Received: 12 February 2022 | Revised: 27 April 2022 | Accepted: 28 April 2022

DOI: 10.1002/jimd.12508

REVIEW

JIMD SSIEM WILEY

How longitudinal observational studies can guide screening strategy for rare diseases

Ulrike Mütze | **Katharina Mengler** | **Nikolas Boy** | **Florian Gleich** | **Thomas Opladen** | **Sven F. Garbade** | **Stefan Kölker**

Division of Child Neurology and Metabolic Medicine, Center for Pediatric and Adolescent Medicine, Heidelberg University Hospital, Heidelberg, Germany

Correspondence
Ulrike Mütze, Division of Child Neurology and Metabolic Medicine, Center for Child and Adolescent Medicine, University Hospital Heidelberg, Im Neuenheimer Feld 430, 69120 Heidelberg, Germany.
Email: ulrike.muetze@med.uni-heidelberg.de

Funding information
Dietmar Hopp Stiftung, Grant/Award Numbers: 2311221, 1DH2011117; Medizinischen Fakultät Heidelberg, Universität Heidelberg, Grant/Award Number: F.206852

Abstract

Newborn screening (NBS) is an important secondary prevention program, aiming to shift the paradigm of medicine to the pre-clinical stage of a disease. Starting more than 50 years ago, technical advances, such as tandem mass spectrometry (MS/MS), paved the way to a continuous extension of NBS programs. However, formal evidence of the long-term clinical benefits in large cohorts and cost-effectiveness of extended NBS programs is still scarce. Although published studies confirmed important benefits of NBS programs, it also unraveled a significant number of limitations. These include an incompletely understood natural history and phenotypic diversity of some screened diseases, unreliable early and precise prediction of individual disease severity, uncertainty about case definition, risk stratification, and indication to treat, resulting in a diagnostic and treatment dilemma in individuals with ambiguous screening and confirmatory test results. Interoperable patient registries are multi-purpose tools that could help to close the current knowledge gaps and to inform further optimization of NBS strategy. Standing at the edge of introducing high throughput genetic technologies to NBS programs with the opportunity to massively extend NBS programs and with the risk of aggravating current limitations of NBS programs, it seems overdue to include mandatory long-term follow-up of NBS cohorts into the list of screening principles and to build an international collaborative framework that enables data collection and exchange in a protected environment, integrating the perspectives of patients, families, and the society.

KEYWORDS

case definition, health benefit, long-term observation, newborn screening, patient registry, rare disease

1 | NEWBORN SCREENING: SHIFTING THE TRADITIONAL PARADIGM OF MEDICINE

The traditional concept of medicine is based on the clinical phenotype. The major advantage of this concept lies in the certainty that the symptomatic individual (hence termed patient) is affected by a disease. The major disadvantages, particularly for rare genetic diseases, can be seen in the often time-consuming path to diagnosis, late introduction of specific treatment, and thus the limited ability to prevent irreversible harm. Diagnostic biomarkers,

which allow to identify affected individuals already during a pre-clinical stage, have become the prerequisite of screening programs, shifting the traditional paradigm of medicine. Especially newborn screening (NBS) programs have the potential to identify individuals with rare treatable conditions at an early stage, to prevent a diagnostic odyssey, to shorten the time to introduction of disease-changing therapies, and to improve health, development, and life expectancy. Therefore, they are considered an important measure of secondary disease prevention and as one of the greatest advances of modern public health with significant individual, socio-economic, and societal benefits.[1] NBS programs were initiated more than 50 years ago, starting with phenylketonuria (PKU) as the first target disease subsequent to the development of a semi-quantitative bacterial inhibition assay[2] and a disease-changing dietary treatment.[3] The success of PKU screening and subsequent technological advances, such as tandem mass spectrometry (MS/MS)[4–7] and genetic screening tests,[8] enabled the extension of NBS disease panels to a growing spectrum of inherited metabolic diseases (IMDs), endocrine, immunologic, hematological and neurological diseases, and other complex disorders such as cystic fibrosis.

Although existing NBS programs still claim to refer to a set of 10 principles for population screening drafted by Wilson and Jungner in 1968[9] and international efforts have been made to develop NBS programs toward a harmonized panel and system,[10–12] the interpretation of the original screening principles has remained controversial at the level of national policy, resulting in a highly variable composition of national NBS disease panels.[13–15] As a consequence, a revision and further extension of the original screening principles as well as the introduction of transparent and objective decision tools for the selection of candidate diseases have been proposed.[16,17] Standing at the edge of introducing large-scale sequencing technologies for the screening of genetic conditions this revision is overdue.[18,19]

2 | CHEMICAL INDIVIDUALITY AND PHENOTYPIC DIVERSITY

New diagnostic opportunities are accompanied by new challenges. An intrinsic challenge of all NBS programs is to identify asymptomatic individuals at risk of a specific disease using biomarkers and to clearly distinguish "physiologic" from "pathologic" conditions. In an ideal world, these biomarkers would be accurate enough for clinicians to provide clinical judgment, particularly since early introduction of treatment aims at preventing or at least delaying or attenuating the disease manifestation. In the real world, there is a fluid transition between unaffected individuals and those with an attenuated phenotype[20] as well as a weak correlation between the concentration of some biomarkers and the assumed disease severity.[21] With the introduction of the concept of chemical individuality in 1902, Sir Archibald Garrod already highlighted this diagnostic challenge long before NBS programs were conceptualized and implemented, recognizing that the individual variation of detectable metabolites reflects the evolutionary need for adaptive variability but does not necessarily indicate a predisposition to disease.[22] Today, the concept of chemical individuality has been significantly extended with the increasing understanding of the (human) metabolism by integration of gene–environment and gene–nutrient interactions, involvement of other metabolic pathways and modifier genes,[23–25] side reactions of enzymes, and the metabolic proof machinery.[26,27] In addition, holo-genomic metabolic cooperation between humans and bacteria[28,29] unravels humans to be metabolic holobionts and shifts IMDs from monogenic to complex diseases.[30] These complex mechanisms not only explain biochemical individuality but also underlie clinical diversity. Since the relationship between genotype, metabolite profile and clinical phenotype might be complex, early phenotypic prediction has remained imprecise for many diseases,[21,31,32] while it is more successful for others.[33–35]

