Data-driven prioritization of genetic disorders for global genomic newborn screening programs

Thomas Minten[1]*, Nina B. Gold[2*†], Sarah Bick[3,4,5], Sophia Adelson[6,7], Nils Gehlenborg[8], Laura M. Amendola, François Boemer[9], Alison J. Coffey[10], Nicolas Encina[11,12,13], Alessandra Ferlini[14], Janbernd Kirschner[15], Bianca E. Russell[16], Laurent Servais[17,18], Kristen L. Sund[19], Ryan J. Taft[10], Petros Tsipouras[20], Hana Zouk[21,22,23], ICoNS Gene List Contributors**, David Bick[24], Robert C. Green[5,12,23,25] for the International Consortium on Newborn Sequencing (ICoNS)***

* These authors contributed equally to the manuscript.
† Corresponding author
** See Supplement A for list of ICoNS Gene List Contributors authors
*** See Supplement B for a list of International Consortium on Newborn Sequencing (ICoNS) authors

Address correspondence to: Nina B. Gold, MD, Mass General Hospital for Children, Division of Medical Genetics and Metabolism, 175 Cambridge Street, Boston, MA 02114, [ngold@mgh.harvard.edu]

[1]KU Leuven; [2]Massachusetts General Hospital, Department of Pediatrics; Harvard Medical School, Department of Pediatrics; [3]Boston Children's Hospital; [4]Massachusetts General Hospital; [5]Harvard Medical School; [6]Brigham and Women's Hospital; [7]Stanford School of Medicine; [8]Harvard Medical School, Department of Biomedical Informatics; [9]University of Liege, CHU Liege; [10]Illumina Inc.; [11]ICoNS; [12]Ariadne Labs; [13]Harvard T.H. Chan School of Public Health; [14]University of Ferrara, Department of Medical Sciences, Department of Medical Sciences, Unit of Medical Genetics; [15]University Medical Center Freiburg, Department of Neuropediatrics and Muscle Disorders; [16]University of California, Los Angeles, David Geffen School of Medicine, Department of Human Genetics, Division of Clinical Genetics; [17]University of Oxford; [18]University of Liege; [19]Nurture Genomics; [20]FirstSteps-BNSI; [21]Massachusetts General Hospital, Department of Pathology, Laboratory for Molecular Medicine; [22]Harvard Medical School, Department of Pathology; [23]Broad Institute; [24]Genomics England; [25]Mass General Brigham

Abstract

Genomic sequencing is poised to expand newborn screening for treatable childhood-onset disorders. Over 30 international research studies and companies are exploring its use, collectively aiming to screen more than 500,000 infants. A key challenge is determining which genes to include in screening. Among 27 newborn sequencing programs, the number of genes analyzed ranged from 134 to 4,299, with only 74 genes included by over 80% of programs. To understand this variability, we assembled a dataset with 25 characteristics of 4,389 genes included in any program and used a multivariate regression analysis to identify characteristics associated with inclusion across programs. These characteristics included presence on the US Recommended Uniform Screening panel, evidence regarding the natural history of disease, and efficacy of treatment. We then used a machine learning model to generate a ranked list of genes, offering a data-driven approach to the future prioritization of disorders for public health newborn screening efforts.

A decade ago, the BabySeq Project piloted newborn and childhood sequencing (NBSeq), a process designed to detect risk for a wide range of genetic disorders in apparently healthy infants.[1–9] Fueled by the recognition that over 700 genetic disorders now have targeted treatments or consensus guidelines for long-term management, enthusiasm for NBSeq has grown significantly.[10,11] Stakeholders, including diverse groups of parents,[12,13] rare disease specialists,[11] primary care physicians,[14] genetic counselors,[11,15,16] and the public[17,18] now support the implementation of genomic newborn screening for at least some disorders.

This growing interest in NBSeq has led to the creation of at least 30 international research programs and companies that are actively exploring this screening approach.[19–21] Most of these programs are exchanging best practices under the International Consortium on Newborn Sequencing (ICoNS).[22] Given the hundreds of treatable disorders that could be candidates for NBSeq, and various interpretations of actionability, selection of the appropriate genes and disorders is a recurring challenge.[20,23,24] Historically, the criteria established by Wilson and Jungner[25] have provided a framework for selecting the disorders to include in public newborn screening programs. These criteria prioritize the inclusion of childhood-onset disorders that are treatable if diagnosed in their earliest stages and require emergent intervention to prevent irreversible damage. However, the technical aspects of genomic sequencing and variant curation, as well as the variable expression and incomplete penetrance of many genetic disorders, present new challenges to this screening paradigm. Many of the international NBSeq research studies and commercial programs around the world are using independent and opaque processes to select genes for screening.

Prior studies have identified discrepancies across the genes being analyzed by a limited number of commercial NBSeq programs[26] and research studies,[27,28] but little is known about the values and variables that underlie these differences. Understanding which genes have high concordance across programs may guide emerging NBSeq research programs as they

select which genes and variants to report to participants. Furthermore, the characteristics of these genes and their associated disorders can be used more empirically to prioritize candidate genes for public health programs.

To understand the variability among newborn sequencing programs, we compared the genes currently selected for analysis by 27 research studies and commercial NBSeq programs. For each gene that was included in any NBSeq program, we assembled a dataset of 25 associated characteristics, then used a multivariate regression analysis to identify which of these characteristics were associated with inclusion across programs. Finally, we used a machine learning model to generate a ranked list of genes, offering a data-driven approach to the prioritization of genetic disorders for population-wide NBSeq.

Results

International Consortium on Newborn Sequencing (ICoNS)

ICoNS, founded in 2022, brings together leaders of global research projects investigating the use of NBSeq. For this project, ICoNS acted as the convening organization requesting gene lists and criteria for gene selection from NBSeq programs.

Description of global NBSeq programs

We identified 35 independent research studies and commercial NBSeq programs (Fig. 1, Supplementary Table 1).[5,29–46] Of these, 10 were located in North America, 10 in Asia, nine in Europe, four in Australia and New Zealand, one in South America and one in Africa. At this time, 26 NBSeq programs are actively recruiting participants and nine are scheduled to begin recruitment. The 27 research programs anticipate a combined total sample size of 519,410 infants, with the intended enrollment in each study varying from 48 to over 100,000 infants.[42,47] Genome or exome sequencing backbones are in use by 13 programs and 17 are using gene panels. Several programs have not yet decided upon either their underlying sequencing platform nor their final gene list.

Positive screening rates of NBSeq programs

Nine of the 20 NBSeq research programs have published or presented the screening results from a collective total of 68,884 infants (Table 1). The percentage of positive screening results ranged from 1.85% in BabyDetect (3,847 infants screened for 405 genes) to 9.43% in BabySeq (159 infants screened for 4,299 genes), with an average of 3.80% positive results across 68,884 infants. There was a significant positive correlation between the percentage of positive screening results in a program and the number of genes they screened (pearson correlation coefficient of 0.653, p=0.041). A majority of the collective 1,937 positive screening results across seven studies for which detailed results were available were due to variants in *G6PD* (56.8%).

Four studies reported the clinical outcomes of infants who had undergone NBSeq, allowing for the calculation of these studies' positive predictive value (PPV). The PPV varied from 12% in 414 infants from the NESTS study to 88% in 142 infants from the GUARDIAN study,[48] with an average across studies of 43% (40% when weighted by sample size).

Description of gene lists across NBSeq programs

Only five programs have published or made available their criteria for selecting genes and disorders for screening (Supplementary Table S2). Of these programs, nearly all indicated the intent to include early-onset, severe, treatable monogenic disorders.

We aggregated the lists of genes being analyzed by 20 NBSeq research and 7 commercial programs. The number of genes included in each program ranged from 134 to 4,299 (median=306). A total of 4,389 genes were included across at least one of the 27 gene lists (Supplementary Table S4). Of these, 4,033 genes (91.8%) were associated with a phenotype in the OMIM database. Additionally, 1,000 genes (22.8%) had corresponding ClinGen gene-disease validity classifications, with 854 gene-disease relationships classified

as having definitive validity and 13 genes from 15 programs with refuted or disputed associations with the disorder being screened. Collectively, genes linked to inherited metabolic disorders (IMDs) (25.7%), neurologic (15.4%), immunologic (12.0%) and endocrine (11.1%) disorders constituted the majority of the gene lists (Fig. 2A, 2B).

Discordance among gene lists used in NBSeq programs

A pairwise Jaccard Index, which compares the content of each of the 27 gene lists with another, indicates that similarity was strongest between gene lists from commercial NBSeq programs (Fig. 2C, Supplementary Fig. 2, 3). Most pairs of gene lists from NBSeq research programs have highly discrepant content. Of the 4,389 genes included in at least one NBSeq program, the vast majority were included by only a small number of NBSeq programs: 4,088 genes (93%) were included by 10 or fewer programs and 3,797 (87%) genes were included by five or fewer programs (Supplementary Fig. 1). A total of 14 out of 27 gene lists included genes not shared by any other study.

The gene lists of the four research studies with the largest intended sample sizes (BabyDetect, the Generation study, the GUARDIAN study and NewbornsInSA) share only 157 genes (19%) out of an aggregated total of 818 genes (Fig. 2D). Across these four programs, 305 genes (37%) were unique to just one of the studies.

Genes with high concordance across NBSeq programs

Despite this variability across gene lists, we found 74 genes (1.7% of 4,389) that were included by over 80% (22 of 27) of NBSeq programs (Fig. 3). Of these 74 genes, 58 were associated with diseases on the US Recommended Uniform Screening Panel (RUSP). A total of 34 genes not linked to disorders on the RUSP appeared on 20 or more lists (Fig. 3C).

Predictors of gene inclusion across NBSeq programs

To understand why certain genes were included by more NBSeq programs than other genes, we assembled a dataset of 25 characteristics for each of the 4,389 genes included in any NBSeq program. We then used univariate regression analyses to identify the factors that influenced the likelihood of a gene being included in multiple programs (Fig. 4A, Supplementary Table 5).[2,10,11,31,49–51] Genes associated with core and secondary conditions on the RUSP were significantly more likely to be included in gene lists compared with disorders not on the RUSP (regression coefficient 74.6%, 95% confidence interval (CI): 0.709-0.783, p<0.01; regression coefficient 60.1%, 95% CI: 0.558-0.644, p<0.01). Additionally, genes that were recommended for inclusion in newborn screening by 80% or more rare disease experts in a recent survey,[11] were 43.5% (95% CI: 37.4%-49.6%, p<0.01) more likely to be included than genes that were recommended by fewer experts (Supplementary Fig. 5).

Among other gene and disease-related characteristics, the strongest predictor of inclusion across NBSeq programs was the evidence base, defined by a combination of gene-disease validity, published descriptions of the natural history of disease, and the availability of expert consensus or professional society practice guidelines for disease diagnosis and management.[49] Genes with the highest evidence base were 29.4% more likely (95% CI: 24.5%-34.3%, p<0.01) to be included in NBSeq programs than those with less available evidence.

