Supplementary Online Content

The NICUSeq Study Group. Effect of whole-genome sequencing on the clinical management of acutely ill infants with suspected genetic disease: a randomized clinical trial. *JAMA Pediatr*. Published online September 27, 2021. doi:10.1001/jamapediatrics.2021.3496

This supplementary material has been provided by the authors to give readers additional information about their work.

eAppendix. NICUSeq Contributors

Clinical Sites

Children's Hospital of Philadelphia
Emma Bedoukian MS, Christopher Gray MS, Jacqueline Leonard MS, Jennifer Tarpinian MS, Cara Skraban MD, Kosuke Izumi MD PhD
University of Nebraska Medical Center
Zahi Zeidan MD
Children's Hospital of Orange County / Rady Children's Institute for Genomic Medicine
Jose Abdenur MD, Irfan Ahmad MD, Amir Ashrafi MD, Kushal Bhakta MD, Phuong Dao JD, Brent Dethlefs, Vijay Dhar MD, Cathy Flores RN, Michel Mikhael MD, Erum Naeem MS, Lisa Salz MS
Washington University / St. Louis Children's Hospital
Katie Shields BS, Dorothy K. Grange MD, Marcia Willing MD, Jorge Granadillo MD, Patti Dickson MD

Illumina Laboratory Services / Illumina Clinical Services Laboratory

Operations: Jon Albay, Joey Flores, Barrick Lo, Jacqueline Mapanao, Rita Pitts
Software & bioinformatics: Sunjay Ravishankar, Andrew Warren, James Avecilla, Felipe Mullen, Basil Juan
Variant Interpretation & Reporting: Erin Thorpe, Carolyn Brown, Alicia Scocchia*, Amanda Buchanan, Hiba Dada*, Maria Laura Cremona, Aditi Chawla, Krista Bluske, Zina Schlachetzki, Alka Malhotra, Vlad Gainullin*, Alison Coffey, Julie Taylor, Anjana Chandrasekhar, Brendan Burns, Nicole Burns, Sarah Schmidt, Revathi Rajkumar, Amanda Clause
**ceased employment at Illumina Inc. over the course of this study*

eMethods 1. Inclusion & Exclusion Criteria

Patient Inclusion Criteria

1. Current admission in a Neonatal Intensive Care Unit/Intensive Care Unit at a participating clinical site at the time of enrollment from day of life 0 to 120 days.
2. A suspected genetic etiology of disease, based on objective clinical findings or other phenotypic defects for which a genetic test would be considered.
3. Must be able to have 1 -1.25 ml tube of whole blood drawn for testing.
4. One parent of the proband must be able to provide written informed consent.
5. At least one biological parent must agree to participate and provide at least 4 ml of whole blood for testing.

Patient Exclusion Criteria

1. Known non-genetic cause(s) of disease, disorder, or phenotypic defect

2. The phenotype is fully explained by complications of prematurity.
3. Trisomy 13, 18 or 21 or Turner Syndrome is the likely diagnosis; such a proband will be eligible if a diagnostic karyotype is normal.
4. Blood transfusion within 48 hours (each proband will be re-eligible 48 hours after the most recent transfusion).
5. The PI decides that the study is not in the best interest of the proband (for example, the neonate or infant is at a high risk of severe morbidity or mortality within the next 7 days and these risks could be mitigated by alternative testing). Subsequent eligibility for enrollment of each proband is at the discretion of the site PI.

Parental Inclusion

1. Biological parent of the enrolled proband.
2. Must be able to have 4-6ml tubes of whole blood collected within 3 working days (Monday-Friday) of the proband.
3. Must be able to provide informed consent.

Parental Exclusion

1. Blood transfusion within 48 hours (participants will be re-eligible 48 hours after the most recent transfusion).
2. The PI decides that the study is not in the best interest of the participant.

Other Affected Family Member Inclusion

1. Must have objective clinical findings or other phenotypic defects that are consistent with those observed in the proband and for which a genetic test would be considered.
2. Must be able to have whole blood collected within 3 working days of the proband (1-1.25 ml < 2 years; 2 ml age 2-18 years; 4-6 ml >18 years).
3. Must be able to provide written informed consent or parental consent and affected family member assent if applicable.

Other Affected Family Member Exclusion

1. Not a biological relative of the enrolled proband.
2. Affected family member has had a blood transfusion within 48 hours (participants will be re-eligible 48 hours after the most recent transfusion).
3. The PI decides that the study is not in the best interest of the participant.

eMethods 2. Clinical Whole Genome Sequencing

Clinical whole genome testing was performed by the Illumina Clinical Services Laboratory (ICSL, San Diego CA USA). Whole genome sequencing was performed on extracted DNA using sequencing-by-synthesis (SBS) next generation sequencing (NGS). The data were aligned and reported using build 37.1 of the Human Reference Genome. The genome was sequenced to an average of 38.9 fold coverage (IQR 37.3-39.77) and an average of 97.9% of the genome was callable (IQR 97.5-98.1)

Based on the quality filters and through the analysis of an extended, multi-generation family set (Platinum Genomes[1]), for SNVs, sensitivity is 98.9% and the analytical Positive Predictive Value (PPV), i.e. true positive/[true positive + false positive], is 99.9%. Small insertion and deletion events were detected and reported for this test. Insertions up to 31 bases and deletions up to 27 bases have a sensitivity and analytical PPV of approximately 80- 85%, determined using Platinum Genomes.

This ICSL WGS test has the capability to detect copy number events greater than 10 kb[2], however sensitivity was only assessed for events greater than 20 kb and was found to be approximately 85%. Boundaries of the CNVs reported cannot be assessed with complete accuracy, and the boundaries were estimated to lie within +/- 1 kb of the event, unless otherwise noted.

For SNVs and small insertion and deletion events, interpretation was limited to variant positions that overlap an exon and 15 base pairs of flanking sequence. For CNVs, interpretation was limited to events that either overlap an exon or have a boundary that lies 1 kb upstream or downstream of an exon. Mitochondrial SNVs detected at an allele fraction greater than or equal to 3% were interpreted for pathogenicity. The percentage of heteroplasmy, however, was not reported. Mitochondrial CNVs and small insertion and deletions were not reported.

