Title
Identifying neonatal intensive care (NICU) admissions using administrative claims data.

Journal
Journal of Neonatal-Perinatal Medicine, 16(4)

Authors
Vance, A
Bell, S
Tilea, A
et al.

Publication Date
2023

DOI
10.3233/NPM-230014

Peer reviewed

HHS Public Access
Author manuscript
J Neonatal Perinatal Med. Author manuscript; available in PMC 2024 March 06.

Published in final edited form as:
J Neonatal Perinatal Med. 2023 ; 16(4): 709–716. doi:10.3233/NPM-230014.

Identifying neonatal intensive care (NICU) admissions using administrative claims data

A.J. Vance[a,b,*], **S. Bell**[c], **A. Tilea**[c], **D. Beck**[d], **K.M. Tabb**[e], **K. Zivin**[f,g,h]

[a]Center for Health Policy and Health Services Research, Henry Ford Health, Detroit, MI, USA

[b]College of Nursing, Michigan State University, East Lansing, MI, USA

[c]Department of Obstetrics and Gynecology, University of Michigan, Ann Arbor, MI, USA

[d]UCLA School of Nursing, Los Angeles, CA, USA

[e]University of Illinois at Urbana-Champaign, School of Social Work, Urbana, IL, USA

[f]Institute for Healthcare Policy and Innovation, University of Michigan, Ann Arbor, MI, USA

[g]Department of Learning Health Sciences, University of Michigan Medical School, Ann Arbor, MI, USA

[h]Department of Psychiatry, University of Michigan Medical School, Ann Arbor, MI, USA

Abstract

BACKGROUND: To define a method for identifying neonatal intensive care unit (NICU) admissions using administrative claims data.

METHODS: This was a retrospective cohort study using claims from Optum's de-identified Clinformatics® Data Mart Database (CDM) from 2016 – 2020. We developed a definition to identify NICU admissions using a list of codes from the *International Classification of Diseases, 10th Revision, Clinical Modification* (ICD-10-CM), Current Procedural Terminology (CPT), and revenue codes frequently associated with NICU admissions. We compared agreement between codes using Kappa statistics and calculated positive predictive values (PPV) and 95% confidence intervals (CI).

RESULTS: On average, revenue codes (3.3%) alone identified more NICU hospitalizations compared to CPT codes alone (1.5%), whereas the use of CPT *and* revenue (8.9%) and CPT *or* revenue codes (13.7%) captured the most NICU hospitalizations, which aligns with rates of preterm birth. Gestational age alone (4.2%) and birthweight codes alone (2.0%) identified the least

*Address for correspondence: Ashlee J. Vance, Center for Health Policy and Health Services Research, Henry Ford Health, Detroit, MI, USA. Tel.: +1 313 799 7225; avance2@hfhs.org.

Conflict of interest disclosure
The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Ethics approval statement
The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the University of Michigan Medical School Institutional Review Board (HUM00188304, approved 9/16/2020).

Supplementary materials
Codes used in NICU Definition.

number of potential NICU hospitalizations. Setting CPT codes as the standard and revenue codes as the "test,", revenue codes resulted in identifying 86% of NICU admissions (sensitivity) and 97% of non-NICU admissions (specificity).

CONCLUSIONS: Using administrative data, we developed a robust definition for identifying neonatal admissions. The identified definition of NICU codes is easily adaptable, repeatable, and flexible for use in other datasets.

Keywords

Claims data; infant; neonatal intensive care unit; validation

1. Introduction

Infants cared for in the Neonatal Intensive Care Unit (NICU) is a heterogenous population consisting of infants born preterm (<34 weeks' gestation), late preterm (34–36 weeks' gestation), and full-term (>37 weeks' gestation) who receive specialized care for a variety of low acuity and highly complex conditions. Although the NICU primarily cares for those born preterm(<37 weeks' gestation) or born very low birthweight (VLBW) (<1500 grams), approximately 40%-49% of NICU admissions include infants born full-term (i.e.,>37 weeks' gestation) [1, 2]. Term infants may be admitted to the NICU for congenital abnormalities, difficulty with birth transition (e.g., birth trauma or respiratory issues), hypoglycemia, or drug exposure in utero [3]. Over the past decade, survival rates have increased for the most vulnerable and medically complex infants [4], regardless of gestational age at birth or birthweight, due to advances in medical and nursing care [5]. Given the increasing prevalence of infant survival beyond the NICU, investing in high-quality methods to correctly identify NICU admissions is vital to our understanding of long-term infant health, epidemiology and outcomes research and quality improvement.