Other disease-related characteristics associated with inclusion across NBSeq programs were high efficacy of disease treatment (16.9%, 95% CI: 12.2%-21.6%, p<0.01), high penetrance (15.4%, 95% CI: 9.7%-21.1%, p<0.01), neonatal- or infantile-onset (15.2%, 95% CI: 10.7%-19.7%, p<0.01), high disease severity (14.8%, 95% CI: 8.5%-21.1%, p<0.01), high acceptability of treatment (with regard to the burdens and risks placed on the individual) (14.8%, 95% CI: 10.1%-19.5%, p<0.01), and the existence of an non-molecular test that could be used to confirm the diagnosis (13.7%, 95% CI: 9.2%-18.2%, p<0.01). Several

quantitative scores previously designed to assess the overall usefulness of disorders for NBSeq were also strong predictors of gene inclusion across NBSeq programs, including the BabySeq Category[2] and ASQM Score[49,50] (pearson correlation coefficients of 0.07, p=0.047; 0.45, p=0.001 respectively, Supplementary Fig. 6).

Measuring evolving knowledge about genes and diseases

We also conducted a multivariate regression analysis to determine how changes in specific variables, such as treatability and evidence base, would individually influence the overall regression (Supplementary Table 6). Notably, the introduction of a new, highly acceptable treatment for a disorder with no previous treatment would increase the likelihood of inclusion in NBSeq programs by 9.7% (95% CI: 0.1%-19.3%, p<0.05). Similarly, improving knowledge related to the natural history of a gene-disorder pair from “none” to “perfect” would increase the likelihood of inclusion in NBSeq programs by 15.0% (95% CI: 5.0%-25.0%, p<0.01).

Predictors of gene inclusion in individual NBSeq programs

To explore whether gene and disease characteristics influence their inclusion differently within individual NBSeq programs, we also conducted separate regressions for each program (Fig. 4B and Supplementary Tables 7). The evidence base of a gene-disease pair was more strongly correlated with gene inclusion for programs such as FORESITE 360 (regression coefficient 53.9%, 95% CI: 44.1%-63.7%, p<0.01) and the commercial genetic test offered by PerkinElmer (47.9%, 95% CI: 38.3%-57.5%, p<0.01), compared with the average correlation across all programs (29.4%, 95% CI: 24.5%-34.3%, p<0.01). In contrast, other programs, such as BabyScreen+, placed more emphasis on the inclusion of early-onset conditions (35.4%, 95% CI: 25.6%-45.2%, p<0.01) compared with the average (15.2%, 95% CI: 10.7%-19.7%, p<0.01).

Machine learning prediction model

We developed a machine learning prediction model to empirically predict and rank the inclusion of novel genes for NBSeq programs. To build this tool, we randomly split the gene list data into 80% training and 20% test sets. Of the previously collected 25 gene and disease characteristics, we selected 13 as features in our model due to availability and minimal overlap (see Methods and Supplementary Table 3). These characteristics were: RUSP category, clinical area, evidence base, treatment efficacy, penetrance, treatment acceptability, age of onset, existence of orthogonal tests, recommendation score, inheritance, prevalence, and ClinGen disease validity and actionability scores.

During the training phase (n=91,291, 80% of all 114,114 potential instances of a 4,389 genes included on a gene list across 26 NBSeq programs), we compared three machine learning methods: linear regression, random forest and boosted trees (see Methods). Boosted trees demonstrated the highest accuracy, with an area under the curve (AUC) of 0.917 and R-squared of 77% on the test set (n=22,823, 20%) (Fig. 4C). The relative importance of all variables in the boosted trees model was highest for characteristics such as the proportion of experts who recommended inclusion of the gene in NBSeq on a recent survey,[11] RUSP classification, and disease prevalence, confirming the results from the regression analysis (Supplementary Fig. 8).

We also used this model to predict the observed inclusion of genes across all NBSeq programs. The result was a list of all genes that had appeared in any NBSeq program, ranked by their predicted inclusion probabilities, which were based on the 13 characteristics included in the model (Table 2). This analysis identified five genes *(ACADSB, PTPRC, NHEJ1, NAGLU,* and *ETFA)* that, despite being highly ranked by the model, were only included in a low proportion of NBSeq programs.

To address missing data for some genes that were included across multiple NBSeq programs, we created a second ranked list. This list combines the rankings generated by our

machine learning model with the proportion of NBSeq programs in which each gene was observed with equal weights (Supplementary Table 8). By integrating these two sources of information, this hybrid list leverages the most comprehensive evidence available to prioritize genes for potential implementation in public health programs.

Discussion

Genomic newborn screening is a rapidly advancing field of global research exploring the impacts of early diagnosis for infants at risk for genetic disorders. With positive screening results in 1.85% to 9.43% of infants and a higher average PPV than some traditional newborn screening techniques,[48,52] findings from NBSeq research programs support the premise that this approach could improve early detection rates for a wide range of treatable disorders. However, selecting the appropriate genes for screening is a critical step toward implementing population-wide NBSeq. This decision will have significant implications for at-risk infants, their families, and pediatric healthcare systems.[19,20] In this study, we compared the genes being analyzed by 27 NBSeq programs. We then collected data on 25 characteristics for all genes that had been included in any program and identified which of these characteristics were associated with inclusion across programs. We then developed a machine learning model to predict the inclusion of a gene across NBSeq programs. By combining this model with observed data from NBSeq programs, we generated a ranked list of genes that offers a data-driven approach to prioritizing genetic disorders for public health programs looking to incorporate NBSeq into their screening strategies.

Similar to the findings of smaller studies,[27,28] our comparison of gene lists from 27 NBSeq programs revealed substantial heterogeneity, which we explored using a series of regression models. We found that the importance of individual gene and disease characteristics varied across studies, potentially due to differences in the international prevalence of disorders, the availability of specialists and treatments in different countries and healthcare systems, or the specific goals of individual NBSeq programs. Furthermore, the rarity of many genetic

disorders often leads to incomplete knowledge of disease characteristics such as penetrance, age of onset, and treatability, which complicates the gene selection process. For example, the penetrance of each disorder may not be well-understood until population-wide genomic screening studies become more routine.[53,54] Unexpectedly, 356 genes with no disease association on OMIM and 52 with limited or refuted gene-disease validity scores from ClinGen were included by some NBSeq programs, demonstrating variation among programs in willingness to include candidate genes or those with new associations to disease.

Despite variations in the gene lists used by NBSeq programs, many share a common focus on certain clinical areas and specific genes. All programs included a substantial proportion of genes associated with disorders that are on the RUSP, reflecting the potential for genomic sequencing to detect cases missed by traditional newborn screening programs.[55–57] Of note, genes associated with some disorders on the RUSP, such as 3-methyl-crotonyl-CoA carboxylase deficiency (*MCCC1*, *MCCC2*), were widely included across lists despite not conforming to the historic Wilson-Jungner criteria. This suggests that some NBSeq programs have anchored their lists around the RUSP even when the disorders are neither severe nor highly treatable.[58] Therefore, the observed concordance of a gene across NBSeq programs alone may not be sufficient to evaluate the suitability of a gene for population-wide screening.

While NBSeq programs expanded the number of disorders included in screening, they largely did so within the same clinical areas already represented in current population-wide newborn screening programs, such as IMDs, inborn errors of immunity, endocrinologic disorders, and hematologic disorders. Several IMD genes with high concordance across NBSeq programs, such as those associated glycogen storage disease types Ia and Ib (*G6PC1*, *SLC37A4*), lack biomarkers that can be accurately assayed on a population scale and therefore are widely recognized as candidates for ascertainment by an NBSeq

approach.[11] In contrast, *F8*, the gene associated with hemophilia A, which shares similar clinical characteristics to hemophilia B and has previously been suggested as a target of NBS, was included only in a minority of lists. This discrepancy may reflect the technical challenge of using genomic sequencing to identify the inversions that most commonly underlie this disorder.[59] Developing the best approach toward complementary genomic and non-molecular screening approaches remains an ongoing challenge in the NBSeq field.

Many gene and disease characteristics emerged as highly associated with the inclusion of genes across NBSeq programs. Although the regression coefficients in our multivariate regression model varied for each characteristic across programs, most of the associations were positive. This indicates that all NBSeq programs value these characteristics, but weigh them differently when curating gene lists. The characteristics that were most strongly associated with inclusion across gene lists included the strength of published data on the natural history of disease, estimated penetrance, and the effectiveness of the associated treatment. The availability of a non-molecular confirmatory diagnostic test also influenced inclusion across NBSeq programs, likely because it offers additional phenotypic information for disorders with incomplete penetrance.[23] Interestingly, despite the inclusion criteria that several NBSeq programs reported, characteristics such as the age of onset and disease severity were weakly associated with inclusion, possibly due to their subjective nature. These results suggest that NBSeq programs are considering a wide range of technical and clinical factors specific to genetic diseases when developing their gene lists. Consequently, the traditional Wilson-Jungner criteria may no longer be adequate for guiding disease selection in public health programs using NBSeq. To address this, ICoNS plans to develop a new set of screening principles tailored to NBSeq by 2026.

The machine learning model developed in this study identifies the disorders that may be most appropriate for genomic newborn screening, based on a set of 13 disorder characteristics and their inclusion across 27 NBSeq programs. On a population scale, it may

not be feasible for every country to implement the screening of hundreds of genes simultaneously. This ranked list, along with the preferences of rare disease experts,[11] could be used to prioritize genes for screening, which could then be manually curated by a team of expert reviewers. Although the list includes many genes associated with disorders that are already routinely screened by many states in the US,[60] it may guide the adoption of screening for these disorders in countries where they are not yet assessed. At this time, the model's predictions reflect a consensus drawn from NBSeq studies and databases, but in the future, these could be combined with hard-coded gene selection criteria. The model's flexibility also allows updates based on regional preferences, new data, or emerging therapeutics. Importantly, this model identified several genes included by only a few NBSeq programs, but which have characteristics that are highly associated with inclusion across programs. For example, although *PTPRC*, a gene associated with severe combined immunodeficiency (SCID), was only included by 12 of 27 NBSeq programs, the model ranked it 122 of 4389 genes. This is likely because *PTPRC* is associated with a severe immunologic disorder that typically presents in childhood and can be treated with an early hematopoietic stem cell transplant, but is a rare cause of SCID.[61] This finding highlights the model's potential to identify genes that may have been overlooked by researchers during the gene selection process.

Our study has several limitations. Although we provided the first overview of screening outcomes across multiple NBSeq programs, detailed data on the percentage of positive screening results were available for only nine research studies. The Jaccard index may exaggerate discrepancies between gene lists of different lengths. For the regression and machine learning models, we consolidated metrics including the ASQM, BabySeq, and ClinGen databases, most of which rely on expert-curated information, such as age of onset, for which definitions vary. There were missing data among the 25 gene and disease characteristics that we collected. These gaps reduced the boosted tree model's prediction accuracy for genes that are not well-characterized, resulting in lower inclusion rates for

genes related to disorders that are rare or have limited published evidence. To mitigate the effects of missing data, we designed the model to be easily updated, and provided a ranked gene list that takes into account observed inclusion in addition to our model estimates. The ranked list may suffer from overfitting, given that it includes genes on which the model was trained. Lastly, the model incorporates data from funded research programs and commercial products, which may not align with the goals or constraints of public health newborn screening programs.

In summary, the growing international interest in genomic newborn screening has prompted urgent questions about which genes and disorders should be prioritized for inclusion. Due to the substantial variation in the genes included by 27 NBSeq programs, we developed an evidence-based approach to prioritizing gene selection that draws from a comprehensive data repository encompassing over 4,000 genes. Our machine learning model uses data from a variety of sources to predict which genes and their related diseases are currently most appropriate for NBSeq. Rather than creating a static list of genes for universal implementation, our dynamic ranking system is adaptable and can be updated as new knowledge about genes, disorders, and therapeutics emerges. This work will support gene selection for both research and public health genomic newborn screening programs and guide ICoNS in updating the Wilson-Jungner criteria to address the unique challenges posed by NBSeq.