Variants were filtered based on multiple factors including population allele frequency, variant consequence, evolutionary conservation, occurrence in a gene with a well- established gene-disease relationship, occurrence in a gene whose disease association overlapswith the patient's reported phenotype, and inheritance mode, as appropriate. Variants were prioritized using a pre-populated list of 2016 early onset conditions with well-established gene- disease relationships, which was reduced or expanded depending on the patient's phenotype. A list of genes associated with key elements of the phenotype generated by searching the Online Mendelian Inheritance In Man (OMIM) database. Variant nomenclature is based on standardized HGVS conventions as referenced in den Dunnen et al[3]. Copy number variants (CNVs) greater than 10 kb were reported using standardized ISCN nomenclature as referenced in McGowan-Jordan et al[4].

During the course of this trial, the ICSL WGS test was validated to identify the absence of the 'C' allele at GRCh37 Chr5:70247773 (NM_000344.3:c.840C>T) in the SMN1 gene on May 22 2018. A total of 60 NICUSeq patients were processed prior to the implementation of this test feature, and 292 afterward. The c.840C/T distinguishes wild-type SMN1 from the SMN2 pseudogene. The c.840C>T variant causes alternate splicing of exon 7 of SMN1 which results in a truncated, unstable protein[5]. An absence of the c.840C allele is consistent with an absence of exon 7 of

SMN1 and in turn absence of wild-type SMN1, and is reported as a positive result for SMA. Over 95% of individuals with SMA have pathogenic variants in SMN1 which result in a biallelic absence of exon 7[6]. This test has not been validated to detect other variants in the SMN1 gene, nor to quantify the number or phase of the SMN1 and SMN2 genes. In addition, the test cannot identify individuals who are carriers of SMA. Only an outcome of 'SMA positive' was reported. The sensitivity of the SMA pipeline was assessed to be 97.5%, with a specificity of 100%.

A secondary findings analysis was offered to each individual tested as part of the family-based analysis (e.g. proband and parents as part of a trio test) in accordance with the with American College of Medical Genetics and Genomics (ACMG) recommendations for the return of secondary findings[7] if approved by the site-specific IRB. For individuals who opted in and were cleared to receive secondary findings analysis by the local IRB, a targeted search was conducted for pathogenic or likely pathogenic variants in the following genes:

ACTA2, *ACTC1*, *APC*, *APOB*, *ATP7B*, *BMPR1A*, *BRCA1*, *BRCA2*, *CACNA1S*, *COL3A1*, *DSC2*, *DSG2*, *DSP*, *FBN1*, *GLA*, *KCNH2*, *KCNQ1*, *LDLR*, *LMNA*, *MEN1*, *MLH1*, *MSH2*, *MSH6*, *MUTYH*, *MYBPC3*, *MYH11*, *MYH7*, *MYL2*, *MYL3*, *NF2*, *OTC*, *PCSK9*, *PKP2*, *PMS2*, *PRKAG2*, *PTEN*, *RB1*, *RET*, *RYR1*, *RYR2*, *SCN5A*, *SDHAF2*, *SDHB*, *SDHC*, *SDHD*, *SMAD3*, *SMAD4*, *STK11*, *TGFBR1*, *TGFBR2*, *TMEM43*, *TNNI3*, *TNNT2*, *TP53*, *TPM1*, *TSC1*, *TSC2*, *VHL*, *WT1*

Variants in genes not included in the above list but with medical actionability that were identified during the standard course of analysis were defined as incidental findings. Reporting of incidental findings was restricted to variants classified as likely pathogenic or pathogenic that appear in a molecular state that corresponds to an expected clinical presentation (e.g. biallelic variants in a gene associated with an autosomal recessive disorder). Incidental findings could be related to pediatric or adult-onset conditions. Reporting of variants in genes related to adult-onset conditions was restricted to conditions in which professional practice guidelines outline condition-specific patient management, surveillance or screening, family management or special circumstances to avoid.

Some regions of the human genome were not covered by this test, including stretches of the human reference genome that have not been completely resolved, or regions where it is difficult to align fragments accurately. Additionally, genes that are associated with regions of high homology are difficult for this test to resolve. These include, but are not limited to, some immunoglobulin (HLA) genes and telomeres.

All reported variants were orthogonally confirmed at an independent laboratory. All but one variant was orthogonal characterized at GeneDx (Gaithersburg, MD). A single CYP21A2 variant was confirmed by Mayo Clinical Laboratories.

eMethods 3. Statistical Analysis

The statistical analysis plan was developed by Precision for Medicine (formerly Agility Clinical) and can be found as a supplemental document and is also available upon request to the corresponding author. Analysis of the primary outcome, difference in COM at 60 days (Visit 3) between the Early and Delayed arm patients, was performed by Jennifer R. Scott, PhD, MPH at Precision for Medicine. All other secondary and exploratory analyses were performed by John Belmont MD, PhD at Illumina Inc in consultation with the PIs of the NICUSeq study group. In addition to those analyses detailed in the manuscript text we note that an assessment of diagnostic accuracy was not performed as detailed in the statistical analysis plan due to elements of the study procedures that involved active discussion with the sites to resolve the interpretation of any potentially ambiguous WGS testing results; that change in care setting was assessed using a Pearson's Chi-squared test with Yates' continuity correction and no significant differences were observed due to high proportion of patients rapidly discharged; and that physician and patient satisfaction survey will be reported in a subsequent manuscript.

eReferences

1. Eberle MA, Fritzilas E, Krusche P, et al. A reference data set of 5.4 million phased human variants validated by genetic inheritance from sequencing a three-generation 17-member pedigree. Genome Res 2017;27(1):157–64.

2. Gross AM, Ajay SS, Rajan V, et al. Copy-number variants in clinical genome sequencing: deployment and interpretation for rare and undiagnosed disease. Genet Med 2019;21(5):1121–30.

3. den Dunnen JT, Dalgleish R, Maglott DR, et al. HGVS Recommendations for the Description of Sequence Variants: 2016 Update. Hum Mutat 2016;37(6):564–9.

4. McGowan-Jordan J, Simons A, Schmid M. ISCN 2016: An International System for Human Cytogenomic Nomenclature (2016). Karger; 2016.

5. Kashima T, Manley JL. A negative element in SMN2 exon 7 inhibits splicing in spinal muscular atrophy. Nat Genet 2003;34(4):460–3.