Health services research often repurposes administrative billing or claims data, allowing for epidemiological surveillance of medical conditions across populations. Historically, neonatal health services research has focused on specific subgroups, such as infants born very low birthweight (VLBW) (i.e., less than 1500 grams), extremely preterm (i.e.,<27 weeks' gestation), or specific diagnoses such as necrotizing enterocolitis, congenital heart defects, or hypoxic-ischemic encephalopathy [6, 7]. There is evidence that events such as very low birthweight, cesarean delivery, or maternal hypertension are coded with a high degree of accuracy in administrative data [8]. Furthermore, several validation studies are available for identifying preterm births using gestational age categories [9–11] or birthweight [12, 13]. Presumably, all infants born less than 35 weeks' gestation will be admitted to the NICU for care; however, algorithms relying on infant age or weight alone fail to account for infants born at or near term (i.e., not preterm) or capture infants whose reason for admission is a medically complex condition. Currently, no definition is available to identify NICU admissions (e.g., any or all infants admitted to the NICU). Such a definition would address a significant gap within the literature, advance perinatal and neonatal health services research, and further our understanding of how NICU care impacts future infant health and development. Thus, this study aimed to define a method for identifying NICU admissions using administrative claims data.

2. Methods

Using Optum's de-identified Clinformatics® Data Mart Database (CDM), we developed a definition to identify NICU admissions. CDM includes a statistically de-identified large claims data warehouse of administrative health claims for approximately 140 million children and adults from all 50 states, including approximately 13 million annual private insurance lives. Data include enrollees covered by employer-sponsored and individual insurance, including Health Insurance Marketplace plans. The University of Michigan Institutional Review Board deemed this project exempt, as it uses de-identified administrative claims data.

2.1. Algorithm sample

Our study population included all hospital deliveries between 2016 and 2020. All birthing individuals were identified who met the following criteria: ages 15–44, with one insurance plan, and continuous enrollment during the entire calendar year of the delivery. These birthing individuals were then linked to their newborns (See Fig. 1).

2.2. Statistical analysis

Hospitals use three standardized coding systems to describe services provided to patients: (1) *International Classification of Diseases, 10th Revision, Clinical Modification* (ICD-10-CM) diagnosis codes (descriptive codes that evaluate a patient's condition, injury), (2) Current Procedural Terminology (CPT) codes to identify medical services and procedures that occurred during a patient's visit, and (3) revenue codes that denote hospital services provided to a patient (e.g., emergency room, intensive care, etc.). We generated a list of ICD-10-CM, CPT, and revenue codes frequently associated with a NICU admission identified from published literature and neonatal clinical research experts (See supplemental materials for lists of ICD-10 and CPT codes used to define conditions). We created flags to identify the presence or absence of (1) preterm birth (i.e., extremely preterm [<27 weeks' gestation] and preterm [28–36 weeks' gestation]), very low birthweight (P07 codes), and small/light for gestational age (P05 codes) using ICD-10-CM codes; (2) type of neonatal care (i.e., critical [99468, 99469] and intermediate care [99477–99480]) using CPT codes; and (3) place of care using neonatal revenue codes (Levels II-IV as defined by the American Academy of Pediatrics(14); see supplemental materials). Revenue codes are used to identify the place of department within a hospital where care was given and useful in identifying type, place of care.

We analyzed the frequency of these codes in the data and compared whether a delivery had evidence of codes in the following categories: (1) Infant codes: gestational age vs. birthweight, (2) NICU codes: CPT codes vs. revenue codes, and (3) NICU codes vs. infant codes. These comparisons helped identify which codes (i.e., infant or NICU) appeared more frequently to better assess a true NICU admission. We estimated agreement between these codes using Cohen's kappa statistic. Kappa values range from 0 to 1, where greater values indicate more agreement. To interpret kappa statistics, we used the following ranges: weak agreement (0.4–0.59), moderate agreement (0.60–0.79), strong agreement (0.8–0.9), and almost perfect (>0.9).(15)

To further explore the value of our definition, we conducted analysis by using CPT codes as the reference group because clinicians document these codes based on services rendered. We assessed the sensitivity, specificity, positive predictive value (PPV), and negative predictive value of using revenue codes (the "test") against CPT codes (the "standard). For deliveries with revenue codes, the positive predictive value provides the probability of such deliveries identifying NICU admissions. For deliveries without revenue codes, the negative predictive value provides the probability of these deliveries identifying non-NICU admissions. All statistics reported 95% confidence intervals (CI). We conducted data management and analyses using SAS v.9.4 (Cary, NC) and R version 4.1.1 (R Core Team, 2021) including the MultinomialCI package (v1.2).(16)

