza A (H1N1) vaccine adverse events. Emerg Infect Dis. 2010;16(9):1500-1501. doi: 10.3201/eid1609.091837; 10.3201/eid1609.091837.
27. Lee CD, Jones TF. Hospital discharge database optimization in Guillain-Barre syndrome surveillance. Muscle Nerve. 2012;46(1):60-62. doi: 10.1002/mus.23261; 10.1002/mus.23261.
28. Polakowski LL, Sandhu SK, Martin DB, et al. Chart-confirmed Guillain-Barre syndrome after 2009 H1N1 influenza vaccination among the Medicare population, 2009-2010. Am J Epidemiol. 2013;178(6):962-973. doi: 10.1093/aje/kwt051; 10.1093/aje/kwt051.
29. Burwen DR, Sandhu SK, MaCurdy TE, et al. Surveillance for Guillain-Barre syndrome after influenza vaccination among the Medicare population, 2009-2010. Am J Public Health. 2012;102(10):1921-1927. doi: 10.2105/AJPH.2011.300510.
30. Armed Forces Health Surveillance Center. Predictive value of surveillance case definitions of Guillain-Barre syndrome in vaccine safety assessment. 2012(19 (9)):8-9.
31. Greene SK, Rett MD, Vellozzi C, et al. Guillain-Barre syndrome, influenza vaccination, and antecedent respiratory and gastrointestinal infections: A case-centered analysis in the vaccine safety datalink, 2009-2011. PLoS One. 2013;8(6):e67185. doi: 10.1371/journal.pone.0067185.
32. Weiss PF, Klink AJ, Hexem K, et al. Variation in inpatient therapy and diagnostic evaluation of children with Henoch-Schönlein purpura. J Pediatr. 2009;155(6):812-818.e1. doi: 10.1016/j.jpeds.2009.05.030; 10.1016/j.jpeds.2009.05.030.
33. Carnahan RM. Mini-sentinel systematic reviews of validated methods for identifying health outcomes using administrative data. analysis of evidence gaps and lessons learned report. 2011.
34. Hudson M, Avina-Zubieta A, Lacaille D, Bernatsky S, Lix L, Jean S. The validity of administrative data to identify hip fractures is high--a systematic review. J Clin Epidemiol. 2013;66(3):278-285. doi: 10.1016/j.jclinepi.2012.10.004; 10.1016/j.jclinepi.2012.10.004.
35. Ray WA, Griffin MR, Fought RL, Adams ML. Identification of fractures from computerized Medicare files. J Clin Epidemiol. 1992;45(7):703-714.
36. Narongroeknawin P, Patkar NM, Shakoory B, et al. Validation of diagnostic codes for subtrochanteric, diaphyseal, and atypical femoral fractures using administrative claims data. J Clin Densitom. 2012;15(1):92-102. doi: 10.1016/j.jocd.2011.09.001; 10.1016/j.jocd.2011.09.001.
37. Jean S, Candas B, Belzile E, et al. Algorithms can be used to identify fragility fracture cases in physician-claims databases. Osteoporos Int. 2012;23(2):483-501. doi: 10.1007/s00198-011-1559-4; 10.1007/s00198-011-1559-4.
38. Tamblyn R, Reid T, Mayo N, McLeod P, Churchill-Smith M. Using medical services claims to assess injuries in the elderly: Sensitivity of diagnostic and procedure codes for injury ascertainment. J Clin Epidemiol. 2000;53(2):183-194.
39. Virnig B, Durham SB, Folsom AR, Cerhan J. Linking the Iowa Women's Health Study cohort to Medicare data: Linkage results and application to hip fracture. Am J Epidemiol. 2010;172(3):327-333. doi: 10.1093/aje/kwq111; 10.1093/aje/kwq111.

40. Lofthus CM, Cappelen I, Osnes EK, et al. Local and national electronic databases in Norway demonstrate a varying degree of validity. J Clin Epidemiol. 2005;58(3):280-285. doi: 10.1016/j.jclinepi.2004.07.003.

41. Sund R, Nurmi-Luthje I, Luthje P, Tanninen S, Narinen A, Keskimaki I. Comparing properties of audit data and routinely collected register data in case of performance assessment of hip fracture treatment in Finland. Methods Inf Med. 2007;46(5):558-566.