6. Prior TW. Perspectives and diagnostic considerations in spinal muscular atrophy. Genet Med 2010;12(3):145–52.

7. Kalia SS, Adelman K, Bale SJ, et al. Recommendations for reporting of secondary findings in clinical exome and genome sequencing, 2016 update (ACMG SF v2.0): a policy statement of the American College of Medical Genetics and Genomics. Genet Med 2017;19(2):249–55.

eTable 1. Change of Management Classification Rubric

Management Classification & Code*	Description
--	Routine supportive clinical care; no classification entry made at time of assessment
Condition-specific management (M)	An alteration to the clinical management based on a primary etiological diagnosis that leads to a change of management specific to the altered molecular pathway.
Condition-specific supportive care (S)	Alteration to the clinical management based on a primary etiological diagnosis that leads to supportive management specific to the condition or that would have not otherwise been implemented.
Palliative or end-of-life management (P)	A change of management associated with the decision to implement or accelerate palliative or end-of-life care due to severity of the genetic diagnosis.

*Single or combinations of codes could be attributed to each case at each study time point. For example, a patient coded with “M,S” would be noted for receiving both condition-specific management and condition-specific supportive care.

eTable 2. Demographic and Clinical Characteristics by Site

Site	1	2	3	4	5
Early arm (n)	21	41	42	31	43
Delayed arm (n)	20	41	42	31	42
Age(days)	20 (7, 37)	19 (6, 39)	16 (7, 35)	15 (8, 29)	14 (7, 22)
Female (%)	11	42	36	30	34
Race					
White	23	53	69	50	55
Black	17	1	4	8	17
Asian	0	11	4	0	3
Other	1	17	7	4	9
Ethnicity					
Latino or Hispanic	4	41	14	3	19
Not Latino or Hispanic	37	39	70	57	66
Unknown	0	2	0	2	0
Family composition for testing					
Trio	30	54	72	48	65
Duo	10	25	12	11	18
Other	1	3	0	2	1

(cont’d on next page)

Indication for testing					
Multiple Congenital Anomalies	34	44	41	43	39
Isolated Major Congenital Anomaly	3	19	20	5	16
Neurological Disorder	1	13	16	11	11
Single Major Clinical Feature	3	6	7	2	18

eTable 3. Enrollment Care Setting

Site	**1**	**2**	**3**	**4**	**5**	**Total**
NICU	37	64	76	52	67*	296
PICU	2	12	8	1	0	23
CVICU	2	6	0	9	18	35
All Patients	**41**	**82**	**84**	**62**	**85**	(354)

NICU = neonatal intensive care unit; PICU = pediatric intensive care unit; CVICU = cardiovascular intensive care unit. *Includes a high acuity 2 week old neonate recruited from the general pediatrics floor and enrolled as an accepted variance (patient 927).

eTable 4. Clinical Classification Definitions

Clinical Classification	Description
Multiple congenital anomalies (MCA)	Two or more unrelated structural malformations that occur during intrauterine development and can be identified prenatally, at birth, or later in infancy*
Isolated congenital major anomaly (ICMA)	A single major structural malformation that occurs during intrauterine development and can be identified prenatally, at birth, or later in infancy
Single major clinical feature (SF)	A single clinical finding, other than a structural malformation, including metabolic disorders or other laboratory and physiological abnormalities
Neurological disorder (ND)	A presentation consistent with functional or structural deficits in the central or peripheral nervous system, including epilepsy

*WHO MCA definition - https://www.who.int/news-room/fact-sheets/detail/congenital-anomalies

eTable 8. Change of Management Outcome by Site at Day 60 (Visit 3)

Site	Early COM	Early no COM	Delayed COM	Delayed no COM
1	3	15	3	17
2	12	26	4	34
3	4	33	3	39
4	6	22	1	28
5	9	31	6	32
All Patients	**34**	**127**	**17**	**150**

eTable 9. Change of Management Outcome by WGS Variant Pathogenicity Classification

	Pathogenic	Mixed[1]	Uncertain	None
Day 60 (Visit 3)				
COM	29	6	9	7
No COM	53	14	40	170
NA[2]	10	4	2	8
Day 90 (Visit 4)				
COM	21	5	5	2
No COM	61	13	43	170
NA[2]	10	6	3	13
Any COM				
COM	46	10	14	8
No COM	36	9	34	165
NA[2]	10	5	3	12

1. Indicates that the case had at least one pathogenic (P) or likely pathogenic (LP) variant
2. Expired before change in management could be evaluated

Note: Day 60 (Visit 3) is the difference in COM state between enrollment and 60 days; Day 90 (Visit 4) indicates the difference between 60 and 90 days. Any COM includes any individual who had COM at any point in the study, and if a patient had COM at 60 days and 90 days this is treated as a single COM. This table includes all patients who received a clinical report.

eTable 13. Diagnostic Efficacy in Preterm Infants

		Early Arm*		Delayed Arm*	
WHO classification	**Age**	**Patients**	**WGS Dx**	**Patients**	**WGS Dx**
Extremely preterm	<28 wk	7	4	6	1
Very preterm	29-32 wk	15	7	17	2
Moderately preterm	33 - 37 wk	73	21	53	16

