ACRM
AMERICAN CONGRESS OF REHABILITATION MEDICINE

Archives of Physical Medicine and Rehabilitation

journal homepage: www.archives-pmr.org

Archives of Physical Medicine and Rehabilitation 2023;000: 1–10

ORIGINAL RESEARCH

Using Billing Codes to Create a Pediatric Functional Status e-Score for Children Receiving Inpatient Rehabilitation

Jennifer P. Lundine, PhD, CCC-SLP,[a,b] Jared D. Huling, PhD,[c] P. David Adelson, MD,[d] Randall S. Burd, MD,[e] Molly Fuentes, MD, MS,[f] Juliet Haarbauer-Krupa, PhD,[g] Kaitlin Hagen, MOT, OTR/L,[h] Cynthia Iske, MS, OTR/L,[i] Christine Koterba, PhD, ABPP,[j] Brad G. Kurowski, MD, MS,[k,l] Stephanie Petrucci, BSN, RN, CRRN,[i] Sean C. Rose, MD,[m] Cristina L. Sadowsky, MD,[h] Jennifer Westendorf, DPT, CBIS,[n] Annie Truelove, MPH,[o] Julie C. Leonard, MD, MPH[o,p]

From the [a]Department of Speech & Hearing Science, The Ohio State University, Columbus, OH; [b]Division of Clinical Therapies & Inpatient Rehabilitation Program, Nationwide Children's Hospital, Columbus, OH; [c]Division of Biostatistics, School of Public Health, University of Minnesota, Minneapolis, MN; [d]Rockefeller Neuroscience Institute and West Virginia, University Medicine Children's Neuroscience Center, Morgantown, WV; [e]Division of Trauma and Burn Surgery, Children's National Hospital, Washington, DC; [f]Department of Rehabilitation Medicine, University of Washington, Seattle, WA; [g]Department of Pediatrics, Emory University School of Medicine, Atlanta, GA; [h]International Center for Spinal Cord Injury, Kennedy Krieger Institute and Johns Hopkins School of Medicine, Baltimore, MD; [i]Inpatient Rehabilitation Program, Nationwide Children's Hospital, Columbus, OH; [j]Division of Pediatric Psychology and Neuropsychology, Department of Pediatrics, The Ohio State University College of Medicine, and Nationwide Children's Hospital, Columbus, OH; [k]Division of Pediatric Rehabilitation Medicine, Cincinnati Children's Hospital Medical Center, Cincinnati, OH; [l]Departments of Pediatrics and Neurology and Rehabilitation Medicine, University of Cincinnati College of Medicine, Cincinnati, OH; [m]Division of Neurology, Department of Pediatrics, Nationwide Children's Hospital and The Ohio State University College of Medicine, Columbus, OH; [n]Division of Occupational Therapy and Physical Therapy, Cincinnati Children's Hospital Medical Center, Cincinnati, OH; [o]Abigail Wexner Research Institute at Nationwide Children's Hospital, Columbus, OH; and [p]Division of Emergency Medicine, Department of Pediatrics, The Ohio State University College of Medicine, and Nationwide Children's Hospital, Columbus, OH.

Abstract

Objective: Provide proof-of-concept for development of a Pediatric Functional Status eScore (PFSeS). Demonstrate that expert clinicians rank billing codes as relevant to patient functional status and identify the domains that codes inform in a way that reliably matches analytical modeling.

Design: Retrospective chart review, modified Delphi, and nominal group techniques.

Setting: Large, urban, quaternary care children's hospital in the Midwestern United States.

Participants: Data from 1955 unique patients and 2029 hospital admissions (2000-2020); 12 expert consultants representing the continuum of rehabilitation care reviewed 2893 codes (procedural, diagnostic, pharmaceutical, durable medical equipment).

Main Outcome Measures: Consensus voting to determine whether codes were associated with functional status at discharge and, if so, what domains they informed (self-care, mobility, cognition/ communication).

Results: The top 250 and 500 codes identified by statistical modeling were mostly composed of codes selected by the consultant panel (78%-80% of the top 250 and 71%-78% of the top 500). The results provide evidence that clinical experts' selection of functionally meaningful codes corresponds with codes selected by statistical modeling as most strongly associated with WeeFIM domain scores. The top 5 codes most strongly related

Research reported in this publication was supported by the *Eunice Kennedy Shriver* National Institute of Child Health & Human Development of the National Institutes of Health under Award Number R03HD101083. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Disclosures: none.

https://doi.org/10.1016/j.apmr.2023.03.025

to functional independence ratings from a domain-specific assessment indicate clinically sensible relationships, further supporting the use of billing data in modeling to create a PFSeS.

Conclusions: Development of a PFSeS that is predicated on billing data would improve researchers' ability to assess the functional status of children who receive inpatient rehabilitation care for a neurologic injury or illness. An expert clinician panel, representing the spectrum of medical and rehabilitative care, indicated that proposed statistical modeling identifies relevant codes mapped to 3 important domains: self-care, mobility, and cognition/communication.

Archives of Physical Medicine and Rehabilitation 2023;000:1−10

More than 18,000 children are hospitalized each year in the United States with acute neurologic injuries or illnesses,[1,2] a group of conditions that includes traumatic and nontraumatic brain, spinal cord, and peripheral nerve diagnoses. Many of these children require specialized medical care and rehabilitation services. The long-term sequelae of pediatric neurologic dysfunction and the potential negative effects on education, employment, health, and quality of life[3,4] pose high costs to families and society.[5,6] For example, an estimated $1 billion is spent managing pediatric traumatic brain injury (TBI)-associated hospitalizations annually in the United States,[7] but the evidence supporting long-term effectiveness of rehabilitation interventions for neurologic injuries (eg, TBI) and illnesses in children is weak.[5,8]

This gap in evidence is due to multiple challenges posed by conventional research methods. Although neurologic injuries and illnesses are a leading cause of morbidity and mortality in children, the overall incidence of each disease requiring hospitalization and/or rehabilitation is low, making the number of children admitted to an individual hospital in a given year relatively small.[2,6,9] Diverse etiologies, patient characteristics, and variability in care provision are also substantial obstacles to comparing the effectiveness of interventions.[8,10] Because of the challenges associated with performing randomized controlled trials on small, heterogeneous samples, researchers struggle to study outcomes and compare treatments.[2,11-13] Another challenge to clinical trials research is the trend toward earlier hospital discharge, leaving a substantial portion of rehabilitation for the home and school environments, where children are often lost to follow-up.[14,15] Thus, we must consider novel methods to track the long-term outcomes of these children to facilitate clinical trials and/or comparative effectiveness research (CER).[10]

CER aims to produce or synthesize evidence to inform best practices and policies for improving health care.[16] CER for pediatric rehabilitation would be advanced if researchers could use large, multicenter databases to overcome the challenges of studying low-incidence diagnoses. However, the lack of meaningful outcome measures within existing datasets limits their current use for CER. Although consensus is growing that institutions should use diagnosis-specific common outcome measures,[17-20] this practice remains inconsistent and extremely challenging to standardize in clinical and research settings.[21] As a result, it is difficult to pool data sources to assess the effectiveness of pediatric rehabilitation interventions.