Tables and Figures

Figure 1. Research and commercial genomic newborn screening (NBSeq) programs. Gene lists from 27 of these programs were included in the analysis (denoted with an asterisk). Intended enrollment sizes are indicated where available.

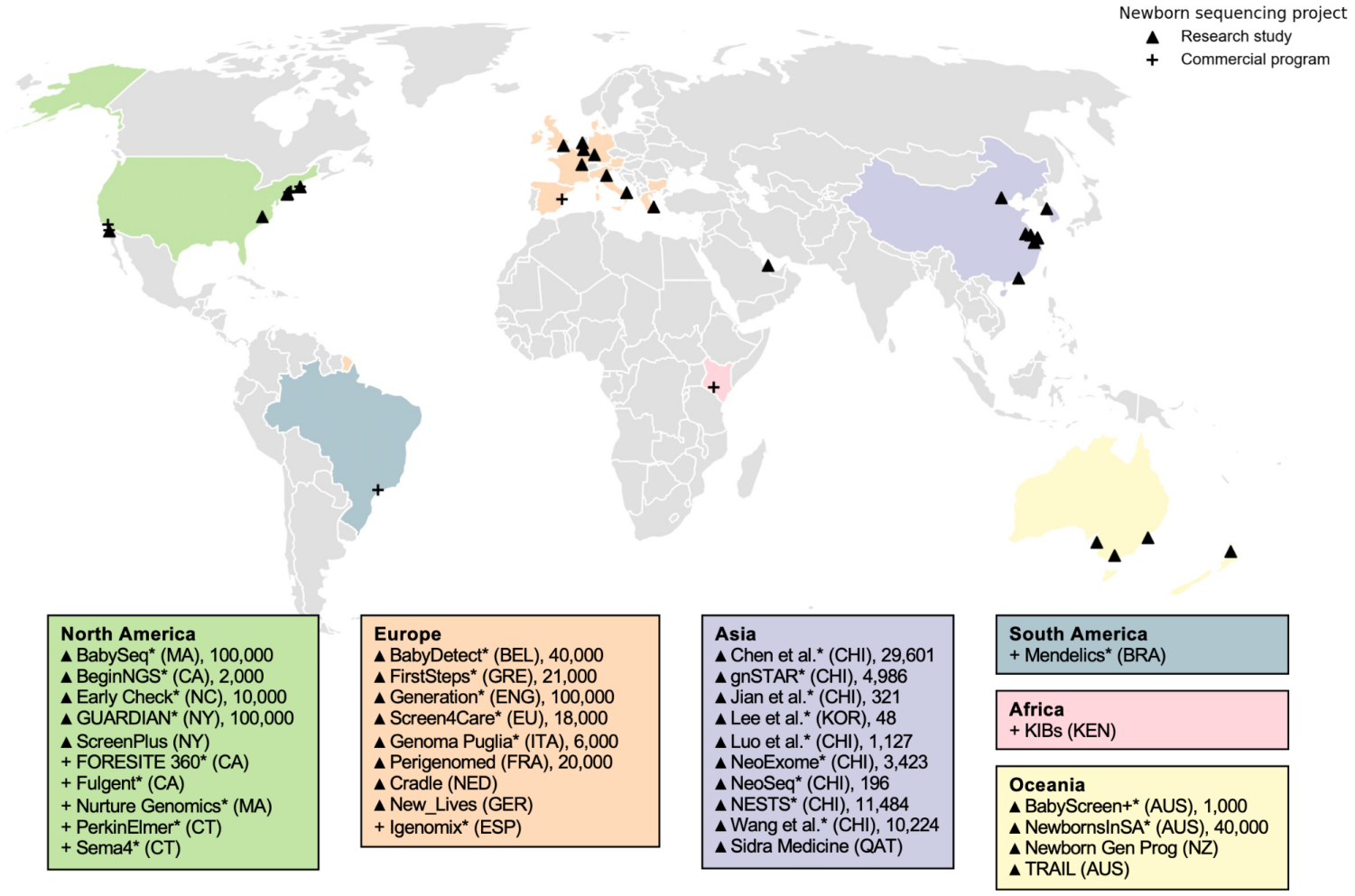
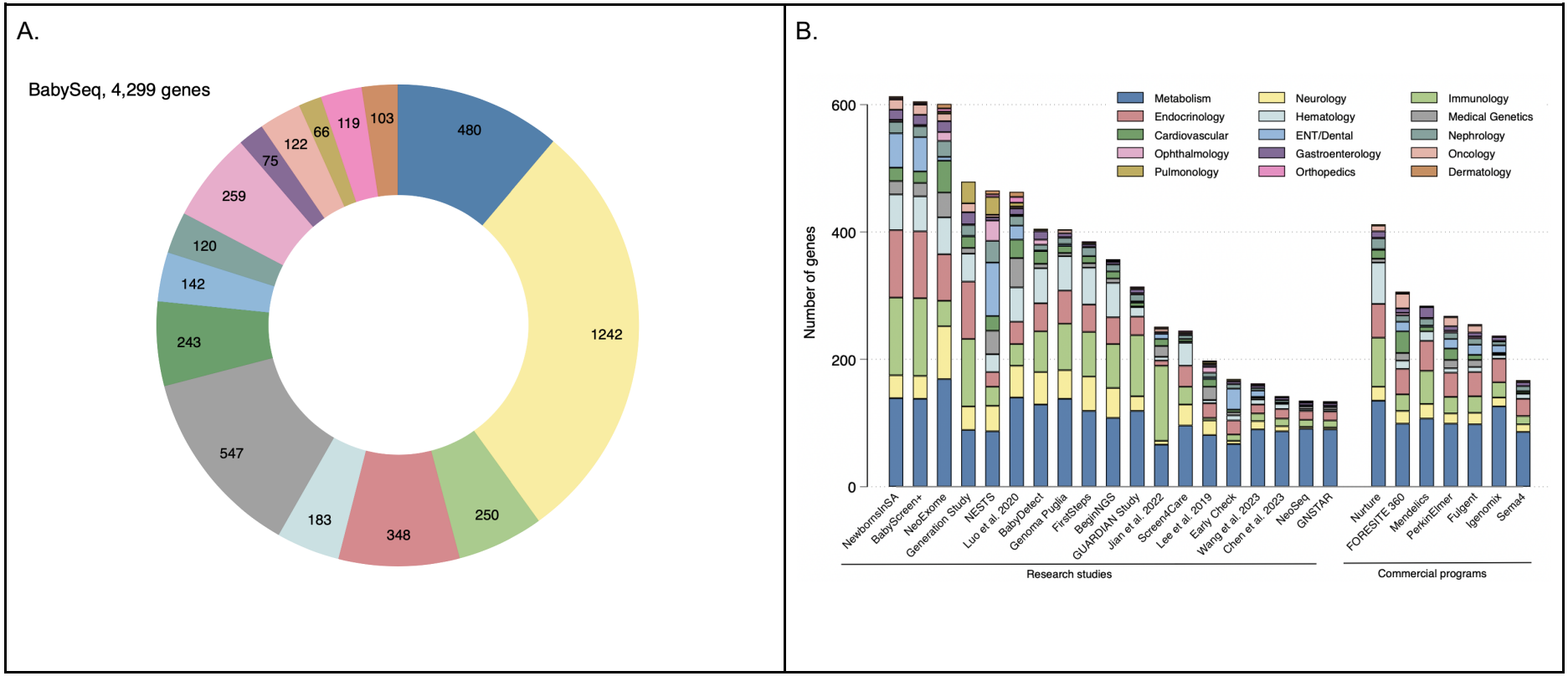
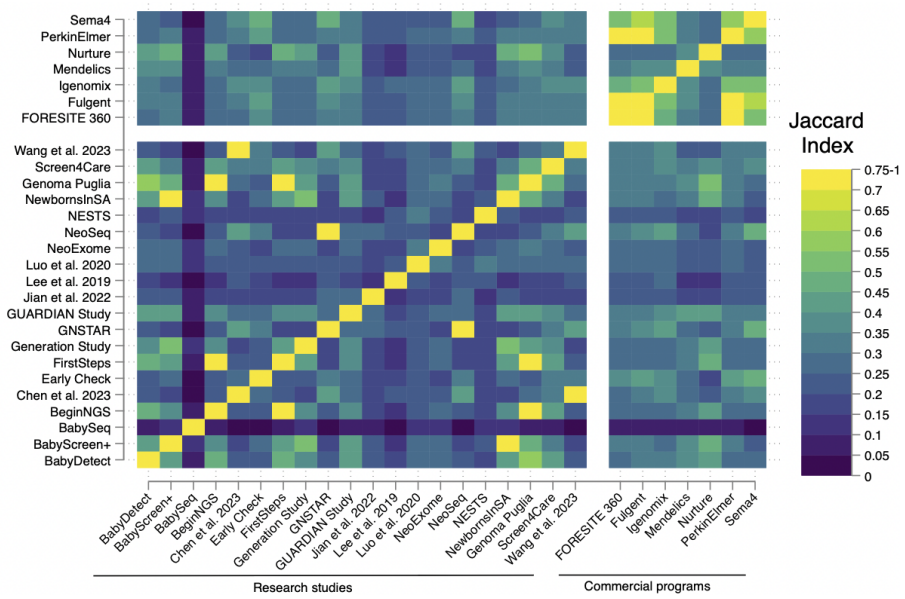


Figure 2. Description and concordance of gene lists of genomic newborn screening programs.

A. Clinical areas of 4,299 genes included in BabySeq.
B. Counts and clinical areas of genes included in 26 research and commercial genomic newborn screening programs (n=4,389).
C. Jaccard similarity index, which offers a quantitative comparison of how closely related the gene lists are.
D. UpSet plot[62] of gene lists of 4 large research studies. The matrix below the bar graph represents each individual study and their intersections (n=818).



C.


Sema4
PerkinElmer
Nurture
Mendelics
Igenomix
Fulgent
FORESITE 360
Wang et al. 2023
Screen4Care
Genoma Puglia
NewbornsInSA
NESTS
NeoSeq
NeoExome
Luo et al. 2020
Lee et al. 2019
Jian et al. 2022
GUARDIAN Study
GNSTAR
Generation Study
FirstSteps
Early Check
Chen et al. 2023
BeginNGS
BabySeq
BabyScreen+
BabyDetect
Jaccard Index
0.75-1
0.7
0.65
0.6
0.55
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
Research studies
Commercial programs

D.

