Title: Predicting Infant Mortality Risk from Information Available at the Time of Birth

Paper (partially incomplete) submitted for 2019 PAA Annual Meeting in Austin, TX

Authors: Antonia Saravanou[1], Clemens Noelke[2,*], Nicholas Huntington[2], Dolores Acevedo-Garcia[2], Dimitrios Gunopulos[1]

***Correspondence to:** Dr. Clemens Noelke
Institute for Child, Youth, and Family Policy
Heller School for Social Policy and Management
Brandeis University
415 South Street, MS 035
Waltham, MA 02453
Phone: 781-736-8748
Email: cnoelke@brandeis.edu

Author Affiliations: [1]Department of Informatics and Telecommunications, University of Athens, Greece; [2]Heller School for Social Policy and Administration, Brandeis University, Waltham, MA, USA.

Funding: This work was supported by grants from the Robert Wood Johnson Foundation and the WK Kellogg Foundation.

Acknowledgements: .

Conflict of Interest: None declared.

Human Participant Protection: This study was approved by the Brandeis University Committee for Protection of Human Subjects, IRB Protocol #18008.

PREDICTING INFANT MORTALITY RISK FROM INFORMATION AVAILABLE AT THE TIME OF BIRTH

Using birth certificate data for all registered US births over a three year period (2000-2002), we explore whether it is possible to reliably identify infants at risk of dying before their first birthday using information that is routinely gathered at the time of birth. We use four classifiers from the machine learning (ML) literature to predict mortality before the first birthday, as well as age at death (early neonatal, late neonatal, and postneonatal) and cause of death. We also explore whether the quality of predictions varies by maternal race/ethnicity. We find that the best-performing classifier correctly identifies, at the time of birth, 3 out of 4 infants who die before their first birthday. The resulting risk scores can potentially be used to allocate more intensive care within and beyond the clinical setting to infants with a high predicted mortality risk.

Word Count Paper: 2,315

Word Count Abstract: 137

With infant mortality rates ranked highest among developed nations, reducing infant mortality remains an important public health goal in the United States(Chen, Oster, and Williams 2016; Thakrar et al. 2018; MacDorman et al. 2014; MacDorman and Mathews 2009). In this study, we explore whether it is possible to reliably identify infants at risk of dying before their first birthday using information that is routinely gathered at the time of birth on the standardized US birth certificate. We assess whether the quality of predictions varies by the age (early neonatal, late neonatal, and postneonatal) and cause of death. Given large inequities in mortality risks that between infants born to mothers with different racial/ethnic backgrounds, we also explore whether the quality of predictions varies by maternal race/ethnicity.

Developing algorithms and screening tools to detect potentially fatal health conditions among newborns have been identified as central priorities for research on newborn health (Yoshida et al. 2014). Alongside universalist programs aimed at improving health for all newborns, prediction models could be used to detect infants at increased mortality risk at the time of birth, who would then be referred to more intensive care within and beyond the clinical setting. Such targeting based on risk scores is desirable when interventions are costly and providing them regardless of need would divert resources from areas of greater need. Furthermore, targeting reduces potential risks associated with unnecessary, potentially harmful, and invasive treatments and interventions. Illustrating the potential use of prediction models in this context, Pan et al. (2017) show how risk scores from machine learning (ML) algorithms can improve the allocation of mothers at risk of adverse birth outcomes to an intensive care program with limited enrolment capacity. Potash et al. (2015) show how risk scores for lead poisoning allow a limited number of inspectors to prioritize high-risk buildings for inspections.