When paired with meaningful, patient-centered outcomes, large data sources such as electronic medical records (EMRs) have the potential to help answer pressing questions in pediatric rehabilitation[11,22-24]; however, widely used, patient-centered outcome measures are lacking.[8] Medical claims data, which is widely available to researchers through existing national databases and partnerships with insurers and health networks, may represent a valuable resource.[25-27] Existing data sources contain information regarding diagnoses, procedures, pharmaceuticals, and durable medical goods related to an individual's health care encounter. Past studies found that medical claims data can be useful in detecting specific diagnoses,[28-30] and that the positive predictive value improves when algorithms use multiple codes.[31,32] Currently such databases are underused for CER because they do not include meaningful measures of patient outcomes.[33]

The primary objective of this study was to demonstrate the feasibility and proof-of-concept for the development of a Pediatric Functional Status Electronic-Score (e-score; PFSeS). The proposed PFSeS will use clinically relevant claims data from the EMR (ie, diagnoses, procedures, pharmaceuticals, and durable medical equipment [DME]) to accurately represent children's functional mobility, self-care, and cognitive/communication status at discharge. The central hypothesis is that claims data specific to a medical encounter, which are typically accurate and complete,[28,29,34] may be used to model children's functional status at the time of that encounter. Modeling functional status from claims data has been done successfully for other health conditions that include pediatric populations.[35-39] As a first step to this goal, consensus methods were used to determine which billing codes are relevant to a patient's functional status and whether these codes were associated with mobility, self-care, and cognitive/communication domain ratings on the criterion standard outcome measure in pediatric rehabilitation.

List of abbreviations:

CER	**comparative effectiveness research**
CPT	**Current Procedural Terminology**
CUI	**Concept Unique Identifier**
DME	**durable medical equipment**
EMR	**electronic medical record**
ICD	***International Classification of Diseases***
PFSeS	**Pediatric Functional Status Electronic-Score**
TBI	**traumatic brain injury**

Methods

Study design

A panel of clinical experts used modified Delphi procedures and nominal group technique to determine whether given billing codes were informative in modeling a patient's functional mobility, self-care, and cognitive/communication status at discharge from inpatient rehabilitation. Then, a data-driven approach was used to validate this process in a cohort of children with neurologic injuries or

illnesses who had WeeFIM ratings. The billing data from this cohort was used to assess whether the consensus process resulted in a selection of codes enriched with strong associations with WeeFIM ratings across the 3 domains.

Principal data source

This study was reviewed and approved by the institutional review board before data collection. The data were obtained from the EMR of a large, quaternary care children's hospital in the Midwestern United States. All children included in this study were admitted to the inpatient rehabilitation unit (within the children's hospital) between January 2000 and July 2020. During the rehabilitation portion of their hospitalization, children were under the care of physiatrists, but other services (eg, neurology, trauma, neurosurgery) were available as consultants or, in some cases, provided continued care during rehabilitation. All admitted patients with verified WeeFIM ratings were retained for this study to allow researchers to analyze the full spectrum of outcomes within this population. One patient was removed because of a parental request that their child's data not be included in research. Extracted billing data associated with their entire hospital admission (acute care through rehabilitation) included codes related to diagnoses (*International Classification of Diseases*; ICD-9 and ICD-10), procedures (Current Procedural Terminology; CPT), pharmaceutical (RxNorm), and DME. ICD is a classification system developed collaboratively between the World Health Organization and 10 international centers to promote similar classification, processing, and presentation of disease and mortality data.[40] CPT is a uniform coding system for medical services and procedures established by the American Medical Association within the United States.[41] The US National Library of Medicine created RxNorm as a normalized naming system for generic and branded drugs.[42] DME includes equipment, prosthetics, orthotics, and supplies, as coded by the Centers for Medicaid and Medicare Services in the United States.

Lead investigators (J.P.L., J.D.H., and J.C.L.) initially reviewed data to identify and exclude codes that were used in more than 90% of admissions. For example, CPT 97110 (therapeutic procedure on 1 or more areas, each lasting 15 minutes) was present in more than 98% of admissions. Similarly, codes that were used less than 10 times across all visits and across all children in the data set also were excluded. Ten was used as a threshold based on previous studies,[43,44] which indicate bias may arise in models including variables with fewer than 10 events. Lead clinical investigators (J.P.L. and J.C.L.) reviewed remaining codes and proposed the exclusion of additional codes as not informative of clinical outcomes. ICD-9, ICD-10, CPT, and RxNorm codes were mapped to Concept Unique Identifiers (CUIs) through the Unified Medical Language System Metathesaurus that groups billing codes according to distinct medical concepts.[45] All data analyses operated on CUIs and DME.

WeeFIM is a validated measure of functional status recommended for use in pediatric rehabilitative medicine[17] that rates degree of independence relative to self-care (8 items), mobility (5 items), and cognition/communication (5 items). In addition to specific demographic data, rehabilitation discharge WeeFIM ratings for all included patients were extracted from the EMR. At our institution, every patient from the inpatient rehabilitation unit receives ratings on 18 WeeFIM items at several points during the admission and after discharge. Each of the 18 items is rated on an ordinal scale, ranging from 1 (total assistance) to 7 (complete independence). WeeFIM at discharge were used in this study (ie, 0-7 ordinal ratings for each of the 18 individual items and a total rating for each of the 3 domains composed of a sum of the items in that category).

Modified Delphi technique to gather group opinion using electronic surveys

Twelve national pediatric experts served as consultants for this phase of the project. The panel included 6 physicians (trauma surgery, neurosurgery, neurology, and 3 physical medicine and rehabilitation) and 6 nonphysicians (speech-language pathology, physical therapy, occupational therapy, rehabilitation nursing, rehabilitation coordinator, and neuropsychology). All panelists had related degrees, training, and active licensure in their fields of practice. Four of the 5 nonphysicians were certified WeeFIM raters for rehabilitation patients. These interdisciplinary clinicians serve the continuum of pediatric neurologic injury, illness, and rehabilitative care. Consultants were divided into 3 groups, each of which included 2 physicians and 2 nonphysicians.