As a consequence, the success of NBS programs critically depends on the precise knowledge about chemical individuality, gene variations, and phenotypic diversity of screened diseases. For instance, individuals with phenylalanine hydroxylase-induced mild hyperphenylalaninemia do not require treatment, unlike those classified to have PKU.[36,37]

To increase methodological accuracy and reduce diagnostic uncertainty of NBS test results significant improvements have been made, such as the identification of objective target ranges for more than 100 biomarkers to be applied for NBS,[38] the application of feature construction methods to disclose information hidden in the whole set of measured values,[39] and the development of biochemical and genetic two- or multiple-tier strategies.[40] Despite these efforts, the vast MS/MS-based extension of NBS programs has revealed significant knowledge gaps, such as (1) the incomplete understanding of the natural history and phenotypic variations of some screened diseases,[41-43] (2) the unreliable early prediction of the individual disease severity,[44] (3) the uncertainty about exact case definition,[43,45] (4) the ambiguity concerning individual risk stratification and indication to treat,[46,47] and (5) the lack of clarity of long-term benefits of extended NBS programs concerning individual health, health economics, and the society.[42,48] Positively speaking, NBS can serve as a major stimulus for rapidly expanding knowledge, compared to the pre-MS/MS period.[49]

To overcome the shortcomings and to continue the success story of NBS systematic long-term observation of NBS cohorts and a careful evaluation of benefits and limitations of NBS programs under real-world conditions are important measures. However, despite this obvious need, and because of their high costs, longitudinal observational studies have remained the neglected part of NBS programs. Formal evidence of the clinical effectiveness and long-term benefits of MS/MS-based NBS in large cohorts with longer follow-up,[42,43,50–52] cross-sectional analysis of patient registries managed by transnational scientific consortia,[53–59] or from meta-analysis[60] is still scarce. Current knowledge is mostly based on short-term follow-up of small- or medium-sized regional cohorts.[51,61–69] In the following, we will discuss major lessons learned from longitudinal observational studies and how they can guide NBS strategy.

3 | LESSONS LEARNED FROM LONG-TERM OBSERVATION OF NEWBORN SCREENING COHORTS

3.1 | The phenotype follows a continuous spectrum

More than 20 years after the start of MS/MS-based NBS programs, a clear-cut case definition is still missing for several screened diseases and disease variants, exposing these individuals and their families to the potential health risk of over- and under-treatment, prognostic uncertainty, stigmatization, and accompanying distress and disruption of family life.

Assuming that the severity of the clinical phenotype of an individual with a genetic disease, such as an IMD, ranges within a continuous spectrum,[20] any dissection of this continuum into "severe," "moderate/intermediate," and "mild/attenuated" phenotypes, or "early/neonatal onset" and "late onset" must be artificial. Despite its intrinsic imprecision, this concept is useful for risk stratification and clinical decision-making and, therefore, is often utilized.[43,53,57,70,71] This concept however, contains pitfalls and the potential for misunderstandings. For instance, the term "mild" may indicate an individual with a biochemical variation without indication to treat, for example, mild hyperphenylalaninemia, or an individual who may succumb to death during a metabolic decompensation, for example, late-onset ornithine transcarbamylase deficiency.[72,73] However, sometimes "mild" simply indicates uncertainty about the clinical significance of a condition. For mild isovaleric aciduria (IVA), which has remained unknown in the pre-NBS era, there is ongoing debate about its clinical significance since the first description of a genetic variant.[41] A recent national long-term outcome study not only showed this disease variant to be most prevalently found, but also demonstrated a normal neurocognitive development without occurrence of metabolic decompensations.[43] Furthermore, it is still difficult to unambiguously distinguish between "mild" and "classic" IVA cases based on biochemical and genetic test results.[43] In analogy to IVA, an increased frequency of attenuated phenotypes has also been found in long-chain acyl-CoA dehydrogenase deficiency and other IMDs since their inclusion into NBS programs.[74,75]

The shift from severe to attenuated phenotypes in NBS cohorts also entails the risk of over-estimating the health impact of NBS programs since the case mixes of NBS and pre-NBS cohorts do not necessarily match. A recent collaborative study has highlighted this problem and has demonstrated a feasible strategy for severity-adjusted evaluation of NBS programs.[76] Utilizing a previously established functional disease prediction model for citrullinemia type 1 and argininosuccinic aciduria,[34,35] which integrates longitudinal follow-up data and patient-related in vitro residual enzyme activities, the authors showed the disease severity of the NBS cohort to be lower than of the pre-NBS cohort. Regardless of this difference, NBS and early start of treatment reduced the initial peak plasma ammonium concentration before start of treatment, particularly in individuals with an attenuated phenotype, but did not reduce the frequency of subsequent hyperammonemic episodes at least with a conservative treatment.[76] Utilizing data from multiple sources, this study highlights the need for accurate and severity-adjusted case definitions and the importance of well-characterized longitudinally followed patient cohorts. In the future, multi-omics data and artificial intelligence-supported diagnostic pathways might help to overcome the current limitations of case definition and risk stratification.[77]

3.2 | To treat or not to treat? And how?