Multivariate predictions models for infant survival have been studied extensively for infants born very prematurely or with very low birth weight. A recent meta-analysis concludes that prediction models can support medical decision making in newborn intensive care units (NICUs) (Medlock et al. 2011). It reports median classifier performance metrics (see Methods, for definitions) across 41 studies that provide a reference for our own analyses: These studies focus on a small subset of newborns with very high mortality risks, and it is unclear how well ML classifiers predict mortality risks in the full population of newborns. Related work has tried to predict the risk of preterm birth, the single most important cause of infant mortality(Lawn et al. 2014), in clinical trial (Vovsha et al. 2016) and administrative data (Pan et al. 2017; Tran et al. 2016), and birth certificate data from North Carolina (Courtney et al. 2008). Studies on preterm birth and survival among infants admitted to NICUs indicate that machine learning classifiers outperform conventional regression or paper & pencil-based screening tools. Finally, research from across the social and medical sciences has identified numerous risk factors for infant mortality many of which are recorded on US birth certificates(Ma and Finch 2010; Almond, Chay, and Lee 2005; Chen, Oster, and Williams 2016; MacDorman and Mathews 2009). These latter studies focus on whether and how specific factors are causing elevated mortality risks.

To our knowledge, this is the first study that tries to leverage information from nearly all variables measured on birth certificates to predict either infant mortality or other markers of newborn health for all US births in a given year. Relying on birth certificate data in this context is appealing because this data is routinely gathered in a standardized format for the entire population at the time of birth. It includes an extensive list of demographic and medical variables that have been shown to both predict infant mortality, and short- and long-term health and socio-economic outcomes. Given that birth certificate data are available for the entire population, they represent a potentially underutilized resource for determining which infants should be considered at higher risk of death and therefore receive additional care and services within and beyond the clinical setting.

In the following, we compare the quality of predictions obtained from conventional logistic regression and machine learning algorithms. Mortality risk is commonly modeled as an additive function of a limited set of predictor variables that includes some interaction and non-linear effects for specific variables. Machine learning algorithms incorporate information from complex non-linear and interactive effects that have been shown to substantially improve predictive accuracy in some health contexts. We also explore whether the quality of predictions varies by age and cause of death and by maternal race.

METHODS

Data

The analyses use data drawn from birth certificates for all registered births over the period from 2000 to 2002 that are linked to death certificates for all deaths occurring before the first birthday, and published by the National Center for Health Statistics (NCHS) as the Linked Birth/lnfant Death Birth Cohort Data Sets. About one percent of infant death certificates cannot be matched to birth certificates, and these observations are excluded from the analysis. We use 2000 and 2001 data for training our models, and 2002 data for testing. Because both outcome and many features have a low prevalence, we pooled two years of data to train our models in order to obtain more robust results.

While more recent data would clearly be desirable, we use 2000-2002 data for the following reasons: This is the most recent set of three consecutive years that the cohort-linked files are published for that (a) does not include a change in the version of the birth certificate the data are drawn from and (b) includes geographic detail that allows us to identify mother's state of residence. To our knowledge, the most recent release of the cohort-linked file is for the 2011 birth cohort, but the publicly available data lack geographic identifiers and observations are based on two different versions of the US birth certificate that introduces many inconsistencies in the availability and coding of variables across years.

The source files include the age at the time of death in days, and information on the cause and manner of death, as well as 158 potential predictor variables. We deleted variables created by the data provider that (a) relate to the quality of measurement of predictor values, such as whether an observation was imputed, or (b) are re-coded versions of other variables, collapsing information from a continuous variable, e.g. birth weight in grams, to a categorical variable. We also generated variables for maternal/paternal race/ethnicity, combining information both on race and Hispanic origin.

The final set of predictors includes 107 variables. Table 1 in the Appendix includes a full list. They include demographic information on mothers, geographic information on maternal place of residence and place of birth, maternal risk factors, medical conditions and information on prior births, as well as, for infants, detailed information on labor and delivery and birth outcomes, such as birth weight, gestational age, and congenital malformations.