REDCap (Research Electronic Data Capture)[46,47] was used as the secure database management system for this project. Proposed excluded codes were sent for consultant review on the first REDCap survey. Consultants reviewed the excluded codes and identified those that they believed had potential in informing patients' functional status. Every code was reviewed by 4 consultants. If any consultant voted that a code should not be excluded, it was returned to the list for review in subsequent surveys.

In 3 subsequent REDCap surveys, consultants were presented with clusters of clinically related codes and asked to indicate whether each code in each cluster informed patients' functional status ("include"). For those marked as "include," consultants then identified the functional outcome domain the code informed: self-care, mobility, cognition/communication. Each round of surveys included codes carried over from previous surveys that had not yet met consensus (at least 3 of 4 consultants) to include, exclude, and determine the domain(s) of relevance (if included).

Nominal group technique

Because of the COVID-19 pandemic, the planned in-person consensus meeting for Fall 2021 was offered remotely to consultants who were not local to the study site. Consultants participated in a half-day meeting, with 4 consultants in-person and 6 consultants participating remotely via Webex. Two consultants were unable to attend the scheduled date. The objective of this consensus process was to discuss and vote on inclusion or exclusion of the codes that had not reached consensus during the modified Delphi procedure. For this phase, 70% agreement was required to meet consensus. Codes were reviewed in clusters of related codes. Using Webex Slido polling, consultants first voted without discussion on the inclusion of each code and/or the functional status domain it informs. If consensus was not reached on the initial vote, consultants discussed the code and its perceived applicability to functional status. Consultants completed one additional round of voting. At that point, a code was determined to either reach consensus or not. Lead clinical investigators (J.P.L. and J.C.L.) reviewed the final list of codes to ensure inclusion/exclusion consistency across comparable codes.

Data analysis

The patient cohort was summarized using descriptive summary measures of the demographics, WeeFIM impairment groups, and WeeFIM ratings. For patients with multiple hospitalizations, only information from the first hospitalization was included in the calculation of these summary measures (table 1). To evaluate how well the consultants identified codes associated with functional status, lead researchers sorted the codes based on the combined strength of their relationship to each of the 3 WeeFIM functional domains (self-care, mobility, cognition/communication). Researchers then assessed how many of the top 250 and top 500 strongest-related codes were selected by the consultants and how many were excluded during the consensus process. The codes were sorted using the ranking method of Henderson and Newton.[48] The ranking method works by accounting for a measure of the strength and uncertainty of the association between the code and the WeeFIM domain rating. For example, one code may have a large, estimated association with a WeeFIM domain rating, but the uncertainty of the estimated association is large. The previously described code would be ranked lower than another code that has an equally strong estimated association but lower uncertainty in the estimated strength of relationship. We estimated the strength of the relationship between each code and the WeeFIM domain ratings by fitting univariate simple linear models regressing the codes on the WeeFIM domain rating; the strength was measured as the estimated difference of the WeeFIM domain rating for those with and without the code. The uncertainty measure is the estimated standard error of the estimated difference. If the top 250 and 500 codes were mostly composed of codes selected during the consensus process, results would provide evidence that

Table 1 Baseline characteristics and WeeFIM outcomes of patients included in study sample

Variable	Overall, N=1955*	Female, N=841*	Male, N=1114*
Age, y	12 (6, 16)	12 (7, 16)	11 (6, 15)
Ethnicity			
Asian	0.9% (18)	1.5% (13)	0.4% (5)
Black	15% (301)	15% (124)	16% (177)
Hispanic	2.3% (44)	1.5% (13)	2.8% (31)
Multiple	0.5% (9)	0.7% (6)	0.3% (3)
Native American	0.1% (2)	0% (0)	0.2% (2)
Other	3.8% (75)	3.8% (32)	3.9% (43)
White	77% (1506)	78% (653)	77% (853)
WeeFIM impairment group†			
Amputations	0.2% (4)	0% (0)	0.4% (4)
Arthritis	0.4% (7)	0.5% (4)	0.3% (3)
Brain dysfunction	44% (867)	39% (329)	48% (538)
Burns	1.5% (30)	1.0% (8)	2.0% (22)
Cardiac disorders	2.6% (50)	2.4% (20)	2.7% (30)
Congenital disorders	0.3% (6)	0.5% (4)	0.2% (2)
Debility	1.9% (38)	2.5% (21)	1.5% (17)
Developmental disabilities	7.5% (146)	8.0% (67)	7.1% (79)
Major multiple trauma	3.9% (77)	3.9% (33)	3.9% (44)
Missing	1.0% (19)	1.5% (13)	0.5% (6)
Neurologic disorders	5.6% (109)	6.3% (53)	5.0% (56)
Orthopedic conditions	2.7% (52)	2.6% (22)	2.7% (30)
Other disabling impairments	5.9% (116)	10.0% (84)	2.9% (32)
Pain syndromes	0.1% (2)	0.1% (1)	<0.1% (1)
Spinal cord dysfunction	16% (320)	16% (132)	17% (188)
Stroke	5.7% (112)	5.9% (50)	5.6% (62)
Admission year			
2000-2004	20% (388)	20% (165)	20% (223)
2005-2009	24% (460)	25% (214)	22% (246)
2010-2014	24% (469)	23% (195)	25% (274)
2015-2020	33% (638)	32% (267)	33% (371)
WeeFIM scores			
Cognition total	28.0 (19.0, 34.0)	29.0 (20.0, 35.0)	27.0 (18.0, 33.0)
Mobility total	24.0 (18.0, 28.0)	24.0 (18.0, 28.0)	24.0 (17.0, 28.0)
Self-care total	40.0 (25.0, 48.0)	40.0 (28.0, 48.0)	39.0 (23.0, 48.0)
Total WeeFIM	91.0 (64.0, 106.0)	93.0 (69.0, 108.0)	89.0 (62.0, 105.0)

NOTE. The total population is shown in male/female subgroups to allow future research with larger groups to consider sex-based differences. For patients with multiple visits, baseline characteristics at their first study visit are reported. Medians and lower/upper quartiles are displayed for continuous variables and percents and counts for discrete variables.

Abbreviation: IQR, interquartile range.

* Median (IQR); % (n)

† WeeFIM Impairment Group is the primary reason for a rehabilitation admission (it does not indicate the exclusion of other diagnoses or impairments).

the statistical modeling reliably identified the codes most clinically relevant and strongly associated with functional status for each of the 3 domains. Using the same ranking approach, we investigated the top 5 codes for each WeeFIM domain.