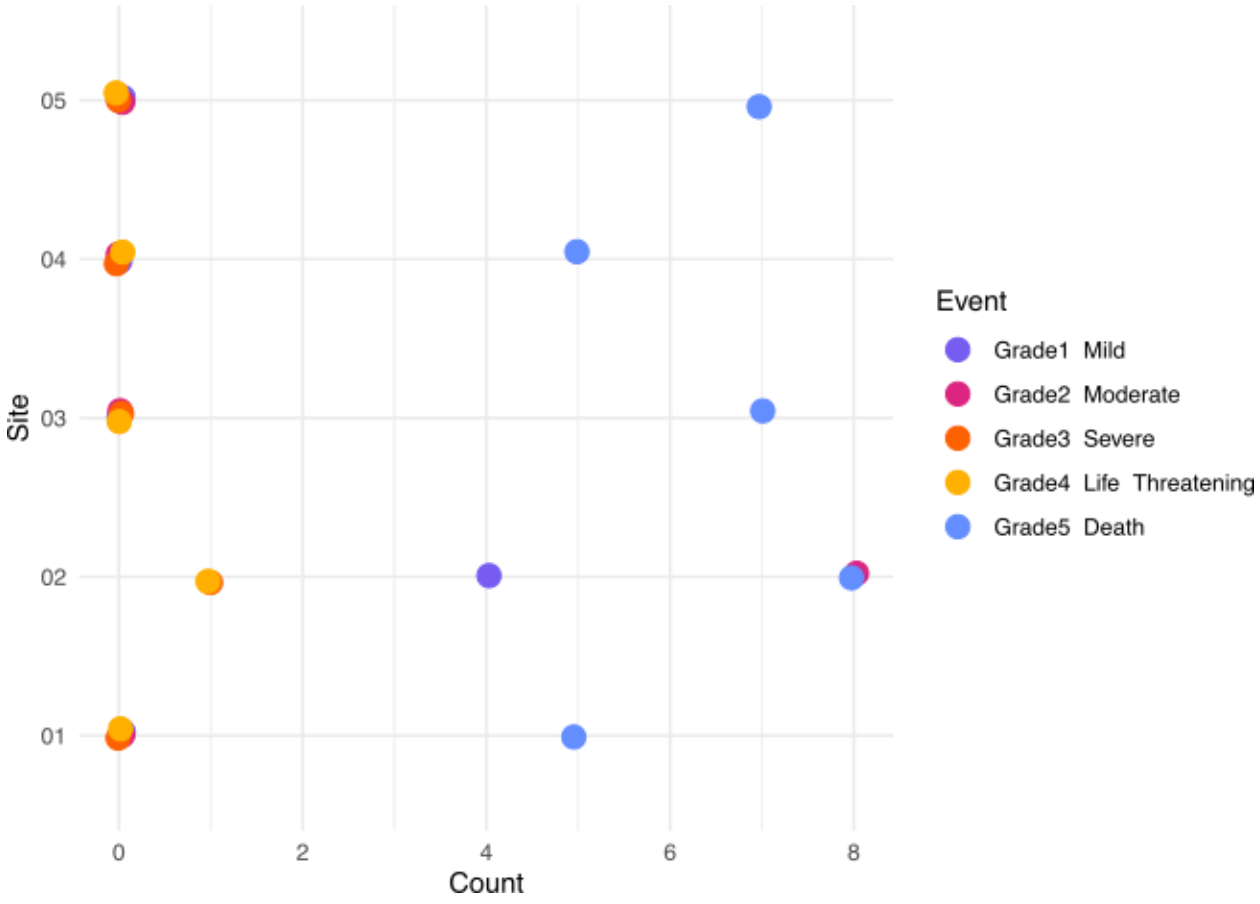
**Cumulative diagnostic efficacy across both arms was 5/13 (38%) in extremely preterm, 9/32 (28%) in very preterm and 37/126 (30%) in moderately preterm.*

eTable 13. Diagnostic Efficacy in Preterm Infants

	WGS - Negative	**WGS - Positive**
UC - Negative/ND	217	58
UC - Positive	10*	53

Please see eTable S5 for details on discordant cases (see column "Test concordance" for cases annotated with "No"). We note that eight of the ten discordant cases were related to laboratory or clinical interpretation of results.

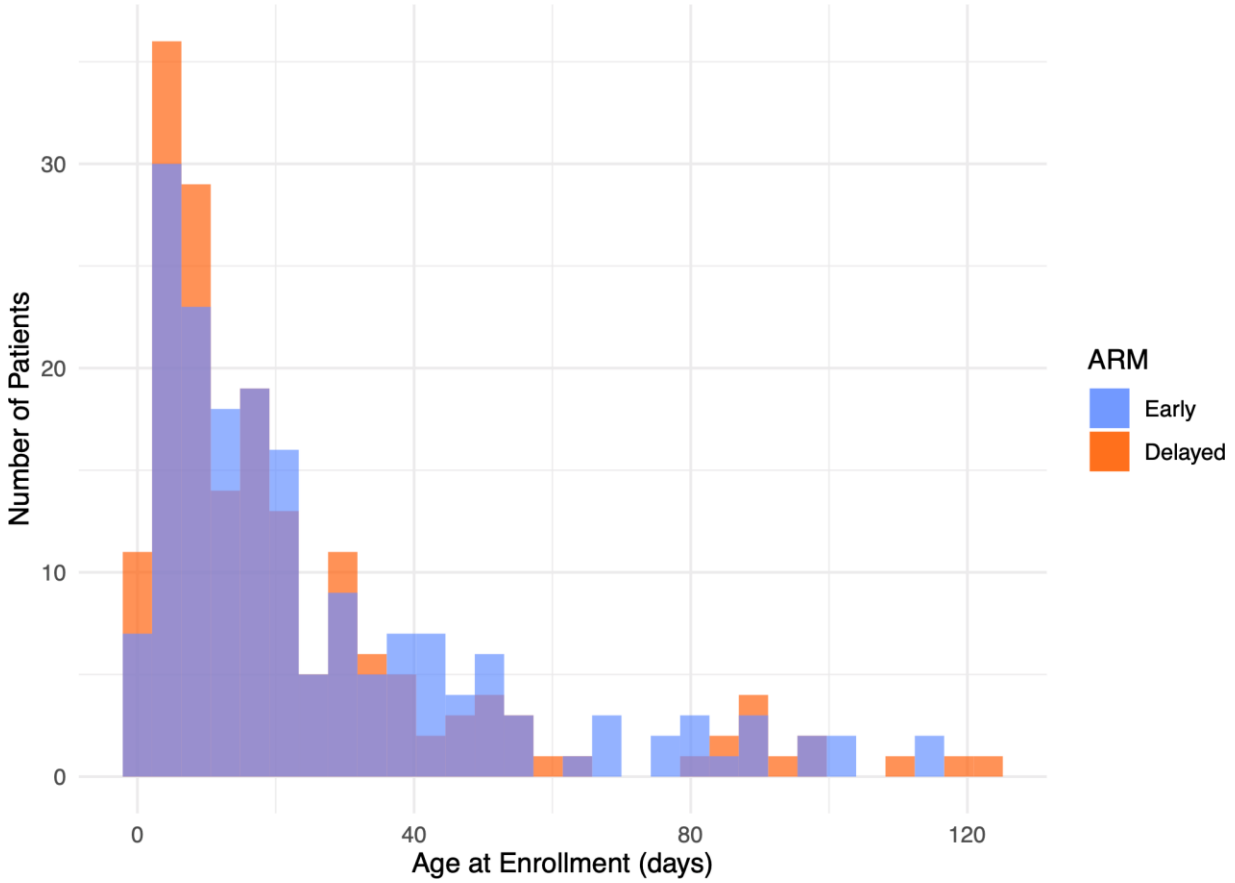
eFigure 1. Adverse Events by Site



eFigure 1. Adverse events by site

The number of events are shown across the x-axis and sites (using anonymized numbers) are shown along the y-axis. Events are colored to represent mild, moderate, severe, life-threatening or death events. Adverse events were rare and none were related to the clinical whole genome test study intervention.