Most variables are categorical by nature, but some, like birth weight, are measured on a continuous scale. We converted all numeric variables to categorical variables. When available, we used information from previous studies to appropriately recode variables so that non-linear relationships can be appropriately captured. For example, we used a relatively detailed classification for birth weight at low and very low birth weights because of the well-known, nonlinear association between birth weight and mortality risk at low levels of birth weight. Otherwise, we grouped continuous variables into bins. After converting all continuous variables to categorical values, we imputed missing values for all variables with the modal category (Pan et al. 2017).

Our main outcome is mortality before the first birthday. The NCHS files include information on the age of death in days, which we use to further differentiate between deaths in the first week of life (days 0-6, early neonatal mortality), late neonatal mortality (days 7-27), and postneonatal (days 8-364) mortality. Furthermore, we explore different causes of death. Based on a 130 cause of death variable generated by NCHS, which is in turn based on ICD-10 Codes, we aggregated different causes of death into six categories with approximately similar prevalence:

- “Diseases/disorders”, which includes infectious and parasitic diseases; neoplasms; diseases of the blood and immune system; endocrine, nutritional and metabolic diseases; diseases of the nervous, circulatory, respiratory, digestive or genitourinary systems (ucod130 codes 1-69).

- “Maternal factors, complications of pregnancy, labor, and delivery” (ucod130 codes 70-85).
- “Length of gestation and fetal malnutrition” (ucod130 codes 86-91).
- “Other perinatal conditions”, which includes birth trauma, hypoxia, asphyxia, respiratory distress/conditions, infections originating in perinatal period, hemorrhagic and hematological disorders of newborn, syndrome of infant of a diabetic mother and neonatal diabetes mellitus, necrotizing enterocolitis, hydrops fetalis not due to hemolytic disease, other perinatal conditions (ucod130 codes 92-117)
- “Congenital malformations, deformations, and chromosomal abnormalities” (ucod130 codes 118-133).
- “External causes, sudden infant death syndrome (SIDS), and causes not elsewhere classified (NEC)” (ucod130 codes 134-158)

STATISTICAL MODELLING

Our main goal is to assess how well infant mortality can be predicted from information available on birth certificates at the time of birth, and how the quality of prediction varies by age, the cause of death, and maternal race. Specifically, the goal is to predict values of a binary or categorical outcome variable from a set of predictor variables, or features. We do not adopt a survival analysis approach because this would require information on the duration of survival, which is not available at birth.

Based on previous studies (Pan et al. 2017; Vovsha et al. 2016; Somanchi et al. 2015), we selected four machine learning classifiers and compared their performance to conventional logistic regression (LOG) without non-linear or interaction effects. The machine learning classifiers are: Gaussian Naive Bayes (GNB), One-class Support Vector Machines (SVMs) with a Radial Basis Function kernel, and Boosted Trees (XGB). Boosted trees are an ensemble method that iteratively using fixed size decision trees, iteratively fitting the residuals of the previous prediction to minimize the loss of the latest prediction.

The low prevalence of the outcome variables, 0.68% of infants do not survive, can degrade the performance of the classifiers we use (Kubat and Matwin 1997). To improve classifier performance, in particular for predicting the less prevalent outcome, i.e., mortality, we subsample the training dataset (all births in 2000 and 2001), keeping all infants who die before their first birthday but only a random sample of those who survive. Following previous research (Somanchi et al. 2015), we explore classifier performance for the binary prediction task, survival until the first birthday, using four different sampling ratios: 1:1, 1:5, 1:10, and 1:145. In the first case, 1:1, we include n_1 infants in the training dataset who die, where n_1 is the number of infants who die and take a random subsample from the infants who survive of size one times n_1. For a 1:5 (1:10) ratio, we include all n_1 infants who die and take a random subsample of five (ten) times n_1 infants who survive, and so forth. 1:145 corresponds to the naturally occurring ratio (no subsampling). Classifiers were evaluated using the (raw, not subsampled) testing data.

We used the Scikit-Learn Library in Python 2.7 (Pedregosa et al. 2011) and a computer with an Intel Core i7-3770 CPU processor (3.40GHz), 16 GB of RAM and 64-bit operating system to perform the analyses. The runtime for the different models varied between a few seconds to a few minutes.