Results

Initial analysis included data from 1955 unique patients and 2029 hospital admissions. Children included in the cohort were predominantly White and male, with a median age of 12 years; nearly one half of the sample had brain dysfunction as their primary WeeFIM impairment group. For the cognition/communication domain, ratings skew toward higher levels of function overall. For the self-care and mobility domains, ratings skew toward higher levels of function with a large group who have very low function (fig 1).

The aforementioned code selection filtered 9306 initial codes to 1382 codes after the consensus process (fig 2). The Henderson and Newton ranking procedure displays the magnitudes of the estimated associations standardized by their standard errors on the y-axis (fig 3). A vast majority of the top 250 (blue) and top 500 (red) codes for each of the 3 domains were composed of codes included by the consultants: self-care (78% and 78%, respectively), mobility (78% and 75%), and cognition (80% and 71%). Among the 1382 total selected codes, 569 were included during the consensus process as being relevant for self-care, 517 as relevant for mobility, and 456 for cognition/communication (table 2). To interpret the first line in table 2, among the 1382 codes included by the consultants, 196 appeared in the top 250 ranked codes for the self-care domain, accounting for approximately 80% of the top 250 codes and accounting for approximately 10% of the total 1382 codes. Similarly, among these 1382 codes, 388 appeared in the top 500 ranked codes for the self-care domain, accounting for 80% of the top 500 and 30% of the 1382 included codes. Further, the proportion of included codes among those in the top 250 and top 500 was greater than the proportion of excluded codes among the top 250 and top 500 for all 3 domains. For self-care, 14% and 28% of all included codes were in the top 250 and top 500, respectively, whereas 7% and 15% of excluded codes were in the top 250 and top 500. Similar trends hold for the mobility and cognition/communication domains. Thus, the included codes overall were more strongly associated with WeeFIM domain ratings and a higher proportion of included codes had a stronger relationship with the WeeFIM domain ratings compared with excluded codes.

The top 5 codes most associated with each of the discharge WeeFIM domain ratings were determined by the ranking procedure (table 3). For example, the use of enteral pump was the most strongly related code to the cognition/communication WeeFIM domain rating and was associated with lower discharge WeeFIM rating of 12.8 (SE=1.2). Although visual reinforcement audiometry had a larger estimated association, its SE was larger and was thus ranked lower by the Henderson and Newton procedure. Baseline characteristics and WeeFIM outcomes of patients included in study sample by highly associative billing codes are shown in supplemental table S3 (available online only at http://www.archives-pmr.org/).

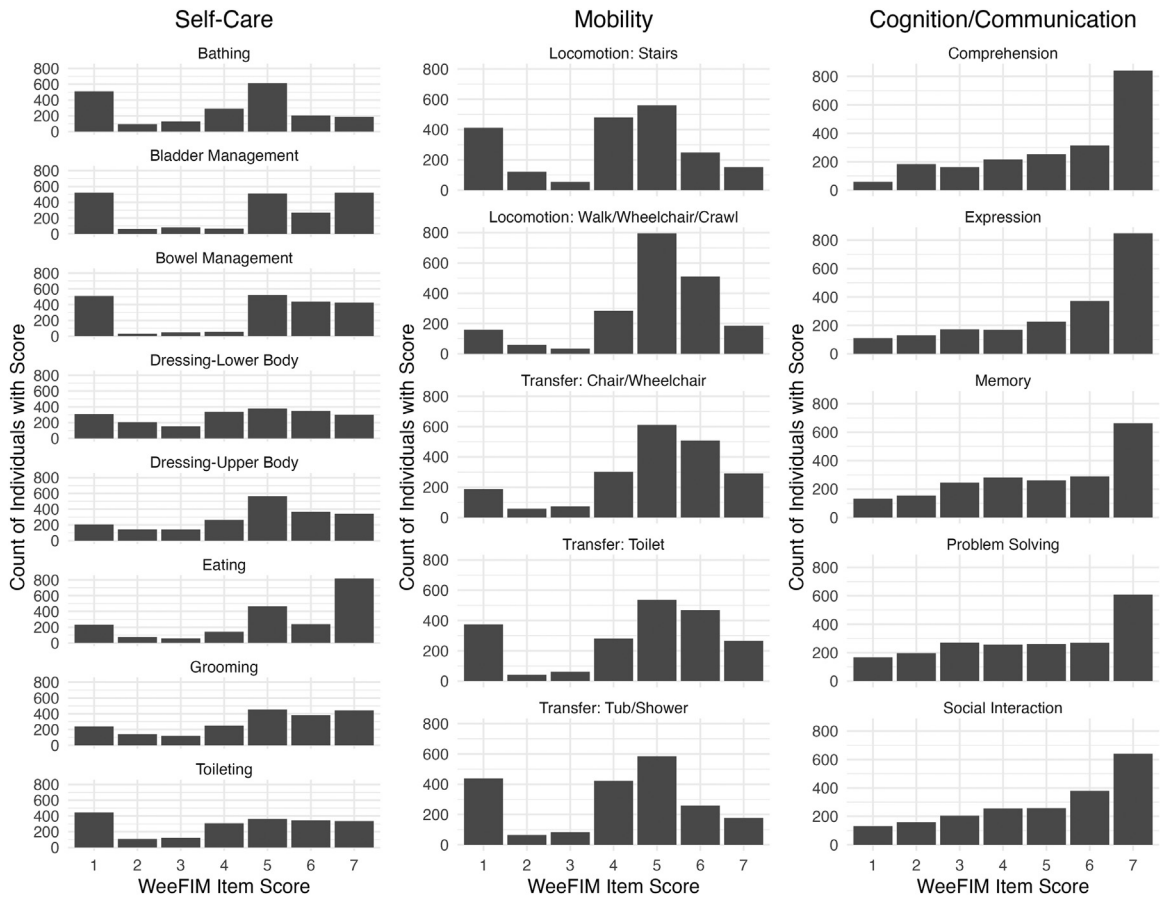

Fig 1 WeeFIM rating distribution across all domains.