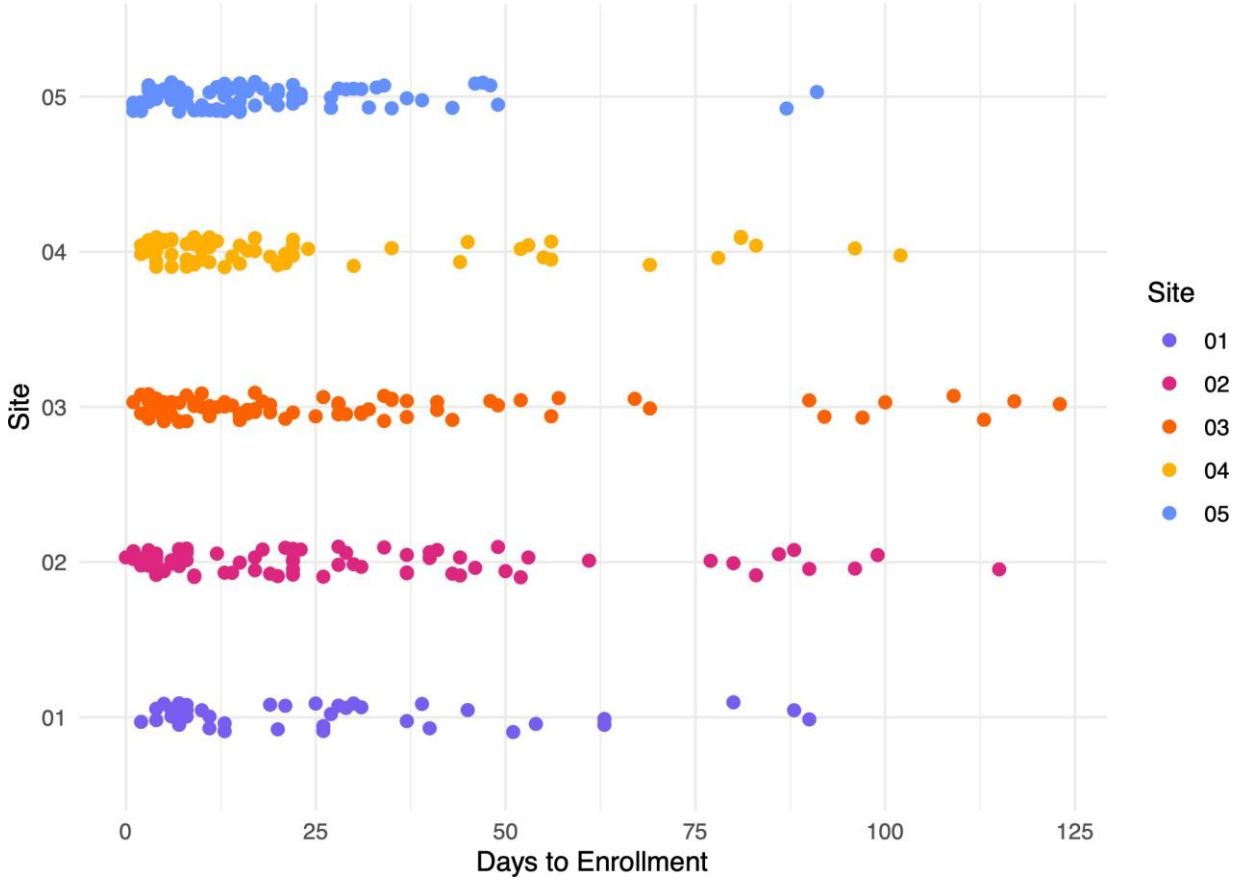
eFigure 2. Patient Age at Enrollment



eFigure 2. Patient age at enrollment

Patient age was skewed left, with the average age ~15-20 days across sites (See Table 1, Table S2). The age distributions in the Early and Delayed arms were not significantly different. A single patient (patient 647) was enrolled at 123 days of life, whose screening was initiated at 119 days of life, as an accepted variance.

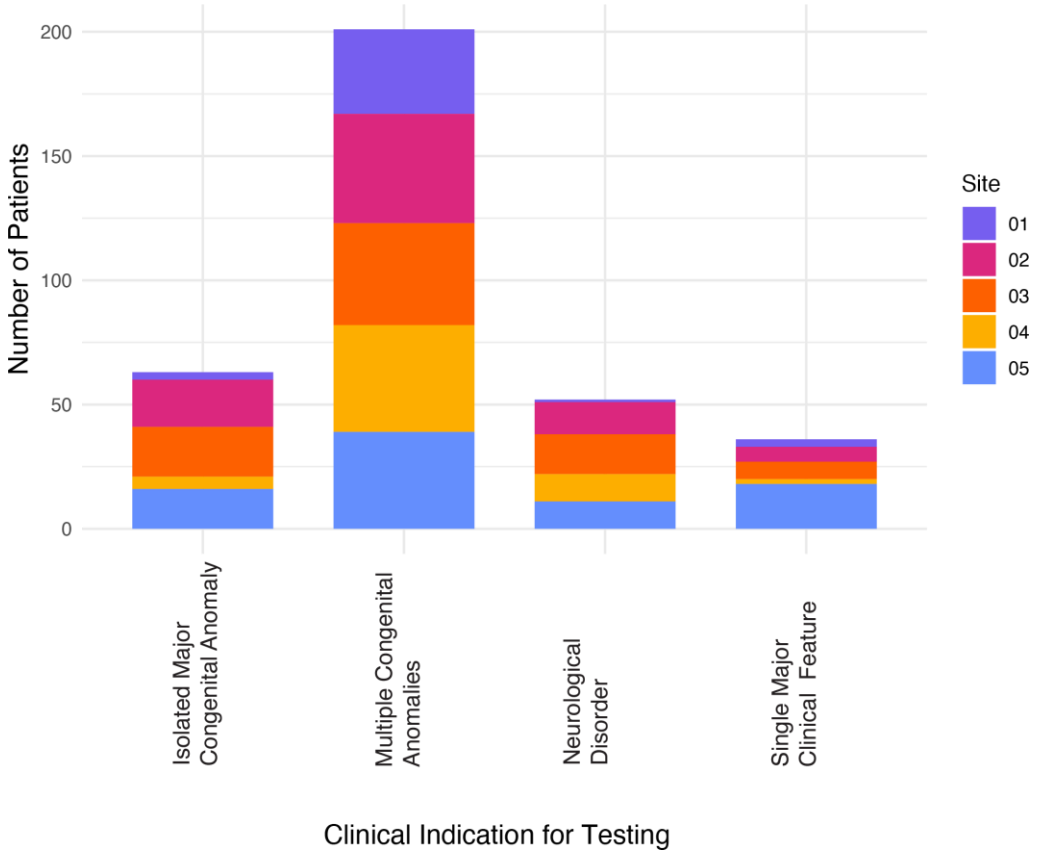
eFigure 3. Time From Admission to NICUSeq Enrollment Across Sites