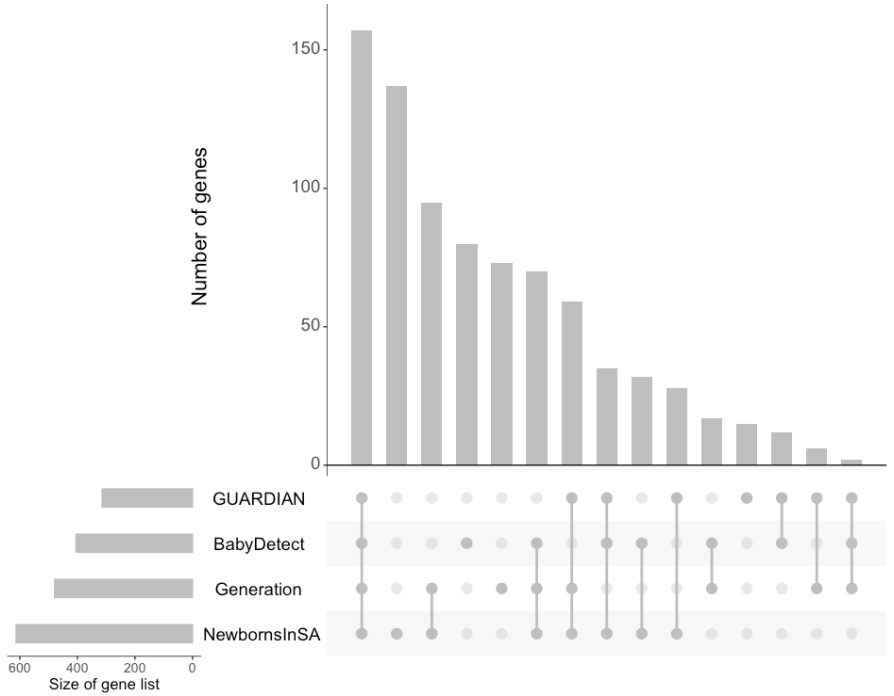

Number of genes
150
100
50
0
GUARDIAN
BabyDetect
Generation
NewbornsInSA
600
400
200
0
Size of gene list

Figure 3. Genes with high concordance across genomic newborn screening programs.

A. Genes associated with core Recommended Uniform Screening Panel (RUSP) conditions.
B. Genes associated with secondary RUSP conditions.
C. Genes on 20 lists or more that are not associated with RUSP conditions.

The x-axis is each genomic newborn screening program and y-axis are individual genes; the corresponding cell is colored if the gene is included on a given list.

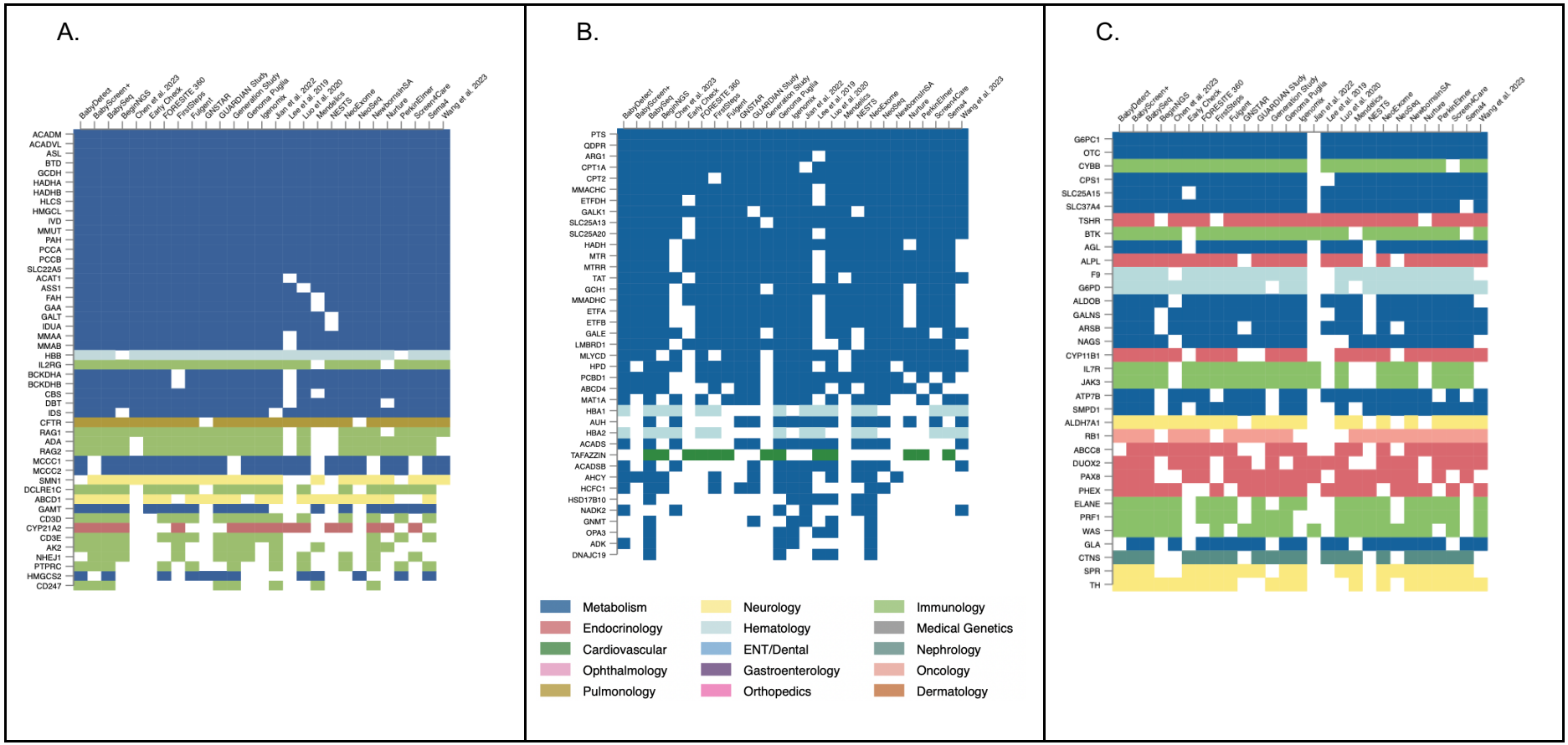
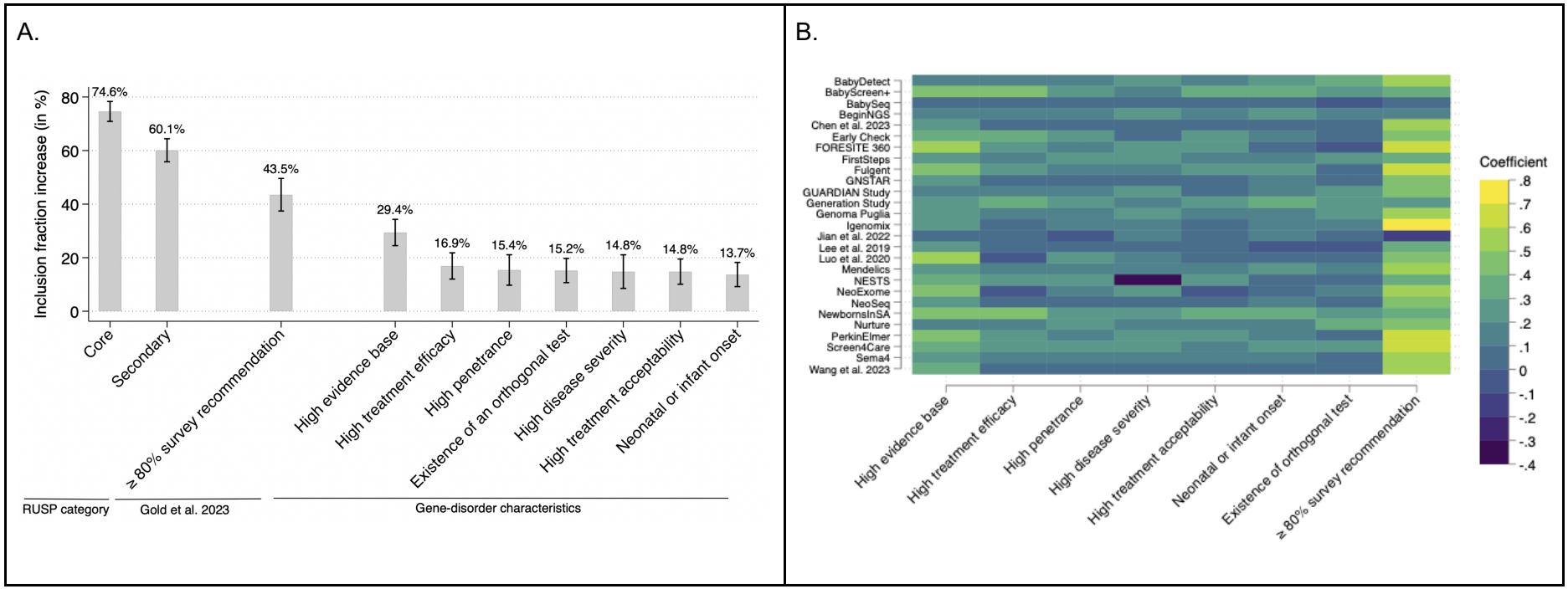


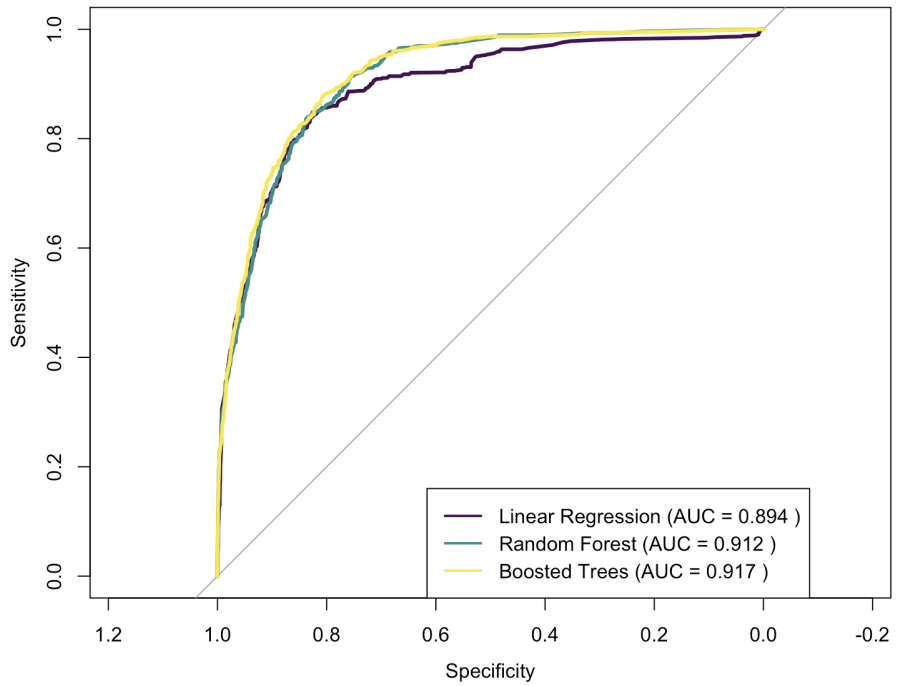
Figure 4. Determinants and prediction model of gene inclusion in genomic newborn screening.

A. Regression coefficients (and confidence intervals) associated with various gene and disease characteristics predicting inclusion across gene lists.
B. Heat map with regression coefficients associated with gene and disease characteristics for each individual genomic newborn screening program.
C. ROC curves for three prediction models in the hold-out test set (n=895 genes).
D. Scatter plot of predicted versus observed gene list inclusion, showing the fit of the boosted trees model on the 20% hold-out set (n=895 genes).