Performance Metrics

We report four metrics that quantify the performance of the classifiers: Sensitivity, Positive Predictive Value (PPV), Accuracy, and Area under the Receiver Operating Characteristic Curve (AUC). Sensitivity is defined as the percentage of infants who die before their first birthday and are correctly classified (true positives) relative to all infants who die before their first birthday (true positives + false negatives), i.e., the percentage of all deaths that were correctly predicted. Sensitivity diminishes both if the number of true positives diminishes and if the number of false negatives increases, i.e., the number of infants predicted to survive when they do not. PPV is defined as the percentage of true positives relative to all infants who are predicted to die (true positives + false positives). PPV diminishes both if the number of true positives diminishes and if the number of false positives increases, i.e., the number of infants predicted to die who survive.

While Sensitivity places a penalty on false negatives, PPV places a penalty on false positives. From an applied perspective, false negatives correspond to infants who die within their first year but are predicted to survive by our model and may therefore not have received potentially life-saving interventions. False positives might correspond to infants who are predicted to die but actually survive and are nevertheless assigned to follow-up care that is costly and may be harmful. In this application, we focus on the costs of false negatives, i.e. not detecting infants who may have survived if detected, and therefore emphasize Sensitivity as the central parameter to judge model quality.

Accuracy is defined as the percentage of infants correctly classified. It is the sum of true positives and true negatives as a percentage of all infants. The area under the receiver operating characteristics curve (AUC) measures the probability that a classifier ranks a randomly chosen positive instance (died) higher than a randomly chosen negative one (survived).

RESULTS

Table 1 reports descriptive statistics for the training data (births in 2000 and 2001) and test data (births in 2002). In both training and test data, about 1 in 145 infants die before their first birthday (0.68%). The distribution of age and cause of death is also similar across datasets. In the test data about 41% of deaths occur within 24 hours of births. Another 27% of deaths occur between the 1 and 27th day. And the remaining 32% of deaths occur after the first month and before the first birthday. The cause of death group with the highest prevalence, “Other perinatal conditions” includes birth trauma, asphyxia, hypoxia, and other perinatal respiratory conditions.

Table 2 reports the best performing classifiers for predicting mortality (death vs. survive) in each of the four subsampled training datasets. Considering Sensitivity, we observe that logistic regression performs best when the data are more balanced (1:1 and 1:5), while Gaussian Naïve Bayes and Boosted Trees perform better in unbalanced data (1:10 and 1:145). Across classifiers and sampling ratios, we obtain the best Sensitivity and AUC using Boosted Trees (XGB) and a 1:10 sampling ratio, i.e., 1 death per 10 survivors in the training data.

Table 3 reports performance metrics based on training data with a 1:10 sampling ratio for four classifiers predicting mortality. Boosted Trees and Gaussian Naïve Bayes achieve the highest and second-highest Sensitivity, but the lowest PPV. Logistic regression and SVMs achieve higher PPV but lower Sensitivity. Overall, Boosted Trees perform best in terms of Sensitivity and AUC, but worst in terms of PPV.

Boosted Trees correctly identify 77% of infants who die before their first birthday from information provided at the time of birth, exceeding the performance of other classifiers substantially in terms of Sensitivity. They also achieve the highest AUC, 0.85. These metrics are similar to the median performance of classifiers predicting mortality among very preterm (or very low birthweight) infants admitted to NICUs reviewed by Medlock et al. (2011). However, PPV for Boosted Trees is low. Only 6% of infants predicted to die actually die. While correctly classifying 3 in 4 infants among those who die, the classifier generates 16 false positives for every true positive.