eFigure 3. Time from admission to NICUSeq enrollment across sites

Patient enrollment showed a wide distribution of total time from admission to first approach and enrollment, with most enrolled within the first 30 days of admission. Similar distributions were observed across sites, shown as colored lanes and anonymously numbered.

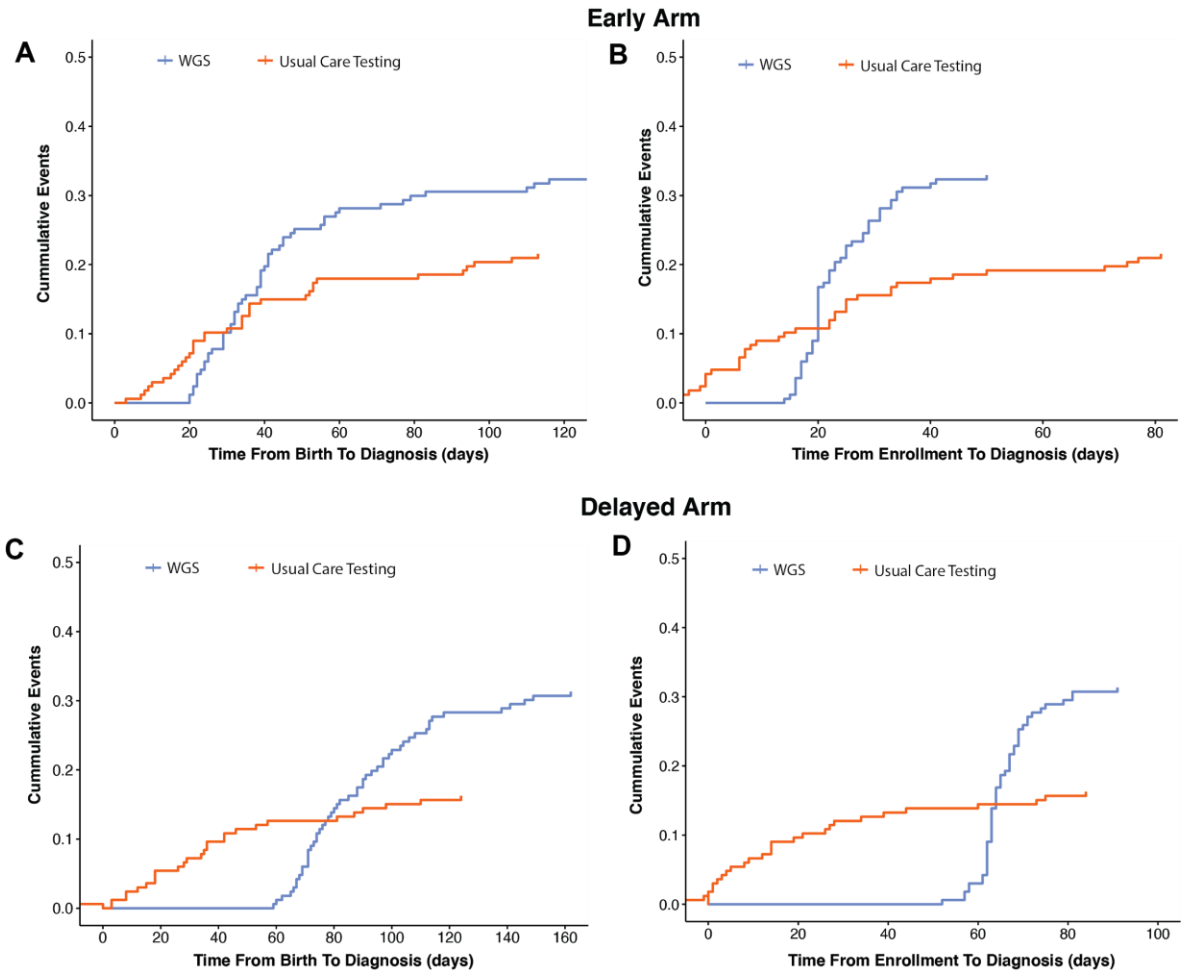
eFigure 4. Clinical Indications for Testing by Site



eFigure 4. Clinical indications for testing by site

All patients were classified to one of four primary indications for testing - isolated major congenital anomaly, multiple congenital anomalies (MCA), neurological disorder or single major clinical features. The majority of patients across the study and at each site presented with MCA.