42. Evans JG, Seagroatt V, Goldacre MJ. Secular trends in proximal femoral fracture, oxford record linkage study area and England 1968-86. J Epidemiol Community Health. 1997;51(4):424-429.

43. Park-Wyllie LY, Mamdani MM, Juurlink DN, et al. Bisphosphonate use and the risk of subtrochanteric or femoral shaft fractures in older women. JAMA. 2011;305(8):783-789. doi: 10.1001/jama.2011.190; 10.1001/jama.2011.190.

44. Kachroo S, Jones N, Reynolds MW. Systematic literature review for evaluation of hip fractures. 2009;A2-8756.

45. Terrell DR, Beebe LA, Vesely SK, Neas BR, Segal JB, George JN. Determining a definite diagnosis of primary immune thrombocytopenia by medical record review. Am J Hematol. 2012;87(9):843-847. doi: 10.1002/ajh.23226; 10.1002/ajh.23226.

46. Heden KE, Jensen AO, Farkas DK, Norgaard M. Validity of a procedure to identify patients with chronic idiopathic thrombocytopenic purpura in the Danish national registry of patients. Clin Epidemiol. 2009;1:7-10.

47. Segal JB, Powe NR. Accuracy of identification of patients with immune thrombocytopenic purpura through administrative records: A data validation study. Am J Hematol. 2004;75(1):12-17. doi: 10.1002/ajh.10445.

48. Watts RG. Idiopathic thrombocytopenic purpura: A 10-year natural history study at the Children's Hospital of Alabama. Clin Pediatr (Phila). 2004;43(8):691-702.

49. Chiao EY, Engels EA, Kramer JR, et al. Risk of immune thrombocytopenic purpura and autoimmune hemolytic anemia among 120 908 US veterans with hepatitis C virus infection. Arch Intern Med. 2009;169(4):357-363. doi: 10.1001/archinternmed.2008.576; 10.1001/archinternmed.2008.576.

50. Galdarossa M, Vianello F, Tezza F, et al. Epidemiology of primary and secondary thrombocytopenia: First analysis of an administrative database in a major Italian institution. Blood Coagul Fibrinolysis. 2012;23(4):271-277. doi: 10.1097/MBC.0b013e328351882d; 10.1097/MBC.0b013e328351882d.

51. Hartsfield CL, Korner EJ, Ellis JL, Raebel MA, Merenich J, Brandenburg N. Painful diabetic peripheral neuropathy in a managed care setting: Patient identification, prevalence estimates, and pharmacy utilization patterns. Popul Health Manag. 2008;11(6):317-328. doi: 10.1089/pop.2008.0015; 10.1089/pop.2008.0015.

52. Berger A, Dukes EM, Oster G. Clinical characteristics and economic costs of patients with painful neuropathic disorders. J Pain. 2004;5(3):143-149. doi: 10.1016/j.jpain.2003.12.004.

53. Shields RW. Peripheral neuropathy. In: Carey W, ed. Disease Management Project. Online ed. Cleveland, OH: The Cleveland Clinic Foundation; 2010.