The aforementioned uncertainty about case definition results in a treatment dilemma. Both the decision to treat or not to treat might harm the individual, either because of adverse effects of unnecessary treatment or the missed opportunity to prevent irreversible disease manifestation in untreated individuals (Figure 1). To aggravate this problem, randomized controlled trials are challenging in rare disease settings and thus are still scarce.[78] Furthermore, clinical outcomes and outcome measurement instruments often vary greatly in the conducted rare disease trials.[79] This hampers the comparison of results and

	Individual with significant health risk	Individual without health risk
Treatment	Individual health: High risk of irreversible harm mitigated or prevented through NBS and early intervention Health economics: Potentially high cost effectiveness (depending on costs and efficacy of therapy) Society: Chance of independent full participation, employment, and tax payments → **Best case 1**	Individual health: No health benefit, but moderate risk of harm through adverse effects of non-indicated therapy and negative psychological impact on the individual and the family Health economics: Low cost effectiveness, no cost savings despite NBS Society: Fair chance of independent full participation, employment, and tax payments (depending on actual harm of non-indicated therapy) → **Second-worst case**
No treatment	Individual health: High risk of irreversible physical and/or cognitive disability Health economics: Very low cost effectiveness, high health costs (like in the pre-NBS era) despite NBS Society: High risk of continuous support and socio-economic burden for family and society → **Worst case**	Individual health: No health benefit, but also low risk of harm (except for positive screening) Health economics: Good cost effectiveness (depending on the acutal costs of the NBS program and cost savings for the national health service) Society: Independent full participation, employment, and tax payments (like the general population) → **Best case 2**

FIGURE 1 Uncertainty about case definition results in a treatment dilemma. Ideally, NBS clearly identifies individuals with increased health risks and distinguishes them from those without. A clear-cut case definition guides the decision to treat (in individuals at risk, best case 1) or not to treat (in individuals not at risk, best case 2). Without a clear-cut case definition, healthcare providers are likely to start treatment to prevent irreversible physical and/or cognitive disability in individuals at risk (worst case); however, this is done at the cost of potentially harming individuals with benign variants through adverse physical effects and psychological burden of non-indicated therapy (second-worst case).

the achievement of robust evidence and underlines the need for the development of core outcome sets as recently published for PKU and medium-chain acyl-CoA dehydrogenase deficiency (MCADD).[80,81]

Longitudinal follow-up of NBS cohorts with careful evaluation of prescribed treatments can fill this important gap, preferentially if combined with systematic evaluation of literature and guideline development. Thus allowing the evaluation of feasibility and health impact of recommended therapy under real-world conditions and gradually improving the evidence level and feasibility of recommendations in an iterative way. A successful example of this approach is glutaric aciduria type 1 (GA1). It is prognostically most important disease manifestation being infantile-onset dystonia due to acute- or insidious-onset striatal damage.[21,82] When first NBS pilot studies and national NBS programs included GA1 in the late 1990s, it was still unknown whether the natural history of this disease could be improved by available therapies. In fact, a meta-analysis on studies reporting 115 symptomatically diagnosed patients concluded that "treatment given after the appearance of symptoms was not associated with a better clinical outcome."[83] Although it was hoped that pre-symptomatic start of treatment might prevent the onset of symptoms, it was not before MS/MS-based NBS started that this could be proven.[21,82] This has been the starting point for international guideline development, evaluating available evidence and prioritizing the multitude of knowledge gaps,[84] and for a set of long-term observational studies whose major results were used for subsequent guideline revisions.[85,86] By this iterative approach transient emergency treatment during putatively threatening catabolic episodes was shown to be the most effective measure to avoid (acute-onset) dystonia[87–89] and carnitine supplementation to reduce mortality, while no evidence supported a positive effect of riboflavin supplementation.[21] Furthermore, untreated individuals with the high and low excreter phenotype of GA1 were both shown to have a high risk to develop irreversible striatal damage,[21] excluding the misconception of low excreter patients to have an attenuated phenotype. The most difficult part was to find solid evidence for the impact of dietary treatment on the outcome. It took several years to demonstrate superiority of low lysine diet with lysine-free, arginine- and tryptophan-enriched amino acid supplements over low protein diet.[87,90] Almost 20 years later, there is now solid evidence that GA1 is a treatable NBS condition. Recommended treatment, if introduced and monitored by an experienced multi-professional team, is safe and the major prerequisite of good neurological outcome, while deviation from recommended treatment increases the risk of striatal damage,[50,87,90,91] highlighting

the importance of the quality of therapy.[60] Some lately recognized aspects of the disease, however, such as chronic kidney disease,[50] progressive white matter changes,[92,93] and slightly reduced IQ in high excreters,[91] do not seem to be impacted by currently available therapies. Besides GA1, this iterative approach of longitudinal observational studies conducted by international scientific consortia and concomitant guideline development is also successfully applied to other IMDs, such as methylmalonic and propionic aciduria,[53,94–96] urea cycle disorders,[57,58,97] cystathionine beta-synthase deficiency,[98] remethylation disorders,[54,99,100] and tetrahydrobiopterin deficiencies.[56,101]

The GA1 example also highlights another important aspect for guiding future extension strategy for NBS programs. Without inclusion of GA1 to NBS pilot studies or NBS programs before having ample evidence that it is a treatable condition the above-described success story would have been inconceivable if not impossible. This stresses the need for conducting pilot studies on candidate diseases with a high potential to be included in NBS programs. One of these candidates is neonatal vitamin B_{12} deficiency due to undiagnosed and hence untreated maternal vitamin B_{12} deficiency. Pilot NBS studies on neonatal vitamin B_{12} deficiency confirmed the feasibility of a combination of previously described second-tier strategies[40,102–104] as well as the positive impact of early treatment on the neurological outcome.[45] The opportunity to prevent irreversible harm in the affected neonate, its mother, and potentially younger siblings let neonatal vitamin B_{12} deficiency appear as very promising NBS candidate. But again, clear-cut case definition appears to be a weak point. Although it is unlikely that all screened individuals with neonatal vitamin B_{12} deficiency would develop neurologic symptoms during infancy without treatment, there is currently no known evidence-based stratification that would allow to distinguish individuals who would benefit from transient treatment from those who would not. Since this distinction is influenced by environmental factors, observational studies unlike in genetically defined IMDs might not be analogously helpful for this disease to improve the case definition and indication to treat unless additional data sources, such as nutrition and the microbiome, are integrated into analysis.