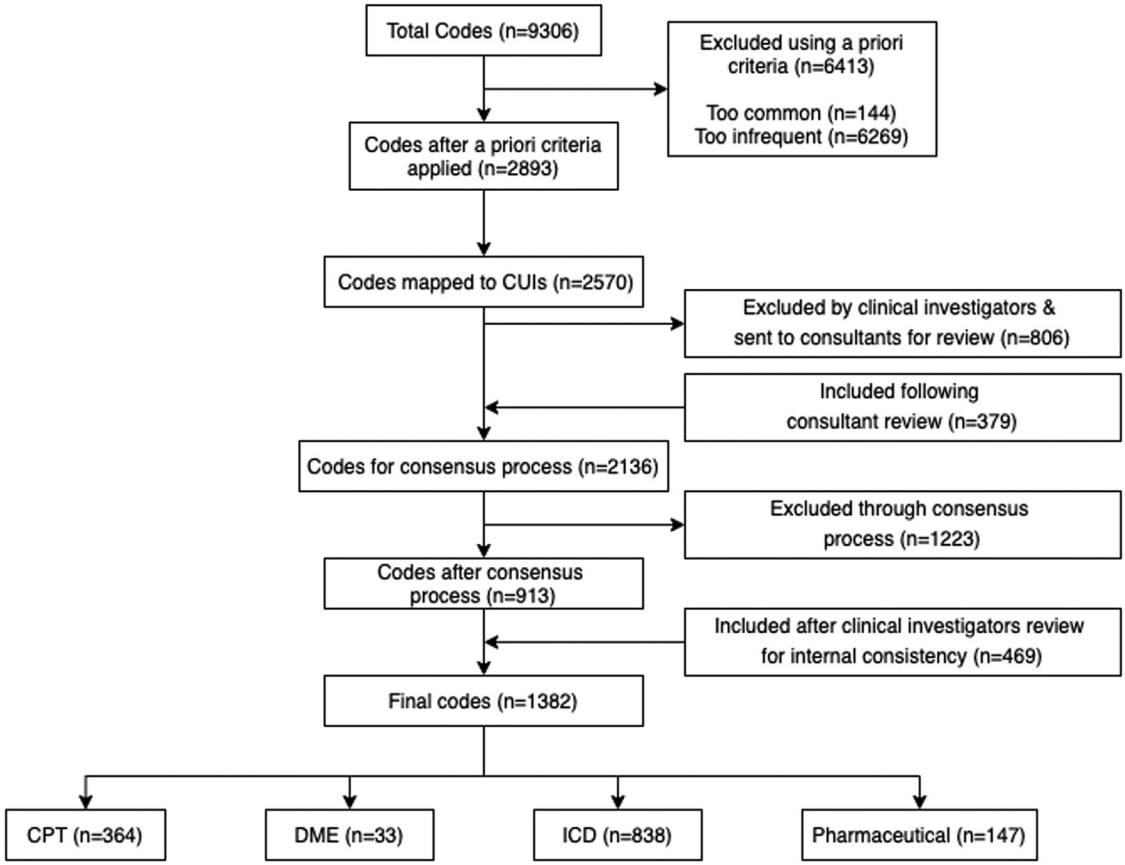

Fig 2 Flow diagram illustrating the process of billing code screening for codes postulated to be related to patients' functional status overall and for domains of self-care, mobility, and cognition/communication.

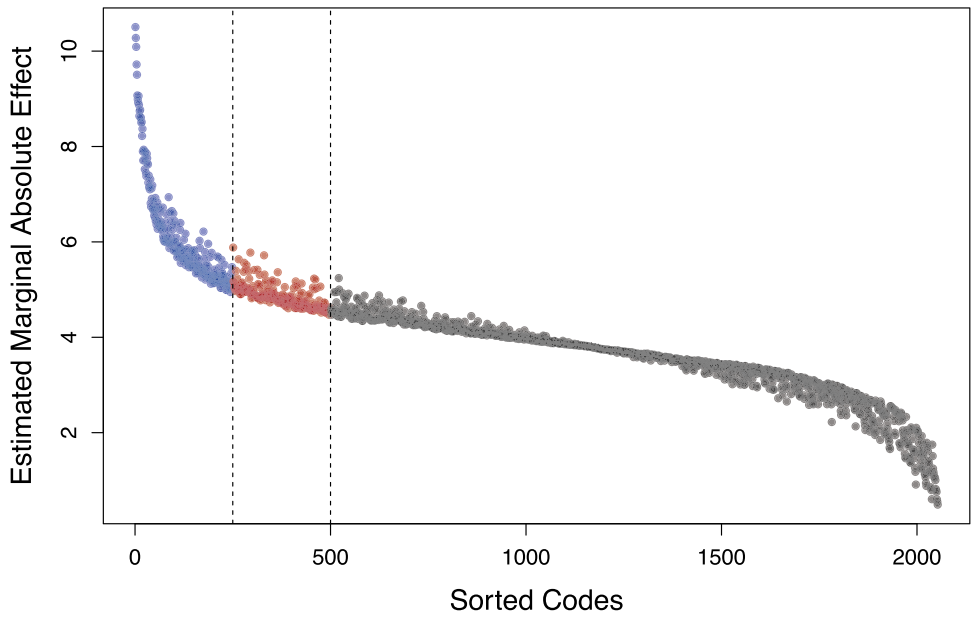

Fig 3 Relationship between the marginal mean difference of overall WeeFIM ratings between those with and without a given code, as measured by the estimated mean difference standardized by the standard deviation of the difference (y-axis) and the ranking of each code as determined by the method of Henderson and Newton (x-axis). Codes in the top 250 are colored blue and the remaining codes in the top 500 are colored red.

Table 2 Top 250 and 500 codes identified by statistical modeling compared to results from the consensus process

Domain	Code Status	Total Codes (N)	Top 250 (n)	Proportion Among Top 250	Proportion in Top 250 Among Total Codes	Top 500 (n)	Proportion Among Top 500	Proportion in Top 500 Among Total Codes
Self-Care	Included by consultants	1382	196	0.8	0.1	388	0.8	0.3
Self-Care	Included for domain	569	91	0.4	0.2	160	0.3	0.3
Self-Care	Excluded by consultants	754	54	0.2	0.1	112	0.2	0.1
Mobility	Included by consultants	1382	196	0.8	0.1	377	0.8	0.3
Mobility	Included for domain	517	63	0.3	0.1	129	0.3	0.3
Mobility	Excluded by consultants	754	54	0.2	0.1	123	0.2	0.2
Cognition/ Communication	Included by consultants	1382	199	0.8	0.1	359	0.7	0.3
Cognition/ Communication	Included for domain	456	80	0.3	0.2	130	0.3	0.3
Cognition/ Communication	Excluded by consultants	754	51	0.2	0.1	141	0.3	0.2

NOTE. The summaries are displayed for codes that were determined to be related to functional status (irrespective of domain) by the consensus process ("Included by consultants"), for codes that were determined to be related to a specific WeeFIM domain ("Included for domain"), and for codes that were determined to not be related to functional status ("Excluded by consultants").