3. Results

Our study identified 359,542 deliveries and the demographic characteristics of the birthing individual appear in Table 1. From 2016 – 2020 on average, less than 1% of deliveries had evidence of extreme preterm birth (<27 weeks), 8% had evidence of preterm birth (28 – 33 weeks), 2% had evidence of low birthweight, 4% had evidence of small/light birthweight, 5% had evidence of critical NICU care (CPT code), 9% had evidence of intermediate NICU care(CPTcode),and12%hadevidenceofNICUcare by revenue codes.

For each year, we evaluated the presence of NICU and infant codes in the following categories: (1) infant: gestational age vs. birthweight, (2) NICU: CPT vs. revenue codes, and (3) NICU (CPT and revenue codes) vs. infant (gestational age and weight) codes (see Table 2). For example, in 2020, the evidence for NICU admissions using CPT codes alone were 1.5%, revenue codes alone were 3.3%, evidence of CPT *and* revenue codes combined was 8.9%, and evidence of CPT *or* revenue codes was 13.7%. In comparison to the infant codes, the evidence of preterm birth only (e.g., gestational age) was 4.2%, birthweight only was 2.0%, evidence of gestational age *and* birthweight was 4.0%, and evidence of gestational age *or* birthweight was 10.7% (see Fig. 2). For each year of data, revenue codes alone identified more potential NICU admissions than CPT codes alone (see Table 2). When combined, the evidence for NICU admissions using revenue *and* CPT codes on average was 9%; however, the evidence for these admissions was greatest when using CPT *or* revenue codes (14% on average). Overall, CPT and revenue codes had the highest agreement among the categories (kappa 0.75), whereas gestational age and birthweight codes, and NICU and infant codes had weak agreement (see Table 3).

Using NICU CPT codes as the standard, we explored sensitivity and specificity compared to revenue codes. Setting CPT as the standard and revenue as the "test," using 2016 data, revenue codes resulted in identifying 86% of NICU admissions (sensitivity) and 97% of non-NICU admissions (specificity) (see Table 4). If revenue codes identified a NICU admission, the delivery would have a 75% PPV of NICU care. If a delivery did not appear to have NICU care using revenue codes, the delivery had a 98% NPV of not having a NICU admission. Thus, revenue codes accurately identified a large portion of NICU admissions and were able to discriminate those hospitalizations that were not NICU related.

4. Discussion

In this study, we developed and tested a definition to identify NICU admissions using a large administrative claims dataset. This effort improves our ability to ascertain a broad NICU cohort, previously limited by existing methods that only identified potential NICU admissions based on preterm status or birthweight. Based on the evidence reported above, revenue and CPT codes offer a greater chance of correctly identifying NICU admissions. Revenue codes alone identified more NICU admissions compared to CPT codes (12% vs. 10% on average, respectively), whereas the use of CPT and/or revenue codes offer the broadest identification. Our analysis suggests that revenue codes are sensitive and specific to identifying infants receiving care in the NICU; thus, making them ideal for use across multiple datasets. Clinicians and/or administrative staff may enter CPT codes into the medical record based on services rendered and thus could be susceptible to variations in clinician or health system practices/norms. Yet, billing departments often rely on CPT codes to assign a revenue code. The decision to use CPT *and* revenue codes compared to CPT *or* revenue codes will likely depend on researcher preference, the research question, and the information available (e.g., presence of all or some of the codes listed above).

4.1. Interpretation and usefulness of a NICU admission definition

With this study being among the first to examine neonatal codes, it provides researchers with a robust and valid definition to study neonatal care and outcomes among infants truly admitted to a Neonatal Intensive Care Unit. The strengths of this study include the large sample size, use of standardized codes for healthcare encounters, and a rich administrative dataset. Based on our analysis, CPT *OR* revenue codes captured the broadest range of infants who may have had a NICU admission, regardless of preterm or birthweight status. Reliance on algorithms that estimate gestational age or birthweight may underestimate the population of interest (e.g., infants admitted to the NICU), and study findings can differ based on the method selected [17]. If future neonatal health services research used the proposed criteria identified in this work (i.e., CPT *OR/AND* revenue codes), the cohort selection process should begin with the most permissive or broadest group, allowing for further refinement using additional criteria to create subgroups. In doing so, we could advance our current understanding of outcomes associated with neonatal care and infant illness trajectories more broadly [18], rather than for specific subgroups (i.e., VLBW, preterm, small for gestational age). It is important to remember that not all infants receiving care in the NICU are born preterm and almost half NICU admission are for infants born full-term with potentially life-threatening or complex conditions (e.g., congenital birth defects, neonatal abstinence syndrome) [19]. Thus, limiting a cohort to those born <37 weeks' gestation or based on birthweight needs to be carefully matched with the research question. Researchers must balance the need for accuracy with data availability and ease of implementation when choosing the best method for cohort selection and subsequent analysis.