3.3 | Time is health: Every day counts

Extended NBS programs include a growing number of IMDs with a risk of neonatal metabolic decompensation pointing on the need for a timely NBS. A recent study has demonstrated that 28 of 191 (14.7%) screened individuals at risk, did actually experience a neonatal metabolic decompensation (median age, 4 days) before the NBS result was known. Fortunately, experiencing a neonatal metabolic decompensation did not necessarily predict a poor long-term health outcome.[42] Noteworthy, none of the neonatal decompensations occurred after the report of a positive NBS result. This highlights the need for excellent diagnostic process quality of extended NBS programs. To achieve this goal, many countries have already shortened the recommended time to NBS sampling. However, improving diagnostic process quality requires the careful evaluation of the NBS program as a complex multi-step system, including the performance of the senders (e.g., hospitals), the carriers (mail services), and the recipients (NBS laboratories) of the NBS sample (Figure 2). A recent evaluation of the NBS process quality in Southwest Germany unraveled improved performance of hospitals and NBS laboratories but increasing shipping intervals (i.e. >48 h). Theoretically, about 25% of acute metabolic decompensations before the first NBS reports could have been prevented by a first NBS report on day 5, as required by the German NBS directive.[42] However, this improvement would still not allow to prevent some fatal neonatal decompensations, such as in MCADD.[105] Because of the above-discussed limitations, there is still uncertainty about the inclusion of other intoxication type IMDs with a supposedly high frequency of individuals with neonatal onset, such as methylmalonic and propionic aciduria, and urea cycle disorders.[106] Analysis of an international cohort of intoxication type IMDs unraveled significant disease-specific variations in the proportion of individuals with disease onset during (EO group) and after (LO group) the neonatal period and the proportion of individuals who could have been identified before the onset of first symptoms.[53,57] In the group of organic acidurias, NBS was shown to clearly reduce the time to diagnosis for GA1 and IVA, while NBS for methylmalonic and propionic aciduria shortened the diagnostic pathway only for the LO group of patients. It was estimated that 78% (disease-specific range, 62%–98%) of individuals with organic acidurias and 70% (disease-specific range, 33%–100%) of those with urea cycle disorders could have been identified pre-symptomatically by NBS within the first week of life.[53] Even if we assume that this cohort, like that of other patient registries, under-represents individuals with a severe phenotype to some extent,[70] these results suggest that an extension of NBS programs to these disease groups might be feasible and could result in a health benefit, particularly for individuals of the LO group.[53,57,58,76,107]

3.4 | Limitations and benefits of NBS: Integration of different perspectives

From the bird's-eye view, screened individuals with IMDs have excellent health outcomes, confirmed by a high

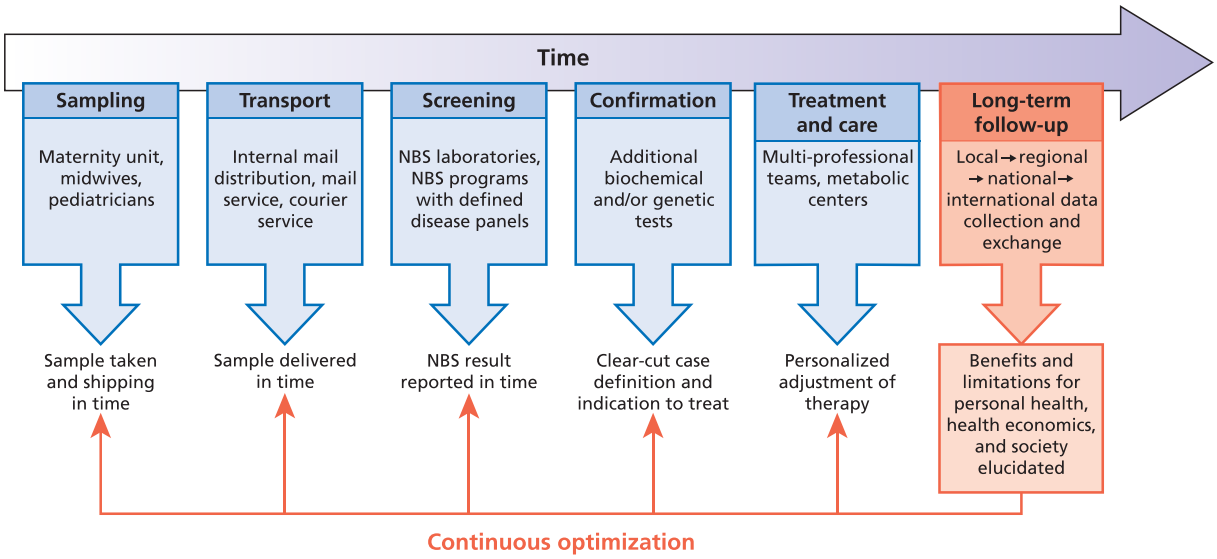

FIGURE 2 The newborn screening (NBS) program is a multi-step system. Every chain is as strong as its weakest link. NBS integrates various successive steps conducted by different professions in a coordinated way. Concerted action is required to minimize the risk of delay, confusion of sample, and incorrect test results. Combined evaluation of diagnostic process quality and long-term clinical outcomes of screened individuals is a prerequisite for continuous optimization of this complex process.

frequency of normal development and normal cognitive outcome (95%), and a high proportion of screened individuals without permanent disease-specific symptoms until last visit (76%).[42] If we dig deeper, we soon identify disease-specific variations and limitations that highlight the need for further optimization on a broad range of aspects. Individuals with maple syrup urine disease, GA1, IVA, and other IMDs might not benefit from NBS in the same way as those with PKU and biotinidase deficiency since some of them might have been missed by NBS, might have developed symptoms before the NBS test results or since available treatments are not effective to prevent the progression of the disease[42] or are invasive such as liver transplantation.[96] Additionally, we could easily over-estimate the health benefits of NBS programs if case mixes and outcomes of NBS and pre-NBS cohorts were not compared within the same country and national health service[10,52,69,76,88,90,108–111] and if analyses were made without consideration of the predicted clinical phenotype of screened individuals.[42,58,76]

Furthermore, the perspectives of patients and their families have to be included into the analysis. Healthcare professionals may still under-estimate that NBS and early treatment, although leading to favorable physical and cognitive outcome, may put significant stress on patients and families because of the life-long risk of decompensations and the need to adhere to burdensome therapies,[112] not to mention the negative psychological effects of false-positive NBS results on family life and parent–child interaction.[113,114] Therefore, parent- and patient-reported outcome and experience measures should be integrated to understand patient and family expectations and current limitations of NBS programs.

