Prevalence and Clinical Significance of Commonly Diagnosed Genetic Disorders in Preterm Infants

Selin S. Everett[a], Miles Bomback[b], Rakesh Sahni, MD[c], Ronald J. Wapner, MD[a], Veeral N. Tolia, MD[d,e,f], Reese H. Clark, MD[d],, Alex Lyford, PhD[g], Thomas Hays, MD, PhD[c]

Affiliations: [a] Division of Maternal Fetal Medicine, Department of Obstetrics and Gynecology, Columbia University Irving Medical Center, New York City
[b] Feinberg School of Medicine, Northwestern University, Chicago
[c] Division of Neonatology, Department of Pediatrics, Columbia University Irving Medical Center, New York City
[d] The Pediatrix Center for Research, Education, Quality and Safety, Sunrise, Florida
[e] Division of Neonatology, Department of Pediatrics, Baylor University Medical Center, Dallas, Texas
[f] Pediatrix Medical Group, Dallas, Texas
[g] Department of Statistics, Middlebury College, Middlebury, Vermont

Address correspondence to: Thomas Hays, Division of Neonatology, Department of Pediatrics, Columbia University Irving Medical Center, 622 West 168th S, Suite PH 17W-303, New York, New York, 10032, th2712@cumc.columbia.edu

Abstract: Preterm infants (<34 weeks gestation) experience high rates of morbidity and mortality before hospital discharge. Genetic disorders substantially contribute to morbidity and mortality in related populations. The prevalence and clinical impact of genetic disorders is unknown in this population. We sought to determine the prevalence of commonly diagnosed genetic disorders in preterm infants, and to determine the association of disorders with morbidity and mortality. This was a retrospective multicenter cohort study of infants born from 23 to 33 weeks gestation between 2000 and 2020. Genetic disorders were abstracted from diagnoses present in electronic health records. We excluded infants transferred from or to other health care facilities prior to discharge or death when analyzing clinical outcomes. We determined the adjusted odds of pre-discharge morbidity or mortality after adjusting for known risk factors. Of 409,704 infants, 5838 (1.4%) had genetic disorders. Infants with trisomy 13, 18, 21, or cystic fibrosis had greater adjusted odds of severe morbidity or mortality. Of the 17,427 infants who died, 566 (3.2%) had genetic disorders. Of the 65,968 infants with a severe morbidity, 1319 (2.0%) had genetic disorders. Our work demonstrates that genetic disorders are prevalent in preterm infants, especially those with life-threatening morbidities. Clinicians should consider genetic testing for preterm infants with severe morbidity and maintain a higher index of suspicion for life-threatening morbidities in preterm infants with genetic disorders. Prospective genomic research is needed to clarify the prevalence of genetic disorders in this population, and the contribution of genetic disorders to preterm birth and subsequent morbidity and mortality.

Introduction

Preterm infants, particularly those born before 34 weeks' gestation, experience high rates of severe morbidity and mortality prior to discharge.[1-3] Genetic disorders have been shown to strongly contribute to morbidity and mortality in related populations including full-term infants and children with critical illness.[4-7] Genetic disorders are particularly prevalent in stillbirth and fetal disorders.[8-10] Despite this, there are limited descriptions of the prevalence of genetic disorders in preterm infants. Boghossian et al. described the prevalence, morbidity, and mortality of very low birthweight infants with common aneuploidies.[11,12] However, no studies to our knowledge have provided a comprehensive description of genetic disorders diagnosed in this population. We therefore sought to achieve the following objectives: 1) determine the prevalence of commonly diagnosed genetic disorders in preterm infants, and 2) determine the association of these genetic disorders with severe morbidity and mortality in preterm infants.

We recently utilized the Pediatrix Clinical Data Warehouse (CDW) in related studies, to investigate the association of congenital renal anomalies with genetic disorders in preterm infants.[13] CDW contains detailed records of 20% of preterm infants born in the United States. Given this utility and broad representation, we sought to employ a similar strategy in this study.

Methods

This study was approved by the institutional review board of Columbia University. This report followed the guidelines for cohort studies as outlined by Strengthening the Reporting of

Observational Studies in Epidemiology (STROBE). This was a retrospective study using the CDW multicenter cohort of infants born prior from 23 to 34 weeks' gestation.

Prevalence of commonly diagnosed genetic disorders in preterm infants.

Multiple genetic disorders, including common aneuploidies, are associated with fetal growth restriction. Therefore, we used a gestational age-based cutoff rather than weight-based. We extracted records from 409,704 infants discharged from 2000 to 2020. Data from NICU admission to discharge were collected, including clinician-noted diagnosis of genetic disorders. This cohort did not undergo prospective, genome-wide genetic testing. Diagnoses were limited to those made during standard care.

Association of commonly diagnosed genetic disorders with severe morbidity and mortality in preterm infants.

We collected infant characteristics (gestational age, birth weight categorization by Fenton criteria,[14] maternal self-identified race, sex, mode of delivery, exposure to antenatal betamethasone, intubation in the first 72 hours of life, year of discharge, and the presence of congenital anomalies) and data regarding clinical outcomes (diagnoses of severe co-morbidities, mortality, and disposition). Severe morbidities were defined as the following: acute kidney injury (AKI) defined by the diagnosis of oliguria or anuria (ICD-9 788.5 or ICD-10: R34); severe intracranial hemorrhage (ICH) defined as grade III or IV intraventricular hemorrhage by Papile criteria or cystic periventricular leukomalacia;[15] medically or surgically treated necrotizing enterocolitis (NEC); severe bronchopulmonary dysplasia (BPD) defined as invasive

mechanical ventilation at 36 weeks' postmenstrual age; severe retinopathy of prematurity (ROP) defined as need for any medical or surgical intervention; culture-positive sepsis defined by any positive blood or urine culture; or shock defined as administration of any vasopressor or inotrope.

We used logistic regression to model the relationship between genetic disorders and severe morbidity or mortality. The following were covariates: year of discharge, gestational age, birth weight (here as a Z-score normalized to gestational age and sex to account for continuous rather than categorical effects), sex, maternal race, mode of delivery, intubation in the first 72 hours of life, presence of a congenital anomaly, and exposure to antenatal betamethasone. We excluded infants with missing covariate data, as well as those transferred after birth (often because of complex illness) and infants transferred prior to discharge (in which the clinical outcome is unknown) to limit ascertainment biases.[16] We determined crude and adjusted odds ratios (OR) after controlling for these covariates. OR were reported with 99.97% confidence intervals (determined as $1 - \frac{\alpha}{n}$) to correct for n comparisons of 9 clinical outcomes in 22 categories of genetic disorders and accepting statistical significance at a Bonferroni corrected, two-tailed α of 0.05. Analyses were made using RStudio (version 2022.02.0; script in **eMethods**).[17]

Results

Prevalence of commonly diagnosed genetic disorders in preterm infants.

Genetic disorders were identified in 5,838 (1.42 %) of the 409,704 infants in this cohort (**Table 1**). Of the 5,838 infants with a genetic diagnosis, 2706 (46.35%) were females, 3121 (53.46%) were males, and 11 (0.19%) were infants with ambiguous genitalia. The disorders included in the dataset were comprised of 19 specific diagnoses, as well as unspecified aneuploidy, unspecified copy number variant, and infants found to have multiple genetic disorders. Aneuploidies (trisomy 13, 18, 21, and 22, Klinefelter syndrome, Turner syndrome) were present in 2,804 infants, with trisomy 21 accounting for more than half (1,463) of these cases. Copy number variants (13q deletion, Cri-du-chat syndrome, DiGeorge syndrome) were present in 155 infants. Unspecified aneuploidy and copy number variants were noted in 856 and 54 infants respectively. Eleven distinct monogenic disorders were identified in 2,859 infants. Most (2,449) of these had hematologic disorders. Thalassemia accounted for 1,485 cases. Sickle cell disease was identified in 252 infants and cystic fibrosis in 274. 20 infants with multiple genetic disorders were identified.

Association of commonly diagnosed genetic disorders with severe morbidity and mortality in preterm infants.

Following exclusions, 320,582 preterm infants remained for analysis of clinical outcomes. Genetic disorders were prevalent in 4,196 (1.3%) of these individuals. The rate of mortality in preterm infants without genetic disorders in this cohort, and following exclusions was 5.3%. Genetic disorders were present in 566 (3.2%) of preterm infants who died, in 1,319 (2.0%) of preterm infants who experienced any severe morbidity or mortality, in 3,630 (1.2%) of preterm

infants who survived to discharge, and in 2,664 (1.1%) of preterm infants who survived to discharge free of severe morbidity. The mortality rate after exclusions for each genetic disorder is provided in Table 1. The highest crude rates of mortality were found in cases of aneuploidy. 84.5 and 83.9% of preterm infants with trisomy 13 and 18 died before discharge. 14.7% of preterm infants with trisomy 21 died before discharge. Mortality rates for preterm infants with copy number variants varied from 12.5% in Cri-du-chat syndrome to 20.0% in cases of unspecified copy number variants. Preterm infants with hematologic disorders had lower observed rates of mortality, with the highest rates found in cases of sickle cell disease (2.6%) and thalassemia (2.2%). Varying rates of mortality were noted in preterm infants with miscellaneous monogenic disorders, including 9.7% mortality for infants with cystic fibrosis.

We then determined the crude and adjusted odds of severe morbidity and mortality for preterm infants with diagnosed genetic disorders as compared to preterm infants without known genetic disorders. The baseline characteristics for covariates used in these analyses are provided for preterm infants with aneuploidies (**Supplemental Table 1**), copy number variants (**Supplemental Table 2**), hematologic disorders (**Supplemental Table 3)**, and miscellaneous monogenic disorders (**Supplemental Table 4**). The number of preterm infants experiencing morbidity and mortality for each diagnosed genetic disorder appears in **Supplemental Table 5.**

Preterm infants with aneuploidies had significantly greater adjusted odds of several outcomes (**Table 2**). Preterm infants with trisomy 13, 18, and 21 had greater adjusted odds of death. Preterm infants with trisomy 13 also had greater adjusted odds of AKI. Preterm infants with

trisomy 21 had greater adjusted odds of AKI, BPD, sepsis, shock, and death. Preterm infants with copy number variants and hematologic disorders were not found to have significantly greater adjusted odds of morbidity or mortality (**Tables 3 and 4**). The only monogenic disorder diagnosis associated with significantly increased adjusted odds of morbidity or mortality was cystic fibrosis **(Table 5)**. Preterm infants with cystic fibrosis had significantly increased adjusted odds of NEC, BPD, and ROP. Elevated risk of morbidities in preterm infants with hypophosphatasia did not persist after adjustment for known risk factors.

Discussion

This study provides the most comprehensive determination to our knowledge of the prevalence of genetic disorders in preterm infants and short-term risk of morbidity and mortality conferred by genetic disorders in this population. We found that genetic disorders, consisting of aneuploidies, copy number variants, and monogenic disorders, were diagnosed in 1.4% of preterm infants. Trisomy 13, 18, and 21, and cystic fibrosis were associated with increased risk of morbidity or mortality after adjustment for known risk factors. Rates of comorbidities for preterm infants with trisomy 13 and 18 may be underestimated given mortality as a competing outcome, particularly following redirection to palliative goals of care.

The overall prevalence of genetic disorders may reflect a contribution of genetic disease to shorter gestation. Several of the most prevalent genetic disorders we analyzed are associated with preterm birth including hemoglobinopathies,[18] cystic fibrosis,[19] and common aneuploidies.[20] This results from a complex interplay of factors including disrupted

placentation,[20] increased risk of fetal growth restriction and stillbirth, closer monitoring of pregnancy, greater maternal age contributing to aneuploidy and risk of preterm birth,[21] as well as idiopathic factors.

We found prevalence rates of trisomy 13 and 18 lower than previously described in preterm infants,[12] however our analysis used a gestational age-based cutoff rather than birth weight. Trisomy 13 and 18 cause fetal growth restriction. Given the same study population, a weight-based cutoff will select for infants with gestational age > 33 weeks with these genetic disorders and estimate a higher prevalence rate. Given these selection differences, it is difficult to compare morbidity and mortality between our cohorts and previous descriptions.[11]

Interestingly, we found increased adjusted odds of NEC, BPD, and ROP in preterm infants with cystic fibrosis. Infants with cystic fibrosis have been found to be born at earlier gestational age and with lower birth weight.[19,22] However, after adjustment for these factors, we found an increased risk of morbidity. Intestinal disease caused by meconium peritonitis and meconium ileus may mimic or predispose infants to NEC.[23,24] BPD and ROP may be attributable to disrupted lung disease, and exposure to oxygen.[25,26] Maternal cystic fibrosis also contributes to neonatal morbidity,[27] and may account for a portion of observed outcomes. However, maternal diagnoses were not available in our cohort.