4.2. Limitations

This study has a few limitations to consider. First, we tested our definition in one administrative dataset representing commercial claims from privately insured individuals.

There may be subtle differences in the completeness of information in other datasets such as IBM MarketScan or Medicaid datasets. Even so, ICD-10, CPT and revenue codes are standard across all healthcare encounters and thus easily testable in any given dataset, including electronic health record, which also contain some degree of missing and incomplete data. Future effort could focus on creating datasets that link claims data to additional information in the electronic health records [8]. Lastly, we acknowledge that our use of CPT codes as the reference group (i.e., gold standard) may appear circular in nature, given that one could simply use NICU CPT codes to identify NICU admissions. Yet, as our analysis revealed the combined use of CPT and Revenue codes (as an *OR* in the definition), had the greatest sensitivity in identification and revenue codes alone identified more NICU admissions than CPT codes (4% vs. 2%, respectively). The widespread use and familiarity of CPT codes make it a feasible and effective standard for our exploration of a potential gold standard. Without a common NICU indicator in claims dataset there are limitations in accurately determining the denominator (e.g., known total of NICU admissions in the dataset), which may result in underestimating NICU admissions. Yet, analysis from this paper offers one potential algorithm (CPT *AND* Revenue codes (8%)) for identifying the closest possible estimate of a denominator in this and other administrative datasets.

5. Conclusion

To our knowledge, this analysis represents one of the first studies to define specific criteria for identifying NICU admissions using administrative claims data in lieu of available electronic health records. We recommend identifying NICU admissions using the broadest or most permissive criteria (i.e., definition: CPT *OR* revenue codes), then refining the group based on gestational age, medical diagnosis, and/or birthweight.

Given the increasing prevalence of infant survival beyond the NICU, investing in high-quality methods to measure and study neonatal (NICU) care is vital to our understanding of long-term infant and family well-being and outcomes [20]. Those who use administrative data to answer questions (e.g., clinicians, researchers, policymakers) should recognize the multifaceted uses of information collected from patients—not just for clinical care or for billing purposes, but to address opportunities for health equity and improvement of health outcomes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Funding statement

This research was supported by the National Institute of Mental Health (R01MH120124) and the National Institute on Minority Health and Health Disparities (R01MD014958).

References

[1]. March of Dimes. Special care nursery admissions. 2011.

[2]. Harrison W, Goodman D. Epidemiologic trends in neonatal intensive care, 2007–2012. JAMA Pediatr. 2015;169(9):855–62. [PubMed: 26214387]

[3]. Haidari ES, Lee HC, Illuzzi JL, Phibbs CS, Lin H, Xu X. Hospital variation in admissions to neonatal intensive care units by diagnosis severity and category. J Perinatol. 2021;41(3):468–77. [PubMed: 32801351]

[4]. Myrhaug HT, Brurberg KG, Hov L, Markestad T. Survival and impairment of extremely premature infants: A meta-analysis. Pediatrics. 2019;143(2).

[5]. Burns KH, Casey PH, Lyle RE, Bird TM, Fussell JJ, Robbins JM. Increasing prevalence of medically complex children in US hospitals. Pediatrics. 2010;126(4):638–46. [PubMed: 20855383]

[6]. Braun D, Braun E, Chiu V, Burgos AE, Gupta M, Volodarskiy M, et al. Trends in neonatal intensive care unit utilization in a large integrated health care system. JAMA Netw Open. 2020;3(6):e205239. [PubMed: 32556257]

[7]. Hannan KE, Bourque SL, Palmer C, Tong S, Hwang SS. Prevalence and predictors of medical complexity in a national sample of VLBW infants. Hosp Pediatr. 2021;11(5):525–35. [PubMed: 33906959]

[8]. Hebert PL, Katon JG. Repurposing of administrative data for research: Still useful but for how much longer? Pediatrics. 2019;143(2).