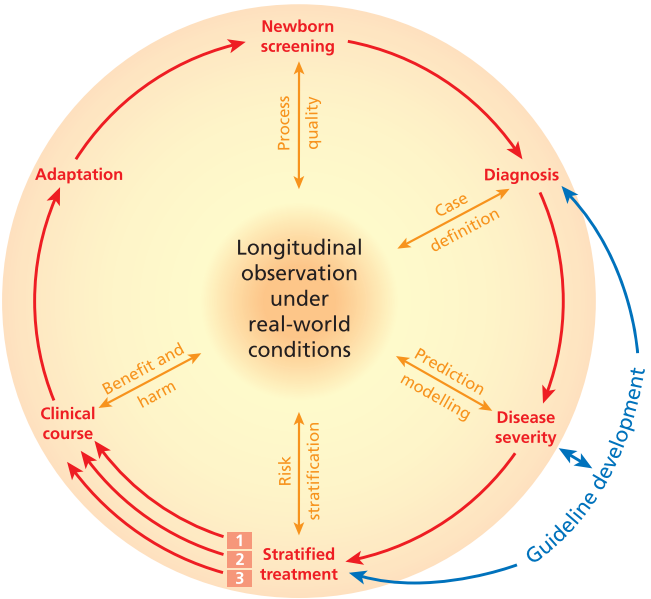

FIGURE 3 Iterative cycle of optimizing newborn screening (NBS) programs through observational studies. Real-world data obtained from longitudinal observational studies provide robust information about current limitations and benefits of the single steps of NBS programs and the programs as a whole, as well as the feasibility, safety, and efficacy of therapy and care. These data inform screening strategy and guideline development in an iterative way, enabling optimization through continuous evaluation and adaptation.

TABLE 1 Summary of the (possible) impact of integrated long-term observational studies on newborn screening (NBS) programs as an 11th screening principle

Goal	Data from long-term observation	Impact on NBS programs
Clear-cut case definition	Prediction of the clinical case in the presymptomatic state by integration of biochemical, enzymatic genetic, natural history, and long-term observational data of NBS cohorts	• Early identification and therapy of true positive patients with high risk for a severe disease course • Reducing the harm on individuals and families by reducing the number of false positives and/or benign disease courses • Development of diagnostic guidelines
Overcome the treatment dilemma (Figure 1)	Long-term observational data including treatment data of NBS cohort Randomized treatment studies in NBS cohorts	• Stratified treatment of severe and attenuated clinical phenotypes • Reduction of overtreatment in individuals with attenuated or benign phenotypes identified by NBS • Development of stratified treatment guidelines
Optimize NBS process	NBS process data combined with the long-term clinical outcome	• Harmonization and optimization of the age at NBS sample and the NBS as a process in regard to the infrastructural possibilities • Further reduction of neonatal metabolic decompensations in NBS conditions • Evaluation of the impact of pre-NBS report metabolic decompensation on the outcome
Expand NBS programs	Measures on the analytical and structural NBS process of new conditions Accompanying measures of the clinical and cognitive outcome Accompanying measures of the treatment	• Implementation accompanying evaluation of new NBS condition • Reevaluation of implemented NBS conditions • Generating data to eliminate unsuccessful conditions from national NBS panels
Reduce family burden	Combination of the above mentioned	• Reducing harm of false positive results • Improving counseling of confirmed cases by better prediction of the clinical course • Early effective treatment of the severly affected • Reduction of overtreatment in attenuated and benign phenotypes • Clustering the medical care in specialized interprofessional centers
Societal benefit	National NBS costs for the health care system and long-term observational data (clinical and cognitive) including educational and socioeconomic data to calculate the savings for the health care system due to prevention of the severe natural history (severe disability) of the disease. Data on treatment costs Extrapolation of societal savings and profits by an unaffected societal participation of the affected individuals	• Cost–benefit evaluation of the NBS program
Interoperable data analysis	Core data sets for screening, outcome, treatment	• Interoperability of different registries to enable combined evaluation to face the rareness of the included diseases
Interoperable networks	Combined data analysis, FAIR principles	• Combined evaluations especially for extreme-rare screening conditions by cooperation in international scientific consortia

Besides the individual health benefit, NBS programs should be feasible and reasonable for the screened population, their societies, and healthcare systems.[9] Costs for screened clinical courses and establishing and maintaining adequate infrastructures have to be compared to unscreened clinical courses[110,115,116] and correlated to the

further costs or savings for the society. To generate this data, education, employment, parenthood, and possible implications of the disease on pregnancies have to be recorded and evaluated in adult IMD cohorts.[117,118] Furthermore, different treatment options and adherence rates need to be considered in economical evaluations. For instance, a recent study demonstrated that NBS for PKU was cost-effective with low phenylalanine diet even if adherence rates were lower than previously assumed, but not with sapropterin dihydrochloride medication.[119] In analogy, the implementation of highly costly gene therapies and other advanced therapy medicinal products for (future) NBS target diseases will require a careful evaluation of economic consequences from a societal perspective, particularly if other therapies are already available. In conclusion, fair evaluation of NBS programs is a true challenge and can succeed only if (1) NBS is understood as a complex system, (2) meaningful indicators and endpoints are chosen, (3) robust clinical health data from NBS and pre-NBS cohorts are available, and (4) different perspectives are integrated into the analysis.

4 | OUTLOOK: MAXIMIZE BENEFIT AND MINIMIZE HARM THROUGH LONGITUDINAL OBSERVATIONAL STUDIES

It is recognized that "all screening programs do harm; some do good as well, and, of these, some do more good than harm at reasonable cost."[48] It is therefore important to acknowledge this at the outset so that we might act cautiously when planning further extensions of NBS programs and work tirelessly to improve the quality of those that already exist.