54. UpToDate. Management of polyneuropathy, treatment of diabetic neuropathy. www.uptodate.com.
55. Facts & comparisons eAnswers. http://online.factsandcomparisons.com/index.aspx?.
56. Compagni A, Melegaro A, Tarricone R. Genetic screening for the predisposition to venous thromboembolism: A cost-utility analysis of clinical practice in the Italian health care system. Value Health. 2013;16(6):909-921. doi: 10.1016/j.jval.2013.05.003; 10.1016/j.jval.2013.05.003.
57. White RH, Murin S, Wun T, Danielsen B. Recurrent venous thromboembolism after surgery-provoked versus unprovoked thromboembolism. J Thromb Haemost. 2010;8(5):987-997. doi: 10.1111/j.1538-7836.2010.03798.x; 10.1111/j.1538-7836.2010.03798.x.
58. White RH, Zhou H, Kim J, Romano PS. A population-based study of the effectiveness of inferior vena cava filter use among patients with venous thromboembolism. Arch Intern Med. 2000;160(13):2033-2041.
59. Venetz C, Jimenez D, Mean M, Aujesky D. A comparison of the original and simplified pulmonary embolism severity index. Thromb Haemost. 2011;106(3):423-428. doi: 10.1160/TH11-04-0263; 10.1160/TH11-04-0263.
60. Venetz C, Labarere J, Jimenez D, Aujesky D. White blood cell count and mortality in patients with acute pulmonary embolism. Am J Hematol. 2013;88(8):677-681. doi: 10.1002/ajh.23484; 10.1002/ajh.23484.
61. Atwood KM, Robitaille CJ, Reimer K, Dai S, Johansen HL, Smith MJ. Comparison of diagnosed, self-reported, and physically-measured hypertension in Canada. Can J Cardiol. 2013;29(5):606-612. doi: 10.1016/j.cjca.2012.11.019; 10.1016/j.cjca.2012.11.019.
62. Lowe BS, Therrien J, Ionescu-Ittu R, Pilote L, Martucci G, Marelli AJ. Diagnosis of pulmonary hypertension in the congenital heart disease adult population impact on outcomes. J Am Coll Cardiol. 2011;58(5):538-546. doi: 10.1016/j.jacc.2011.03.033; 10.1016/j.jacc.2011.03.033.
63. Aujesky D, Obrosky DS, Stone RA, et al. A prediction rule to identify low-risk patients with pulmonary embolism. Arch Intern Med. 2006;166(2):169-175. doi: 10.1001/archinte.166.2.169.
64. Hyduk A, Croft JB, Ayala C, Zheng K, Zheng ZJ, Mensah GA. Pulmonary hypertension surveillance--united states, 1980-2002. MMWR Surveill Summ. 2005;54(5):1-28.
65. Cziraky MJ, Willey VJ, McKenney JM, et al. Statin safety: An assessment using an administrative claims database. Am J Cardiol. 2006;97(8A):61C-68C. doi: 10.1016/j.amjcard.2005.12.011.
66. Setoguchi S, Higgins JM, Mogun H, Mootha VK, Avorn J. Propranolol and the risk of hospitalized myopathy: Translating chemical genomics findings into population-level hypotheses. Am Heart J. 2010;159(3):428-433. doi: 10.1016/j.ahj.2009.12.008; 10.1016/j.ahj.2009.12.008.
67. Floyd JS, Heckbert SR, Weiss NS, Carrell DS, Psaty BM. Use of administrative data to estimate the incidence of statin-related rhabdomyolysis. JAMA. 2012;307(15):1580-1582. doi: 10.1001/jama.2012.489; 10.1001/jama.2012.489.
68. Andrade SE, Graham DJ, Staffa JA, et al. Health plan administrative databases can efficiently identify serious myopathy and rhabdomyolysis. J Clin Epidemiol. 2005;58(2):171-174. doi: 10.1016/j.jclinepi.2004.10.004.

69. Racoosin JA, Ryan P. Implications of health outcomes of interest definitions: Acute liver injury case study. http://omop.fnih.org/sites/default/files/OMOP%202011%20Symposium%20HOI%20Racoosin%20Final.pdf.

70. Hansen RA, Gray MD, Fox BI, et al. Expert panel assessment of acute liver injury identification in observational data. Res Social Adm Pharm. 2013. doi: 10.1016/j.sapharm.2013.04.012; 10.1016/j.sapharm.2013.04.012.

71. Lo Re V,3rd, Haynes K, Goldberg D, et al. Validity of diagnostic codes to identify cases of severe acute liver injury in the US Food and Drug Administration's Mini-Sentinel Distributed Database. Pharmacoepidemiol Drug Saf. 2013;22(8):861-872. doi: 10.1002/pds.3470; 10.1002/pds.3470.

72. Chung CP, Murray KT, Stein CM, Hall K, Ray WA. A computer case definition for sudden cardiac death. Pharmacoepidemiol Drug Saf. 2010;19(6):563-572. doi: 10.1002/pds.1888.

73. Iribarren C, Crow RS, Hannan PJ, Jacobs DR,Jr, Luepker RV. Validation of death certificate diagnosis of out-of-hospital sudden cardiac death. Am J Cardiol. 1998;82(1):50-53.