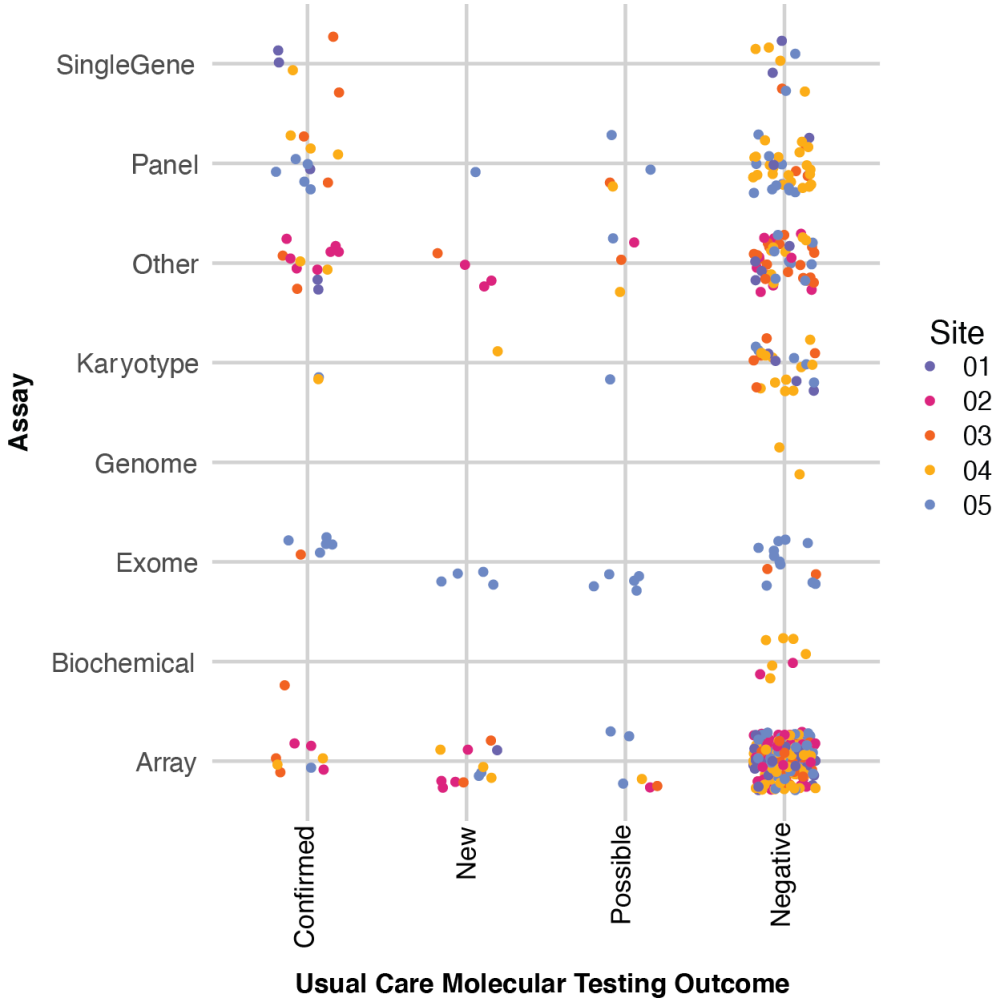
eFigure 5. Time From Enrollment or Birth to Diagnosis



eFigure 5. Time from enrollment or birth to diagnosis.

Cumulative diagnoses by each test type is shown for WGS (blue) and usual care testing (orange). (A,B) Time from birth or enrollment to diagnosis in the early arm. (C,D) Time from birth or enrollment to diagnosis in the delayed arm.

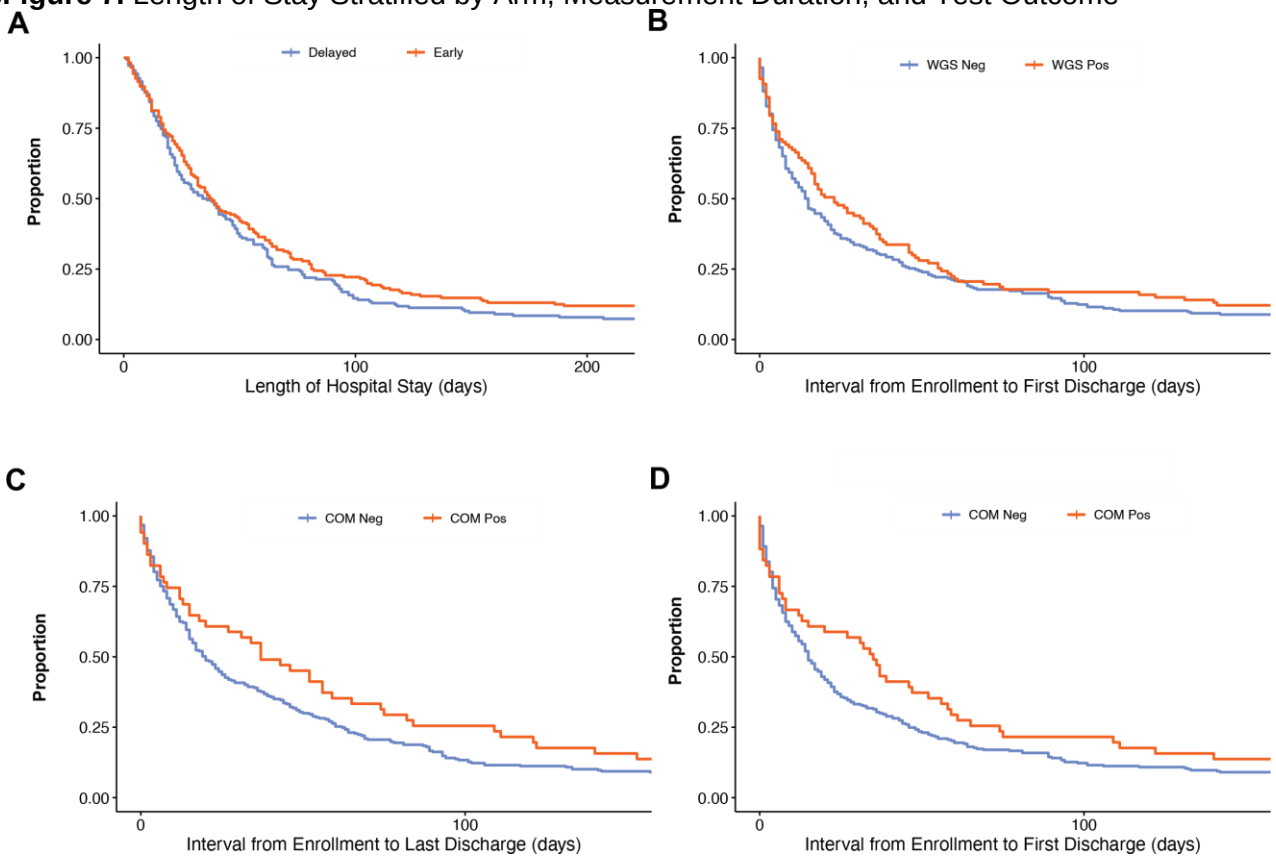
eFigure 6. Variation in Usual Care Testing by Assay, Outcome, and Site



eFigure 6. Variation in usual care testing by assay, outcome and site

The distribution of UC tests performed during the study observation window and their outcomes is shown. Assays are shown along the y-axis and test outcomes are shown along the x-axis. Anonymized sites are shown as colored dots. The majority of UC tests were negative microarrays.