This retrospective study has significant strengths and important weaknesses. It encompasses a large number of infants, cared for in multiple institutions, and is highly generalizable. The depth

of data, including patient characteristics, diagnoses of genetic disorders, and co-morbidities should enable clinicians to better counsel families and inform clinical decision-making in rare presentations. It encompasses a broad period, during which genetic testing has undergone considerable change. This may contribute to heterogeneity in the cohort with respect to year of birth. As genome-wide testing becomes more common, genetic disorders may be diagnosed more frequently in this population. This dataset lacks information regarding social determinants of health and maternal access to care (including genetic testing). Multiple factors could have contributed to over- or underestimation of the prevalence of genetic disorders. Preterm infants experience high rates of morbidity and mortality, which may prompt testing for genetic disease in some settings. Genetic disorders often present non-specifically in preterm infants,[4,28] compared to more complete phenotypes found in term infants or pediatric or adult populations. Therefore, genetic disorders may have gone undetected in this cohort. Unfortunately, we do not have data regarding the extent of pre- and postnatal genetic testing, including negative results, in this cohort. Crucially, these data do not include broad, prospective screening for genetic disorders, but rather only the retrospectively provided data from electronic health records. Given the prevalence of genetic disorders we found, particularly in preterm infants who experienced morbidity or mortality, there is a crucial need for broad, prospective studies to establish the true prevalence of genetic disorders in preterm infants and to determine the portion of morbidity and mortality attributable to genetic disorders in this population.

In summary, using a multicenter retrospective cohort, we found that genetic disorders are present in 1.4% of infants born before 34 weeks' gestation, and that trisomy 13, 18, 21, and cystic fibrosis are associated with greater adjusted odds of death or severe morbidity. We determined the prevalence of 19 distinct genetic disorders, and the crude and adjusted risk of mortality and severe morbidities. These data may help clinicians provide more informed guidance to families and lead to better clinical decision-making in rare presentations.[29] These data indicate the need for a greater index of suspicion for multiple critical illnesses, such as sepsis in the setting of trisomy 21 and NEC in the setting of cystic fibrosis. Given the trend towards greater adjusted odds of multiple morbid outcomes, clinicians should also consider genetic evaluation when caring for preterm infants with critical illnesses. Prospective, genome-wide screening is needed to determine the true prevalence of genetic disorders in this high-risk population, the contribution of genetic disorders to preterm birth, and the portion of disease in preterm infants attributable to genetic disorders.

Acknowledgments

The authors wish to gratefully acknowledge Dr. Ali G. Gharavi (Columbia University Irving Medical Center) for their mentorship and thoughtful review of this project.

Author TH is supported by two career development awards: the *Thrasher Research Fund Early Career Award* and the *NIH National Center for Advancing Translational Sciences* (KL2TR001874).

Table 1. Prevalence of genetic disorders in preterm infants and associated rates of mortality.

Genetic Disorder	n (Prevalence per 100,000 Preterm Infants)	n* (Mortality Rate %)
Aneuploidy	2804 (684.4)	480/1863 (25.76)
Trisomy 13	149 (36.4)	93/110 (84.55)
Trisomy 18	229 (55.9)	130/155 (83.87)
Trisomy 21	1463 (357.1)	143/973 (14.7)
Trisomy 22	3 (0.7)	1/1 (100)
Klinefelter syndrome	45 (11)	2/35 (5.71)
Turner syndrome	59 (14.4)	5/41 (12.2)
Unspecified Aneuploidy	856 (208.9)	106/548 (19.34)
Copy number variation disorder	155 (37.8)	14/87 (16.09)
13q deletion syndrome	11 (2.7)	1/7 (14.29)
Cri-du-chat syndrome	22 (5.4)	2/16 (12.5)
DiGeorge syndrome	68 (16.6)	4/29 (13.79)
Unspecified CNV	54 (13.2)	7/35 (20)
Monogenic disorder	2859 (697.8)	55/2068 (2.66)
Hematologic disorder	2449 (597.7)	39/1973 (1.98)
Thalassemia	1485 (362.5)	26/1174 (2.21)
Hemoglobin bart's	451 (110.1)	4/384 (1.04)
Sickle cell disease	252 (61.5)	5/194 (2.58)
G6PD deficiency	241 (58.8)	2/206 (0.97)
Hemophilia A	14 (3.4)	2/10 (20)
Hemophilia B	5 (1.2)	0/5 (0)
Von Willebrand disease	1 (0.2)	0
Miscellaneous	410 (100.1)	32/260 (12.31)
Cystic fibrosis	274 (66.9)	16/165 (9.7)
Fragile X syndrome	8 (2)	0/7 (0)
Hypophosphatasia	125 (30.5)	15/86 (17.44)
Fanconi Anemia	3 (0.7)	1/2 (50)
Multiple Disorders	20 (4.9)	1/13 (7.69)

*n provided as number of cases of mortality over the number of preterm infants with each genetic disorder, following exclusions.

Table 2. Odds of severe morbidity or mortality in preterm infants with aneuploidies[1]

Outcome	Trisomy 13 (n=110)		Trisomy 18 (n=155)		Trisomy 21 (n=973)		Klinefelter syn. (n=35)		Turner syn. (n=41)	
	Crude OR	Adjusted OR	Crude OR	Adjusted OR	Crude OR	Adjusted OR	Crude OR	Adjusted OR	Crude OR	Adjusted OR
Mortality or any severe morbidity	53.42 (15.61, 356.05)*	61.87 (16.15,437.72)*	29.47 (12.34, 90.49)	28.09 (10.56,93.09)*	2.04 (1.58, 2.62)	2.39 (1.75,3.26)*	1.66 (0.35, 6.18)	0.98 (0.16,4.75)	1.88 (0.48, 6.27)	2.43 (0.47,10.51)
Mortality	97.18 (39.3, 303.91)*	265.24 (93.12,916.46)*	92.37 (43.1, 231.52)	196.84 (81.82,539.66)*	3.06 (2.14, 4.26)	5.93 (3.83,8.95)*	1.08 (0.01, 8.2)	0.81 (0.01,8.14)	2.47 (0.22, 11.21)	4.17 (0.3,29.12)
AKI	6.67 (2.19, 16.27)*	4.17 (1.24,11.54)*	3.12 (0.84, 8.17)	1.91 (0.5,5.34)	3.39 (2.15, 5.07)	4.02 (2.43,6.35)*	2.06 (0.02, 15.82)	1.47 (0.01,13.72)	4.73 (0.43, 21.61)	6.91 (0.55,40.06)
ICH	—	—	0.32 (0, 2.17)	0.28 (0,1.99)	1.09 (0.56, 1.9)	1.29 (0.64,2.36)	1.5 (0.01, 11.48)	1.12 (0.01,10.61)	1.95 (0.06, 11.49)	2.2 (0.07,17.02)
NEC	1.33 (0.12, 5.36)	0.77 (0.07,3.31)	0.93 (0.09, 3.68)	0.58 (0.05,2.36)	1.55 (0.85, 2.57)	1.44 (0.78,2.45)	—	—	1.43 (0.01, 10.69)	1.55 (0.01,12.75)
BPD	5.37 (0.8, 18.46)	1.84 (0.26,7.04)	5.45 (1.19, 15.76)	1.78 (0.37,5.5)	4.1 (2.21, 6.91)	3.32 (1.71,5.94)*	—	—	—	—
ROP	—	—	—	—	0.34 (0.02, 1.44)	0.47 (0.03,2.24)	—	—	—	—
Sepsis	0.89 (0.17, 2.73)	0.49 (0.09,1.65)	0.76 (0.18, 2.11)	0.44 (0.1,1.28)	1.49 (1.02, 2.11)	1.56 (1.04,2.29)*	2.95 (0.47, 11.77)	2.31 (0.34,10.54)	2.8 (0.51, 10.24)	3.73 (0.6,16.07)
Shock	2.11 (0.74, 5)	1.08 (0.34,2.89)	2.38 (1.04, 4.85)	1.44 (0.58,3.25)	1.8 (1.28, 2.48)	2.49 (1.64,3.7)*	1.23 (0.08, 6.41)	0.89 (0.05,6.22)	1.63 (0.19, 6.86)	2.03 (0.2,11.97)

Abbreviations: AKI (acute kidney injury), BPD (bronchopulmonary dysplasia), ICH (intracerebral hemorrhage), OR (odds ratio) NEC (necrotizing enterocolitis), ROP (retinopathy of prematurity)

[1]odd ratio (99.97% confidence interval), * indicates two-tailed, Bonferroni-corrected p-value < 0.05

Table 3. Odds of severe morbidity or mortality in preterm infants with copy number variants[1]

Outcome	DiGeorge syndrome (n=29)		Cri-du-chat syndrome (n=16)	
	Crude OR	**Adjusted OR**	**Crude OR**	**Adjusted OR**
Mortality or any severe morbidity	2.56 (0.54, 10.68)	2.02 (0.29,12.01)	1.65 (0.13, 11.54)	1.64 (0.09,17.42)
Mortality	2.84 (0.17, 15.59)	2.59 (0.12,23.83)	2.54 (0.02, 24.38)	4.6 (0.04,73.4)
AKI	—	—	4.87 (0.05, 47.1)	5.67 (0.05,79.46)
ICH	—	—	—	—
NEC	—	—	3.99 (0.04, 38.63)	3.93 (0.04,44.41)
BPD	5.85 (0.06, 46.55)	2.62 (0.02,32.83)	—	—
ROP	—	—	—	—
Sepsis	2.6 (0.3, 12.16)	2.22 (0.22,12.98)	3.32 (0.19, 24.16)	3.91 (0.19,35.63)
Shock	1.52 (0.09, 8.35)	0.98 (0.05,8.17)		

Abbreviations: AKI (acute kidney injury), BPD (bronchopulmonary dysplasia), ICH (intracerebral hemorrhage), OR (odds ratio) NEC (necrotizing enterocolitis), ROP (retinopathy of prematurity)

[1]odd ratio (99.97% confidence interval), * indicates two-tailed, Bonferroni-corrected p-value < 0.05

Table 4. Odds of severe morbidity or mortality in preterm infants with hematologic disorders[1]

Outcome	Thalassemia (n=1174)		Hemoglobin Bart's (n=384)		Sickle cell disease (n=194)		G6PD deficiency (n=206)	
	Crude OR	Adjusted OR	Crude OR	Adjusted OR	Crude OR	Adjusted OR	Crude OR	Adjusted OR
Mortality or severe morbidity	1.28 (0.99, 1.64)	0.83 (0.6,1.14)	0.75 (0.44, 1.23)	0.86 (0.45,1.57)	0.91 (0.44, 1.73)	0.64 (0.26,1.45)	0.72 (0.33, 1.39)	0.78 (0.31,1.77)
Mortality	0.4 (0.17, 0.79)*	0.24 (0.1,0.49)*	0.19 (0.01, 0.8)*	0.21 (0.01,1)	0.47 (0.04, 1.83)	0.32 (0.03,1.47)	—	—
AKI	1.42 (0.76, 2.4)	1.04 (0.54,1.82)	0.36 (0.02, 1.55)	0.4 (0.03,1.86)	1.66 (0.33, 4.94)	1.4 (0.26,4.64)	0.85 (0.08, 3.31)	1.05 (0.1,4.5)
ICH	1.36 (0.79, 2.16)	1.07 (0.61,1.77)	0.53 (0.09, 1.61)	0.63 (0.11,2.09)	0.79 (0.1, 2.82)	0.67 (0.08,2.64)	0.87 (0.13, 2.89)	1.09 (0.16,4.04)
NEC	1.37 (0.78, 2.24)	0.98 (0.54,1.62)	0.67 (0.13, 1.95)	0.71 (0.14,2.13)	0.29 (0, 1.94)	0.23 (0,1.59)	1.43 (0.31, 4.06)	1.54 (0.33,4.58)
BPD	1.37 (0.51, 2.92)	0.99 (0.36,2.19)	1.04 (0.1, 4)	1.26 (0.12,5.23)	2.09 (0.2, 8.18)	1.82 (0.17,7.81)	1.96 (0.19, 7.67)	2.58 (0.24,10.96)
ROP	1.77 (0.74, 3.53)	1.42 (0.56,3.02)	—	—	0.85 (0.01, 5.65)	1.04 (0.01,8.08)	—	—
Sepsis	1.53 (1.09, 2.09)	1.14 (0.78,1.61)	1 (0.48, 1.85)	1.18 (0.53,2.31)	1.21 (0.45, 2.66)	1.02 (0.36,2.46)	0.78 (0.24, 1.91)	0.84 (0.24,2.2)
Shock	1.13 (0.78, 1.6)	0.75 (0.49,1.11)	0.66 (0.27, 1.34)	0.86 (0.32,1.96)	0.63 (0.16, 1.66)	0.44 (0.1,1.34)	0.43 (0.09, 1.29)	0.54 (0.1,1.87)

Abbreviations: AKI (acute kidney injury), BPD (bronchopulmonary dysplasia), ICH (intracerebral hemorrhage), OR (odds ratio) NEC (necrotizing enterocolitis), ROP (retinopathy of prematurity)
[1]odd ratio (99.97% confidence interval), * indicates two-tailed, Bonferroni-corrected p-value < 0.05

Table 5. Odds of severe morbidity or mortality in preterm infants with miscellaneous genetic disorders[1]