74. Fox CS, Evans JC, Larson MG, et al. A comparison of death certificate out-of-hospital coronary heart disease death with physician-adjudicated sudden cardiac death. Am J Cardiol. 2005;95(7):856-859. doi: 10.1016/j.amjcard.2004.12.011.

75. Hennessy S, Leonard CE, Freeman CP, et al. Validation of diagnostic codes for outpatient-originating sudden cardiac death and ventricular arrhythmia in Medicaid and Medicare claims data. Pharmacoepidemiol Drug Saf. 2010;19(6):555-562. doi: 10.1002/pds.1869.

76. McDonald KM, Hlatky MA, Saynina O, Geppert J, Garber AM, McClellan MB. Trends in hospital treatment of ventricular arrhythmias among Medicare beneficiaries, 1985 to 1995. Am Heart J. 2002;144(3):413-421.

77. Hennessy S, Leonard CE, Newcomb C, Kimmel SE, Bilker WB. Cisapride and ventricular arrhythmia. Br J Clin Pharmacol. 2008;66(3):375-385. doi: 10.1111/j.1365-2125.2008.03249.x; 10.1111/j.1365-2125.2008.03249.x.

78. Mines D. Commentary on the validation studies of sudden cardiac death and ventricular arrhythmia by Hennessey et al. and Chung et al. Pharmacoepidemiol Drug Saf. 2010;19(6):573-5; author reply 576; author reply 577-8. doi: 10.1002/pds.1913.

79. Leonard CE, Freeman CP, Deo R, et al. Response to commentary entitled "commentary on the validation studies of sudden cardiac death and ventricular arrhythmia by Hennessey et al. and Chung et al.". Pharmacoepidemiol Drug Saf. 2010;19(6):576-576. doi: 10.1002/pds.1956.

80. Kawai VK, Murray KT, Stein CM, et al. Validation of a computer case definition for sudden cardiac death in opioid users. BMC Res Notes. 2012;5:473-0500-5-473. doi: 10.1186/1756-0500-5-473; 10.1186/1756-0500-5-473.

81. Schelleman H, Bilker WB, Strom BL, et al. Cardiovascular events and death in children exposed and unexposed to ADHD agents. Pediatrics. 2011;127(6):1102-1110. doi: 10.1542/peds.2010-3371; 10.1542/peds.2010-3371.

82. Ray WA, Meredith S, Thapa PB, Meador KG, Hall K, Murray KT. Antipsychotics and the risk of sudden cardiac death. Arch Gen Psychiatry. 2001;58(12):1161-1167.

83. Staffa JA, Jones JK, Gable CB, Verspeelt JP, Amery WK. Risk of selected serious cardiac events among new users of antihistamines. Clin Ther. 1995;17(6):1062-1077.

84. Tamariz L, Harkins T, Nair V. A systematic review of validated methods for identifying ventricular arrhythmias using administrative and claims data. Pharmacoepidemiol Drug Saf. 2012;21 Suppl 1:148-153. doi: 10.1002/pds.2340; 10.1002/pds.2340.

85. Walkup JT, Townsend L, Crystal S, Olfson M. A systematic review of validated methods for identifying suicide or suicidal ideation using administrative or claims data. Pharmacoepidemiol Drug Saf. 2012;21 Suppl 1:174-182. doi: 10.1002/pds.2335; 10.1002/pds.2335.

86. Shevchenko IP, Lynch JT, Mattie AS, Reed-Fourquet LL. Verification of information in a large medical database using linkages with external databases. Stat Med. 1995;14(5-7):511-530.

87. Weis MA, Bradberry C, Carter LP, Ferguson J, Kozareva D. An exploration of human services system contacts prior to suicide in South Carolina: An expansion of the South Carolina violent death reporting system. Inj Prev. 2006;12 Suppl 2:ii17-ii21. doi: 10.1136/ip.2006.012427.

88. Simon GE, Savarino J. Suicide attempts among patients starting depression treatment with medications or psychotherapy. Am J Psychiatry. 2007;164(7):1029-1034. doi: 10.1176/appi.ajp.164.7.1029.