Discussion

Researchers need transformative tools to accelerate the integration of big data and data science into rehabilitation research. The development, expansion, and use of multiple data resources and research methods are needed to assess how we diagnose, treat, and monitor long-term effects of pediatric neurologic injuries and illnesses.[49] This study sought to demonstrate that expert clinicians could categorize billing codes as relevant (or not) to patient functional status and functional domains in a way that reliably matched analytical modeling. This work is the first step in developing a pediatric functional status electronic score (that is, PFSeS) that uses clinically relevant claims data from the EMR (that is, diagnoses, procedures, pharmaceuticals, and DME) to model children's functional mobility, self-care, and cognitive/communication status. Using modified Delphi and nominal group consensus techniques, findings indicate that experts agree with analytical modeling of the PFSeS from billing data. Findings suggest that the PFSeS modeling identifies appropriate and clinically relevant codes that inform a patient's status after inpatient rehabilitation. In addition, statistical modeling appropriately identifies the specific domains that codes inform, as demonstrated by the numerous clinician-included codes with strong associations (as measured by univariable linear models) with WeeFIM ratings.

Table 3 Top 5 codes for each WeeFIM domain as ranked by the method of Henderson and Newton

Domain	Type	Description	Effect	SE
Cognition/ Communication	DME	DME - Enteral Pump	−12.8	1.2
Cognition/ Communication	DME	DME - Enteral Feedings/Supplies	−11.6	0.9
Cognition/ Communication	CPT	Visual reinforcement audiometry	−16.0	2.0
Cognition/ Communication	RXNORM	Levetiracetam 100 mg/mL oral solution	−11.0	1.0
Cognition/ Communication	CPT	Sensory integrative techniques to enhance sensory processing and promote adaptive responses to environmental demands, direct (one-on-one) patient contact, each 15 minutes	−12.0	1.4
Mobility	ICD	Conversion disorder	12.6	0.9
Mobility	DME	DME - Patient Lifts	−11.6	1.1
Mobility	DME	DME - Enteral Pump	−10.8	1.1
Mobility	DME	DME - Enteral Feedings/Supplies	−9.8	0.8
Mobility	RXNORM	Diazepam 1 mg/mL oral solution	−13.1	1.6
Self-care	DME	DME - Enteral Feedings/Supplies	−20.2	1.5
Self-care	DME	DME - Enteral Pump	−21.9	2.0
Self-care	ICD	Conversion disorder	19.8	1.7
Self-care	CPT	Visual reinforcement audiometry	−25.4	3.2
Self-care	CPT	Subsequent inpatient pediatric critical care, per day, for the evaluation and management of a critically ill infant or young child, 2 through 5 years of age	−19.6	2.3

NOTE. Also displayed are the marginal mean difference of the WeeFIM domain rating between those with and without each code and standard errors of each estimated mean difference. These codes rank among the top 5 most strongly related codes for the given domain. The ranking considers the strength of the association of the code with the domain specific WeeFIM score and the standard error of this estimated association. Codes with large associations and small standard errors of the estimated associations will generally rank higher.
Abbreviations: CPT, Current Procedural Terminology; DME, durable medical equipment; ICD, *International Classification of Diseases*.

As the findings illustrate, the top 5 codes most strongly associated with domain-specific discharge WeeFIM ratings hold clinically sensible relationships. Results also reflect how individual codes are associated with different patient phenotypes and their functional status, consistent with previous work demonstrating that claims data can detect specific diagnoses.[28-30] Although some codes appeared across multiple domains, other codes were domain specific. Further, codes could have both positive and negative univariate associations with the PFSeS model. For example, within the self-care domain, 2 codes for DME (enteral feeding supplies and enteral pump) show the strongest relationships to total self-care domain ratings at discharge but also relate to the mobility and cognition/communication total domain ratings. Oral self-feeding is an activity requiring motor strength and coordination and cognition of the numerous implements and steps. Forty-four percent of children in our sample were classified within the brain dysfunction WeeFIM impairment group, so it is not surprising that this important and complex skill was affected in our cohort. Common conditions for enteric feeding include neurologic diseases and injuries, motor neuron diseases, brain tumors, and disorders of consciousness.[50] An example would be a patient with severe TBI resulting in static encephalopathy or patients with progressive neurodegenerative processes such as Lennox-Gastaut syndrome. Those who have an enteral feeding supplies DME code have a 20.2 point lower average discharge WeeFIM rating compared with those who do not. Patients with an enteral pump DME code have a 21.9 point lower WeeFIM rating than those without.

Another illustration is the code for conversion disorder. Patients with conversion disorder had greater overall functioning and no cognitive-communication limitations. In conversion disorder, symptoms of altered voluntary motor, cognitive, or sensory function are inconsistent with demonstrated preserved physiological testing.[51] Statistical modeling identified that self-care discharge WeeFIM ratings for patients with conversion disorder were on average 19.8 points greater compared with patients with other diagnoses. Because the etiology of conversion disorder is not physiological, we would expect that after therapy, they would have return to their functional baseline, which should be greater than that of a patient with an organic cause for their deficit.

The proposed PFSeS is modeled using WeeFIM ratings at discharge from inpatient rehabilitation. While the National Institute of Neurological Disorders and Stroke includes WeeFIM in its recommended common data elements for pediatric brain and spinal cord injuries,[52] WeeFIM data are not widely available across institutions. Typically, only large pediatric care centers with inpatient rehabilitation units use WeeFIM and it is not used universally to track outcomes in children. Even within the institution studied, patients discharged without an inpatient rehabilitation admission have no WeeFIM ratings to track outcomes. In addition, a known limitation of the WeeFIM is that it has a ceiling where it is difficult to distinguish among greater levels of function and among older children. To compare the effectiveness of interventions across institutions, diagnoses, and developmental stages, we need to develop outcome models that are more appropriate for a broader range of children. Given larger datasets, it may be possible to identify other methods to categorize outcomes for children with neurologic injuries and illnesses, which could include patient-reported outcome measures or a tool like the PFSeS that uses data from medical encounters.[33]

Multiple data elements are available thanks to advances in technology, health informatics, and standardized outcome measures that can and should be used to improve the effectiveness of rehabilitation interventions.[33] Unfortunately, our current systems of care lack connections that allow us to track outcomes from inpatient to outpatient settings.[33] Gathering key billing data available in a patient's admission record and modeling that data using the WeeFIM allows us to create a PFSeS. A PFSeS is valuable in that it can model a patient's mobility, self-care, and cognition/communication status using CPT, pharmaceutical, DME, and diagnostic codes. The use of discharge WeeFIM to construct the PFSeS allows researchers to model functional status outcomes at the point of discharge. Ultimately, the goal of this work is to create a PFSeS to assess patient status for any admission that includes associated billing data.