Despite their extreme high cost, long-term observational studies using patient registries are multi-purpose tools for rare disease research and guiding screening strategy. Since long-term data for a rare disease might not be assessable or very limited prior to the planned extension of NBS programs, premature decisions on the implementation of new NBS diseases potentially cause a treatment dilemma for individuals identified by NBS (Figure 1). Therefore, a concomitant evaluation of NBS programs by long-term observational studies seems to be indispensable to continuously fill the knowledge gaps and optimize the NBS program in an iterative way (Figure 3), comparable to the indispensable post-authorization safety studies for drugs.[120] As a consequence, we propose to add an 11th screening principle to the original list[9]: "Cohorts of individuals identified by NBS should be systematically followed." A follow-up of screened individuals would help (1) to better understand the natural history and phenotypic diversity of rare diseases, (2) to early and precisely predict individual disease courses, (3) to reduce the uncertainty about case definition, risk stratification, and indication to treat, and (4) to evaluate the individual health benefits as well as the economical and societal benefits of NBS programs (Table 1). This can also be a cost-saver in the long run (Table 1).

The future success and further optimization of NBS programs therefore depend on the establishment and maintenance of an international collaborative framework that enables long-term data collection of screened individuals and data exchange in a protected environment. Particularly paying attention to compliance with ethical and legal requirements, such as the general data protection regulation (GDPR, https://gdpr.eu), and the FAIR data principles, which are indispensable to make collected data findable, accessible, interoperable, and reusable (https://www.go-fair.org/fair-principles/), and integrating the perspectives of patients, families, and the society. The most efficient and feasible way to achieve this goal is the development of a federated network of existing registry infrastructures with suitable design, such as regional and national observational studies of NBS cohorts,[13,42,121] and the official registry of the European Reference Network MetabERN (U-IMD[100];), and to agree on a limited subset of meaningful core outcome sets. If we establish this collaborate NBS framework, we would overcome the current data fragmentation and duplication, uncoordinated parallel activities and research in small sample sizes, and other limitations that still hamper rare disease research and progress. For rare disease research, the world may be not enough.

AUTHOR CONTRIBUTIONS

All authors designed the concept of this review. Ulrike Mütze and Stefan Kölker produced the first draft of the manuscript, while all authors revised it thoroughly. Ulrike Mütze is the corresponding author.

ACKNOWLEDGMENTS

We are indebted to all individuals with inherited metabolic diseases (IMDs) and their families who over the last two decades significantly contributed to various regional, national, and international observational studies and provided immeasurable insights into the current benefits and limitations of newborn screening programs, long-term management, and care under real-world conditions. We are thankful to all colleagues who are committed to improving health and quality of life of patients with IMDs and their families by sharing experience and data, contributing to scientific consortia, particularly E-IMD (https://www.e-imd.org/; https://www.eimd-registry.org/), E-HOD (https://www.e-hod.org/; https://www.ehod-registry.org/

about-ehod), iNTD (https://www.intd-registry.org/), Meta bERN (https://metab.ern-net.eu/), NBS 2020/2025, UCDC (https://www1.rarediseasesnetwork.org/cms/ucdc), and U-IMD (https://u-imd.org/; https://www.u-imd-registry.org/index.php?id=about), and conducting longitudinal observational studies, improving the knowledgebase about individuals with rare diseases, and evaluating the safety and efficacy of diagnostic and therapeutic measures. We thank Bettina Haase and Elena Boyd for carefully revising the manuscript.

Open Access funding enabled and organized by Projekt DEAL.

FUNDING INFORMATION

The long-term observational study ("Long-term outcome of patients with inherited metabolic diseases after diagnosis by expanded newborn screening") is generously supported by the Dietmar Hopp Foundation, St. Leon-Rot, Germany (grant numbers 2311221, 1DH2011117 to SK). UM has been further supported by an Olympia Morata research fellowship of the Medical Faculty of the Ruprecht Karl University of Heidelberg (grant number F.206852). The authors confirm independence from the sponsors; the content of the article has not been influenced by the sponsors.

CONFLICT OF INTEREST

Ulrike Mütze has received research grants from the Medical Faculty of the Ruprecht Karl University of Heidelberg (grant number F.206852). Stefan Kölker has received research grants from Dietmar Hopp Foundation, St. Leon-Rot, Germany (grant numbers 2311221, 1DH2011117). Katharina Mengler, Nikolas Boy, Florian Gleich, Thomas Opladen declare that they have no conflict of interest. The authors confirm independence from the sponsors; the content of the article has not been influenced by the sponsors.

DATA AVAILABILITY STATEMENT

The manuscript has no associated data.