[9]. Eworuke E, Hampp C, Saidi A, Winterstein AG. An algorithm to identify preterm infants in administrative claims data. Pharmacoepidemiol Drug Saf. 2012;21(6):640–50. [PubMed: 22504840]

[10]. Fitzpatrick T, Wilton AS, Guttmann A. Development and validation of a simple algorithm to estimate common gestational age categories using standard administrative birth record data in Ontario, Canada. J Obstet Gynaecol. 2021;41(2):207–11. [PubMed: 32590915]

[11]. Moll K, Wong HL, Fingar K, Hobbi S, Sheng M, Burrell TA, et al. Validating claims-based algorithms determining pregnancy outcomes and gestational age using a linked claims-electronic medical record database. Drug Saf. 2021;44(11):1151–64. [PubMed: 34591264]

[12]. Biks GA, Blencowe H, Hardy VP, Geremew BM, Angaw DA, Wagnew A, et al. Birthweight data completeness and quality in population-based surveys: EN-INDEPTH study. Popul Health Metr. 2021;19(Suppl 1):17. [PubMed: 33557859]

[13]. Chomistek AK, Phiri K, Doherty MC, Calderbank JF, Chiuve SE, McIlroy BH, et al. Development and validation of ICD-10-CM-based algorithms for date of last menstrual period, pregnancy outcomes, and infant outcomes. Drug Saf. 2023;46(2):209–22. [PubMed: 36656445]

[14]. American Academy of Pediatrics Committee on Fetus and Newborn. Levels of neonatal care. Pediatrics. 2012;130(3):587–97. [PubMed: 22926177]

[15]. McHugh ML. Interrater reliability: The kappa statistic. Biochem Med (Zagreb). 2012;22(3):276–82. [PubMed: 23092060]

[16]. May WL, Johnson WD. Constructing two-sided simultaneous confidence intervals for multinomial proportions for small counts in a large number of cells. Journal of Statistical Software. 2000;5(6):1 – 24.

[17]. Lynch CD, Zhang J. The research implications of the selection of a gestational age estimation method. Paediatr Perinat Epidemiol. 2007;21 Suppl 2:86–96. [PubMed: 17803622]

[18]. Azar R. A concept analysis of children with complex health conditions: Implications for research and practice. BMC Pediatr. 2021.

[19]. Braun D, Edwards EM, Schulman J, Profit J, Pursley DM, Goodman DC. Choosing wisely for the other 80%: What we need to know about the more mature newborn and NICU care. Semin Perinatol. 2021;45(3):151395. [PubMed: 33573773]

[20]. Coller RJ, Berry JG, Kuo DZ, Kuhlthau K, Chung PJ, Perrin JM, et al. Health System research priorities for children and youth with special health care needs. Pediatrics. 2020;145(3).