The remainder of the analyses will examine the quality of predictions for different ages (early neonatal, late neonatal, and postneonatal) and causes of death. We will also examine whether the quality of predictions varies by race/ethnicity of mothers. Finally, we will conduct supplementary analyses to better understand both false negative and false positive predictions. The large number of false positives generated by the Boosted Trees present an interesting challenge, because they share many of the features of infants who do not survive, but yet they do. It will be interesting to see whether these differ systematically in their distribution across different US states and by maternal education, and whether they are predicted to die of preventable causes. Furthermore, while the quality of prediction is expected *to vary across different causes, we intend to identify potentially preventable causes for which the classifiers perform particularly well.*

REFERENCES

Almond, Douglas, Kenneth Y. Chay, and David S. Lee. 2005. “The Costs of Low Birth Weight.” *The Quarterly Journal of Economics* 120 (3): 1031–83. https://doi.org/10.1093/qje/120.3.1031.

Chen, Alice, Emily Oster, and Heidi Williams. 2016. “Why Is Infant Mortality Higher in the United States Than in Europe?” *American Economic Journal. Economic Policy* 8 (2): 89–124. https://doi.org/10.1257/pol.20140224.

Courtney, Karen L., Sara Stewart, Mihail Popescu, and Linda K. Goodwin. 2008. “Predictors of Preterm Birth in Birth Certificate Data.” *Studies in Health Technology and Informatics* 136: 555–60.

Kubat, Miroslav, and Stan Matwin. 1997. “Addressing the Curse of Imbalanced Training Sets: One-Sided Selection.” In *In Proceedings of the Fourteenth International Conference on Machine Learning*, 179–186. Morgan Kaufmann.

Lawn, Joy E., Hannah Blencowe, Shefali Oza, Danzhen You, Anne C. C. Lee, Peter Waiswa, Marek Lalli, et al. 2014. “Every Newborn: Progress, Priorities, and Potential beyond Survival.” *Lancet (London, England)* 384 (9938): 189–205. https://doi.org/10.1016/S0140-6736(14)60496-7.

Ma, Sai, and Brian Karl Finch. 2010. “Birth Outcome Measures and Infant Mortality.” *Population Research and Policy Review* 29 (6): 865–91. https://doi.org/10.1007/s11113-009-9172-3.

MacDorman, Marian F., and T.J. Mathews. 2009. “The Challenge of Infant Mortality: Have We Reached a Plateau?” *Public Health Reports* 124 (5): 670–81.

MacDorman, Marian F., T. J. Matthews, Ashna D. Mohangoo, and Jennifer Zeitlin. 2014. “International Comparisons of Infant Mortality and Related Factors: United States and Europe, 2010.” *National Vital Statistics Reports: From the Centers for Disease Control and Prevention, National Center for Health Statistics, National Vital Statistics System* 63 (5): 1–6.

Medlock, Stephanie, Anita C. J. Ravelli, Pieter Tamminga, Ben W. M. Mol, and Ameen Abu-Hanna. 2011. “Prediction of Mortality in Very Premature Infants: A Systematic Review of Prediction Models.” *PloS One* 6 (9): e23441. https://doi.org/10.1371/journal.pone.0023441.

Pan, Ian, Laura B. Nolan, Rashida R. Brown, Romana Khan, Paul van der Boor, Daniel G. Harris, and Rayid Ghani. 2017. “Machine Learning for Social Services: A Study of Prenatal Case Management in Illinois.” *American Journal of Public Health* 107 (6): 938–44. https://doi.org/10.2105/AJPH.2017.303711.

Pedregosa, Fabian, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, et al. 2011. “Scikit-Learn: Machine Learning in Python.” *J. Mach. Learn. Res.* 12 (November): 2825–2830.

Potash, Eric, Joe Brew, Alexander Loewi, Subhabrata Majumdar, Andrew Reece, Joe Walsh, Eric Rozier, Emile Jorgenson, Raed Mansour, and Rayid Ghani. 2015. “Predictive Modeling for Public Health: Preventing Childhood Lead Poisoning.” In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2039–2047. KDD ’15. New York, NY, USA: ACM. https://doi.org/10.1145/2783258.2788629.