ORCID

Ulrike Mütze https://orcid.org/0000-0002-4251-7257

REFERENCES

1. Moliner AM, Waligora J. The European Union policy in the field of rare diseases. *Rare Diseases Epidemiology: Update and Overview*. Springer; 2017:561-587.
2. Guthrie R, Susi A. A simple phenylalanine method for detecting phenylketonuria in large populations of newborn infants. *Pediatrics*. 1963;32:338-343.
3. Bickel H, Gerrard J, Hickmans EM. Influence of phenylalanine intake on phenylketonuria. *Lancet*. 1953;265(6790):812-813. doi:10.1016/s0140-6736(53)90473-5
4. Millington DS, Kodo N, Norwood DL, Roe CR. Tandem mass spectrometry: a new method for acylcarnitine profiling with potential for neonatal screening for inborn errors of metabolism. *J Inherit Metab Dis*. 1990;13(3):321-324. doi:10.1007/bf01799385
5. Naylor EW, Chace DH. Automated tandem mass spectrometry for mass newborn screening for disorders in fatty acid, organic acid, and amino acid metabolism. *J Child Neurol*. 1999;14(Suppl 1):S4-S8. doi:10.1177/088307389901400102 1
6. Schulze A, Lindner M, Kohlmuller D, Olgemoller K, Mayatepek E, Hoffmann GF. Expanded newborn screening for inborn errors of metabolism by electrospray ionization-tandem mass spectrometry: results, outcome, and implications. *Pediatrics*. 2003;111(6 Pt 1):1399-1406.
7. Wilcken B, Wiley V, Hammond J, Carpenter K. Screening newborns for inborn errors of metabolism by tandem mass spectrometry. *N Engl J Med*. 2003;348(23):2304-2312. doi:10.1056/NEJMoa025225
8. Vill K, Kölbel H, Schwartz O, et al. One year of newborn screening for SMA – results of a German pilot project. *J Neuromuscul Dis*. 2019;6(4):503-515. doi:10.3233/jnd-190428
9. Wilson J, Jungner G. *Principles of Screening for Disease*. World Health Organization; 1968.
10. Newborn screening: toward a uniform screening panel and system. *Genet Med*. 2006;8(Suppl 1):1s-252s. doi:10.1097/01.gim.0000223891.82390.ad
11. Burgard P, Rupp K, Lindner M, et al. Newborn screening programmes in Europe; arguments and efforts regarding harmonization. Part 2. From screening laboratory results to treatment, follow-up and quality assurance. *J Inherit Metab Dis*. 2012;35(4):613-625. doi:10.1007/s10545-012-9484-z
12. Loeber JG, Burgard P, Cornel MC, et al. Newborn screening programmes in Europe; arguments and efforts regarding harmonization. Part 1. From blood spot to screening result. *J Inherit Metab Dis*. 2012;35(4):603-611. doi:10.1007/s10545-012-9483-0
13. Loeber JG, Platis D, Zetterström RH, et al. Neonatal screening in Europe revisited: an ISNS perspective on the current state and developments since 2010. *Int J Neonatal Screen*. 2021;7(1):15. doi:10.3390/ijns7010015
14. Pollitt RJ. International perspectives on newborn screening. *J Inherit Metab Dis*. 2006;29(2–3):390-396. doi:10.1007/s10545-006-0259-2
15. Therrell BL, Padilla CD, Loeber JG, et al. Current status of newborn screening worldwide: 2015. *Semin Perinatol*. 2015;39(3):171-187. doi:10.1053/j.semperi.2015.03.002
16. Andermann A, Blancquaert I, Beauchamp S, Déry V. Revisiting Wilson and Jungner in the genomic age: a review of screening criteria over the past 40 years. *Bull World Health Organ*. 2008;86(4):317-319. doi:10.2471/blt.07.050112
17. Petros M. Revisiting the Wilson-Jungner criteria: how can supplemental criteria guide public health in the era of genetic screening? *Genet Med*. 2012;14(1):129-134. doi:10.1038/gim.0b013e31823331d0
18. Adhikari AN, Gallagher RC, Wang Y, et al. The role of exome sequencing in newborn screening for inborn errors of metabolism. *Nat Med*. 2020;26(9):1392-1397. doi:10.1038/s41591-020-0966-5
19. Schaaf CP, Kölker S, Hoffmann GF. Genomic newborn screening: proposal of a two-stage approach. *J Inherit Metab Dis*. 2021;44(3):518-520. doi:10.1002/jimd.12381