89. Iribarren C, Sidney S, Jacobs DR,Jr, Weisner C. Hospitalization for suicide attempt and completed suicide: Epidemiological features in a managed care population. Soc Psychiatry Psychiatr Epidemiol. 2000;35(7):288-296.

90. Blanc PD, Jones MR, Olson KR. Surveillance of poisoning and drug overdose through hospital discharge coding, poison control center reporting, and the drug abuse warning network. Am J Emerg Med. 1993;11(1):14-19.

91. Rhodes AE, Links PS, Streiner DL, Dawe I, Cass D, Janes S. Do hospital E-codes consistently capture suicidal behavior?. Chronic Dis Can. 2002;23(4):139-145.

92. National Center for Injury Prevention and Control. Recommended actions to improve external-cause-of-injury coding in state-based hospital discharge and emergency department data systems. 2009.

93. Poordad F, Theodore D, Sullivan J, Grotzinger K. Medical resource utilization and healthcare costs in patients with chronic hepatitis C viral infection and thrombocytopenia. J Med Econ. 2011;14(2):194-206. doi: 10.3111/13696998.2011.562266; 10.3111/13696998.2011.562266.

94. Poordad F, Theodore D, Sullivan J, Grotzinger K. Evaluating medical resource utilization and costs associated with thrombocytopenia in chronic liver disease patients. J Med Econ. 2012;15(1):112-124. doi: 10.3111/13696998.2011.632463; 10.3111/13696998.2011.632463.

95. Handler SM, Altman RL, Perera S, et al. A systematic review of the performance characteristics of clinical event monitor signals used to detect adverse drug events in the hospital setting. J Am Med Inform Assoc. 2007;14(4):451-458. doi: 10.1197/jamia.M2369.

96. Glance LG, Dick AW, Osler TM, Mukamel DB. Accuracy of hospital report cards based on administrative data. Health Serv Res. 2006;41(4 Pt 1):1413-1437. doi: 10.1111/j.1475-6773.2006.00554.x.

97. ten Berg MJ, van Solinge WW, van den Bemt PM, Huisman A, Schobben AF, Egberts TC. Platelet measurements versus discharge diagnoses for identification of patients with potential drug-induced thrombocytopenia: A cross-sectional study in the netherlands. Drug Saf. 2009;32(1):69-76. doi: 10.2165/00002018-200932010-00006.

98. Zgibor JC, Orchard TJ, Saul M, et al. Developing and validating a diabetes database in a large health system. Diabetes Res Clin Pract. 2007;75(3):313-319. doi: 10.1016/j.diabres.2006.07.007.

99. O'Connor PJ, Rush WA, Pronk NP, Cherney LM. Identifying diabetes mellitus or heart disease among health maintenance organization members: Sensitivity, specificity, predictive value, and cost of survey and database methods. Am J Manag Care. 1998;4(3):335-342.

100. Hebert PL, Geiss LS, Tierney EF, Engelgau MM, Yawn BP, McBean AM. Identifying persons with diabetes using Medicare claims data. Am J Med Qual. 1999;14(6):270-277.

101. Solberg LI, Engebretson KI, Sperl-Hillen JM, Hroscikoski MC, O'Connor PJ. Are claims data accurate enough to identify patients for performance measures or quality improvement? the case of diabetes, heart disease, and depression. Am J Med Qual. 2006;21(4):238-245. doi: 10.1177/1062860606288243.

102. Tu K, Manuel D, Lam K, Kavanagh D, Mitiku TF, Guo H. Diabetics can be identified in an electronic medical record using laboratory tests and prescriptions. J Clin Epidemiol. 2011;64(4):431-435. doi: 10.1016/j.jclinepi.2010.04.007; 10.1016/j.jclinepi.2010.04.007.

103. Humphries KH, Rankin JM, Carere RG, Buller CE, Kiely FM, Spinelli JJ. Co-morbidity data in outcomes research: Are clinical data derived from administrative databases a reliable alternative to chart review?. J Clin Epidemiol. 2000;53(4):343-349.

104. Birman-Deych E, Waterman AD, Yan Y, Nilasena DS, Radford MJ, Gage BF. Accuracy of ICD-9-CM codes for identifying cardiovascular and stroke risk factors. Med Care. 2005;43(5):480-485.