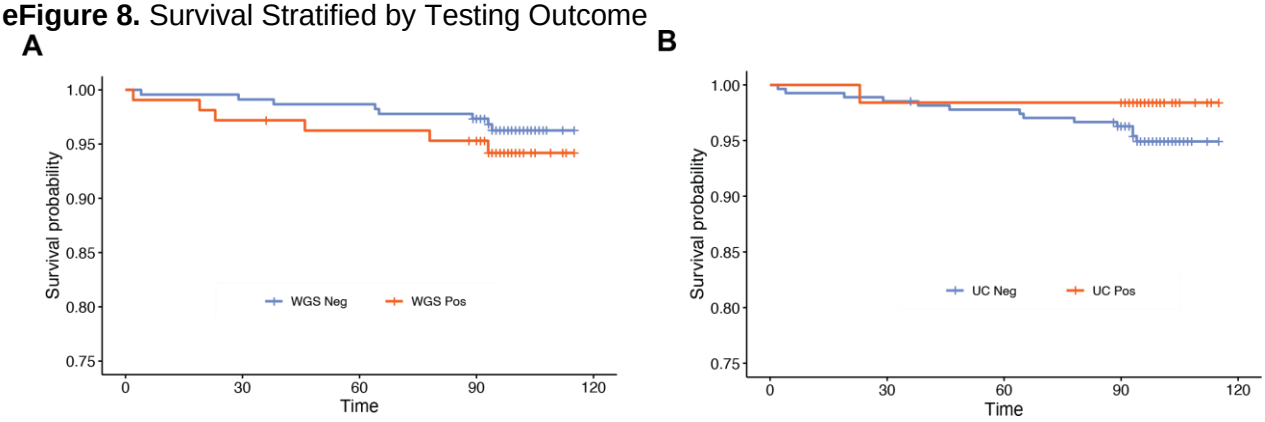
eFigure 7. Length of Stay Stratified by Arm, Measurement Duration, and Test Outcome



eFigure 7. Length of stay stratified by arm, measurement duration and test outcome

Length of stay was assessed by (A) examining the total length of hospitals stay stratified by arm, or the interval from NICUSeq study enrollment to first discharge or last discharge stratified by (B) WGS test outcome across both study arms or (C,D) change in management. Stratification by study arm can be found in Figure 3C in the main text. No significant differences were observed.

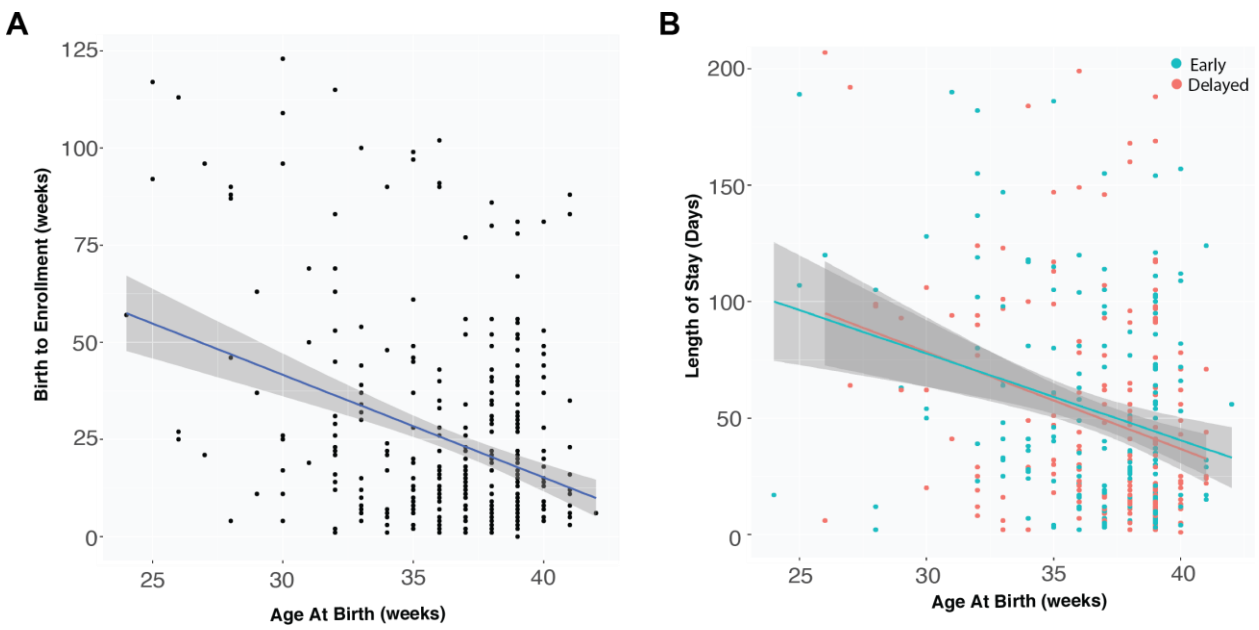
eFigure 8. Survival Stratified by Testing Outcome



eFigure 8. Survival stratified by testing outcome

Patient survival was assessed by examining the survival probability across the duration of the study observation period stratified by (A) WGS testing outcome or (B) UC testing outcome. Stratification by study arm can be found in Figure 3D in the main text. No significant differences were observed.

eFigure 9. Gestational Age and Time to Enrollment or Length of Stay



eFigure 9. Gestational age and time to enrollment or length of stay

Age at birth (in weeks) compared to either time to enrollment (A) or length of stay (B) is shown. A negative correlation is observed for both, with premature infants enrolled later and staying in hospital longer.