In a and b, RUSP category (n=4,474), survey recommendation and orthogonal test (n=649), evidence base, efficacy, penetrance, disease severity, treatment acceptability and neonatal or infant onset (n=749). ROC, Receiver Operating Characteristic; AUC, Area Under the Curve.



C.


Sensitivity
1.0
0.8
0.6
0.4
0.2
0.0
1.2
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
Specificity
Linear Regression (AUC = 0.894)
Random Forest (AUC = 0.912)
Boosted Trees (AUC = 0.917)

D.

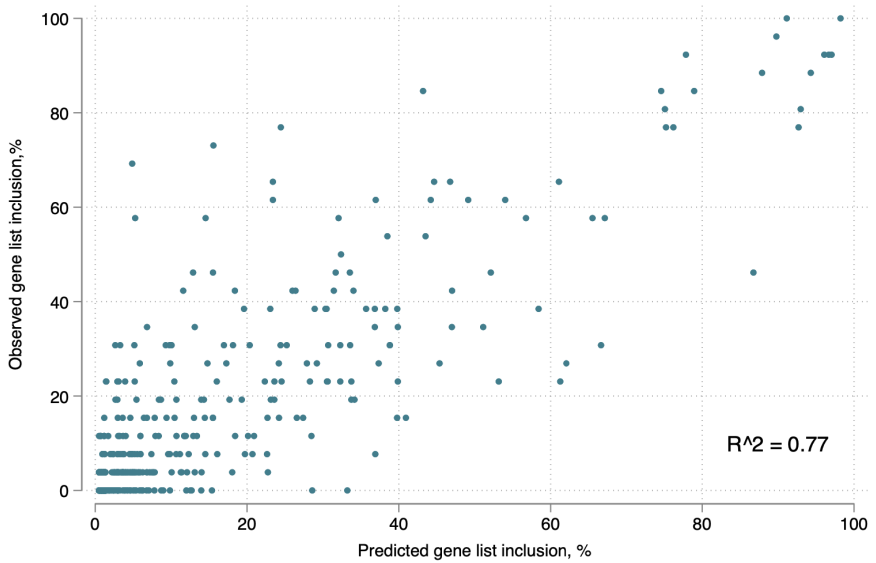

Observed gene list inclusion,%
100
80
60
40
20
0
0
20
40
60
80
100
Predicted gene list inclusion, %
R^2 = 0.77

Table 1. Percentage of positive results from genomic newborn screening research programs. Positive cases of *G6PD,* as well as follow-up data for positively screened infants reported where available. PPV, positive predictive value.

		Sequencing data				Follow-up data		
NBSeq research study	Genes	Infants sequenced	G6PD deficiency	Total positive	Total positive (in %)	Total follow-up	Total diagnosed	PPV
BabyDetect	405	3,847	44	71	1.85%			
PerkinElmer Panel	268	606	4	13	2.15%			
Wang et al. 2023	164	10,334		232	2.25%	231	50	22%
gnSTAR	134	4,986	40	113	2.27%			
Chen et al. 2023	142	29,601	689	813	2.75%	797	402	50%
GUARDIAN Study	148	4,256		144	3.38%	142	125	88%
NESTS	465	11,484	338	902	7.85%	414	50	12%
PerkinElmer WGS	6000	562	4	46	8.19%			
NeoExome	601	3,049		271	8.89%			
BabySeq1	4299	159	1	15	9.43%			
TOTAL		**68,884**		**2,620**	**3.80%**			

Table 2. List of 50 genes with highest predicted inclusion across NBSeq programs. RUSP, Recommended Uniform Screening Panel; ASQM, Age-Based Semi Quantitative Metric.

Rank	Gene	Disorder	Clinical area	RUSP	BabySeq category	ClinGen validity	ClinGen Actionability (/12)	ASQM score (/15)	Expert rec.	Observed list inclusion (/27)	Predicted list inclusion
1	*ACADVL*	VLCAD deficiency	Metabolism	Core	A	Definitive	9DD	15		27	98%
2	*ACADM*	Acyl-CoA dehydrogenase, medium chain, deficiency of	Metabolism	Core	A	Definitive	11NB	15		27	98%
3	*HADHA*	LCHAD deficiency	Metabolism	Core	A	Definitive	9CC	15		27	98%
4	*GAA*	Glycogen storage disease II	Metabolism	Core	A	Definitive	9CA	14		26	97%
5	*GCDH*	Glutaricaciduria, type I	Metabolism	Core	A	Definitive	10DD	13		27	97%
6	*BTD*	Biotinidase deficiency	Metabolism	Core	A	Definitive	11CC	13		27	97%
7	*PAH*	Phenylketonuria	Metabolism	Core	A	Definitive	10AA	13		27	97%
8	*GALT*	Galactosemia	Metabolism	Core	A	Definitive		13		26	96%
9	*MMUT*	Methylmalonic aciduria, mut(0) type	Metabolism	Core	A	Definitive	10CB	13		27	96%
10	*MMAB*	Methylmalonic aciduria, vitamin B12-responsive, cblB type	Metabolism	Core	A	Definitive	10CB	13		26	96%
11	*GALK1*	Galactokinase deficiency with cataracts	Metabolism	Secondary	A	Definitive		13		25	96%
12	*DBT*	Maple syrup urine disease, type II	Metabolism	Core	A	Definitive	9DB	12		25	96%
13	*IVD*	Isovaleric acidemia	Metabolism	Core	A	Definitive		13		27	96%
14	*FAH*	Tyrosinemia, type I	Metabolism	Core	A	Definitive		13		26	96%
15	*CBS*	Homocystinuria, B6-responsive and nonresponsive types	Metabolism	Core	A	Definitive	9CB	13		25	96%
16	*HLCS*	Holocarboxylase synthetase deficiency	Metabolism	Core	A	Definitive		13		27	96%
17	*MMAA*	Methylmalonic aciduria, vitamin B12-responsive, cblA type	Metabolism	Core	A	Definitive	10CB	13		26	96%
18	*ASS1*	Citrullinemia	Metabolism	Core	A	Definitive	10CB	12		26	95%
19	*ASL*	Argininosuccinic aciduria	Metabolism	Core	A	Definitive	9NC	12		27	95%
20	*OTC*	Ornithine transcarbamylase deficiency	Metabolism		A	Definitive	10AB	14	98%	26	95%
21	*PCCA*	Propionic acidemia	Metabolism	Core	A	Definitive		13		27	95%
22	*IDS*	Mucopolysaccharidosis II	Metabolism	Core	A	Definitive	9CB	11	89%	25	95%
23	*SLC25A20*	Carnitine-acylcarnitine translocase deficiency	Metabolism	Secondary	A	Definitive		13		25	95%
24	*QDPR*	Hyperphenylalaninemia, BH4-deficient, C	Metabolism	Secondary	A	Definitive		13		27	95%
25	*PTS*	Hyperphenylalaninemia, BH4-deficient, A	Metabolism	Secondary	A	Definitive		13		27	95%
26	*HADHB*	Mitochondrial trifunctional protein deficiency 2	Metabolism	Core	A	Definitive	9CC	13		27	95%
27	*G6PC1*	Glycogen storage disease Ia	Metabolism		A			13	93%	26	94%
28	*BCKDHA*	Maple syrup urine disease, type Ia	Metabolism	Core	A	Definitive	9DB	12		25	94%
29	*BCKDHB*	Maple syrup urine disease, type Ib	Metabolism	Core	A	Definitive	9DB	12		25	94%
30	*CPT2*	CPT II deficiency, infantile	Metabolism	Secondary	A	Definitive	8CN	13		26	94%
31	*PCCB*	Propionic acidemia	Metabolism	Core	A	Definitive		13		27	94%
32	*SLC22A5*	Carnitine deficiency, systemic primary	Metabolism	Core	A	Definitive	10BA	10		27	94%
33	*IDUA*	Mucopolysaccharidosis I	Metabolism	Core	A	Definitive	8AB	11		26	94%
34	*MMACHC*	Methylmalonic aciduria and homocystinuria, cblC type	Metabolism	Secondary	A	Definitive		13		26	94%
35	*SLC25A13*	Citrullinemia, type II, neonatal-onset	Metabolism	Secondary	A	Definitive	6DN	10		25	93%
36	*SLC37A4*	Glycogen storage disease Ib	Metabolism		A			13	93%	25	92%
37	*CPS1*	Carbamoylphosphate synthetase I deficiency	Metabolism		A	Definitive	9AB	13	86%	25	92%
38	*HADH*	3-hydroxyacyl-CoA dehydrogenase deficiency	Metabolism	Secondary	C	Definitive		13		24	92%
39	*ACAT1*	Alpha-methylacetoacetic aciduria	Metabolism	Core	A			13	57%	26	92%
40	*HMGCL*	HMG-CoA lyase deficiency	Metabolism	Core	A	Definitive		11		27	90%
41	*ETFA*	Glutaric acidemia IIA	Metabolism	Secondary	A	Definitive		9		22	90%
42	*CPT1A*	CPT deficiency, hepatic, type IA	Metabolism	Secondary	A	Definitive		12		26	90%
43	*CFTR*	Cystic fibrosis	Pulmonology	Core	A			12		25	90%
44	*HBB*	Sickle cell disease	Hematology	Core	A	Definitive		12		25	89%
45	*ARG1*	Argininemia	Metabolism	Secondary	A	Definitive		10		26	89%
46	*LMBRD1*	Methylmalonic aciduria and homocystinuria, cblF type	Metabolism	Secondary	A	Definitive		11		22	88%
47	*RAG1*	Severe combined immunodeficiency, B cell-negative	Immunology	Core	A	Definitive		14		23	88%
48	*ADA*	Severe combined immunodeficiency due to ADA deficiency	Immunology	Core	A	Definitive		14		22	88%
49	*IL2RG*	Combined immunodeficiency, X-linked, moderate	Immunology	Core	A	Definitive		15		25	88%
50	*SMN1*	Spinal muscular atrophy-1	Neurology	Core	A			13		22	88%

Methods

Identification of lists of genes from research studies and commercial programs

Research studies' and commercial programs' gene panels related to NBSeq were identified based on inclusion in the International Consortium on Newborn Sequencing (ICoNS) and through an online search using terms related to genomic sequencing of infants. In total, 35 programs were identified, of which 27 provided gene lists (Supplementary Table S1, Fig. 1).

We included gene lists from 27 NBseq programs, including 20 research studies: BabyDetect[29,30], BabyScreen+[27], BabySeq[2], BeginNGS[31,32], Chen et al. 2023[33], Early Check[34], FirstSteps, the Generation study, gnSTAR[35,37], GUARDIAN study[41], Jian et al. 2022[36], Lee et al. 2019[42], Luo et al. 2020[43], NeoExome[46], NeoSeq[39], NESTS[40], NewbornsInSA, Progetto Genoma Puglia, Screen4Care[44] and Wang et al. 2023[38]. In two studies (GUARDIAN[41] and Early Check[34]), all infants receive testing for a gene list focused on conditions with effective treatments and parents have the option to be tested for an expanded gene list. For both of these studies, we included only the core gene list focused on treatable genetic conditions. Seven lists of genes from commercial firms that offer products related to genomic newborn screening were included: FORESITE 360, Fulgent, Igenomix, Mendelics, Nurture Genomics, PerkinElmer[45], Sema4[29]. Of note, the Sema4 product is no longer commercially available.