105. Wilson C, Susan L, Lynch A, Saria R, Peterson D. Patients with diagnosed diabetes mellitus can be accurately identified in an Indian health service patient registration database. Public Health Rep. 2001;116(1):45-50.

106. Miller DR, Safford MM, Pogach LM. Who has diabetes? best estimates of diabetes prevalence in the department of veterans affairs based on computerized patient data. Diabetes Care. 2004;27 Suppl 2:B10-21.

107. Hux JE, Ivis F, Flintoft V, Bica A. Diabetes in Ontario: Determination of prevalence and incidence using a validated administrative data algorithm. Diabetes Care. 2002;25(3):512-516.

108. Bobo WV, Cooper WO, Stein CM, et al. Positive predictive value of a case definition for diabetes mellitus using automated administrative health data in children and youth exposed to antipsychotic drugs or control medications: A Tennessee Medicaid study. BMC Med Res Methodol. 2012;12:128-2288-12-128. doi: 10.1186/1471-2288-12-128; 10.1186/1471-2288-12-128.

109. Rhodes ET, Laffel LM, Gonzalez TV, Ludwig DS. Accuracy of administrative coding for type 2 diabetes in children, adolescents, and young adults. Diabetes Care. 2007;30(1):141-143. doi: 10.2337/dc06-1142.

110. Klompas M, Eggleston E, McVetta J, Lazarus R, Li L, Platt R. Automated detection and classification of type 1 versus type 2 diabetes using electronic health record data. Diabetes Care. 2013;36(4):914-921. doi: 10.2337/dc12-0964; 10.2337/dc12-0964.

111. Vanderloo SE, Johnson JA, Reimer K, et al. Validation of classification algorithms for childhood diabetes identified from administrative data. Pediatr Diabetes. 2012;13(3):229-234. doi: 10.1111/j.1399-5448.2011.00795.x; 10.1111/j.1399-5448.2011.00795.x.

112. Abbott KC, Agodoa LY. Hospitalizations for valvular heart disease in chronic dialysis patients in the united states. Nephron. 2002;92(1):43-50.

113. De Bruin ML, Langendijk PN, Koopmans RP, Wilde AA, Leufkens HG, Hoes AW. In-hospital cardiac arrest is associated with use of non-antiarrhythmic QTc-prolonging drugs. Br J Clin Pharmacol. 2007;63(2):216-223. doi: 10.1111/j.1365-2125.2006.02722.x.

114. Sundstrom J, Ingelsson E, Berglund L, et al. Cardiac troponin-I and risk of heart failure: A community-based cohort study. Eur Heart J. 2009;30(7):773-781. doi: 10.1093/eurheartj/ehp047; 10.1093/eurheartj/ehp047.

115. Weill A, Paita M, Tuppin P, et al. Benfluorex and valvular heart disease: A cohort study of a million people with diabetes mellitus. Pharmacoepidemiol Drug Saf. 2010;19(12):1256-1262. doi: 10.1002/pds.2044; 10.1002/pds.2044.

116. Zadikoff C, Duong-Hua M, Sykora K, Marras C, Lang A, Rochon P. Pergolide associated cardiac valvulopathy based on Ontario administrative data. Can J Neurol Sci. 2008;35(2):173-178.

VIII. APPENDIX

Appendix Table 1. Operational definitions and observed or derived algorithms

Please refer back to the body of the paper for the workgroup's review and impression of the algorithm performance. The appearance of an algorithm here does not mean that it will perform well.
If the algorithm has not been validated or has a PPV less than 75%, it is denoted with an "" in the table below.*