Study limitations

This study has limitations that should be considered as this work moves forward. The switch from ICD-9 to ICD-10 presents challenges for billing data analyzed over long periods of time. Our analyses accounted for this by mapping to CUIs. We note that this may not always result in perfect conceptual matches or may miss relationships between ICD-9 and ICD-10 codes. Individual practices in billing and coding may also change over time, resulting in heterogeneity in the relationships of specific codes to WeeFIM ratings; again, mapping to CUIs can help account for billing and coding changes by linking codes to medical concepts that may be more distinct over time. We acknowledge that the generalizability of our findings also may be limited by possible institution-specific coding/billing practices. When attempting to generalize outside of the United States, other than ICD (which is an international coding system adopted by the World Health Organization), analyses will need to account for different coding standards and payor systems. While this study describes the patient age of our sample, it does not account for age in assessing for functional status. Future studies with larger datasets that include the WeeFIM can adjust for age and consider using a converted score like the Developmental Functional Quotient.[53]

Conclusions

This project offers a first step in efforts to improve the ability of researchers to assess the functional status of children who receive inpatient care for neurologic injuries or illnesses. The long-term goal of this research is to create a tool that can be applied to quality improvement projects and in CER. Consensus agreement by an expert clinician panel, representing the spectrum of medical and rehabilitative care received by these patients, indicates that the statistical modeling identifies relevant codes mapped to 3 important domains: mobility, self-care, and cognition/communication. Next steps for this work include further analysis to create the PFSeS and then validation of the scoring system on a larger sample of retrospective patient data from various institutions and on a prospective cohort. Adding patient-reported outcomes measures to prospective validation of the PFSeS will further strengthen the scoring system. Development of the PFSeS for use in CER has the potential to inform the multi-site development and comparison of rehabilitation interventions. Ultimately, this may allow us to track and compare outcomes more effectively for children with neurologic illnesses or injuries. By building an outcome measurement that can be applied to data beyond the limited number of children who do receive inpatient rehabilitation, we can increase the number of children whose care can be assessed and increase evidence-based services.

Keywords

Child; Comparative effectiveness research; Delphi technique; Function; Pediatrics; Rehabilitation

Corresponding author

Jennifer Lundine, 110 Pressey Hall, 1070 Carmack Rd., Columbus, OH 43214. *E-mail address:* lundine.4@osu.edu.

Acknowledgments

We gratefully acknowledge Nicole Thompson for her work on this project.

References

1. Taylor CA, Bell JM, Breiding MJ, Xu L. Traumatic brain injury–related emergency department visits, hospitalizations, and deaths—United States, 2007 and 2013. MMWR Surveill Summ 2017;66:1–16.
2. Dhillon JK, Shi J, Janezic A, Wheeler KK, Xiang H, Leonard JC. U.S. estimates of pediatric spinal cord injury: implications for clinical care and research planning. J Neurotrauma 2017;34:2019–26.
3. Corrigan JD, Hammond FM. Traumatic brain injury as a chronic health condition. Arch Phys Med Rehabil 2013;94:1199–201.
4. Corrigan JD, Bogner JA. Neighborhood characteristics and outcomes after traumatic brain injury. Arch Phys MedRehabil 2008;89:912–21.
5. Centers for Disease Control and Prevention. Report to Congress: The management of traumatic brain injury in children. National Center for Injury Prevention and Control; Division of Unintentional Injury Prevention, 2018. Available at: https://www.cdc.gov/traumaticbraininjury/pdf/reportstocongress/managementoftbiinchildren/TBI-Reportto-Congress-508.pdf. Accessed February 25, 2020.
6. Saunders LL, Selassie A, Cao Y, Zebracki K, Vogel LC. Epidemiology of pediatric traumatic spinal cord injury in a population-based cohort: 1998-2012. Topics Spinal Cord Injury Rehabil 2015;21:325–32.
7. Schneier AJ, Shields BJ, Hostetler SG, Xiang H, Smith GA. Incidence of pediatric traumatic brain injury and associated hospital resource utilization in the United States. Pediatrics 2006;118:483–92.
8. Forsyth R, Kirkham F. Predicting outcome after childhood brain injury. CMAJ 2012;184:1257–64.
9. Stanley RM, Bonsu BK, Zhao W, Ehrlich PF, Rogers AJ, Xiang H. US estimates of hospitalized children with severe traumatic brain injury: implications for clinical trials. Pediatrics 2012;129:e24–30.
10. Forsyth RJ, Salorio CF, Christensen JR. Modelling early recovery patterns after paediatric traumatic brain injury. Arch Dis Child 2010;95:266–70.
11. Mayo CS, Matuszak MM, Schipper MJ, Jolly S, Hayman JA. Ten Haken RK. Big data in designing clinical trials: opportunities and challenges. Front Oncol 2017;7:187.
12. Neuman HB, Greenberg CC. Comparative effectiveness research: opportunities in surgical oncology. Semin Radiat Oncol 2014;24:43–8.
13. Luce BR, Kramer JM, Goodman SN, et al. Rethinking randomized clinical trials for comparative effectiveness research: the need for transformational change. Ann Intern Med 2009;151:206–9.
14. Knippel T, Houtrow A. 10-year trends in pediatric inpatient rehabilitation. 2016 AAP Annual Meeting, Feb-2016; American Journal of Physical Medicine & Rehabilitation, Feb-2016. Available at: https://www.physiatry.org/mpage/10YearTrendsinPe. Accessed September 10, 2019.
15. Spaw AJ, Lundine JP, Johnson SA, et al. Follow-up care adherence after hospital discharge in children with traumatic brain injury. J Head Trauma Rehabil 2018;33:E1–E10.
16. Sox HC, Greenfield S. Comparative effectiveness research: a report from the institute of medicine. Ann Intern Med 2009;151:203.
17. McCauley SR, Wilde EA, Anderson VA, et al. Recommendations for the use of common outcome measures in pediatric traumatic brain injury research. J Neurotrauma 2012;29:678–705.
18. Grinnon ST, Miller K, Marler JR, et al. National Institute of Neurological Disorders and Stroke Common Data Element Project: approach and methods. Clin Trials 2012;9:322–9.
19. Zumsteg JM, Ennis SK, Jaffe KM, et al. Quality of care indicators for the structure and organization of inpatient rehabilitation care of children with traumatic brain injury. Arch Phys Med Rehabil 2012;93:386–93. e1.
20. Pollack MM, Holubkov R, Glass P, et al. Functional status scale: a new pediatric outcome measure. Pediatrics 2009;124:e18–28.
21. Greene NH, Kernic MA, Vavilala MS, Rivara FP. Variation in pediatric traumatic brain injury outcomes in the United States. Arch Phys Med Rehabil 2014;95:1148–55.
22. Curtis LH, Brown J, Platt R. Four health data networks illustrate the potential for a shared national multipurpose big-data network. Health Aff (Millwood) 2014;33:1178–86.
23. Westra BL, Latimer GE, Matney SA, et al. A national action plan for sharable and comparable nursing data to support practice and translational research for transforming health care. J Am Med Inform Assoc 2015;22:600–7.
24. Faroqi-Shah Y. The rise of big data in neurorehabilitation. Semin Speech Language 2016;37:3–9.
25. Steiner C, Elixhauser A, Schnaier J. The healthcare cost and utilization project: An overview. Eff Clin Pract 2002;5:143–51.
26. Agency for Healthcare Research and Quality. Healthcare Cost and Utilization Project (HCUP), 2019. Published Available at: http://www.ahrq.gov/data/hcup/index.html. Accessed February 19, 2019.
27. Children's Hospital Association. Data Analytics & Research: PHIS. Available at: https://www.childrenshospitals.org/programs-and-services/data-analytics-and-research/pediatric-analytic-solutions/pediatric-health-information-system. Accessed February 19, 2019.
28. Cohen S, Jannot AS, Iserin L, Bonnet D, Burgun A, Escudié JB. Accuracy of claim data in the identification and classification of adults with congenital heart diseases in electronic medical records. Arch Cardiovasc Dis 2019;112:31–43.
29. Remontet L, Mitton N, Couris CM, et al. Is it possible to estimate the incidence of breast cancer from medico-administrative databases? Eur J Epidemiol 2008;23:681–8.
30. Wang Y, Sharpe-Stimac M, Cross PK, Druschel CM, Hwang SA. Improving case ascertainment of a population-based birth defects registry in New York State using hospital discharge data. Birth Defects Res Part A Clin Mol Teratol 2005;73:663–8.
31. Chung CP, Rohan P, Krishnaswami S, McPheeters ML. A systematic review of validated methods for identifying patients with rheumatoid arthritis using administrative or claims data. Vaccine 2013;31(Suppl 10):K41–61.
32. Parker JP, McCombs JS, Graddy EA. Can pharmacy data improve prediction of hospital outcomes? Comparisons with a diagnosis-based comorbidity measure. Med Care 2003;41:407–19.
33. Keeney T, Kumar A, Erler KS, Karmarkar AM. Making the case for patient-reported outcome measures in big-data rehabilitation research: implications for optimizing patient-centered care. Arch Phys Med Rehabil 2022;103:S140–5.
34. Cronk CE, Malloy ME, Pelech AN, et al. Completeness of state administrative databases for surveillance of congenital heart disease. Birth Defects Res Part A Clin Mol Teratol 2003;67:597–603.
35. Alessandrini EA, Alpern ER, Chamberlain JM, et al. Developing a diagnosis-based severity classification system for use in emergency medical services for children. Acad Emerg Med 2012;19:70–8.
36. Alessandrini EA, Alpern ER, Chamberlain JM, Shea JA, Gorelick MH. A new diagnosis grouping system for child emergency department visits. Acad Emerg Med 2010;17:204–13.
37. Donzé J, Aujesky D, Williams D, Schnipper JL. Potentially avoidable 30-day hospital readmissions in medical patients: derivation and validation of a prediction model. JAMA Intern Med 2013;173:632–8.