Rates of positive screening results

We obtained data provided by all included studies for results of their NBSeq activities. A total of nine studies had published or presented results as of August 2024. As studies had different approaches to participant selection, Table 1 only reports results of NBSeq programs that screen apparently healthy infants. We excluded results from samples where specifically at-risk infants (such as those in the neonatal intensive care unit) were sequenced. From the Wang *et al* study, only the results on healthy infants were retained.[38] From the BabySeq study, we included both the healthy and NICU infant sample, as only unanticipated results

unconnected with NICU clinical presentations were described for the NICU sample.[5] We excluded a sample from the NeoEXOME study of neonates that had positive results in conventional NBS.[46]

Aggregation of gene lists

Gene names were converted to the current nomenclature set forth by the HUGO Gene Nomenclature Committee (HGNC) based on an available online multi-symbol checker.[63] For purposes of analysis, each gene was linked to one condition. The multistep process for linking genes to a single disorder began by first identifying the phenotypes associated with each gene on Online Inheritance in Man (OMIM).[64] If only one disease name was associated with the gene on OMIM, a gene-disease pair was formed. If the gene was known to be associated with multiple OMIM disorders, we used the ClinGen gene-disease validity resource to select only the disorder with definitive classification when available.[51] For genes with more than one disorder with definitive validity or for genes without any disorder with definite validity, one disorder was selected based on the highest number of programs in this study that indicated it as screening target. For example, for *RYR1*, which has a definitive association with both susceptibility to malignant hyperthermia and myopathy,[51] susceptibility to malignant hyperthermia was selected as the target disorder. Susceptibility to malignant hyperthermia was indicated as a target disease by five of seven programs with disease information available screening for this gene, compared with myopathy which was listed as a target disease by only two of seven programs.

A total of 264 genes were not annotated with disorder information from either ClinGen or OMIM databases, possibly because they were candidate genes or had very recently been substantiated as disease genes. These were then manually matched to diseases by search in HGMD (https://www.hgmd.cf.ac.uk/ac/index.php). Two gene names, *GTM* and *CD1*, which could not be linked to HGNC approved gene names, were omitted from the analysis.

Establishing a data repository of characteristics of genes-disorder pairs

We established a data repository consisting of 25 characteristics for all gene-disorder pairs, sourced from ten different references: five research papers and five existing databases.[2,10,11,31,49–51] Matching of data from these databases was based on finding an exact match with the gene-disorder pair. Matches were then manually checked for correctness. Three variables (prevalence, clinical area, and RUSP category) were newly constructed for this study and were based on consolidated information from varying sources. Supplementary Table 3 provides an overview of all variables, as well as a description of all metrics and their respective sources.

To determine whether each gene-disease pair was associated with a disorder listed on the United States Recommended Uniform Screening Panel (RUSP), we cross-referenced the genes identified by Owen *et al*[32] with the diseases listed on the RUSP section of the United States Health Resources and Services Administration (HRSA) website (https://www.hrsa.gov/advisory-committees/heritable-disorders/rusp). A gene was considered associated with a RUSP-listed disease if it appeared in the "Cause" section of the corresponding disease-specific HRSA webpage.

Each gene and its associated condition was assigned to one of 12 clinical areas, determined by the clinical specialty most likely to treat patients with those disorders. The clinical areas include cardiovascular, dermatology, ENT/Dental, endocrinology, gastroenterology, hematology, immunology, medical genetics, metabolism, nephrology, neurology, oncology, ophthalmology, orthopedics, and pulmonology. For 649 disorders, clinical area assignments were based on a previously published paper.[11] For all other gene-disorder pairs, one co-author assigned the clinical area, and another co-author verified this assignment (S.A., S.B., N.B.G.).

Prevalence estimates for gene-disease pairs were obtained from four sources: Orphanet (https://www.orpha.net/), RX-genes (https://www.rx-genes.com/), and two previously published studies.[11,31] Gene-disorder links in Orphanet were established through Orphacode, and only global prevalence estimates were retained. Consolidated prevalence estimates were selected based on the following preference: RX-genes, Orphanet, Gold *et al* (2023), and Kingsmore *et al* (2022). Prevalence data were obtained from these sources for 1,364 gene-disease pairs, covering 31% of the 4,389 genes. Notably, 68% of these disorders had a prevalence of 1 per 1,000,000 or less, highlighting the ultra-rare nature of many disorders included in newborn sequencing programs.

Additional gene and disorder characteristics, including disease penetrance, severity, treatment acceptability and efficacy, age of onset, evidence base (which refers to the level of knowledge about the natural history of the disorder and its treatments), inheritance patterns and the existence of an orthogonal confirmatory non-diagnostic test, were derived from five previously published studies.[2,11,49,50] When two modes of inheritance were implicated for the same gene and disease in these studies, such as for *MYO6*, a cause of non-syndromic deafness, dominant inheritance was selected as it was expected to lead to the most inclusive reporting criteria.

Gene-disorder pairs were also matched with ClinGen (https://clinicalgenome.org/) gene-disease curations, which were evaluated using a standardized approach to assess the strength of evidence linking a gene to a monogenic disease. Additionally, ClinGen clinical pediatric and adult actionability curations were obtained, with the highest actionability score retained when multiple curations were available.[8,50] For gene-disorder pairs with both pediatric and adult curations, only the pediatric score was retained.

Three existing metrics that address the suitability of genes for newborn sequencing were also included. The age-based semi-quantitative metric (ASQM) score[49,50] is a metric which

assigns a number between 0 and 15 to a gene-disease pair to denote overall appropriateness for newborn screening based on several sub-scores. The BabySeq Category is another metric scored by the BabySeq Project,[2] where BabySeq Category A is designated as the category of genes most amenable to newborn screening, while Category C is considered less amenable to newborn screening. Finally, we incorporated the proportion of 238 rare disease experts who recommended screening newborns for 649 genes, as determined by an online survey.[11]

Statistical analysis

All statistical analyses were carried out in Stata 18 (College Station, TX), R version 4.3.1 (Vienna, Austria) and Python version 3.11.2 (Python Software Foundation, Beaverton, OR). Descriptive statistics for each gene list, including the length of the list, proportion of genes in each clinical category, and the number of genes associated with RUSP conditions were calculated. The distribution of genes across BabySeq categories, average ASQM score, and the proportion of rare disease experts recommending screening for the gene were calculated within each program (Supplementary Table 4).

To represent overlapping parts of different gene lists, an UpSet plot was plotted using the UpSet[62] library in R (Fig. 2D), as well as Venn diagrams (Supplementary Fig. 3). To provide information on the concordance across all lists of genes, Jaccard similarity indices were calculated using the Jaccard command in Stata (Fig. 2C and Supplementary Fig. 4). This index measures the number of genes in the intersection set divided by the number of genes in the union set of two gene lists.

A linear regression model was used to identify factors associated with inclusion in multiple gene lists. Two types of regressions were performed: (1) regressions in which the outcome variable is the proportion of gene list inclusion across all genomic newborn screening programs and the independent variables are factors related to each gene and its associated

condition (Supplementary Table 5, 6), and (2) regressions in which inclusion of a gene for each *individual* study was explored (Supplementary Table 7). These regressions were implemented using the reg command in Stata with default standard error settings. For Fig. 4A, the coefficient on each characteristic was measured in a separate regression, where the only other control was the RUSP category. For Fig. 4B, coefficients were obtained for each program separately running several regressions for each program, one for each program-characteristic combination. Again, these regressions were controlled for the RUSP category. Standard errors, while not reported in 4B, can be found in the Supplementary Table 7 where the full results are reported.

Machine learning prediction analysis

The prediction analysis was implemented in R using linear regression, random forest and boosted trees (glm, randomforest and gbm packages). Our machine learning algorithm was developed to predict a binary variable: whether a gene was included in an NBSeq program's gene list. We also used the proportion of 27 NBSeq programs that included the gene as an outcome measure, but results were similar. Since the BabySeq project uses an "elective exome approach" that includes nearly all genes associated with human disease, we excluded this list when training the model. The overall dataset used for prediction consists of 114,114 program-gene observations. We randomly assigned 80% of genes (91,291 observations) to the training set, and 20% to the hold-out test set.

Out of 25 potential gene and disease characteristics, we selected 13 for inclusion in our model. These included the RUSP category, clinical area, evidence base, treatment efficacy, penetrance, treatment acceptability, age of onset, existence of an orthogonal test, recommendation score, inheritance, prevalence, and the ClinGen Disease Validity and Actionability scores. The remaining 12 gene and disease characteristics were excluded due to a high amount of missing data or overlapping evidence. For example, ClinGen actionability scores were not used due to their availability for only 242 genes. For composite

scores like the ASQM Score, NC Nexus category, or BabySeq category, we instead included the individual variables used to create these scores. Additionally, when characteristics described similar concepts (for example, both ASQM and BabySeq address the inheritance pattern of disease), we selected the characteristic with data for the most genes. For the 13 gene and disease characteristics included in the model, missing values were handled by adding dummy variables in the regression model and setting the missing predictor to zero. In the random forest and boosted trees models, missing values were set to -1 or labeled as 'missing' in the case of categorical variables.

We conducted a grid search using cross-validation to optimize the hyperparameters in the boosted trees model using the caret package. We optimized interaction depth (interaction.depth: 1, 5), number of trees (n.trees: 100, 500), learning rate (shrinkage: 0.01, 0.1), and minimum observations per node (n.minobsinnode: 10, 20). The final model configuration was chosen based on its performance during three-fold cross-validation, aiming to maximize predictive accuracy.

Model performance was evaluated using Receiver Operating Characteristic (ROC) curves and the Area Under the Curve (AUC) metric for each model. These metrics were computed to assess the models' discriminative ability. Additionally, we plotted a calibration plot, showing the observed versus predicted inclusion probabilities for each gene-disorder pair in the test set (Supplementary Fig.7). We assessed the importance of gene and disease characteristics for the random forest and boosted trees models using the randomForest and gbm packages, where importance was measured by the mean decrease in accuracy for random forest and reduction in deviance for boosted trees.

References

1. Waisbren, S. E. *et al.* Parents are interested in newborn genomic testing during the early postpartum period. *Genet. Med.* **17**, 501–504 (2015).
2. Ceyhan-Birsoy, O. *et al.* A curated gene list for reporting results of newborn genomic sequencing. *Genet. Med.* **19**, 809–818 (2017).
3. Genetti, C. A. *et al.* Parental interest in genomic sequencing of newborns: enrollment experience from the BabySeq Project. *Genet. Med.* **21**, 622–630 (2019).
4. Holm, I. A. *et al.* The BabySeq project: implementing genomic sequencing in newborns. *BMC Pediatr.* **18**, 225 (2018).
5. Ceyhan-Birsoy, O. *et al.* Interpretation of Genomic Sequencing Results in Healthy and Ill Newborns: Results from the BabySeq Project. *Am. J. Hum. Genet.* **104**, 76–93 (2019).
6. Pereira, S. *et al.* Psychosocial Effect of Newborn Genomic Sequencing on Families in the BabySeq Project: A Randomized Clinical Trial. *JAMA Pediatr.* **175**, 1132–1141 (2021).
7. Pereira, S. *et al.* Perceived Benefits, Risks, and Utility of Newborn Genomic Sequencing in the BabySeq Project. *Pediatrics* **143**, S6–S13 (2019).
8. Green, R. C. *et al.* Actionability of unanticipated monogenic disease risks in newborn genomic screening: Findings from the BabySeq Project. *Am. J. Hum. Genet.* **110**, 1034–1045 (2023).
9. Smith HS, Zettler B, Genetti CA, Hickingbotham MR, Coleman TF, Lebo M, Nagy A, Zouk H, Mahanta L, Christensen KD, Pereira S, Shah ND, Gold NB, Walmsley S, Edwards S, Homayouni R, Krasan GP, Hakonarson H, Horowitz CR, Gelb BD, Korf BR, McGuire AL, Holm IA, Green RC. The BabySeq Project: A Clinical Trial of Genome Sequencing in a Diverse Cohort of Infants. *Am. J. Hum. Genet., In Press* (2024).
10. Bick, D. *et al.* An online compendium of treatable genetic disorders. *Am. J. Med. Genet. C Semin. Med. Genet.* **187**, 48–54 (2021).
11. Gold, N. B. *et al.* Perspectives of Rare Disease Experts on Newborn Genome

Sequencing. *JAMA Netw Open* **6**, e2312231 (2023).