Outcome	Cystic Fibrosis (n=165)		Hypophosphatasia (n=86)	
	Crude OR	**Adjusted OR**	**Crude OR**	**Adjusted OR**
Mortality or severe morbidity	5.18 (2.86, 9.57)*	2.97 (1.38, 6.47)*	7.14 (3.09, 17.88)*	2.08 (0.7, 6.74)
Mortality	1.91 (0.6, 4.62)	0.69 (0.2, 1.96)	3.75 (1.09, 9.97)*	0.9 (0.24, 2.7)
AKI	4.17 (1.4, 9.77)*	1.92 (0.6, 5.03)	7.79 (2.35, 20.38)*	2.43 (0.69, 7)
ICH	2.29 (0.66, 5.83)	1.07 (0.29, 3)	3.25 (0.69, 9.91)	1.21 (0.25, 4.02)
NEC	7.52 (3.44, 14.88)*	4.03 (1.73, 8.59)*	4.53 (1.12, 12.95)*	1.62 (0.39, 4.92)
BPD	9.67 (3.25, 22.68)*	4.63 (1.45, 12.03)*	7 (1.03, 24.62)*	1.92 (0.27, 7.29)
ROP	9.36 (3.04, 22.35)*	3.95 (1.13, 11.45)*	11.95 (2.76, 35.24)*	3.16 (0.67, 10.7)
Sepsis	3.73 (1.85, 7.05)*	1.91 (0.87, 4.01)	6.2 (2.57, 14.14)*	2.05 (0.78, 5.21)
Shock	4.26 (2.18, 7.89)*	1.93 (0.87, 4.17)	6.22 (2.6, 14.16)*	1.82 (0.67, 4.89)

Abbreviations: AKI (acute kidney injury), BPD (bronchopulmonary dysplasia), ICH (intracerebral hemorrhage), OR (odds ratio) NEC (necrotizing enterocolitis), ROP (retinopathy of prematurity)
[1]odd ratio (99.97% confidence interval), * indicates two-tailed, Bonferroni-corrected p-value < 0.05

Supplemental Content

Supplemental Table 1. Baseline characteristics of preterm infants diagnosed with aneuploidies.
Supplemental Table 2. Baseline characteristics of preterm infants diagnosed with copy number variants.
Supplemental Table 3. Baseline characteristics of preterm infants diagnosed with hematologic disorders.
Supplemental Table 4. Baseline characteristics of preterm infants diagnosed with miscellaneous genetic disorders.
Supplemental Table 5. Instances of severe morbidity and mortality in preterm infants with genetic disorders.
Methods. R script

References

1. Ray JG, Park AL, Fell DB. Mortality in Infants Affected by Preterm Birth and Severe Small-for-Gestational Age Birth Weight. *Pediatrics*. Dec 2017;140(6)doi:10.1542/peds.2017-1881
2. Horbar JD, Carpenter JH, Badger GJ, et al. Mortality and neonatal morbidity among infants 501 to 1500 grams from 2000 to 2009. *Pediatrics*. Jun 2012;129(6):1019-26. doi:10.1542/peds.2011-3028
3. Baer RJ, Rogers EE, Partridge JC, et al. Population-based risks of mortality and preterm morbidity by gestational age and birth weight. *J Perinatol*. Nov 2016;36(11):1008-1013. doi:10.1038/jp.2016.118
4. Hays T, Wapner RJ. Genetic testing for unexplained perinatal disorders. *Curr Opin Pediatr*. Apr 1 2021;33(2):195-202. doi:10.1097/MOP.0000000000000999
5. Kingsmore SF, Henderson A, Owen MJ, et al. Measurement of genetic diseases as a cause of mortality in infants receiving whole genome sequencing. *NPJ Genom Med*. 2020;5(1):49. doi:10.1038/s41525-020-00155-8
6. Owen MJ, Wright MS, Batalov S, et al. Reclassification of the Etiology of Infant Mortality With Whole-Genome Sequencing. *JAMA Network Open*. 2023;6(2):e2254069-e2254069. doi:10.1001/jamanetworkopen.2022.54069
7. Clark MM, Hildreth A, Batalov S, et al. Diagnosis of genetic diseases in seriously ill children by rapid whole-genome sequencing and automated phenotyping and interpretation. *Sci Transl Med*. Apr 24 2019;11(489):eaat6177-eaat6177. doi:10.1126/scitranslmed.aat6177
8. Stanley KE, Giordano J, Thorsten V, et al. Causal Genetic Variants in Stillbirth. *N Engl J Med*. Sep 17 2020;383(12):1107-1116. doi:10.1056/NEJMoa1908753
9. Petrovski S, Aggarwal V, Giordano JL, et al. Whole-exome sequencing in the evaluation of fetal structural anomalies: a prospective cohort study. *Lancet*. Feb 23 2019;393(10173):758-767. doi:10.1016/S0140-6736(18)32042-7
10. Lord J, McMullan DJ, Eberhardt RY, et al. Prenatal exome sequencing analysis in fetal structural anomalies detected by ultrasonography (PAGE): a cohort study. *Lancet*. Feb 23 2019;393(10173):747-757. doi:10.1016/S0140-6736(18)31940-8
11. Boghossian NS, Hansen NI, Bell EF, et al. Mortality and morbidity of VLBW infants with trisomy 13 or trisomy 18. *Pediatrics*. Feb 2014;133(2):226-35. doi:10.1542/peds.2013-1702
12. Boghossian NS, Horbar JD, Carpenter JH, Murray JC, Bell EF. Major chromosomal anomalies among very low birth weight infants in the Vermont Oxford Network. *J Pediatr*. May 2012;160(5):774-780.e11. doi:10.1016/j.jpeds.2011.11.005
13. Hays T, Thompson MV, Bateman DA, et al. The Prevalence and Clinical Significance of Congenital Anomalies of the Kidney and Urinary Tract in Preterm Infants. *JAMA Netw Open*. Sep 1 2022;5(9):e2231626. doi:10.1001/jamanetworkopen.2022.31626
14. Fenton TR, Kim JH. A systematic review and meta-analysis to revise the Fenton growth chart for preterm infants. *BMC Pediatr*. Apr 20 2013;13:59. doi:10.1186/1471-2431-13-59
15. Papile LA, Burstein J, Burstein R, Koffler H. Incidence and evolution of subependymal and intraventricular hemorrhage: a study of infants with birth weights less than 1,500 gm. *J Pediatr*. Apr 1978;92(4):529-34. doi:10.1016/s0022-3476(78)80282-0

16. Tolia VN, Clark RH. The Denominator Matters! Lessons from Large Database Research in Neonatology. *Children (Basel)*. Nov 7 2020;7(11):216-216. doi:10.3390/children7110216
17. *R: A language and environment for statistical computing.* R Foundation for Statistical Computing; 2022. https://www.R-project.org/
18. ACOG Practice Bulletin No. 78: hemoglobinopathies in pregnancy. *Obstet Gynecol*. Jan 2007;109(1):229-37. doi:10.1097/00006250-200701000-00055
19. Festini F, Taccetti G, Repetto T, et al. Gestational and neonatal characteristics of children with cystic fibrosis: a cohort study. *J Pediatr*. Sep 2005;147(3):316-20. doi:10.1016/j.jpeds.2005.04.031
20. Bianco K, Gormley M, Farrell J, et al. Placental transcriptomes in the common aneuploidies reveal critical regions on the trisomic chromosomes and genome-wide effects. *Prenat Diagn*. Sep 2016;36(9):812-22. doi:10.1002/pd.4862
21. Savva GM, Walker K, Morris JK. The maternal age-specific live birth prevalence of trisomies 13 and 18 compared to trisomy 21 (Down syndrome). *Prenat Diagn*. Jan 2010;30(1):57-64. doi:10.1002/pd.2403
22. Schlüter DK, Griffiths R, Adam A, et al. Impact of cystic fibrosis on birthweight: a population based study of children in Denmark and Wales. *Thorax*. 2019;74(5):447-454. doi:10.1136/thoraxjnl-2018-211706
23. Neu J. Necrotizing Enterocolitis: The Future. *Neonatology*. 2020;117(2):240-244. doi:10.1159/000506866
24. Sathe M, Houwen R. Meconium ileus in Cystic Fibrosis. *J Cyst Fibros*. Nov 2017;16 Suppl 2:S32-s39. doi:10.1016/j.jcf.2017.06.007
25. Sly PD, Brennan S, Gangell C, et al. Lung disease at diagnosis in infants with cystic fibrosis detected by newborn screening. *Am J Respir Crit Care Med*. Jul 15 2009;180(2):146-52. doi:10.1164/rccm.200901-0069OC
26. Meyerholz DK, Stoltz DA, Gansemer ND, et al. Lack of cystic fibrosis transmembrane conductance regulator disrupts fetal airway development in pigs. *Lab Invest*. Jun 2018;98(6):825-838. doi:10.1038/s41374-018-0026-7
27. Jelin AC, Sharshiner R, Caughey AB. Maternal co-morbidities and neonatal outcomes associated with cystic fibrosis(). *J Matern Fetal Neonatal Med*. Jan 2017;30(1):4-7. doi:10.3109/14767058.2016.1161747
28. Petrikin JE, Willig LK, Smith LD, Kingsmore SF. Rapid whole genome sequencing and precision neonatology. *Semin Perinatol*. Dec 2015;39(8):623-31. doi:10.1053/j.semperi.2015.09.009
29. Hays T, Hernan R, Disco M, et al. Implementation of Rapid Genome Sequencing for Critically Ill Infants With Complex Congenital Heart Disease. *Circ Genom Precis Med*. Jul 7 2023:e004050. doi:10.1161/circgen.122.004050

Supplemental Table 1. Baseline characteristics of preterm infants diagnosed with aneuploidies.

	Trisomy 13 (n=110)	Trisomy 18 (n=155)	Trisomy 21 (n=973)	Trisomy 22 (n=1)	Klinefelter syndrome (n=35)	Turner syndrome (n=41)
Gestational age (weeks)						
Mean (SD)	30.9 (2.30)	31.0 (2.16)	31.0 (2.19)	32.0 (NA)	29.6 (2.57)	30.4 (2.63)
Median [Min, Max]	32.0 [24.0, 33.0]	32.0 [24.0, 33.0]	32.0 [23.0, 33.0]	32.0 [32.0, 32.0]	30.0 [23.0, 33.0]	31.0 [24.0, 33.0]
Birth weight						
SGA	27 (24.5%)	101 (65.2%)	190 (19.5%)	0 (0%)	9 (25.7%)	9 (22.0%)
AGA	77 (70.0%)	50 (32.3%)	666 (68.4%)	1 (100%)	26 (74.3%)	25 (61.0%)
LGA	6 (5.5%)	4 (2.6%)	117 (12.0%)	0 (0%)	0 (0%)	7 (17.1%)
Race						
Asian	1 (0.9%)	2 (1.3%)	30 (3.1%)	0 (0%)	2 (5.7%)	2 (4.9%)
Black	27 (24.5%)	29 (18.7%)	134 (13.8%)	0 (0%)	7 (20.0%)	2 (4.9%)
Hispanic	31 (28.2%)	54 (34.8%)	238 (24.5%)	1 (100%)	9 (25.7%)	7 (17.1%)
Other	8 (7.3%)	12 (7.7%)	81 (8.3%)	0 (0%)	1 (2.9%)	2 (4.9%)
White	43 (39.1%)	58 (37.4%)	490 (50.4%)	0 (0%)	16 (45.7%)	28 (68.3%)
Sex						
Female	47 (42.7%)	96 (61.9%)	425 (43.7%)	1 (100%)	8 (22.9%)	41 (100%)
Male	63 (57.3%)	59 (38.1%)	548 (56.3%)	0 (0%)	27 (77.1%)	0 (0%)
Delivery mode						
Cesarean Section	81 (73.6%)	119 (76.8%)	766 (78.7%)	1 (100%)	28 (80.0%)	32 (78.0%)
Vaginal Delivery	29 (26.4%)	36 (23.2%)	207 (21.3%)	0 (0%)	7 (20.0%)	9 (22.0%)
Antenatal betamethasone						
Absent	40 (36.4%)	72 (46.5%)	214 (22.0%)	0 (0%)	7 (20.0%)	9 (22.0%)
Present	70 (63.6%)	83 (53.5%)	759 (78.0%)	1 (100%)	28 (80.0%)	32 (78.0%)
Intubation within 72 hours						
Absent	31 (28.2%)	54 (34.8%)	628 (64.5%)	0 (0%)	18 (51.4%)	24 (58.5%)
Present	79 (71.8%)	101 (65.2%)	345 (35.5%)	1 (100%)	17 (48.6%)	17 (41.5%)

	Trisomy 13 (n=110)	Trisomy 18 (n=155)	Trisomy 21 (n=973)	Trisomy 22 (n=1)	Klinefelter syndrome (n=35)	Turner syndrome (n=41)
Discharge Year						
2000 to 2004	27 (24.5%)	31 (20.0%)	144 (14.8%)	0 (0%)	0 (0%)	10 (24.4%)
2005 to 2009	23 (20.9%)	48 (31.0%)	257 (26.4%)	0 (0%)	7 (20.0%)	11 (26.8%)
2010 to 2014	31 (28.2%)	35 (22.6%)	258 (26.5%)	0 (0%)	13 (37.1%)	7 (17.1%)
2015 to 2020	29 (26.4%)	41 (26.5%)	314 (32.3%)	1 (100%)	15 (42.9%)	13 (31.7%)
Anomaly						
Absent	38 (34.5%)	64 (41.3%)	579 (59.5%)	0 (0%)	28 (80.0%)	32 (78.0%)
Present	72 (65.5%)	91 (58.7%)	394 (40.5%)	1 (100%)	7 (20.0%)	9 (22.0%)
Morbidity						
Absent	7 (6.4%)	17 (11.0%)	623 (64.0%)	0 (0%)	24 (68.6%)	27 (65.9%)
Present	103 (93.6%)	138 (89.0%)	350 (36.0%)	1 (100%)	11 (31.4%)	14 (34.1%)
Disposition						
Alive	17 (15.5%)	25 (16.1%)	830 (85.3%)	0 (0%)	33 (94.3%)	36 (87.8%)
Died	93 (84.5%)	130 (83.9%)	143 (14.7%)	1 (100%)	2 (5.7%)	5 (12.2%)

Abbreviations: SGA (small for gestational age), AGA (appropriate for gestational age), LGA (large for gestational age)

Supplemental Table 2. Baseline characteristics of preterm infants diagnosed with copy number variants.