38. Yurkovich M, Avina-Zubieta JA, Thomas J, Gorenchtein M, Lacaille D. A systematic review identifies valid comorbidity indices derived from administrative health data. J Clin Epidemiol 2015;68:3–14.
39. Fishman PA, Shay DK. Development and estimation of a pediatric chronic disease score using automated pharmacy data. Med Care 1999;37:874–83.
40. National Center for Health Statistics. International Classification of Diseases 10th Revision (ICD-10). .Centers for Disease Control and Prevention, Department of Health and Human Services; 2017. Available at: https://www.cdc.gov/nchs/data/dvs/icd10fct.pdf Accessed December 2, 2022.
41. American Medical Association. CPT® overview and code approval. . PublishedAvailable at https://www.ama-assn.org/practice-management/cpt/cpt-overview-and-code-approval#:~:text=CPT%C2%AE%20code%3F-,What%20is%20a%20CPT%C2%AE%20code%3F,reporting%2C%20increase%20accuracy%20and%20efficiency. Accessed December 2, 2022.
42. National Library of Medicine. RxNorm Overview. Unified Medical Language System (UMLS). . Published Available at https://www.nlm.nih.gov/research/umls/rxnorm/overview.html Accessed December 2, 2022.
43. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol 1996;49:1373–9.
44. Peduzzi P, Concato J, Feinstein AR, Holford TR. Importance of events per independent variable in proportional hazards regression analysis. II. Accuracy and precision of regression estimates. J Clin Epidemiol 1995;48:1503–10.
45. UMLS® Reference Manual [Internet]. Metathesaurus, 2009. National Library of Medicine; 2009. Available at: https://www.ncbi.nlm.nih.gov/books/NBK9684/. Accessed December 15, 2022.
46. Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019;95:103208.
47. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap): a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009;42:377–81.
48. Henderson NC, Newton MA. Making the cut: improved ranking and selection for large-scale inference. J R Stat Soc B 2016;78:781–804.
49. Federal Coordinating Council for Comparative Effectiveness Research. Report to the President and Congress. Washington, DC: US Department of Health and Government Statistics; 2009.
50. Rahnemai-Azar AA, Rahnemaiazar AA, Naghshizadian R, Kurtz A, Farkas DT. Percutaneous endoscopic gastrostomy: indications, technique, complications and management. World J Gastroenterol 2014;20:7739–51.
51. O'Neal MA, Baslet G. Treatment for patients with a functional neurological disorder (conversion disorder): an integrated approach. Am J Psychiatry 2018;175:307–14.
52. Adelson PD, Pineda J, Bell MJ, et al. Common data elements for pediatric traumatic brain injury: recommendations from the working group on demographics and clinical assessment. J Neurotrauma 2012;29:639–53.
53. Suskauer SJ, Slomine BS, Inscore AB, Lewelt AJ, Kirk JW, Salorio CF. Injury severity variables as predictors of WeeFIM scores in pediatric TBI: time to follow commands is best. J Pediatr Rehabil Med 2009;2:297–307.