Somanchi, Sriram, Samrachana Adhikari, Allen Lin, Elena Eneva, and Rayid Ghani. 2015. “Early Prediction of Cardiac Arrest (Code Blue) Using Electronic Medical Records.” In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2119–2126. KDD ’15. New York, NY, USA: ACM. https://doi.org/10.1145/2783258.2788588.

Thakrar, Ashish P., Alexandra D. Forrest, Mitchell G. Maltenfort, and Christopher B. Forrest. 2018. “Child Mortality In The US And 19 OECD Comparator Nations: A 50-Year Time-Trend Analysis.” *Health Affairs (Project Hope)* 37 (1): 140–49. https://doi.org/10.1377/hlthaff.2017.0767.

Tran, Truyen, Wei Luo, Dinh Phung, Jonathan Morris, Kristen Rickard, and Svetha Venkatesh. 2016. “Preterm Birth Prediction: Deriving Stable and Interpretable Rules from High Dimensional Data.” *ArXiv:1607.08310 [Stat]*, July. http://arxiv.org/abs/1607.08310.

Vovsha, Ilia, Ansaf Salleb-Aouissi, Anita Raja, Thomas Koch, Alex Rybchuk, Axinia Radeva, Ashwath Rajan, et al. 2016. “Using Kernel Methods and Model Selection for Prediction of Preterm Birth.” *ArXiv:1607.07959 [Cs, Stat]*, July. http://arxiv.org/abs/1607.07959.

Yoshida, Sachiyo, Igor Rudan, Joy E. Lawn, Stephen Wall, João Paulo Souza, José Martines, and Rajiv Bahl. 2014. “Newborn Health Research Priorities beyond 2015.” *The Lancet* 384 (9938): e27–29. https://doi.org/10.1016/S0140-6736(14)60263-4.

TABLES

Table 1. Descriptive Statistics for Training and Testing Data, Cohort-Linked NCHS Natality Infant Mortality Files, 2000-2002.

	Training Data Births in 2000-2001		**Test Data Births in 2002**	
	Births	**% of births**	**Births**	**% of births**
Number of live births	8,084,950	100.00	4,021,825	100.00
Number of deaths	54,849	0.68	27,500	0.68
Age at death				
Under 1 hour	8,133	0.10	4,111	0.10
1-23 hours	13,505	0.17	7,060	0.18
1-6 days	7,479	0.09	3,610	0.09
7-27 days	7,500	0.09	3,690	0.09
28 days and over	18,232	0.23	9,029	0.22
Cause of death				
Diseases/disorders	6,678	0.08	3,383	0.08
Maternal factors, complications of pregnancy, labor and delivery	5,487	0.07	3,007	0.07
Length of gestation and fetal malnutrition	8,848	0.11	4,646	0.12
Other perinatal conditions	12,952	0.16	6,228	0.15
Congenital malformations, deformations and chromosomal abnormalities	11,168	0.14	5,554	0.14
External causes, SIDS, not elsewhere classified	9,716	0.12	4,682	0.12

Table 2. Best performing classifiers for different sampling ratios.

Sampling Ratios Deaths:Survivors	Model	Sensitivity	PPV	Accuracy	AUC
[1:1]	Logistic Regression	0.75	0.05	0.89	0.83
[1:5]	Logistic Regression	0.60	0.19	0.98	0.79
[1:10]	Boosted Trees	0.77	0.06	0.92	0.85
[1:145]	Gaussian Naïve Bayes	0.67	0.07	0.94	0.80

Table 3. Performance metrics for different classifiers using 1:10 sampling ratio.