	DiGeorge syndrome (n=29)	Cri-du-chat syndrome (n=16)	13q-deletion syndrome (n=7)
Gestational age (weeks)			
Mean (SD)	30.6 (2.40)	31.3 (2.65)	31.1 (1.35)
Median [Min, Max]	31.0 [25.0, 33.0]	32.0 [24.0, 33.0]	31.0 [29.0, 33.0]
Birth weight			
SGA	3 (10.3%)	2 (12.5%)	3 (42.9%)
AGA	23 (79.3%)	14 (87.5%)	4 (57.1%)
LGA	3 (10.3%)	0 (0%)	0 (0%)
Race			
Black	9 (31.0%)	6 (37.5%)	4 (57.1%)
Hispanic	9 (31.0%)	4 (25.0%)	3 (42.9%)
Other	2 (6.9%)	1 (6.3%)	0 (0%)
White	8 (27.6%)	5 (31.3%)	0 (0%)
Asian	1 (3.4%)	0 (0%)	0 (0%)
Sex			
Female	17 (58.6%)	12 (75.0%)	3 (42.9%)
Male	12 (41.4%)	4 (25.0%)	4 (57.1%)
Delivery mode			
Cesarean Section	24 (82.8%)	12 (75.0%)	5 (71.4%)
Vaginal Delivery	5 (17.2%)	4 (25.0%)	2 (28.6%)
Antenatal betamethasone			
Absent	9 (31.0%)	1 (6.3%)	2 (28.6%)
Present	20 (69.0%)	15 (93.8%)	5 (71.4%)
Intubation within 72 hours			
Absent	14 (48.3%)	8 (50.0%)	4 (57.1%)
Present	15 (51.7%)	8 (50.0%)	3 (42.9%)
Discharge Year			
2000 to 2004	1 (3.4%)	2 (12.5%)	0 (0%)
2005 to 2009	8 (27.6%)	5 (31.3%)	0 (0%)
2010 to 2014	6 (20.7%)	2 (12.5%)	1 (14.3%)
2015 to 2020	14 (48.3%)	7 (43.8%)	6 (85.7%)

	DiGeorge syndrome (n=29)	Cri-du-chat syndrome (n=16)	13q-deletion syndrome (n=7)
Anomaly			
Absent	13 (44.8%)	8 (50.0%)	5 (71.4%)
Present	16 (55.2%)	8 (50.0%)	2 (28.6%)
Morbidity			
Absent	17 (58.6%)	11 (68.8%)	5 (71.4%)
Present	12 (41.4%)	5 (31.3%)	2 (28.6%)
Disposition			
Alive	25 (86.2%)	14 (87.5%)	6 (85.7%)
Died	4 (13.8%)	2 (12.5%)	1 (14.3%)

Abbreviations: SGA (small for gestational age), AGA (appropriate for gestational age), LGA (large for gestational age)

Supplemental Table 3. Baseline characteristics of preterm infants diagnosed with hematologic disorders.

	Thalassemia (n=1174)	Hemoglobin Bart's (n=384)	Sickle cell anemia (n=194)	G6PD deficiency (n=206)	Hemophilia A (n=10)	Hemophilia B (n=5)
Gestational age (weeks)						
Mean (SD)	29.5 (2.73)	30.1 (2.43)	29.7 (2.81)	30.3 (2.48)	29.6 (3.06)	29.4 (3.36)
Median [Min, Max]	30.0 [23.0, 33.0]	31.0 [23.0, 33.0]	30.0 [23.0, 33.0]	31.0 [23.0, 33.0]	31.0 [25.0, 33.0]	28.0 [26.0, 33.0]
Birth weight						
SGA	137 (11.7%)	32 (8.3%)	35 (18.0%)	25 (12.1%)	1 (10.0%)	0 (0%)
AGA	996 (84.8%)	332 (86.5%)	152 (78.4%)	173 (84.0%)	9 (90.0%)	5 (100%)
LGA	41 (3.5%)	20 (5.2%)	7 (3.6%)	8 (3.9%)	0 (0%)	0 (0%)
Race						
Black	692 (58.9%)	296 (77.1%)	180 (92.8%)	169 (82.0%)	2 (20.0%)	0 (0%)
Asian	93 (7.9%)	22 (5.7%)	2 (1.0%)	9 (4.4%)	0 (0%)	0 (0%)
Hispanic	88 (7.5%)	15 (3.9%)	6 (3.1%)	9 (4.4%)	1 (10.0%)	1 (20.0%)
Other	81 (6.9%)	22 (5.7%)	1 (0.5%)	13 (6.3%)	0 (0%)	0 (0%)
White	220 (18.7%)	29 (7.6%)	5 (2.6%)	6 (2.9%)	7 (70.0%)	4 (80.0%)
Sex						
Female	569 (48.5%)	208 (54.2%)	104 (53.6%)	18 (8.7%)	0 (0%)	0 (0%)
Male	605 (51.5%)	176 (45.8%)	90 (46.4%)	188 (91.3%)	10 (100%)	5 (100%)
Delivery mode						
Cesarean Section	795 (67.7%)	280 (72.9%)	144 (74.2%)	138 (67.0%)	8 (80.0%)	3 (60.0%)
Vaginal Delivery	379 (32.3%)	104 (27.1%)	50 (25.8%)	68 (33.0%)	2 (20.0%)	2 (40.0%)
Antenatal betamethasone						
Present	970 (82.6%)	325 (84.6%)	152 (78.4%)	178 (86.4%)	8 (80.0%)	4 (80.0%)
Absent	204 (17.4%)	59 (15.4%)	42 (21.6%)	28 (13.6%)	2 (20.0%)	1 (20.0%)

	Thalassemia (n=1174)	Hemoglobin Bart's (n=384)	Sickle cell anemia (n=194)	G6PD deficiency (n=206)	Hemophilia A (n=10)	Hemophilia B (n=5)
Intubation within 72 hours						
Absent	659 (56.1%)	274 (71.4%)	114 (58.8%)	152 (73.8%)	6 (60.0%)	2 (40.0%)
Present	515 (43.9%)	110 (28.6%)	80 (41.2%)	54 (26.2%)	4 (40.0%)	3 (60.0%)
Discharge Year						
2000 to 2004	290 (24.7%)	4 (1.0%)	23 (11.9%)	1 (0.5%)	2 (20.0%)	0 (0%)
2005 to 2009	337 (28.7%)	117 (30.5%)	49 (25.3%)	51 (24.8%)	1 (10.0%)	4 (80.0%)
2010 to 2014	229 (19.5%)	107 (27.9%)	58 (29.9%)	93 (45.1%)	2 (20.0%)	0 (0%)
2015 to 2020	318 (27.1%)	156 (40.6%)	64 (33.0%)	61 (29.6%)	5 (50.0%)	1 (20.0%)
Anomaly						
Present	86 (7.3%)	28 (7.3%)	10 (5.2%)	12 (5.8%)	1 (10.0%)	0 (0%)
Absent	1088 (92.7%)	356 (92.7%)	184 (94.8%)	194 (94.2%)	9 (90.0%)	5 (100%)
Morbidity						
Absent	868 (73.9%)	318 (82.8%)	155 (79.9%)	172 (83.5%)	6 (60.0%)	2 (40.0%)
Present	306 (26.1%)	66 (17.2%)	39 (20.1%)	34 (16.5%)	4 (40.0%)	3 (60.0%)
Disposition						
Alive	1148 (97.8%)	380 (99.0%)	189 (97.4%)	204 (99.0%)	8 (80.0%)	5 (100%)
Died	26 (2.2%)	4 (1.0%)	5 (2.6%)	2 (1.0%)	2 (20.0%)	0 (0%)

Abbreviations: SGA (small for gestational age), AGA (appropriate for gestational age), LGA (large for gestational age)

Supplemental Table 4. Baseline characteristics of preterm infants diagnosed with miscellaneous genetic disorders.

	Cystic fibrosis (n=165)	Hypophosphatasia (n=86)	Fragile X syndrome (n=7)	Fanconi anemia (n=2)
Gestational age (weeks)				
Mean (SD)	29.0 (3.12)	27.6 (2.84)	30.1 (2.34)	29.5 (4.95)
Median [Min, Max]	29.0 [23.0, 33.0]	27.0 [23.0, 33.0]	31.0 [26.0, 32.0]	29.5 [26.0, 33.0]
Birth weight				
SGA	26 (15.8%)	32 (37.2%)	0 (0%)	2 (100%)
AGA	126 (76.4%)	52 (60.5%)	5 (71.4%)	0 (0%)
LGA	13 (7.9%)	2 (2.3%)	2 (28.6%)	0 (0%)
Race				
Asian	1 (0.6%)	4 (4.7%)	0 (0%)	2 (100%)
Black	17 (10.3%)	21 (24.4%)	1 (14.3%)	0 (0%)
Hispanic	14 (8.5%)	22 (25.6%)	0 (0%)	0 (0%)
Other	6 (3.6%)	6 (7.0%)	0 (0%)	0 (0%)
White	127 (77.0%)	33 (38.4%)	6 (85.7%)	0 (0%)
Sex				
Female	71 (43.0%)	40 (46.5%)	2 (28.6%)	2 (100%)
Male	94 (57.0%)	46 (53.5%)	5 (71.4%)	0 (0%)
Delivery mode				
Cesarean Section	115 (69.7%)	69 (80.2%)	5 (71.4%)	2 (100%)
Vaginal Delivery	50 (30.3%)	17 (19.8%)	2 (28.6%)	0 (0%)
Antenatal betamethasone				
Absent	33 (20.0%)	18 (20.9%)	2 (28.6%)	2 (100%)
Present	132 (80.0%)	68 (79.1%)	5 (71.4%)	0 (0%)
Intubation within 72 hours				
Absent	52 (31.5%)	21 (24.4%)	3 (42.9%)	1 (50.0%)
Present	113 (68.5%)	65 (75.6%)	4 (57.1%)	1 (50.0%)
Discharge Year				
2000 to 2004	68 (41.2%)	27 (31.4%)	2 (28.6%)	0 (0%)
2005 to 2009	56 (33.9%)	26 (30.2%)	0 (0%)	0 (0%)
2010 to 2014	24 (14.5%)	19 (22.1%)	1 (14.3%)	1 (50.0%)
2015 to 2020	17 (10.3%)	14 (16.3%)	4 (57.1%)	1 (50.0%)

	Cystic fibrosis (n=165)	Hypophosphatasia (n=86)	Fragile X syndrome (n=7)	Fanconi anemia (n=2)
Anomaly				
Absent	126 (76.4%)	78 (90.7%)	5 (71.4%)	2 (100%)
Present	39 (23.6%)	8 (9.3%)	2 (28.6%)	0 (0%)
Morbidity				
Absent	68 (41.2%)	29 (33.7%)	6 (85.7%)	1 (50.0%)
Present	97 (58.8%)	57 (66.3%)	1 (14.3%)	1 (50.0%)
Disposition				
Alive	149 (90.3%)	71 (82.6%)	7 (100%)	1 (50.0%)
Died	16 (9.7%)	15 (17.4%)	0 (0%)	1 (50.0%)

SGA: small for gestational age; AGA: appropriate for gestational age; LGA: large for gestational age.

Supplemental Table 5. Instances of severe morbidity and mortality in preterm infants with genetic disorders.