Health Outcome of Interest	Observed or Derived Algorithm	Comments
1. Achilles tendon rupture (ATR)		
Primary	Any claim type, any position >= 1 ICD-9 code: 727.67 AND >= 1 CPT code: 27605, 27606, 27650, 27652, 27654, or 01472	Seeger et al.
2. Erythema multiforme major (EMM), including Stevens Johnson syndrome (SJS) and toxic epidermal necrolysis (TEN)		
Primary*	Any claim type, any position >= 1 ICD-9 code: 695.1, 695.10, 695.11, 695.12, 695.13, 695.14, 695.15, or 695.19	No validation statistics reported on the newly defined ICD-9-CM coding.
Secondary*	Inpatient, any position >= 1 ICD-9 code: 695.1, 695.10, 695.11, 695.12, 695.13, 695.14, 695.15, or 695.19	For serious EM. No validation statistics reported on the newly defined ICD-9-CM coding.
3. Febrile seizure		
Primary	ED or inpatient , any position >= 1 ICD-9 code: 780.31 or 780.32	To-be published PRISM febrile seizure report, Tse et al.
4. Guillain-Barre syndrome (GBS)		
Primary*	Inpatient, primary position >= 1 ICD-9 code: 357.0 AND Outpatient, any position >= 1 ICD-9 code: 357.0 EXCLUDE patients with the following in the 365 day baseline: Any claim type, any position >= 1 ICD-9 code: 357.0 or 357.81	PRISM GBS review
Secondary*	Inpatient, primary position >= 1 ICD-9 code: 357.0 EXCLUDE patients with the following in the 365 day baseline: Any claim type, any position >= 1 ICD-9 code: 357.0 or 357.81	
5. Henoch-Schönlein purpura (HSP)		
Primary*	Inpatient, any position >= 1 ICD-9 code: 287.0 in patients <18 years	Based on an unvalidated algorithm by Weiss et al.

Health Outcome of Interest	Observed or Derived Algorithm	Comments
6. Hip fracture		
Primary	Inpatient, primary position >= 1 ICD-9 code: 820.00, 820.01, 820.02, 820.03, 820.09, 820.20, 820.21, or 820.8	Narongroeknawin et al. algorithm to define "typical hip fracture"
Secondary	Inpatient, any position >= 1 ICD-9 code: 820.00, 820.01, 820.02, 820.03, 820.09, 820.20, 820.21, or 820.8	
7. Idiopathic thrombocytopenia purpura (ITP)		
Primary*	Any claim type, any position >= 1 ICD-9 code: 287.31	Modified version of an algorithm presented by Galdarossa et al.
Secondary*	Any claim type, any position >= 1 ICD-9 code: 287.30, 287.31, 287.32, or 287.39	Composite algorithm developed by the workgroup
Secondary*	Any claim type, any position >= 1 ICD-9 code: 287.31 AND >= 1 CPT code: 85025, 85027, 85032, 85049, 85060, or 85097 on the same day or one day prior	Algorithm developed by FDA and HealthCore provided through a personal communication
8. Peripheral neuropathy (PN)		
Primary*	*Definition #1:* Any claim type, any position >= 1 ICD-9 code (broad): 250.6, 250.60, 250.61, 250.62, 250.63, 337.0, 337.00, 337.01, 337.09, 337.1, 337.9, 349.9, 356.4, 356.8, 356.9, 357, 357.0, 357.1, 357.2, 357.3, 357.4, 357.5, 357.6, 357.7, 357.8, 357.81, 357.82, 357.89, 357.9, or 729.2 AND >= PN drug *Definition #2:* Any claim type, any position >= 1 ICD-9 code (broad): 250.6, 250.60, 250.61, 250.62, 250.63, 337.0, 337.00, 337.01, 337.09, 337.1, 337.9, 349.9, 356.4, 356.8, 356.9, 357, 357.0, 357.1, 357.2, 357.3, 357.4, 357.5, 357.6, 357.7, 357.8, 357.81, 357.82, 357.89, 357.9, or 729.2 without mention of a PN drug *Definition #3:* Any claim type, any position >= 1 ICD-9 code (narrow): 356.4, 356.8, 356.9, 357, 357.0, 357.1, 357.6, 357.7, 357.8, 357.81, 357.89, 357.9, or 729.2 AND >= 1 PN drug *Definition #4:*	The workgroup is offering definitions for four unvalidated PN cohorts (in no particular order), intended to be used at the discretion of the researcher on a per study basis with regard to files used and coding position Code lists to define PN drugs are not provided.