Model	Sensitivity	PPV	Accuracy	AUC
Logistic Regression	0.56	0.28	0.99	0.77
Gaussian Naïve Bayes	0.68	0.07	0.94	0.81
Support Vector Machines	0.54	0.24	0.99	0.77
Boosted Trees	0.77	0.06	0.92	0.85

APPENDIX

Table A1. List of Features

Variable Name	Label
birmon	month of birth
weekdayb	day of week child born
gestat	gestation - in weeks
dbirwt	birth weight - in grams, converted to percentiles
missval	number of variables missing, generated
dtotord	detailed total birth order
dnliv	number of previous live births no longer living, generated
dmage	age of mother
dfage	age of father
wtgain	weight gain during pregnancy
nprevist	total number of prenatal visits
drink	average number of drinks per week
cigar	average number of cigarettes per day
stoccfipb	state of occurrence - FIPS code
stresfipb	state of residence (FIPS)
resstatb	resident status
mracehisp	father's race/ethnicity
dmeduc	education of mother detail
married	mother's marital status (recode of dmar)
mplbir	place of birth of mother
mpcb	detail month of pregnancy prenatal care began
nodad	no information on father on any variable, generated
fracehisp	father's race/ethnicity
birattnd	attendant at delivery
pldel	place or facility of delivery
female	infant sex (recode of csex)
fmaps	five minute APGAR score
dplural	plurality
vacuum	vacuum
forcep	forceps
repeac	repeat C-section
primac	primary C-section
vbac	vaginal birth after previous C-section
vaginal	vaginal
phyper	hypertension – pregnancy-associated
chyper	hypertension – chronic
hemo	hemoglobinopathy
hydra	hydramnios/oligohydramnios
herpes	genital herpes
diabetes	diabetes
lung	acute or chronic lung disease
cardiac	cardiac disease
anemia	anemia (hct.<30/hgb.<10)
othermr	other medical risk factors
uterine	uterine bleeding
rh	rh sensitization
renal	renal disease
preterm	previous preterm or small-for-gestational-age infant
pre4000	previous infant 4000+ grams

incervix	incompetent cervix
eclamp	eclampsia
otherob	other obstetric procedures
ultras	ultrasound
tocol	tocolysis
stimula	stimulation of labor
induct	induction of labor
monitor	electronic fetal monitoring
amnio	amniocentesis
otherlb	other complication of labor and/or delivery
distress	fetal distress
anesthe	anesthetic complications
cord	cord prolapse
cephalo	cephalopelvic disproportion
breech	breech/malpresentation
dysfunc	dysfunctional labor
prolong	prolonged labor (>20 hours)
precip	precipitous labor (<3 hours)
seizure	seizures during labor
excebld	other excessive bleeding
preplace	placenta previa
abruptio	abruptio placenta
rupture	premature rupture of membrane (>12 hours)
meconium	meconium
febrile	febrile (>100 degrees F or >38 degrees C)
otherab	other abnormal conditions of the newborn
nseiz	seizures
ven30m	assisted ventilation - 30 minutes or more
venl30	assisted ventilation - less than 30 minutes
meconsyn	meconium aspiration syndrome
hyaline	hyaline membrane disease
alcosyn	fetal alcohol syndrome
injury	birth injury
nanemia	anemia (hct.<39/hgb.<13)
gastro	other gastrointestinal anomalies
omphalo	omphalocele/gastroschisis
tracheo	tracheo-esophageal fistula/esophageal atresia
rectal	rectal atresia/stenosis
circul	other circulatory/respiratory anomalies
heart	heart malformations
nervous	other central nervous system anomalies
microce	microcephalus
hydro	hydrocephalus
spina	spina bifida/meningocele
anen	anencephalus
othercon	other congenital anomalies
chromo	other chromosomal anomalies
downs	Down's syndrome
musculo	other musculoskeletal/integumental anomalies
hernia	diaphragmatic hernia
clubfoot	club foot
adactyly	polydactyly/syndactyly/adactyly
cleftlp	cleft lip/palate
Urogen	other urogenital anomalies
Renalage	renal agenesis
Genital	malformed genitalia