Genetic Disorder	Mortality or Any Severe Morbidity	Mortality	AKI	ICH	NEC	BPD	ROP	Sepsis	Shock
No genetic diagnosis	*68326*	*16861*	*9021*	*12289*	*10933*	*3953*	*3836*	*28874*	*30087*
Any genetic diagnosis	*1526*	*565*	*279*	*204*	*232*	*172*	*72*	*585*	*615*
Aneuploidy	883	480	171	93	101	105	16	258	341
Trisomy 13	103	93	18	0	5	7	1	9	20
Trisomy 18	138	130	13	2	5	10	0	11	31
Trisomy 21	350	143	88	41	51	48	4	127	155
Trisomy 22	1	1	1	0	0	0	0	0	0
Klinefelter syndrome	11	2	2	2	0	1	0	8	4
Turner syndrome	14	5	5	3	2	1	0	9	6
Unspecified Aneuploidy	266	106	44	45	38	38	11	94	125
Copy number variation disorder	35	14	7	1	8	5	0	16	13
13q deletion syndrome	2	1	0	1	0	0	0	0	0
Cri-du-chat syndrome	5	2	2	0	2	1	0	4	2
DiGeorge syndrome	12	4	1	0	1	2	0	6	4
Unspecified CNV	16	7	4	0	5	2	0	6	7
Monogenic disorder	608	71	101	110	123	62	56	311	261
Hematologic disorder	452	39	66	85	76	36	28	232	174
Thalassemia	306	26	47	61	55	20	25	156	125
Hemoglobin bart's	66	4	4	8	9	5	0	35	25
Sickle cell disease	39	5	9	6	2	5	2	21	12
G6PD deficiency	34	2	5	7	10	5	1	15	9
Hemophilia A	4	2	1	2	0	1	0	3	2
Hemophilia B	3	0	0	1	0	0	0	2	1
Miscellaneous	156	32	35	25	47	26	28	79	87
Cystic fibrosis	97	16	18	14	35	18	17	45	51
Fragile X syndrome	1	0	0	0	0	0	0	1	1
Hypophosphatasia	57	15	16	10	12	7	11	33	34
Fanconi Anemia	1	1	1	1	0	1	0	0	1
Multiple Disorders	6	1	2	1	3	2	1	5	2

Abbreviations: AKI (acute kidney injury), BPD (bronchopulmonary dysplasia), ICH (intracerebral hemorrhage), OR (odds ratio) NEC (necrotizing enterocolitis), ROP (retinopathy of prematurity)