12. Joseph, G. *et al.* Parental Views on Expanded Newborn Screening Using Whole-Genome Sequencing. *Pediatrics* **137 Suppl 1**, S36–46 (2016).
13. Timmins, G. T., Wynn, J., Saami, A. M., Espinal, A. & Chung, W. K. Diverse Parental Perspectives of the Social and Educational Needs for Expanding Newborn Screening through Genomic Sequencing. *Public Health Genomics* 1–8 (2022).
14. Acharya, K., Ackerman, P. D. & Ross, L. F. Pediatricians' attitudes toward expanding newborn screening. *Pediatrics* **116**, e476–84 (2005).
15. Cao, M., Notini, L., Ayres, S. & Vears, D. F. Australian healthcare professionals' perspectives on the ethical and practical issues associated with genomic newborn screening. *J. Genet. Couns.* **32**, 376–386 (2023).
16. del Rosario, M. C. *et al.* Genetic counselors' perspectives on genomic screening of apparently healthy newborns in the United States. *Genetics in Medicine Open* 101885 (2024).
17. Bombard, Y. *et al.* Public views on participating in newborn screening using genome sequencing. *Eur. J. Hum. Genet.* **22**, 1248–1254 (2014).
18. Lynch, F. *et al.* Australian Public Perspectives on Genomic Newborn Screening: Risks, Benefits, and Preferences for Implementation. *Screening* **10**, (2024).
19. Stark, Z. & Scott, R. H. Genomic newborn screening for rare diseases. *Nat. Rev. Genet.* **24**, 755–766 (2023).
20. Downie, L., Halliday, J., Lewis, S. & Amor, D. J. Principles of Genomic Newborn Screening Programs: A Systematic Review. *JAMA Netw Open* **4**, e2114336 (2021).
21. Bros-Facer, V., Taylor, S. & Patch, C. Next-generation sequencing-based newborn screening initiatives in Europe: an overview. *Rare Dis Orphan Drugs J* **2**, 21 (2023).
22. International Consortium on Newborn Sequencing. https://www.iconseq.org/.
23. Bick, D. *et al.* Newborn Screening by Genomic Sequencing: Opportunities and Challenges. *Screening* **8**, (2022).
24. Baple, E. L. *et al.* Exploring the benefits, harms and costs of genomic newborn

screening for rare diseases. *Nat. Med.* **30**, 1823–1825 (2024).

25. Wilson, J. M. G. & Jungner, G. *The Principles and Practice of Screening for Disease*. (1966).
26. DeCristo, D. M. *et al.* Actionability of commercial laboratory sequencing panels for newborn screening and the importance of transparency for parental decision-making. *Genome Med.* **13**, 50 (2021).
27. Downie, L. *et al.* Gene selection for genomic newborn screening: moving towards consensus? *Genet. Med.* 101077 (2024).
28. Betzler, I. R. *et al.* Comparative analysis of gene and disease selection in genomic newborn screening studies. *J. Inherit. Metab. Dis.* (2024) doi:10.1002/jimd.12750.
29. Dangouloff, T. *et al.* Baby detect: Universal genomic newborn screening for early, treatable, and severe conditions. *J. Neurol. Sci.* **455**, (2023).
30. Website. 'Population-Based, First-Tier Genomic Newborn Screening in a Single Maternity Ward in Belgium: Results of Babydetect Project.' n.d. Accessed August 8, 2024. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4896054.
31. Kingsmore, S. F. *et al.* A genome sequencing system for universal newborn screening, diagnosis, and precision medicine for severe genetic diseases. *Am. J. Hum. Genet.* **109**, 1605–1619 (2022).
32. Owen, M. J. *et al.* An automated 13.5 hour system for scalable diagnosis and acute management guidance for genetic diseases. *Nat. Commun.* **13**, 4057 (2022).
33. Chen, T. *et al.* Genomic Sequencing as a First-Tier Screening Test and Outcomes of Newborn Screening. *JAMA Netw Open* **6**, e2331162 (2023).
34. Bailey, D. B., Jr *et al.* Early Check: translational science at the intersection of public health and newborn screening. *BMC Pediatr.* **19**, 238 (2019).
35. Huang, X. *et al.* Application of a next-generation sequencing (NGS) panel in newborn screening efficiently identifies inborn disorders of neonates. *Orphanet J. Rare Dis.* **17**, 66 (2022).
36. Jian, M. *et al.* A pilot study of assessing whole genome sequencing in newborn

screening in unselected children in China. *Clin. Transl. Med.* **12**, e843 (2022).

37. Yang, R.-L. *et al.* A multicenter prospective study of next-generation sequencing-based newborn screening for monogenic genetic diseases in China. *World J. Pediatr.* **19**, 663–673 (2023).
38. Wang, X. *et al.* Newborn genetic screening is highly effective for high-risk infants: A single-centre study in China. *J. Glob. Health* **13**, 04128 (2023).
39. Wang, H. *et al.* NeoSeq: a new method of genomic sequencing for newborn screening. *Orphanet J. Rare Dis.* **16**, 481 (2021).
40. Hao, C. *et al.* Newborn screening with targeted sequencing: a multicenter investigation and a pilot clinical study in China. *J. Genet. Genomics* **49**, 13–19 (2022).
41. Chung, W. K., Kanne, S. M. & Hu, Z. An Opportunity to Fill a Gap for Newborn Screening of Neurodevelopmental Disorders. *Screening* **10**, (2024).
42. Lee, H. *et al.* Implementation of a Targeted Next-Generation Sequencing Panel for Constitutional Newborn Screening in High-Risk Neonates. *Yonsei Med. J.* **60**, 1061–1066 (2019).
43. Luo, X. *et al.* A pilot study of expanded newborn screening for 573 genes related to severe inherited disorders in China: results from 1,127 newborns. *Ann Transl Med* **8**, 1058 (2020).
44. Ferlini, A., Gross, E. S., Garnier, N. & Screen4Care consortium. Rare diseases' genetic newborn screening as the gateway to future genomic medicine: the Screen4Care EU-IMI project. *Orphanet J. Rare Dis.* **18**, 310 (2023).
45. Balciuniene, J. *et al.* At-Risk Genomic Findings for Pediatric-Onset Disorders From Genome Sequencing vs Medically Actionable Gene Panel in Proactive Screening of Newborns and Children. *JAMA Netw Open* **6**, e2326445 (2023).
46. Cao, Z. *et al.* Targeted exome sequencing strategy (NeoEXOME) for Chinese newborns using a pilot study with 3423 neonates. *Mol Genet Genomic Med* **12**, e2357 (2024).
47. Spiekerkoetter, U. *et al.* Genomic newborn screening: Are we entering a new era of screening? *J. Inherit. Metab. Dis.* **46**, 778–795 (2023).

48. Chung, W. *et al.* O35: Feasibility of expanded newborn screening using genome sequencing for early actionable conditions in a diverse city. *Genetics in Medicine Open* **2**, 101369 (2024).
49. Milko, L. V. *et al.* An Age-Based Framework for Evaluating Genome-Scale Sequencing Results in Newborn Screening. *J. Pediatr.* **209**, 68–76 (2019).
50. Berg, J. S. *et al.* A semiquantitative metric for evaluating clinical actionability of incidental or secondary findings from genome-scale sequencing. *Genet. Med.* **18**, 467–475 (2016).
51. Rehm, H. L. *et al.* ClinGen--the Clinical Genome Resource. *N. Engl. J. Med.* **372**, 2235–2242 (2015).
52. Kwon, C. & Farrell, P. M. The magnitude and challenge of false-positive newborn screening test results. *Arch. Pediatr. Adolesc. Med.* **154**, 714–718 (2000).
53. Gold, N. B. *et al.* Low frequency of treatable pediatric disease alleles in gnomAD: An opportunity for future genomic screening of newborns. *HGG Adv* **3**, 100059 (2022).
54. Gold, J. I. *et al.* Phenotypes of undiagnosed adults with actionable OTC and GLA variants. *HGG Adv* **4**, 100226 (2023).
55. Wojcik, M. H. *et al.* Discordant results between conventional newborn screening and genomic sequencing in the BabySeq Project. *Genet. Med.* **23**, 1372–1375 (2021).
56. Adhikari, A. N. *et al.* The role of exome sequencing in newborn screening for inborn errors of metabolism. *Nat. Med.* **26**, 1392–1397 (2020).
57. Cook, S. *et al.* Molecular testing in newborn screening: VUS burden among true positives and secondary reproductive limitations via expanded carrier screening panels. *Genet. Med.* **26**, 101055 (2023).
58. Forsyth, R. *et al.* Outcomes of cases with 3-methylcrotonyl-CoA carboxylase (3-MCC) deficiency - Report from the Inborn Errors of Metabolism Information System. *Mol. Genet. Metab.* **118**, 15–20 (2016).
59. Johnsen, J. M. *et al.* Results of genetic analysis of 11 341 participants enrolled in the My Life, Our Future hemophilia genotyping initiative in the United States. *J. Thromb.*

Haemost. **20**, 2022–2034 (2022).

60. Singh, S., Ojodu, J., Kemper, A. R., Lam, W. K. K. & Grosse, S. D. Implementation of Newborn Screening for Conditions in the United States First Recommended during 2010-2018. *Screening* **9**, (2023).
61. Roberts, J. L. *et al.* CD45-deficient severe combined immunodeficiency caused by uniparental disomy. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 10456–10461 (2012).
62. Conway, J. R., Lex, A. & Gehlenborg, N. UpSetR: an R package for the visualization of intersecting sets and their properties. *Bioinformatics* **33**, 2938–2940 (2017).
63. Seal, R. L. *et al.* Genenames.org: the HGNC resources in 2023. *Nucleic Acids Res.* **51**, D1003–D1009 (2023).
64. Hamosh, A., Scott, A. F., Amberger, J., Valle, D. & McKusick, V. A. Online Mendelian Inheritance in Man (OMIM). *Hum. Mutat.* **15**, 57–61 (2000).

Acknowledgements

This work was supported by the following grants: T32GM007748 (S.B.), R01HG011773 (N.G.), K08HG012811-01 (N.B.G.), TR003201 (N.B.G.), HD077671 (R.C.G.) TR003201 (R.C.G.), and EU-IMI H2020 GRANT 101034427 (A.F., J.K.).