Health Outcome of Interest	Observed or Derived Algorithm	Comments
	Any claim type, any position >= 1 ICD-9 code (narrow): 356.4, 356.8, 356.9, 357, 357.0, 357.1, 357.6, 357.7, 357.8, 357.81, 357.89, 357.9, or 729.2 without mention of a PN drug	
9. Pulmonary hypertension (PH)		
Primary*	Any claim type, any position >= 1 ICD-9 code: 416.0, 416.8, or 416.9	Composite algorithm developed by the workgroup; has not been validated.
10. Rhabdomyolysis		
Primary*	Any claim type, any position >= 1 ICD-9 code: 728.88	Code has not been well studied. Algorithm presented by Floyd et al.
11. Severe acute liver injury (SALI)		
Primary*	Inpatient, any position >= 2 ICD-9 code: 573.3, 573.8, 570, 572.2, 572.4, 572.8, or V42.7 OR Inpatient, any position >= 1 ICD-9 code: 573.3, 573.8, 570, 572.2, 572.4, 572.8, or V42.7 AND >= 1 ICD-9 procedure code: 50.1x, 50.9x OR >= 1 CPT code: 47000, 47001, or 47100	Lo Re et al. Mini-Sentinel Validation Report on SALI
12. Sudden cardiac death (SCD) and Ventricular Arrhythmia (VA)		
Primary	ED or Inpatient, primary/first-listed position >= 1 ICD-9 code: 427.5, 798, 798.1, 798.2, 427.1, 427.4, 427.41, or 427.42 EXCLUDE patients with a prior SCD/VA diagnosis: Any claim type, any position >= 1 ICD-9 code: 427.5, 798, 798.1, 798.2, 427.1, 427.4, 427.41, or 427.42	Hennessy et al.
13. Attempted suicide		
Primary	Inpatient or outpatient, any position >= 1 ICD-9 E-code: E950, E951, E952, E953, E954, E955, E956, E957, E958 90 days before or 180 days after initial antidepressant prescription or psychotherapy visit related to a depression code	Simon et al. The workgroup does not know how this algorithm will perform outside of those with depression or who are on an antidepressant.
Secondary	Inpatient, any position >= 1 ICD-9 E-code: E950, E951, E952, E953, E954, E955, E956, E957, E958, or E959	Iribarren et al.
14. Thrombocytopenia		
Primary*	Inpatient, any position >= 1 ICD-9 code: 287.1, 287.30, 287.31, 287.32, 287.33, 287.39, 287.4, 287.5, or 289.84	The workgroup developed this modified algorithm based on Galdarossa et al. Includes code for heparin induced thrombocytopenia.
Secondary	Inpatient, any position >= 1 ICD-9 code: 287.1, 287.30, 287.31, 287.32,	Galdarossa et al.

Health Outcome of Interest	Observed or Derived Algorithm	Comments
	287.33, 287.39, 287.4, or 287.5	
15. Type I diabetes mellitus		
Primary	Inpatient or outpatient, any position >= 1 ICD-9 code: 250.01, 250.03, 250.11, 250.13, 250.21, 250.23, 250.31, 250.33, 250.41, 250.43, 250.51, 250.53, 250.61, 250.63, 250.71, 250.73, 250.81, 250.83, 250.91, or 250.93	Rhodes et al.
Secondary	Outpatient, any position >= 2 ICD-9 codes: 250.01, 250.03, 250.11, 250.13, 250.21, 250.23, 250.31, 250.33, 250.41, 250.43, 250.51, 250.53, 250.61, 250.63, 250.71, 250.73, 250.81, 250.83, 250.91, or 250.93 on two or more occasions AND a current prescription for insulin AND no prescriptions for oral hypoglycemic	Klompas et al. Code lists to define insulin and oral hypoglycemic are not provided.
16. Valvulopathy		
Primary*	Inpatient, any position >=1 ICD-9 code: V42.2 , V43.3, 415, or 428 AND >= 1 ICD-9 code: 394.x-397.x , 398.9, or 424.x	Unvalidated, composite algorithm developed by the workgroup to define severe valvulopathy.
Secondary	Inpatient, any position >= 1 ICD-9 code:394.x-397.x, 398.9, 424.x, V42.2, or V43.3	Birman-Deych et al.
* algorithm has not been validated or has a PPV less than 75%		