eMethods

The following script was used to generate tables and figures. The script was created using RStudio (version 2022.07.0).

```
setwd("~/Library/Mobile Documents/com~apple~CloudDocs/Desktop/Preterm R")
library(readxl)
library(tidyverse)
library(dplyr)

prevalence <- read_csv("Prevalence_data.csv")
outcomes <- read_csv("Outcome_data.csv")

#updated list with hematological disorders
up_prev <- read_csv("with_hematology.csv")

#join original GxDx's with hematology disorder
updated.prev <- merge(prevalence, up_prev, by = "PatientRandomSeqID")

#LGA, AGA, SGA column from Miles
from.miles <- read_csv("Prevalence_Data_FOR_SELIN.csv")

#join full GD dataset with SGA/LGA/AGA column
full.prev <- merge(updated.prev, from.miles, by = "PatientRandomSeqID")

#rename column with comprehensive list of GDs
full.prev$GxDx.x <- full.prev$GxDx.y


######TABLE 1

order2 <- c("t13",
        "t18",
        "t21",
        "trisomy 22",
        "klinefelter syndrome",
        "turner syndrome",
        "unspecified chromosomal abnormality",
        "13q-syndrome",
        "cri-du-chat _syndrome",
        "dgs",
        "unspecified microdeletion",
        "cf",
```

```
        "fragile x syndrome",
        "hypophosphatasia",
        "fanconi anemia",
        "thalassemia",
        "hemoglobin barts",
        "sickle cell anemia",
        "g6pd deficiency",
        "hemophilia a",
        "hemophilia b",
        "von willebrand disease",
        "multiple",
        "absent")

ordered2 <- agg_tbl_w_prev2 %>%
  slice(match(order2, Genetic.Disorder))

#Split up the table into 3 to get the 3 totals
aneuploidy2 <- ordered2[1:7,] %>%
  bind_rows(summarise_all(., ~if(is.numeric(.)) sum(.) else "Aneuploidy"))

CNV2 <- ordered2[8:11,] %>%
  bind_rows(summarise_all(., ~if(is.numeric(.)) sum(.) else "Copy number variation disorder"))

monogenic2 <- ordered2[12:22, ] %>%
  bind_rows(summarise_all(., ~if(is.numeric(.)) sum(.) else "Monogenic disorder"))

hematological <- ordered2[16:22, ] %>%
  bind_rows(summarise_all(., ~if(is.numeric(.)) sum(.) else "Hematologic disorder"))

miscellaneous <- ordered2[12:15, ] %>%
  bind_rows(summarise_all(., ~if(is.numeric(.)) sum(.) else "Miscellaneous"))


other <- ordered2[23:24, ]

ordered2$Genetic.Disorder <-unique(full.prev$GxDx.y)

#Now merge the tables
table.full <- add_row(aneuploidy2, CNV2)
table.full <- add_row(table.full, monogenic2)
table.full <- add_row(table.full, hematological)
table.full <- add_row(table.full, miscellaneous)
table.full <- table.full %>% rbind(other)
```

```
# table.full$Genetic.Disorder <-unique(table.full$Genetic.Disorder) ##this is achieved by
slice/match
# full.table.1 <- add_row(table2, ordered[14,]) %>%
#   select(1:14)

#Rearrange so that Any Aneuploidy/CNV/Monogenic is at the top of each section

order.full <- c("Aneuploidy",
            "t13",
            "t18",
            "t21",
            "trisomy 22",
            "klinefelter syndrome",
            "turner syndrome",
            "unspecified chromosomal abnormality",
            "Copy number variation disorder",
            "13q-syndrome",
            "cri-du-chat _syndrome",
            "dgs",
            "unspecified microdeletion",
            "Monogenic disorder",
            "Hematologic disorder",
            "thalassemia",
            "hemoglobin barts",
            "sickle cell anemia",
            "g6pd deficiency",
            "hemophilia a",
            "hemophilia b",
            "von willebrand disease",
            "Miscellaneous",
            "cf",
            "fragile x syndrome",
            "hypophosphatasia",
            "fanconi anemia",
            "multiple",
            "absent")

ordered.full <- table.full %>%
  slice(match(order.full, Genetic.Disorder))

#QUESTION: where do we want to put absent in this table/how is that number going to be
shared -- figure footnote?

#Fix the Genetic Disorder names upper/lowercase
```

```
ordered.full$Genetic.Disorder[2] <-"Trisomy 13"
ordered.full$Genetic.Disorder[3] <- "Trisomy 18"
ordered.full$Genetic.Disorder[4] <- "Trisomy 21"
ordered.full$Genetic.Disorder[5] <- "Trisomy 22"
ordered.full$Genetic.Disorder[6] <- "Klinefelter syndrome"
ordered.full$Genetic.Disorder[7] <- "Turner syndrome"
ordered.full$Genetic.Disorder[8] <- "Unspecified Aneuploidy"
ordered.full$Genetic.Disorder[10] <- "13q deletion syndrome"
ordered.full$Genetic.Disorder[11] <- "Cri-du-chat syndrome"
ordered.full$Genetic.Disorder[12] <- "DiGeorge syndrome"
ordered.full$Genetic.Disorder[13] <- "Unspecified CNV"
ordered.full$Genetic.Disorder[16] <- "Thalassemia"
ordered.full$Genetic.Disorder[17] <- "Hemoglobin bart's" #should i name this as Alpha
Thalassemia?
ordered.full$Genetic.Disorder[18] <- "Sickle cell disease"
ordered.full$Genetic.Disorder[19] <- "G6PD deficiency"
ordered.full$Genetic.Disorder[20] <- "Hemophilia A"
ordered.full$Genetic.Disorder[21] <- "Hemophilia B"
ordered.full$Genetic.Disorder[22] <- "Von Willebrand disease"
ordered.full$Genetic.Disorder[24] <- "Cystic fibrosis"
ordered.full$Genetic.Disorder[25] <- "Fragile X syndrome"
ordered.full$Genetic.Disorder[26] <- "Hypophosphatasia"
ordered.full$Genetic.Disorder[27] <- "Fanconi Anemia"
ordered.full$Genetic.Disorder[28] <- "Multiple Disorders"
ordered.full$Genetic.Disorder[29] <- "Absent"


ordered.full
colnames(ordered.full)[2] <- "Total Count"
colnames(ordered.full)[3] <- "Prevalence per 100,000 preterm infants"

library(stringr)
ordered.full$combined <- str_c(ordered.full$`Total Count`, " (", ordered.full$`Prevalence per
100,000 preterm infants` %>% round(1), ")")

#let's now add mortality rate as a column
#first we have to calculate it for each  of the 19 genetic disorders

#let's now calculate the n and mortality rate for babies with each GD after excluding babies that
were
#transferred in/out of the hospital
full.prev.w.deliv <- read_csv("fullprev2_with_Deliv_Mode.csv")

#rename column with comprehensive list of GDs
```

```
full.prev.w.deliv$GxDx.x <- full.prev.w.deliv$GxDx.y


#filter out transfers (n=40,263) and mode of delivery NAs (n=4,582) and transfers (n=51,774)
full.prev.w.deliv.no.na.or.transfers <- full.prev.w.deliv %>% filter(Dispo == "Alive" | Dispo ==
"Died")%>%
  filter(`Inborn Outborn` == "Inborn") %>%
  filter(Delivery_Mode == "Cesarean Section" | Delivery_Mode == "Vaginal Delivery") %>%
  filter(BW_CAT == "SGA" | BW_CAT == "LGA" | BW_CAT == "AGA")

full.prev.w.deliv.no.na.or.transfers %>% nrow() #320,582


#now calculate mortality rate and n
full.prev.w.deliv.no.na.or.transfers$GxDx.y <- full.prev.w.deliv.no.na.or.transfers$GxDx.y %>%
  tolower()

mortality.rate  <- full.prev.w.deliv.no.na.or.transfers %>%
  count(Dispo, GxDx.y) %>%
  pivot_wider(names_from = Dispo, values_from = n) %>%
  mutate(Alive = case_when(is.na(Alive) ~ 0,
                           TRUE ~ as.numeric(Alive)),
         Died = case_when(is.na(Died) ~ 0,
                          TRUE ~ as.numeric(Died))) %>%
  mutate(mortality = (Died/(Alive + Died)*100))


mortality.rate$n <- mortality.rate$Alive + mortality.rate$Died


colnames(mortality.rate)[1] = "Genetic.Disorder"

order2 <- c("t13",
            "t18",
            "t21",
            "trisomy 22",
            "klinefelter syndrome",
            "turner syndrome",
            "unspecified chromosomal abnormality",
            "13q-syndrome",
            "cri-du-chat _syndrome",
            "dgs",
            "unspecified microdeletion",
            "cf",
```

```
        "fragile x syndrome",
        "hypophosphatasia",
        "fanconi anemia",
        "thalassemia",
        "hemoglobin barts",
        "sickle cell anemia",
        "g6pd deficiency",
        "hemophilia a",
        "hemophilia b",
        "von willebrand disease",
        "multiple",
        "absent")

# mr <- mortality.rate %>% select(`Genetic.Disorder`, `Mortality Rate % (n)`)

mr.order <- mortality.rate %>%
  slice(match(order2, Genetic.Disorder))



#Split up the table into 3 to get the 3 totals
aneuploidy3 <- mr.order[1:7,] %>%
  bind_rows(summarise_all(., ~if(is.numeric(.)) sum(.) else "Aneuploidy"))

CNV3 <- mr.order[8:11,] %>%
  bind_rows(summarise_all(., ~if(is.numeric(.)) sum(.) else "Copy number variation disorder"))

monogenic3 <- mr.order[13:21, ] %>%
  bind_rows(summarise_all(., ~if(is.numeric(.)) sum(.) else "Monogenic disorder"))

hematological2 <- mr.order[16:21, ] %>%
  bind_rows(summarise_all(., ~if(is.numeric(.)) sum(.) else "Hematologic disorder"))

miscellaneous2 <- mr.order[12:15, ] %>%
  bind_rows(summarise_all(., ~if(is.numeric(.)) sum(.) else "Miscellaneous"))


multiple <- mr.order[22, ]

#calculate mortality rate for aneuploidy/cnv/hematological/monogenic/misc

updated.mr <- add_row(aneuploidy3, CNV3)
updated.mr <- add_row(updated.mr, monogenic3)
updated.mr <- add_row(updated.mr, hematological2)
```

```
updated.mr <- add_row(updated.mr, miscellaneous2)
updated.mr <- updated.mr %>% rbind(multiple)



corrected.mr <- updated.mr %>%
  mutate(mr = (Died/n)*100) %>%
  print(n=26)

library(dplyr)

corrected.mr$new <- corrected.mr$mr %>% round(digits=1)


corrected.mr$`n (Mortality Rate %)` <- str_c(corrected.mr$Died, "/",corrected.mr$n, " (",
corrected.mr$new %>% round(1), ")")


order.full <- c("Aneuploidy",
            "t13",
            "t18",
            "t21",
            "trisomy 22",
            "klinefelter syndrome",
            "turner syndrome",
            "unspecified chromosomal abnormality",
            "Copy number variation disorder",
            "13q-syndrome",
            "cri-du-chat _syndrome",
            "dgs",
            "unspecified microdeletion",
            "Monogenic disorder",
            "Hematologic disorder",
            "thalassemia",
            "hemoglobin barts",
            "sickle cell anemia",
            "g6pd deficiency",
            "hemophilia a",
            "hemophilia b",
            "von willebrand disease",
            "Miscellaneous",
            "cf",
            "fragile x syndrome",
            "hypophosphatasia",
```

```
            "fanconi anemia",
            "multiple",
            "absent")

mr.full <- corrected.mr %>%
  slice(match(order.full, Genetic.Disorder))

r#Fix the Genetic Disorder names upper/lowercase
mr.full$Genetic.Disorder[2] <-"Trisomy 13"
mr.full$Genetic.Disorder[3] <- "Trisomy 18"
mr.full$Genetic.Disorder[4] <- "Trisomy 21"
mr.full$Genetic.Disorder[5] <- "Trisomy 22"
mr.full$Genetic.Disorder[6] <- "Klinefelter syndrome"
mr.full$Genetic.Disorder[7] <- "Turner syndrome"
mr.full$Genetic.Disorder[8] <- "Unspecified Aneuploidy"
mr.full$Genetic.Disorder[10] <- "13q deletion syndrome"
mr.full$Genetic.Disorder[11] <- "Cri-du-chat syndrome"
mr.full$Genetic.Disorder[12] <- "DiGeorge syndrome"
mr.full$Genetic.Disorder[13] <- "Unspecified CNV"
mr.full$Genetic.Disorder[16] <- "Thalassemia"
mr.full$Genetic.Disorder[17] <- "Hemoglobin bart's" #should i name this as Alpha Thalassemia?
mr.full$Genetic.Disorder[18] <- "Sickle cell disease"
mr.full$Genetic.Disorder[19] <- "G6PD deficiency"
mr.full$Genetic.Disorder[20] <- "Hemophilia A"
mr.full$Genetic.Disorder[21] <- "Hemophilia B"


mr.full$Genetic.Disorder[23] <- "Cystic fibrosis"
mr.full$Genetic.Disorder[24] <- "Fragile X syndrome"
mr.full$Genetic.Disorder[25] <- "Hypophosphatasia"
mr.full$Genetic.Disorder[26] <- "Fanconi Anemia"

mr.full$Genetic.Disorder[27] <- "Multiple Disorders"
# mr.full$Genetic.Disorder[29] <- "Absent"


colnames(mr.full)[1] <-  "Genetic Disorder"
colnames(ordered.full)[1] <- "Genetic Disorder"

prev.values <- ordered.full %>% select(c(`Genetic Disorder`,
                                         combined))

prev.and.mortality <- full_join(prev.values, mr.full, by = "Genetic Disorder") %>%
  select(`Genetic Disorder`,
```

```
        combined,
        `n (Mortality Rate %)`)

prev.and.mortality <- prev.and.mortality[-29,]

#put a dash in the place of NA for von willebrand's mortality rate because the 1 baby with that
diagnosis was transferred in/out of the hoppital and therefore excluded
prev.and.mortality[is.na(prev.and.mortality)] <- "  —"

library(stringr)
colnames(prev.and.mortality)[2] <- "n (Prevalence per 100,000 Preterm Infants)"
#str_split_i(colnames(prev.and.mortality)[2], "\n", c(1, 2))

names_spaced <- c(
  'Genetic Disorder', 'n<br/> (Prevalence per 100,000 Preterm Infants)',              ## add <br/>
  'n* <br/> (Mortality Rate %)')


t1.temp <- dust(prev.and.mortality) %>%
  sprinkle(col = 1:3,
           round = 0) %>%
  #bold = T, part = "head")
  sprinkle(col = 2,
           halign = "center") %>%
  kable(align=c(rep('lcc')),
        col.names = names_spaced,
        escape = F) %>%
  kable_styling(font_size = 20) %>%
  kable_classic(full_width = F, html_font = "Cambria") #%>%
#kable("html")


t1.tmp2 <- add_indent(t1.temp, c(2:8, 10:13, 15, 23),
                      level_of_indent = 1,
                      all_cols = FALSE)

indented.more.t2 <- add_indent(t1.tmp2, c(16:22, 24:27),
                               level_of_indent = 2,
                               all_cols = FALSE)

t1.real <- indented.more.t2 %>%
  row_spec(0,bold=TRUE) %>%
  column_spec(1, width = "20em")
```

```
t1 <- t1.real %>% kable_styling(bootstrap_options = "striped", full_width = T, position = "left")
t1



##################   Tables 2, 3, 4 & 5 - crude and adjusted odds ratios
##################

#Load in data that includes delivery mode (CS or vaginal) so we can use it as a covariate for
adjusted odds ratio calculations
full.prev.w.deliv <- read_csv("fullprev2_with_Deliv_Mode.csv")

#filter out transfers (n=40,263) and mode of delivery NAs (n=4,582) and transfers (n=51,774)
full.prev.w.deliv.no.na.or.transfers <- full.prev.w.deliv %>% filter(Dispo == "Alive" | Dispo ==
"Died") %>%
  filter(`Inborn Outborn` == "Inborn") %>%
  filter(Delivery_Mode == "Cesarean Section" | Delivery_Mode == "Vaginal Delivery") %>%
  filter(BW_CAT == "SGA" | BW_CAT == "LGA" | BW_CAT == "AGA")

full.prev.w.deliv.no.na.or.transfers %>% nrow() #320,765

# no.transfers.deliv.mode <- full.prev.w.deliv.no.na %>%
#   filter(Delivery_Mode == "Cesarean Section" | Delivery_Mode == "Vaginal Delivery")

##add column specifically for present/absence of congenital anomaly instead of having the
specific ones (CHD, CAKUT, CNS, GI, Hydrops, Pulm, Multiple)
no.transfers.deliv.mode <- full.prev.w.deliv.no.na.or.transfers %>% mutate(Anomaly =
case_when(Any_Anom == "Absent" ~ "Absent",
                                                                TRUE ~ "Present"))

no.transfers.deliv.mode %>% nrow() #320,765

no.transfers.deliv.mode$Sex <- factor(no.transfers.deliv.mode$Sex)
no.transfers.deliv.mode$Race <- factor(no.transfers.deliv.mode$Race)
no.transfers.deliv.mode$Delivery_Mode <- factor(no.transfers.deliv.mode$Delivery_Mode)
no.transfers.deliv.mode$Anomaly <- factor(no.transfers.deliv.mode$Anomaly)

no.transfers.deliv.mode$`Antenatal Steroids` <-
as.character(no.transfers.deliv.mode$`Antenatal Steroids`)
no.transfers.deliv.mode["Antenatal Steroids"][no.transfers.deliv.mode["Antenatal Steroids"] ==
0] <- "Absent"
no.transfers.deliv.mode["Antenatal Steroids"][no.transfers.deliv.mode["Antenatal Steroids"] ==
1] <- "Present"
```

```