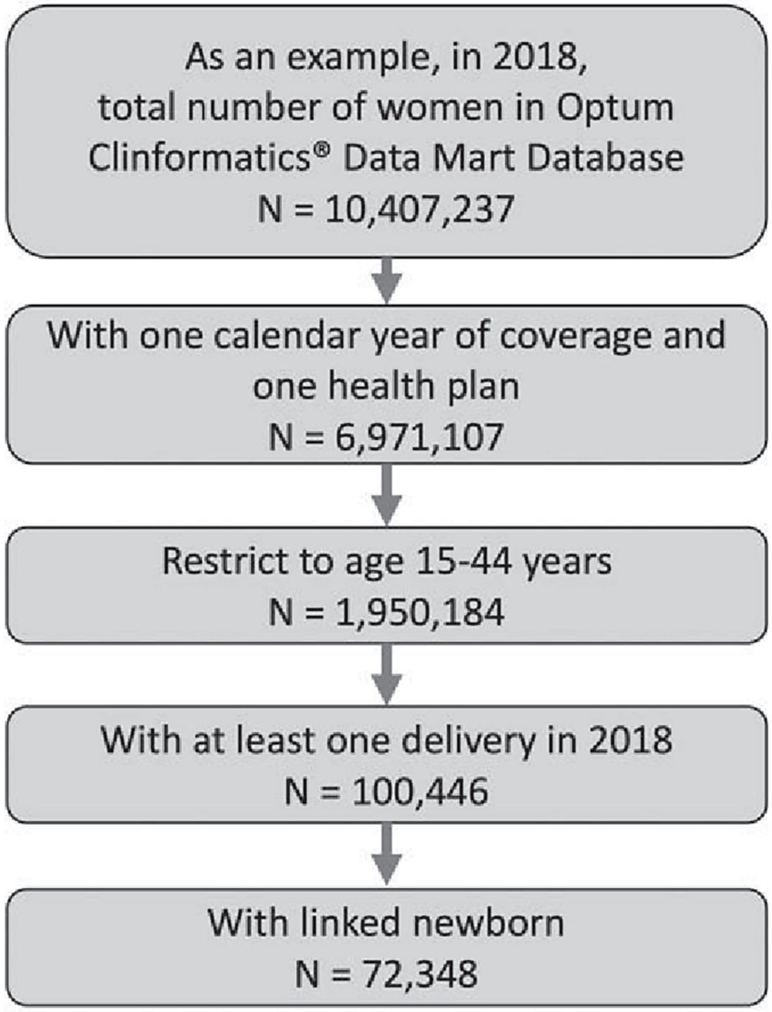
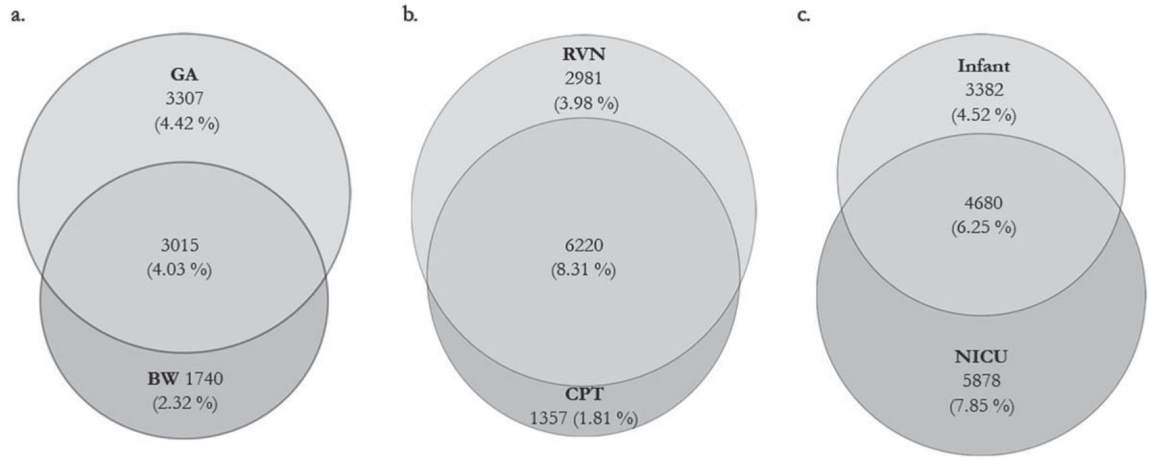


Fig. 1.
Flow chart of the study cohort.



In the deliveries studied, evidence for NICU hospitalizations was 10.8% using low GA or BW codes (a), 14% using CPT or revenue codes (b), and 11% using infant or NICU codes (c).

Fig. 2.

Evidence of defined codes in delivery cohort, 2020.

Table 1

Cohort characteristics, overall and by Infant and NICU codes

	Overall		**Infant[1] Codes**		**NICU[2] Codes**	
Number of delivering individuals	329,977		34,465		47,219	
Number of infants	373,230		39,200		51,087	
	N	% (95% CI)	N	% (95% CI)	N	% (95% CI)
Age						
≤24	13,975	4.24 (4.23,4.24)	1,346	3.91 (3.91, 3.91)	1,816	3.85 (3.57, 4.12)
25–39	296,663	89.9 (89.9, 89.9)	30,172	87.54 (87.5, 87.5)	41,809	88.54 (88.2, 88.8)
≥40	19,339	5.86 (5.86, 5.86)	2,947	8.55 (8.55, 8.55)	3,594	7.61 (7.34, 7.89)
Race						
Asian	27,773	8.42 (8.25, 8.59)	3,321	9.64 (9.11, 10.17)	4,187	8.87 (8.42, 9.32)
Black	24,518	7.43 (7.26,7.6)	3,197	9.28 (8.75, 9.81)	3,994	8.46 (8.01, 8.91)
Hispanic	41,523	12.58 (12.41, 12.75)	4,464	12.95 (12.42, 13.49)	5,954	12.61 (12.16, 13.06)
White	190,302	57.67 (57.5, 57.84)	18,709	54.28 (53.76, 54.82)	26,696	56.54 (56.09, 56.99)
Unknown	45,861	13.9 (13.73, 14.07)	4,774	13.85 (13.32, 14.38)	6,388	13.53 (13.08, 13.98)
Insurance						
HMO	74,004	22.43 (22.43, 22.43)	7,577	21.98 (21.53, 22.44)	10,674	22.61 (22.21, 23)
POS	249,858	75.72 (75.72, 75.72)	26,306	76.33 (75.87, 76.78)	35,643	75.48 (75.09, 75.88)
PPO	3,599	1.09 (1.09, 1.09)	349	1.01 (0.56, 1.47)	487	1.03 (0.64, 1.43)
Other	2,516	0.76 (0.76, 0.76)	233	0.68 (0.22, 1.13)	415	0.88 (0.48, 1.27)
Region						
Midwest	85,134	25.8 (25.61, 25.99)	8,675	25.17 (24.6, 25.74)	12,475	26.42 (25.93, 26.91)
Northeast	36,288	11 (10.81, 11.18)	4,131	11.99 (11.42, 12.56)	5,475	11.59 (11.11, 12.08)
South	133,673	40.51 (40.32, 40.7)	14,103	40.92 (40.35, 41.49)	19,004	40.25 (39.76, 40.74)
West	73,992	22.42 (22.24, 22.61)	7,467	21.67 (21.1, 22.24)	10,151	21.5 (21.01, 21.99)
Unknown	890	0.27 (0.08, 0.45)	89	0.26 (0, 0.83)	114	0.24 (0, 0.73)