Declaration of interests

A.J.C. and R.J.T. are employees and shareholders at Illumina Inc. N.G. is co-founder and equity owner of Datavisyn. N.B.G. provides occasional consulting services to RCG Consulting and receives honoraria from Ambry Genetics. R.C.G. has received compensation for advising the following companies: Allelica, Atria, Fabric, Genomic Life and Juniper Genomics; and is co-founder of Genome Medical and Nurture Genomics. B.E.R. and K.L.S. are consultants at Nurture Genomics. L.S. received personal compensation from Zentech and Illumina Inc. P.T. is a co-founder of PlumCare RWE, LLC.

Author contributions

Conceptualization: L.M.A., D.B., F.B., A.J.C., N.E., A.F., N.B.G., R.C.G., J.K., T.M., B.E.R., K.L.S., L.S., R.J.T. Data curation: S.B., A.J.C., T.M., K.L.S, H.Z. Formal analysis: N.B.G., T.M. Funding acquisition: R.C.G, L.S., P.T. Investigation: S.A., S.B., N.B.G., T.M. Methodology: N.G., N.B.G., R.C.G., T.M. Resources: H.Z. Project administration: T.M. Software: T.M. Supervision: N.B.G., R.C.G., L.S. Visualization: N.G., N.B.G., T.M. Writing-original draft: S.A., N.B.G., T.M. Writing-review & editing: all authors.

Code and data availability statement

At the time of publication, all datasets generated and/or analyzed in this study, along with the Stata, R, and Python code necessary to replicate the results, will be made publicly available in a Github repository.

Web resources

BabyDetect, https://babydetect.com/en/
BabyScreen+, https://babyscreen.mcri.edu.au/
BabySeq, https://www.genomes2people.org/research/babyseq/
BeginNGS, https://radygenomics.org/begin-ngs-newborn-sequencing/
EarlyCheck, https://earlycheck.org/
FirstSteps, https://www.firststeps-ngs.gr/
FORESITE 360, https://foresite360.com/
Fulgent Genetics, https://www.fulgentgenetics.com/
Genomics England Newborn Genomes Programme, https://www.genomicsengland.co.uk/initiatives/newborns
GUARDIAN Study, https://guardian-study.org/
International Consortium on Newborn Sequencing (ICoNS), https://www.iconseq.org/
Igenomix, https://www.igenomix.eu/
Kenya Bioinformatics Institute, https://www.kibs.co.ke/
Mendelics, https://mendelics.com.br/
NEW_LIVES, https://www.klinikum.uni-heidelberg.de/en/new-lives-genomic-newborn-screening-programs
NewbornsInSA, https://www.wch.sa.gov.au/research/other-research-projects/newbornsinsa-research-study
Nurture Genomics, https://nurturegenomics.com/
Screen4Care, https://screen4care.eu/
ScreenPlus, https://www.einsteinmed.edu/research/screenplus/
PerkinElmer/Revvity, https://www.revvity.com/be-en/category/newborn-screening

Supplement A: ICoNS Gene List Contributors

Programs:

BabyDetect, BabyScreen+, BabySeq, BeginNGS, Chen et al. 2023, Early Check, FirstSteps, the Generation study, gnSTAR, GUARDIAN study, Jian et al. 2022, Lee et al. 2019, Luo et al. 2020, NeoExome, NeoSeq, NESTS, NewbornsInSA, Screen4Care, Progetto Genoma Puglia, Wang et al. 2023, FORESITE 360, Fulgent, Igenomix, Nurture Genomics, PerkinElmer, Sema4.

Individuals:

First name and initials	**Last name**	**Degrees**	**Affiliation**	**Program**
Mattia	Gentile			Progetto Genoma Puglia
Paola	Orsini			Progetto Genoma Puglia
Romina	Ficarella			Progetto Genoma Puglia
Maria Luisa	Valente			Progetto Genoma Puglia
Emanuela	Ponzi			Progetto Genoma Puglia
Athina	Ververi	MD, PhD, MSc. PgCert	Papageorgiou General Hospital, Thessaloniki, Greece	FirstSteps
Maria	Koutsogianni	MD, MSc.	Institute of Child Health Great Ormond Street University College London.	FirstSteps
Huang	Xinwen			gnSTAR
Xiao	Rui			gnSTAR
Zhao	Zhengyan			gnSTAR
Matthew J.	Pelo			FORESITE 360 by Fore Genomics

Jovanka	King	PhD FRACP FRCPA BMBS (Hons) BPod	Genetics and Molecular Pathology, SA Pathology, Adelaide SA 5000, Australia Adelaide Medical School, The University of Adelaide, Adelaide SA 5000, Australia The Women's and Children's Health Network, Adelaide SA 5000, Australia	NewbornsInSA
Carol	Siu	MBBS, PhD, FRCPA, FRCPath	Genetics and Molecular Pathology, SA Pathology, Adelaide SA 5000, Australia Adelaide Medical School, The University of Adelaide, Adelaide SA 5000, Australia	NewbornsInSA
Karin	Kassahn	PhD, FFSc(RCPA)	Genetics and Molecular Pathology, SA Pathology, Adelaide SA 5000, Australia Adelaide Medical School, The University of Adelaide, Adelaide SA 5000, Australia	NewbornsInSA
Sansen	Stefaan		Sanofi	Screen4Care
Enrico	Bertini		Paediatric Hospital Bambino Gesù Rome	Screen4Care
Aldona	Zygmunt		Pfizer	Screen4Care

Supplement B: International Consortium on Newborn Sequencing (ICoNS) authors

Sophia Adelson
Emanuele Agolini
Aljazi Al-Maraghi
Karla Alex
Fowzan Alkuraya
Ammira Alshabeeb Akil
Munira Alshehri
Derek Ansel
Niki Armstrong
Matthew Aujla
Don Bailey
Mei Baker
Jorune Balciuniene
Andrew Barry
Bruce Bennetts
Melissa Berenger
Jonathan Berg
Donna Bernstein
Arindam Bhatatcharjee
Sucheta Bhatt
David Bick
Tracey Bishop
Asaf Bitton
François Boemer
Natasha Bonhomme
George Bowley
Brenna Boyd
Heiko Brennenstuhl
Steven Brenner
Mairead Bresnahan
Thomas Brewster
PJ Brooks
Katya Broomberg
Amy Brower
Gemma Brown
James Buchanan
Caleb Bupp
Candance Cameron
Mattia Gentile
Jessica Giordano
Ulrich Glumer Jensen
David Godler
Nina Gold
Aaron Goldenberg
Katie GoldenGrant
Cassie Goldman
Chema González de Aledo
Daniel Gottlieb
Robert Green
Christopher Greene
Brooke Greenstein
Scott Grosse
Annette Grueters
Gulcin Gumus
Kelly Hagman
Kevin Hall
Aymeric Harmant
Sally Hartmanis
Robin Hayeems
Rose Heald
Madhuri Hegde
Rebecca Heiner-Fokkema
Lidewij Henneman
Becca Hernan
Charlotte Hobbs
Ingrid Holm
Layla Horwitz
Zhanzhi Hu
Maria Iascone
Ken Irvine
Guanjun Jin
Kelsey Kalbfleisch
Ines Kander
Lucy Kaplun
Dalia Kasperaviciute
Karin Kassahn
Mette Nyegaard
Justin O'Sullivan
Jelili Ojodu
Paola Orsini
Andrea Oza
Katrina Paleologos
Richard Parad
Holly Peay
Matthew Pelo
Carolyn Philstrom
Dominique Pichard
Amanda Pichini
Holly Pickering
Michelle Pirreca
Malgorzata Ponikowska
Amy Ponte
Andreas Posch
Cynthia Powell
Liana Protopsaltis
Yeyson Quevedo
Marianna Raia
Rebecca Reimers
Andy Rohrwasser
Paul Rollier
Lene Rottensten
Irakli Rtskhiladze
Nabihah Sachedina
George Sahyoun
Aditi Satija
Christian Schaaf
Jennifer Schleit
Richard Scott
Lauren Scully
Stacey Seeloff
Laurent Servais
Nidhi Shah
Maija Siitonen
Sikha Singh

Lauren Capacchione
Diana Carli
Onassis Castillo Ceballo
Kee Chan
Jillian Chance
Georgia Charalambidou
Winnie Chen
Yun-Ru Chen
Wendy Chung
Brian Chung
Megan Clarke
Susan Clasper
F. Sessions Cole
Heidi Cope
Stephanie Coury
Tony Cox
Tamara Dangouloff
Earnest James Paul Daniel
Katrin Eivindardottir Danielsen
Emeline Davoine
Tom Defay
Geethanjali Devadoss Gandhi
Joseph Dewulf
Lisa Diller
Pakhi Dixit
Martijn Dolle
Lilian Downie
Erin Drake
Suzanne Drury
Annelotte Duintjer
Bugrahan Duz
David Eckstein
Matthew Ellinwood
Katarzyna Ellsworth
Sarah Elsea
Nicolas Encina
Harriet Etheredge
Laurence Faivre
Alessandra Ferlini

Leni Kauko
Riina Kaukonen
Nicole Kelly
Dhayo Khangsar
Jovanka King
Clare Kingsley
Stephen Kingsmore
Brian Kirmse
Rachel Klein
Stefan Koelker
Youssef Kousa
Elizaveta Krupoderova
Paul Kruszka
Katherine Langley
Ciara Leckie
Emmanuelle Lecommandeur
David Ledbetter
Pamela Lee
Beomhee Lee
Camille Level
Celine Lewis
Anna Lewis
Ruby Liu
Mauro Longoni
Alberte Lundquist
Sebastian Lunke
Kate MacDuffie
Ankit Malhotra
Lionel Marcelis
Maria Martinez-Fresno
Gert Matthijs
Roberts Melbardis
Jessica Merritt
Radja Messai Badji
Lou Metherell
Nanna Balle Mikkelsen
Laura Milko
Nicole Miller
Thomas Minten

Carol Siu
Hadley Smith
Lisa Sniderman King
Neal Sondheimer
Lourdes St George
Zornitza Stark
Robert Steiner
Ulrik Stoltze
Asbjørg Stray-Pedersen
Kristen Sund
Paris Tafas
Polakit Teekakirikul
Dimitrios Thanos
Audrey Thurm
Meekai To
Petros Tsipouras
Alice Tuff-Lacey
Heather Turner
Philip Twiss
Fiona Ulph
Daniel Uribe
Tiina Urv
Cora Vacher
Kris Van Den Bogaert
Mirjam van der Burg
Eva Van Steijvoort
Yiota Veloudi
Elizabeth Vengoechea
Els Voorhoeve
Martin Vu
Melissa Wasserstein
Michael Watson
Bryn Webb
Anna Wedell
Thomas Westover
Alexandra Wiedemann
Meredith Wright
Cindy Wu
Julie Yeo

Monica Ferrie
Terri Finkel
Petra Furu
Jamie Galarza-Cornejo
Ya Gao
Judit Garcia-Villoria
Liz Gardner
Amy Gaviglio
Michael Gelb
Sian Morgan
Katarzyna Mosiewicz
Ulrike Mütze
Sukhvinder Nicklen
Minna Niemela
Dau-Ming Niu
Sarah Norris
Antonio Novelli
Arwa Nusair
Nancy Yin-Hsiu Chien
Shamila Yusuff
Tomasz Zemojtel
Bethany Zettler
Zhengyan Zhao
Joanna Ziff
Rebekah Zimmerman
Michela Zuccolo