no.transfers.deliv.mode$`MVD0-2` <- as.character(no.transfers.deliv.mode$`MVD0-2`)
no.transfers.deliv.mode["MVD0-2"][no.transfers.deliv.mode["MVD0-2"] == "Y"] <- "Yes"
no.transfers.deliv.mode["MVD0-2"][no.transfers.deliv.mode["MVD0-2"] == "N"] <- "No"


no.transfers.deliv.mode$GA <- as.numeric(no.transfers.deliv.mode$GA)
no.transfers.deliv.mode$BW_CAT <- factor(no.transfers.deliv.mode$BW_CAT)
no.transfers.deliv.mode$BW_Z_Score <- as.numeric(no.transfers.deliv.mode$BW_Z_Score)
no.transfers.deliv.mode$CritIllorDeath <- factor(no.transfers.deliv.mode$CritIllorDeath)
#full.prev.copy.21$T21 <- factor(full.prev.copy.21$T21)


no.transfers.deliv.mode$Sex <- relevel(no.transfers.deliv.mode$Sex, ref = "Female")
no.transfers.deliv.mode$Race <- relevel(no.transfers.deliv.mode$Race, ref = "White")
no.transfers.deliv.mode$Delivery_Mode <- relevel(no.transfers.deliv.mode$Delivery_Mode, ref
= "Vaginal Delivery") #use whats lower risk or what's numerous

#full.prev.copy.21$T21 <- relevel(full.prev.copy.21$T21, ref = "absent")

library(expss)
no.transfers.deliv.mode <- apply_labels(no.transfers.deliv.mode, DischargeYear = "Discharge
Year",
                                        Delivery_Mode = "Mode of Delivery",
                                        `Antenatal Steroids` = "Antenatal Steroids",
                                        GA = "Gestational Age (weeks)",
                                        BW_CAT =
                                          "Birthweight",
                                        `MVD0-2` = "MV within 72 hours",
                                        BW_Z_Score = "Birth Weight (Z-score)",
                                        Any_Anom = "Major Congenital Anomaly")



###Calculate the crude and adjusted odds ratios for babies with GDs developing a severe illness
or dying

#Which genetic disorders were diagnosed in more than 10 babies that were non-transfers?
no.transfers.deliv.mode %>% group_by(GxDx.y) %>% count() %>% arrange(desc(n)) %>% print(n
= 24)
#answer: thalassemia, t21, hb barts, g6pd def, sickle cell, cf, t18, t13, hypophosphatasia, turner
syndorme, klinefelter, dgs, cri-du-chat, hemophilia a
#(excluded unspecified aneuploidy, unspecified CNV, and babies with multiple diagnoses)
```

```
#First, filter data for just the genetic disorders and the outcomes we're interested in


#Let's make sure every outcome is of class factor. Dialysis was too rare (n = 78) in this dataset
to analyze as an outcome for 360k babies.
no.transfers.deliv.mode <- no.transfers.deliv.mode %>%
  mutate(across(c(21:27, 29, 31), ~as.factor(.x))) #.x does as factor every column in this interval

#Next, initialize empty df of answer (OR and CIs) and choose the 8 most prevalent GDs in the
dataset
#initializing empty df

no.transfers.deliv.mode <- no.transfers.deliv.mode %>% mutate(GxDx.y = recode(GxDx.y, "13q-
syndrome" = "thirteen-q syndrome")) #rename 13q-syndrome because model reads in
alphabetical order so 13q is read before absent which makes it operate incorrectly
genetic_disorder <- unique(no.transfers.deliv.mode$GxDx.y)[c(2, 3, 4, 6, 7, 8, 9, 10, 12, 13, 14,
16, 17, 18, 19, 20, 21, 22, 23)]
outcome <- colnames(no.transfers.deliv.mode)[c(21:27, 29, 31)]
adjusted.ORs <- data.frame(matrix(ncol = length(`genetic_disorder`), nrow = length(outcome)))


#Run the loop which takes each of the outcomes of severe illness/death adjusting for
#the year the baby was discharged, presence of antenatal steroids, sex, race, gestational age,
birthweight Z-score, delivery mode,
#mechanical ventilation within 2 hours of life, and presence of any of the genetic disorders
where n>10 in the dataset.
for(j in 1:length(outcome)){
  for(i in 1:length(genetic_disorder)){

    adjusted.OR.data.filtered <- no.transfers.deliv.mode %>% filter(GxDx.y %in% c("absent",
genetic_disorder[i])) #within

    LogReg <- glm(as.formula(paste(outcome[j], "~ DischargeYear +
                                   `Antenatal Steroids` +
                                   Delivery_Mode +
                                   Sex +
                                   Race +
                                   GA +
                                   BW_Z_Score +
                                   Anomaly +
                                   `MVD0-2` +
                                   GxDx.y")), ##as.formula makes it a formula object - have to do this
in order to pass a string into glm. j is the jth outcome we're trying to predict. every model is
calculating OR of having outcome j given that you have GD i
```

```
            data = adjusted.OR.data.filtered,
            family = "binomial") #take out sample_n when run on whole set

   interval <- exp(confint(LogReg, level = 1-
0.05/((length(outcome)*length(`genetic_disorder`))*2)))[names(LogReg$coefficients) ==
paste0("GxDx.y", genetic_disorder[i])] %>% round(2) #CI for only GD/outcome OR -- plausible
values of OR are within this interval
   estimate <- exp(coef(LogReg, level = 1-
0.05/((length(outcome)*length(`genetic_disorder`))*2)))[names(LogReg$coefficients) ==
paste0("GxDx.y", genetic_disorder[i])] %>% round(2) #point estimate = single best guess of OR.
exp function turning log odds (calculated in LogReg) into odds.

   adjusted.ORs[i,j] <- paste0(estimate, " ", "(", interval[1], ",", interval[2], ")" )
   print(genetic_disorder[i])
  }
  print(outcome[j])

}


#new width of CI is: (1-.05/176) so that we take into account the number of intervals we're
building.

# adjusted.ORs.clean <- adjusted.ORs %>% na.omit()
copy.adjusted.ORs <- adjusted.ORs
colnames(copy.adjusted.ORs) <- paste0(outcome)
rownames(copy.adjusted.ORs) <- paste0(genetic_disorder)


adjusted.ORs
colnames(adjusted.ORs) <- paste0(outcome)
rownames(adjusted.ORs) <- paste0(genetic_disorder)
fixed.axes.aOR <- t(adjusted.ORs)


#transpose so GDs are the columns and clinical outcome is the row
fixed.axes.aOR <- t(copy.adjusted.ORs)

aOR.table <- fixed.axes.aOR %>% as.data.frame()

#convert to df
table.it <- fixed.axes.aOR %>% na.omit() %>% as.data.frame()

#export to excel to format & later combine with crude OR values
```

```
library(openxlsx)
#write.xlsx(aOR.table, file = "~/Library/Mobile
Documents/com~apple~CloudDocs/Desktop/Preterm R/v4.adjustedORs.xlsx", colNames = T,
borders = "surrounding", rowNames = T)



##Calculate the crude ORs
no.transfers.deliv.mode <- no.transfers.deliv.mode %>%
  mutate(across(c(21:27, 29, 31), ~as.factor(.x))) #.x does as factor every column in this interval

#Next, initialize empty df of answer (OR and CIs) and choose the 8 most prevalent GDs in the
dataset
#initializing empty df

no.transfers.deliv.mode <- no.transfers.deliv.mode %>% mutate(GxDx.y = recode(GxDx.y, "13q-
syndrome" = "thirteen-q syndrome")) #rename 13q-syndrome because model reads in
alphabetical order so 13q is read before absent which makes it operate incorrectly
genetic_disorder <- unique(no.transfers.deliv.mode$GxDx.y)[c(2, 3, 4, 6, 7, 8, 9, 10, 12, 13, 14,
16, 17, 18, 19, 20, 21, 22)] #not including t22 because n=1
outcome <- colnames(no.transfers.deliv.mode)[c(21:27, 29, 31)]
#adjusted.ORs <- data.frame(matrix(ncol = length(`genetic_disorder`), nrow =
length(outcome)))

crude.ORs <- data.frame(matrix(ncol = length(`genetic_disorder`), nrow = length(outcome)))
#initializing empty df


for(j in 1:length(outcome)){
  for(i in 1:length(genetic_disorder)){

    crude.OR.data.filtered <- no.transfers.deliv.mode %>% filter(GxDx.y %in% c("absent",
`genetic_disorder`[i])) # filtering for only absent and each of the 8 GDs we're interested in
within the GD column


    LogReg <- glm(as.formula(paste(outcome[j], "~ GxDx.y")), ##as.formula makes it a formula
object - have to do this in order to pass a string into glm. j is the jth outcome we're trying to
predict. every model is calculating OR of having outcome j given that you have GD i
                data = crude.OR.data.filtered,
                family = "binomial") #take out sample_n when run on whole set

    interval <- exp(confint(LogReg, level= 1-
0.05/((length(outcome)*length(`genetic_disorder`))*2)))[names(LogReg$coefficients) ==
```

```
paste0("GxDx.y", genetic_disorder[i])] %>% round(2) #CI for only GD/outcome OR -- plausible
values of OR are within this interval
   estimate <- exp(coef(LogReg, level= 1-
0.05/((length(outcome)*length(`genetic_disorder`))*2)))[names(LogReg$coefficients) ==
paste0("GxDx.y", genetic_disorder[i])] %>% round(2) #point estimate = single best guess of OR.
exp function turning log odds (calculated in LogReg) into odds.

   crude.ORs[i,j] <- paste0(estimate, " ", "(", interval[1], ", ", interval[2], ")" )
   print(`genetic_disorder`[i])
  }
  print(outcome[j])

}

crude.ORs
colnames(crude.ORs) <- paste0(outcome)
rownames(crude.ORs) <- paste0(genetic_disorder)

copy.crude.ORs <- crude.ORs
colnames(copy.crude.ORs) <- paste0(outcome)
rownames(copy.crude.ORs) <- paste0(genetic_disorder)


fixed.axes.cOR <- t(crude.ORs)

#convert to df
crude.table <- fixed.axes.cOR %>% as.data.frame()
write.xlsx(crude.table, file = "~/Library/Mobile
Documents/com~apple~CloudDocs/Desktop/Preterm R/v6.crudeORs.xlsx", colNames = T,
borders = "surrounding", rowNames = T)

#################    Supplementary tables 1, 2, 3, 4   ################
##want to show the demographic breakdown of:
#GA (mean and median), Race, % SGA (fenton), Sex, BW in grams (mean and median),
#presence of antenatal steroids, mechanical ventilation within 72 hours of life, discharge year,
presence of anomalies, delivery mode CS or vaginal)
##for the preterm infants excluding those that were transferred in/out of the hospital and that
have NAs for delivery mode and SGA/LGA/AGA

library(dplyr)
full.prev.w.deliv <- read_csv("fullprev2_with_Deliv_Mode.csv")

#rename column with comprehensive list of GDs
full.prev.w.deliv$GxDx.x <- full.prev.w.deliv$GxDx.y
```

```
#filter out transfers (n=40,263),  mode of delivery NAs (n=4,582),  transfers (n=51,774), and
SGA/LGA/AGA classification NAs (n=305)
full.prev.w.deliv.no.na.or.transfers <- full.prev.w.deliv %>% filter(Dispo == "Alive" | Dispo ==
"Died")%>%
  filter(`Inborn Outborn` == "Inborn") %>%
  filter(Delivery_Mode == "Cesarean Section" | Delivery_Mode == "Vaginal Delivery") %>%
  filter(BW_CAT == "SGA" | BW_CAT == "LGA" | BW_CAT == "AGA")

full.prev.w.deliv.no.na.or.transfers %>% nrow() #320,582

full.prev.w.deliv.no.na.or.transfers$GxDx.y <- full.prev.w.deliv.no.na.or.transfers$GxDx.y %>%
  tolower()


outcomes <- full.prev.w.deliv.no.na.or.transfers


aneuploidy.GDs <- c("t13",
                    "t18",
                    "t21",
                    "trisomy 22",
                    "klinefelter syndrome",
                    "turner syndrome")

cnv.GDs <- c("13q-syndrome",
             "cri-du-chat _syndrome",
             "dgs")

hema.GDs <- c("thalassemia",
              "hemoglobin barts",
              "sickle cell anemia",
              "g6pd deficiency",
              "hemophilia a",
              "hemophilia b",
              "fanconi anemia")

misc.GDs <- c("cf",
              "fragile x syndrome",
              "hypophosphatasia")


library(tidyverse)
```

```
library(dplyr)
outcome.aneuploidy <- outcomes %>% filter(GxDx.y %in% aneuploidy.GDs)

dem.aneuploidies <- outcome.aneuploidy %>% mutate(GxDx.y = recode(GxDx.y, "klinefelter
syndrome" = "Klinefelter syndrome",
                                                  "t13" = "Trisomy 13",
                                                  "t18" = "Trisomy 18",
                                                  "t21" = "Trisomy 21",
                                                  "trisomy 22" = "Trisomy 22",
                                                  "turner syndrome" = "Turner syndrome"))


dem.aneuploidies$DischargeYear <- factor(ifelse(dem.aneuploidies$DischargeYear < 2005,
"2000 to 2004",
                                        ifelse(dem.aneuploidies$DischargeYear < 2010, "2005 to 2009",
                                            ifelse(dem.aneuploidies$DischargeYear < 2015, "2010 to 2014",
"2015 to 2020"))))

#Change Any Anomaly to be Present and Absent
#Change Antenatal Steroids to be present/absent
#Change mechanical ventilation to be yes or no

dem.aneuploidies <- dem.aneuploidies %>%
  mutate(Anomaly = case_when(Any_Anom == "Absent" ~ "Absent",
                       TRUE ~ "Present")) %>%
  mutate(`Antenatal betamethasone` = case_when(`Antenatal Steroids` == "0" ~ "Absent",
                                       TRUE ~ "Present")) %>%
  mutate(`Intubation within 72 hours` = case_when(`MVD0-2` == "N" ~ "Absent",
                                         TRUE ~ "Present")) %>%
  mutate(Morbidity  = case_when(CritIllorDeath == "N" ~ "Absent",
                         TRUE ~ "Present"))


library(sjPlot)
colnames(dem.aneuploidies)[3] = "Delivery mode"
colnames(dem.aneuploidies)[4] = "Discharge Year"
colnames(dem.aneuploidies)[6] = "Disposition"
colnames(dem.aneuploidies)[7] = "Gestational age (weeks)"
colnames(dem.aneuploidies)[33] = "Birth weight"



library(table1)
```

```
dem.an.table <- table1(~`Gestational age (weeks)` +
                 `Birth weight` +
                 Race +
                 Sex +
                 `Delivery mode`+
                 `Antenatal betamethasone` +
                 `Intubation within 72 hours` +
                 `Discharge Year` +
                 Anomaly +
                 Morbidity +
                 Disposition | factor(GxDx.y),
                data = dem.aneuploidies,
                overall=F)


#demographics for infants with copy number variation disorders

outcome.cnv <- outcomes %>% filter(GxDx.y %in% cnv.GDs)

dem.cnv <- outcome.cnv %>% mutate(GxDx.y = recode(GxDx.y, 'cri-du-chat _syndrome' = 'Cri-
du-chat syndrome',
                                  'dgs' = 'DiGeorge syndrome',
                                  '13q-syndrome'= "13q-deletion syndrome"))


dem.cnv$DischargeYear <- factor(ifelse(dem.cnv$DischargeYear < 2005, "2000 to 2004",
                            ifelse(dem.cnv$DischargeYear < 2010, "2005 to 2009",
                                 ifelse(dem.cnv$DischargeYear < 2015, "2010 to 2014", "2015 to
2020"))))

#Change Any Anomaly to be Present and Absent
#Change Antenatal Steroids to be present/absent
#Change mechanical ventilation to be yes or no

dem.cnv <- dem.cnv %>%
  mutate(Anomaly = case_when(Any_Anom == "Absent" ~ "Absent",
                     TRUE ~ "Present")) %>%
  mutate(`Antenatal betamethasone` = case_when(`Antenatal Steroids` == "0" ~ "Absent",
                                  TRUE ~ "Present")) %>%
  mutate(`Intubation within 72 hours` = case_when(`MVD0-2` == "N" ~ "Absent",
                                   TRUE ~ "Present")) %>%
  mutate(Morbidity  = case_when(CritIllorDeath == "N" ~ "Absent",
                       TRUE ~ "Present"))
```

```
library(sjPlot)
colnames(dem.cnv)[3] = "Delivery mode"
colnames(dem.cnv)[4] = "Discharge Year"
colnames(dem.cnv)[6] = "Disposition"
colnames(dem.cnv)[7] = "Gestational age (weeks)"
colnames(dem.cnv)[33] = "Birth weight"



library(table1)

dem.cnv.table <- table1(~`Gestational age (weeks)` +
                  `Birth weight` +
                  Race +
                  Sex +
                  `Delivery mode` +