NICU: Neonatal Intensive Care Unit, **CI:** Confidence Intervals, **HMO:** Health maintenance plan, **POS:** Point of service, **PPO:** preferred provider organization

[1]Infant codes include ICD-10 codes for gestational age (GA) and birthweight (BW).

[2]NICU codes include current procedural terminology (CPT) and revenue codes.

Table 2

Evidence of NICU and Infant codes by year (n[1], % (95% CI))

Codes	2016	2017	2018	2019	2020
GA only	3,008	2,932	3,159	3,170	3,307
	3.95 (3.75, 4.15)	3.98 (3.78, 4.19)	4.2 (4, 4.41)	4.32 (4.11, 4.53)	4.42 (4.21, 4.63)
BW only	1,377	1,406	1,474	1,544	1,740
	1.81 (1.61, 2.01)	1.91 (1.7, 2.12)	1.96 (1.76, 2.17)	2.1 (1.89, 2.31)	2.32 (2.12, 2.53)
GA *AND* BW	3,314	3,277	3,281	3,196	3,015
	4.35 (4.15, 4.56)	4.45 (4.24, 4.66)	4.37 (4.16, 4.57)	4.35 (4.14, 4.56)	4.03 (3.82, 4.24)
GA *OR* BW	7,699	7,615	7,914	7,910	8,062
	10.1 (9.9, 10.33)	10.3 (10.12, 10.56)	10.5 (10.31, 10.75)	10.7 (10.55, 10.99)	10.7 (10.55, 11)
CPT only	1,067	1,090	1,165	1,100	1,357
	1.4 (1.17, 1.63)	1.48 (1.24, 1.72)	1.55 (1.31, 1.79)	1.5 (1.25, 1.74)	1.81 (1.58, 2.05)
Revenue only	2,186	2,049	2,308	2,624	2,981
	2.87 (2.64, 3.1)	2.78 (2.55, 3.02)	3.07 (2.84, 3.31)	3.57 (3.33, 3.82)	3.98 (3.75, 4.22)
CPT *AND* revenue	6,549	6,707	6,768	6,916	6,220
	8.6 (8.37, 8.83)	9.11 (8.87, 9.34)	9.01 (8.77, 9.24)	9.42 (9.17, 9.66)	8.31 (8.07, 8.55)
CPT *OR* revenue	9,802	9,846	10,241	10,640	10,558
	12.88 (12.64, 13.12)	13.3 (13.12, 13.62)	13.6 (13.38, 13.87)	14.4 (14.23, 14.74)	14.1 (13.86, 14.36)
NICU[2] only	5,303	5,335	5,575	5,887	5,878
	6.97 (6.71, 7.22)	7.24 (6.98, 7.51)	7.42 (7.16, 7.68)	8.01 (7.74, 8.29)	7.85 (7.59, 8.12)
Infant[3] only	3,200	3,104	3,248	3,157	3,382
	4.2 (3.95, 4.46)	4.21 (3.95, 4.48)	4.32 (4.06, 4.58)	4.3 (4.03, 4.57)	4.52 (4.25, 4.79)
NICU *AND* infant	4,499	4,511	4,666	4,753	4,680
	5.91 (5.66, 6.17)	6.12 (5.86, 6.39)	6.21 (5.95, 6.47)	6.47 (6.2, 6.74)	6.25 (5.99, 6.52)
NICU *OR* infant	13,002	12,950	13,489	13,797	13,940
	17.08 (16.82, 17.35)	17.5 (17.31, 17.86)	17.9 (17.68, 18.22)	18.7 (18.5, 19.07)	18.6 (18.35, 18.91)

GA: Gestational age (ICD-10 codes), **BW:** Birthweight (ICD-10 codes), **CPT:** Current procedural terminology codes.

[1] Refers to the total number of babies with evidence of the code on claims.

[2] NICU codes included all CPT and revenue codes.

[3]Infant codes included all gestational age and weight codes.

Table 3

Agreement between NICU and Infant codes by year

Comparison of categories[1]	Year	Kappa coefficient	Kappa LCL	Kappa UCL	Proportion of negative agreement[2]	Proportion of positive agreement[2]
Gestational age vs. birthweight	2016	0.5715	0.5602	0.5829	0.9803	0.5242
CPT vs. revenue		0.7772	0.7699	0.7846	0.9681	0.8599
NICU codes vs. infant codes		0.4521	0.4423	0.4618	0.9518	0.4590
Gestational age vs. birthweight	2017	0.5706	0.5592	0.5820	0.9792	0.5278
CPT vs. revenue		0.7864	0.7793	0.7936	0.9689	0.8602
NICU codes vs. infant codes		0.4529	0.4431	0.4627	0.9514	0.4582
Gestational age vs. birthweight	2018	0.5537	0.5423	0.5651	0.9786	0.5095
CPT vs. revenue		0.7699	0.7626	0.7772	0.9657	0.8532
NICU codes vs. infant codes		0.4485	0.4389	0.4581	0.9500	0.4556
Gestational age vs. birthweight	2019	0.5416	0.5301	0.5532	0.9770	0.5020
CPT vs. revenue		0.7593	0.7520	0.7667	0.9599	0.8628
NICU codes vs. infant codes		0.4437	0.4342	0.4533	0.9497	0.4467
Gestational age vs. birthweight	2020	0.5087	0.4970	0.5205	0.9746	0.4769
CPT vs. revenue		0.7092	0.7010	0.7173	0.9557	0.8209
NICU codes vs. infant codes		0.4335	0.4239	0.4431	0.9474	0.4433

NICU: Neonatal Intensive Care Unit, **CPT:** Current Procedural Terminology, **LCL: l**ower confidence limit, **UCL:** upper confidence limit.

[1] See supplemental table for list of codes used in each category.

[2] The proportion of negative or positive agreement provides a measure of agreement between the two methods in identifying non-NICU stays (negative agreement) or NICU stays (positive agreement). We treated gestational age (GA), CPT, and NICU codes as the non-reference standard and birthweight (BW), revenue, and infant codes as the 'new test.' For instance, in comparing GA and BW methods of NICU care identification, the proportion of negative agreement was calculated as number of deliveries with neither GA or BW codes divided by number of deliveries without GA codes (with or without BW codes), and positive agreement was number of deliveries with both GA and BW codes divided by number of deliveries with GA codes (with or without BW codes).

Table 4

Sensitivity and Specificity of NICU[1] CPT[2] and revenue codes

Year	Sensitivity (95%CI)	Specificity (95%CI)	PPV[3] (95%CI)	NPV[4] (95%CI)
2016	0.86 (0.85, 0.87)	0.97 (0.97, 0.97)	0.75 (0.74, 0.76)	0.98 (0.98, 0.99)
2017	0.86 (0.85, 0.87)	0.97 (0.97, 0.97)	0.77 (0.76, 0.77)	0.98 (0.98, 0.98)
2018	0.85 (0.85, 0.86)	0.97 (0.96, 0.97)	0.75 (0.74, 0.75)	0.98 (0.98, 0.98)
2019	0.86 (0.86, 0.87)	0.96 (0.96, 0.96)	0.72 (0.72, 0.73)	0.98 (0.98, 0.98)
2020	0.82 (0.81, 0.83)	0.96 (0.95, 0.96)	0.68 (0.67, 0.69)	0.98 (0.98, 0.98)

NICU: Neonatal Intensive Care Unit, **CPT:** Current Procedural Terminology, **PPV:** Positive Predictive Value, **NPV:** Negative Predictive Value.