                  `Antenatal betamethasone` +
                  `Intubation within 72 hours` +
                  `Discharge Year` +
                  Anomaly +
                  Morbidity +
                  Disposition | factor(GxDx.y),
                 data = dem.cnv,
                 overall=F)



#dem table for babies with hematological disorders

outcome.hematological <- outcomes %>% filter(GxDx.y %in% hema.GDs)

dem.hematological <- outcome.hematological %>% mutate(GxDx.y = recode(GxDx.y,
"thalassemia" = "Thalassemia",
                                                 'g6pd deficiency' = "G6PD deficiency",
                                                 'hemoglobin barts' = "Hemoglobin bart's",
                                                 'hemophilia a' = "Hemophilia A",
                                                 'hemophilia b' = "Hemophilia B",
                                                 'fanconi anemia' = "Fanconi anemia",
                                                 "sickle cell anemia" = "Sickle cell anemia"))
```

```
dem.hematological$DischargeYear <- factor(ifelse(dem.hematological$DischargeYear < 2005,
"2000 to 2004",
                                    ifelse(dem.hematological$DischargeYear < 2010, "2005 to 2009",
                                        ifelse(dem.hematological$DischargeYear < 2015, "2010 to
2014", "2015 to 2020"))))

#Change Any Anomaly to be Present and Absent
#Change Antenatal Steroids to be present/absent
#Change mechanical ventilation to be yes or no

dem.hematological <- dem.hematological %>%
  mutate(Anomaly = case_when(Any_Anom == "Absent" ~ "Absent",
                        TRUE ~ "Present")) %>%
  mutate(`Antenatal betamethasone` = case_when(`Antenatal Steroids` == "0" ~ "Absent",
                                    TRUE ~ "Present")) %>%
  mutate(`Intubation within 72 hours` = case_when(`MVD0-2` == "N" ~ "Absent",
                                     TRUE ~ "Present")) %>%
  mutate(Morbidity  = case_when(CritIllorDeath == "N" ~ "Absent",
                          TRUE ~ "Present"))


library(sjPlot)
colnames(dem.hematological)[3] = "Delivery mode"
colnames(dem.hematological)[4] = "Discharge Year"
colnames(dem.hematological)[6] = "Disposition"
colnames(dem.hematological)[7] = "Gestational age (weeks)"
colnames(dem.hematological)[33] = "Birth weight"




library(table1)

dem.hema.table <- table1(~`Gestational age (weeks)` +
                    `Birth weight` +
                    Race +
                    Sex +
                    `Delivery mode` +
                    `Antenatal betamethasone` +
                    `Intubation within 72 hours` +
                    `Discharge Year` +
                    Anomaly +
                    Morbidity +
                    Disposition | factor(GxDx.y),
```

```
                    data = dem.hematological,
                    overall=F)



#dem table for babies with miscellaneous monogenic disorders
outcome.monogenic <- outcomes %>% filter(GxDx.y %in% misc.GDs)

dem.monogenic <- outcome.monogenic %>% mutate(GxDx.y = recode(GxDx.y, 'fragile x
syndrome' = "Fragile X syndrome",
                                              'cf' = "Cystic fibrosis",
                                              'hypophosphatasia' = "Hypophosphatasia"))


dem.monogenic$DischargeYear <- factor(ifelse(dem.monogenic$DischargeYear < 2005, "2000
to 2004",
                                     ifelse(dem.monogenic$DischargeYear < 2010, "2005 to 2009",
                                          ifelse(dem.monogenic$DischargeYear < 2015, "2010 to 2014",
"2015 to 2020"))))

#Change Any Anomaly to be Present and Absent
#Change Antenatal Steroids to be present/absent
#Change mechanical ventilation to be yes or no

dem.monogenic <- dem.monogenic %>%
  mutate(Anomaly = case_when(Any_Anom == "Absent" ~ "Absent",
                        TRUE ~ "Present")) %>%
  mutate(`Antenatal betamethasone` = case_when(`Antenatal Steroids` == "0" ~ "Absent",
                                     TRUE ~ "Present")) %>%
  mutate(`Intubation within 72 hours` = case_when(`MVD0-2` == "N" ~ "Absent",
                                       TRUE ~ "Present")) %>%
  mutate(Morbidity  = case_when(CritIllorDeath == "N" ~ "Absent",
                          TRUE ~ "Present"))


library(sjPlot)
colnames(dem.monogenic)[3] = "Delivery mode"
colnames(dem.monogenic)[4] = "Discharge Year"
colnames(dem.monogenic)[6] = "Disposition"
colnames(dem.monogenic)[7] = "Gestational age (weeks)"
colnames(dem.monogenic)[33] = "Birth weight"
```

```
library(table1)

dem.monogenic.table <- table1(~`Gestational age (weeks)` +
                        `Birth weight` +
                        Race +
                        Sex +
                        `Delivery mode` +
                        `Antenatal betamethasone` +
                        `Intubation within 72 hours` +
                        `Discharge Year` +
                        Anomaly +
                        Morbidity +
                        Disposition | factor(GxDx.y),
                       data = dem.monogenic,
                       overall=F)




#################    Supplementary table 5   ################

few.gds <- no.transfers.deliv.mode

#%>% filter(GxDx.y == "fanconi anemia" | GxDx.y == "trisomy 22" | GxDx.y == "hemophilia a" |
GxDx.y == "hemophilia b" | GxDx.y == "thirteen-q syndrome")

few.gds <- few.gds %>% mutate(across(c(21:31), ~as.character(.x)))


tmp <- few.gds %>% mutate(AKI = case_when(AKI == "Y" ~ "1",
                                TRUE ~ "0")) %>%
  mutate(IVHPVL = case_when(IVHPVL == "Y" ~ "1",
                      TRUE ~ "0")) %>%
  mutate(`NECSIP` = case_when(NECSIP == "Y" ~ "1",
                       TRUE ~ "0")) %>%
  mutate(S3BPD  = case_when(S3BPD == "Y" ~ "1",
                      TRUE ~ "0")) %>%
  mutate(ROP = case_when(ROP == "Y" ~ "1",
                     TRUE ~ "0")) %>%
  mutate(Sepsis = case_when(Sepsis == "Y" ~ "1",
                      TRUE ~ "0")) %>%
  mutate(Shock = case_when(Shock == "Y" ~ "1",
```

```
                    TRUE ~ "0")) %>%
  mutate(Dialysis = case_when(Dialysis == "Y" ~ "1",
                      TRUE ~ "0")) %>%
  mutate(Death = case_when(Death == "Died" ~ "1",
                    TRUE ~ "0")) %>%
  mutate(CritIllorDeath = case_when(CritIllorDeath == "Y" ~ "1",
                            TRUE ~ "0"))


summary.table <- tmp %>% mutate(across(c(21:31), ~as.numeric(.x))) %>%
  group_by(GxDx.y) %>%
  summarise(across(where(is.numeric), ~ sum(.x, na.rm = TRUE))) %>%
  select(GxDx.y,
       AKI,
       IVHPVL,
       NECSIP,
       S3BPD,
       ROP,
       Sepsis,
       Shock,
       Death,
       CritIllorDeath)


order.outcomes <- c("absent",
                "t13",
                "t18",
                "t21",
                "trisomy 22",
                "klinefelter syndrome",
                "turner syndrome",
                "unspecified chromosomal abnormality",
                "thirteen-q syndrome",
                "cri-du-chat _syndrome",
                "dgs",
                "unspecified microdeletion",
                "cf",
                "fragile x syndrome",
                "hypophosphatasia",
                "fanconi anemia",
                "thalassemia",
                "hemoglobin barts",
                "sickle cell anemia",
                "g6pd deficiency",
```

```
                "hemophilia a",
                "hemophilia b",
                "multiple")

colnames(summary.table)[1] <-  "Genetic Disorder"

tmp.3 <- summary.table %>%
  slice(match(order.outcomes, `Genetic Disorder`))



#Split up the table into 3 to get the 3 totals

absent <- tmp.3[1,]

any.GD <- tmp.3[2:22,] %>%
  bind_rows(summarise_all(., ~if(is.numeric(.)) sum(.) else "Any genetic diagnosis"))

aneuploidy.outcome <- tmp.3[2:8,] %>%
  bind_rows(summarise_all(., ~if(is.numeric(.)) sum(.) else "Aneuploidy"))

CNV.outcome <- tmp.3[9:12,] %>%
  bind_rows(summarise_all(., ~if(is.numeric(.)) sum(.) else "Copy number variation disorder"))

monogenic.outcome <- tmp.3[13:22, ] %>%
  bind_rows(summarise_all(., ~if(is.numeric(.)) sum(.) else "Monogenic disorder"))

hematological.outcome <- tmp.3[17:22, ] %>%
  bind_rows(summarise_all(., ~if(is.numeric(.)) sum(.) else "Hematologic disorder"))

miscellaneous.outcome <- tmp.3[13:16, ] %>%
  bind_rows(summarise_all(., ~if(is.numeric(.)) sum(.) else "Miscellaneous"))


multiple.outcome <- tmp.3[23, ]


#calculate mortality rate for aneuploidy/cnv/hematological/monogenic/misc

outcomes.n <- add_row(absent, any.GD)
outcomes.n <- add_row(outcomes.n, aneuploidy.outcome)
outcomes.n <- add_row(outcomes.n, CNV.outcome)
outcomes.n <- add_row(outcomes.n, monogenic.outcome)
outcomes.n <- add_row(outcomes.n, hematological.outcome)
```

```
outcomes.n <- add_row(outcomes.n, miscellaneous.outcome)
outcomes.n <- outcomes.n %>% rbind(multiple.outcome)


order.full <- c("absent",
            "Any genetic diagnosis",
            "Aneuploidy",
            "t13",
            "t18",
            "t21",
            "trisomy 22",
            "klinefelter syndrome",
            "turner syndrome",
            "unspecified chromosomal abnormality",
            "Copy number variation disorder",
            "thirteen-q syndrome",
            "cri-du-chat _syndrome",
            "dgs",
            "unspecified microdeletion",
            "Monogenic disorder",
            "Hematologic disorder",
            "thalassemia",
            "hemoglobin barts",
            "sickle cell anemia",
            "g6pd deficiency",
            "hemophilia a",
            "hemophilia b",
            "von willebrand disease",
            "Miscellaneous",
            "cf",
            "fragile x syndrome",
            "hypophosphatasia",
            "fanconi anemia",
            "multiple")

outcome.full <- outcomes.n %>%
  slice(match(order.full, `Genetic Disorder`))

#Fix the Genetic Disorder names upper/lowercase
outcome.full$`Genetic Disorder`[1] <-"No genetic diagnosis"
outcome.full$`Genetic Disorder`[4] <-"Trisomy 13"
outcome.full$`Genetic Disorder`[5] <- "Trisomy 18"
outcome.full$`Genetic Disorder`[6] <- "Trisomy 21"
outcome.full$`Genetic Disorder`[7] <- "Trisomy 22"
```

```
outcome.full$`Genetic Disorder`[8] <- "Klinefelter syndrome"
outcome.full$`Genetic Disorder`[9] <- "Turner syndrome"
outcome.full$`Genetic Disorder`[10] <- "Unspecified Aneuploidy"
outcome.full$`Genetic Disorder`[12] <- "13q deletion syndrome"
outcome.full$`Genetic Disorder`[13] <- "Cri-du-chat syndrome"
outcome.full$`Genetic Disorder`[14] <- "DiGeorge syndrome"
outcome.full$`Genetic Disorder`[15] <- "Unspecified CNV"
outcome.full$`Genetic Disorder`[18] <- "Thalassemia"
outcome.full$`Genetic Disorder`[19] <- "Hemoglobin bart's" #should i name this as Alpha
Thalassemia?
outcome.full$`Genetic Disorder`[20] <- "Sickle cell disease"
outcome.full$`Genetic Disorder`[21] <- "G6PD deficiency"
outcome.full$`Genetic Disorder`[22] <- "Hemophilia A"
outcome.full$`Genetic Disorder`[23] <- "Hemophilia B"


outcome.full$`Genetic Disorder`[25] <- "Cystic fibrosis"
outcome.full$`Genetic Disorder`[26] <- "Fragile X syndrome"
outcome.full$`Genetic Disorder`[27] <- "Hypophosphatasia"
outcome.full$`Genetic Disorder`[28] <- "Fanconi Anemia"

outcome.full$`Genetic Disorder`[29] <- "Multiple Disorders"
# outcome.full$`Genetic Disorder`[29] <- "Absent"


colnames(outcome.full)[3] <- "ICH"
colnames(outcome.full)[4] <- "NEC"
colnames(outcome.full)[5] <- "BPD"
colnames(outcome.full)[9] <- "Mortality"
colnames(outcome.full)[10] <- "Mortality or Any Severe Morbidity"


corrected.outcomes <- outcome.full[, c(1, 10, 9, 2, 3, 4, 5, 6, 7, 8)]


library(dplyr)
library(tidyverse)
library(pixiedust)
library(kableExtra)
library(knitr)

outcome.counts <- dust(corrected.outcomes, rownames = F) %>%
  sprinkle(col = 1:9,
         round = 0) %>%
```

```
  #bold = T, part = "head")
  sprinkle(col = 10,
         halign = "center") %>%
  kable(align=c(rep('lcc'))) %>%
  kable_styling(font_size = 20) %>%
  kable_classic(full_width = F, html_font = "Cambria")



tmp.1 <- add_indent(outcome.counts, c(1:2, 4:10, 12:15, 17, 24),
                 level_of_indent = 1,
                 all_cols = FALSE)

tmp.2<- add_indent(tmp.1, c(18:23, 25:28),
                level_of_indent = 2,
                all_cols = FALSE)

outcome.finals <- tmp.2 %>%
  row_spec(0,bold = T) %>%
  row_spec(1:2, italic = T) %>%
  column_spec(1, width = "20em")

outcome.n.table <- outcome.finals %>% kable_styling(bootstrap_options = "striped", full_width
= T, position = "left") #final table
outcome.n.table



###################### OVERALL STATS
################################################

###use prevalence data
#number of infants with a GD
full.prev %>% filter(GxDx.y != "absent") %>% nrow() #5838

#divide by total # of infants to get rate of genetic diagnoses in the population
#(5838/409704) * 100 = 1.42%


###use outcomes data (no transfers/no NAs for delivery mode and birth weight)
#out of the babies who died, calculate the number of babies who had a GD diagnosis and then
convert to a percentage
```

```
# 1-above value = out of the babies who died, calculate the % of babies who died that didn't
have a genetic diagnosis

#number of babies who died
full.prev.w.deliv.no.na.or.transfers %>% filter(Dispo == "Died") %>% nrow() #17427 -- #566
babies with GD died and 16861 babies with no GD died

#% of babies that died
full.prev.w.deliv.no.na.or.transfers %>% filter(Dispo == "Died") %>% count(GxDx.y) %>%
mutate((Freq = n/sum(n))*100)
#96.8

#96.8% of babies without a GD died. 100-96.8 = 3.2% of babies who died had a GD
#566 babies with GD died
#16861 babies with no GD died

#number of babies who had a genetic diagnosis AND a morbidity
full.prev.w.deliv.no.na.or.transfers %>% filter(AnyCritIllness == "Y") %>% count(GxDx.y) %>%
mutate((Freq = n/sum(n))*100)
#of the babies who developed a severe morbidity, 98% had no genetic diagnosis and 2% had a
genetic diagnosis
#65968 total babies developed severe morbidity, 1319 of them had a GD.

#number of infants who died who did NOT have a genetic diagnosis
#calculate mortality rate



#number of babies with a GD and and experienced severe morbidity
full.prev.w.deliv.no.na.or.transfers %>% filter(GxDx.y != "absent") %>% filter(AnyCritIllness ==
"Y") %>% nrow()
full.prev.w.deliv.no.na.or.transfers %>% filter(AnyCritIllness == "Y") %>% nrow()

#number of babies with a GD that had a severe morbidity and survived to discharge
full.prev.w.deliv.no.na.or.transfers  %>% filter(GxDx.y != "absent") %>% filter(Dispo == "Alive")
%>% nrow()
full.prev.w.deliv.no.na.or.transfers %>% filter(Dispo == "Alive") %>% nrow()

#number of babies with a GD who made it to discharge alive and who had no morbidity
full.prev.w.deliv.no.na.or.transfers  %>% filter(GxDx.y != "absent") %>% filter(Dispo == "Alive")
%>% filter(AnyCritIllness == "N") %>% nrow()
full.prev.w.deliv.no.na.or.transfers %>% filter(Dispo == "Alive") %>% filter(AnyCritIllness == "N")
%>% nrow()
```

```
#
full.prev.w.deliv.no.na.or.transfers %>% filter(GxDx.y != "absent")  %>% nrow()

##############################################################################
#############
```