20. Simon E, Flaschker N, Schadewaldt P, Langenbeck U, Wendel U. Variant maple syrup urine disease (MSUD)—the entire spectrum. *J Inherit Metab Dis*. 2006;29(6):716-724. doi:10.1007/s10545-006-0276-1
21. Kölker S, Garbade SF, Greenberg CR, et al. Natural history, outcome, and treatment efficacy in children and adults with glutaryl-CoA dehydrogenase deficiency. *Pediatr Res*. 2006;59(6):840-847. doi:10.1203/01.pdr.0000219387.79887.86
22. Garrod AJTL. The incidence of alkaptonuria: a study in chemical individuality. *Mol Med*. 1902;160(4137):1616-1620.
23. Bezerra GA, Foster WR, Bailey HJ, et al. Crystal structure and interaction studies of human DHTKD1 provide insight into a mitochondrial megacomplex in lysine catabolism. *IUCrJ*. 2020;7(Pt 4):693-706. doi:10.1107/s205225252000696x
24. Biagosch C, Ediga RD, Hensler SV, et al. Elevated glutaric acid levels in Dhtkd1−/Gcdh− double knockout mice challenge our current understanding of lysine metabolism. *Biochim Biophys Acta Mol Basis Dis*. 2017;1863(9):2220-2228. doi:10.1016/j.bbadis.2017.05.018
25. Leandro J, Khamrui S, Wang H, et al. Inhibition and crystal structure of the human DHTKD1-thiamin diphosphate complex. *ACS Chem Biol*. 2020;15(8):2041-2047. doi:10.1021/acschembio.0c00114
26. Van Schaftingen E, Veiga-da-Cunha M, Linster CL. Enzyme complexity in intermediary metabolism. *J Inherit Metab Dis*. 2015;38(4):721-727. doi:10.1007/s10545-015-9821-0
27. Veiga-da-Cunha M, Van Schaftingen E, Bommer GT. Inborn errors of metabolite repair. *J Inherit Metab Dis*. 2020;43(1):14-24. doi:10.1002/jimd.12187
28. Parras-Moltó M, Aguirre de Cárcer D. A comprehensive human minimal gut metagenome extends the host's metabolic potential. *Microb Genom*. 2020;6(11). doi:10.1099/mgen.0.000466
29. Shin SY, Fauman EB, Petersen AK, et al. An atlas of genetic influences on human blood metabolites. *Nat Genet*. 2014;46(6):543-550. doi:10.1038/ng.2982
30. Lanpher B, Brunetti-Pierri N, Lee B. Inborn errors of metabolism: the flux from Mendelian to complex diseases. *Nat Rev Genet*. 2006;7(6):449-460. doi:10.1038/nrg1880
31. Kemp S, Pujol A, Waterham HR, et al. ABCD1 mutations and the X-linked adrenoleukodystrophy mutation database: role in diagnosis and clinical correlations. *Hum Mutat*. 2001;18(6):499-515. doi:10.1002/humu.1227
32. Moser HW, Mahmood A, Raymond GV. X-linked adrenoleukodystrophy. *Nat Clin Pract Neurol*. 2007;3(3):140-151. doi:10.1038/ncpneuro0421
33. Hillert A, Anikster Y, Belanger-Quintana A, et al. The genetic landscape and epidemiology of phenylketonuria. *Am J Hum Genet*. 2020;107(2):234-250. doi:10.1016/j.ajhg.2020.06.006
34. Zielonka M, Garbade SF, Gleich F, et al. From genotype to phenotype: early prediction of disease severity in argininosuccinic aciduria. *Hum Mutat*. 2020;41(5):946-960. doi:10.1002/humu.23983
35. Zielonka M, Kölker S, Gleich F, et al. Early prediction of phenotypic severity in Citrullinemia type 1. *Ann Clin Transl Neurol*. 2019;6(9):1858-1871. doi:10.1002/acn3.50886
36. van Spronsen FJ, van Wegberg AM, Ahring K, et al. Key European guidelines for the diagnosis and management of patients with phenylketonuria. *Lancet Diabetes Endocrinol*. 2017;5(9):743-756. doi:10.1016/s2213-8587(16)30320-5
37. Weglage J, Ullrich K, Pietsch M, Fünders B, Zass R, Koch HG. Untreated non-phenylketonuric-hyperphenylalaninaemia: intellectual and neurological outcome. *Eur J Pediatr*. 1996;155(Suppl 1):S26-S28. doi:10.1007/pl00014244
38. McHugh D, Cameron CA, Abdenur JE, et al. Clinical validation of cutoff target ranges in newborn screening of metabolic disorders by tandem mass spectrometry: a worldwide collaborative project. *Genet Med*. 2011;13(3):230-254. doi:10.1097/GIM.0b013e31820d5e67
39. Ho S, Lukacs Z, Hoffmann GF, Lindner M, Wetter T. Feature construction can improve diagnostic criteria for high-dimensional metabolic data in newborn screening for medium-chain acyl-CoA dehydrogenase deficiency. *Clin Chem*. 2007;53(7):1330-1337. doi:10.1373/clinchem.2006.081802
40. Monostori P, Klinke G, Richter S, et al. Simultaneous determination of 3-hydroxypropionic acid, methylmalonic acid and methylcitric acid in dried blood spots: Second-tier LC-MS/MS assay for newborn screening of propionic acidemia, methylmalonic acidemias and combined remethylation disorders. *PLoS One*. 2017;12(9):e0184897. doi:10.1371/journal.pone.0184897
41. Ensenauer R, Vockley J, Willard JM, et al. A common mutation is associated with a mild, potentially asymptomatic phenotype in patients with isovaleric acidemia diagnosed by newborn screening. *Am J Hum Genet*. 2004;75(6):1136-1142. doi:10.1086/426318
42. Mütze U, Garbade SF, Gramer G, et al. Long-term outcomes of individuals with metabolic diseases identified through newborn screening. *Pediatrics*. 2020;146(5):e20200444. doi:10.1542/peds.2020-0444
43. Mütze U, Henze L, Gleich F, et al. Newborn screening and disease variants predict neurological outcome in isovaleric aciduria. *J Inherit Metab Dis*. 2021;44(4):857-870. doi:10.1002/jimd.12364
44. van Rijt WJ, Ferdinandusse S, Giannopoulos P, et al. Prediction of disease severity in multiple acyl-CoA dehydrogenase deficiency: a retrospective and laboratory cohort study. *J Inherit Metab Dis*. 2019;42(5):878-889. doi:10.1002/jimd.12147
45. Mütze U, Walter M, Keller M, et al. Health outcomes of infants with vitamin B(12) deficiency identified by newborn screening and early treated. *J Pediatr*. 2021;235:42-48. doi:10.1016/j.jpeds.2021.02.009
46. Maier EM, Liebl B, Röschinger W, et al. Population spectrum of ACADM genotypes correlated to biochemical phenotypes in newborn screening for medium-chain acyl-CoA dehydrogenase deficiency. *Hum Mutat*. 2005;25(5):443-452. doi:10.1002/humu.20163
47. Spiekerkoetter U, Lindner M, Santer R, et al. Treatment recommendations in long-chain fatty acid oxidation defects: consensus from a workshop. *J Inherit Metab Dis*. 2009;32(4):498-505. doi:10.1007/s10545-009-1126-8
48. Gray JA, Patnick J, Blanks RG. Maximising benefit and minimising harm of screening. *BMJ*. 2008;336(7642):480-483. doi:10.1136/bmj.39470.643218.94
49. Levy HL. Newborn screening conditions: what we know, what we do not know, and how we will know it. *Genet Med*. 2010;12-(12 Suppl):S213-S214. doi:10.1097/GIM.0b013e3181fe5d77
50. Boy N, Mengler K, Thimm E, et al. Newborn screening: a disease-changing intervention for glutaric aciduria type 1. *Ann Neurol*. 2018;83(5):970-979. doi:10.1002/ana.25233

51. Lindner M, Gramer G, Haege G, et al. Efficacy and outcome of expanded newborn screening for metabolic diseases—report of 10 years from South-West Germany. *Orphanet J Rare Dis*. 2011;6:44. doi:10.1186/1750-1172-6-44
52. Strauss KA, Carson VJ, Soltys K, et al. Branched-chain α-ketoacid dehydrogenase deficiency (maple syrup urine disease): treatment, biomarkers, and outcomes. *Mol Genet Metab*. 2020;129(3):193-206. doi:10.1016/j.ymgme.2020.01.006
53. Heringer J, Valayannopoulos V, Lund AM, et al. Impact of age at onset and newborn screening on outcome in organic acidurias. *J Inherit Metab Dis*. 2016;39(3):341-353. doi:10.1007/s10545-015-9907-8
54. Keller R, Chrastina P, Pavlikova M, et al. Newborn screening for homocystinurias: recent recommendations versus current practice. *J Inherit Metab Dis*. 2019;42(1):128-139. doi:10.1002/jimd.12034
55. Kölker S, Gleich F, Mütze U, Opladen T. Rare disease registries are key to evidence-based personalized medicine: highlighting the European experience. *Front Endocrinol (Lausanne)*. 2022;13:832063. doi:10.3389/fendo.2022.832063
56. Kuseyri Hübschmann O, Horvath G, Cortès-Saladelafont E, Yıldız Y, Mastrangelo M, Pons R, et al. Insights into the expanding phenotypic spectrum of inherited disorders of biogenic amines. Nat Commun 2021;12(1):5529. doi: 10.1038/s41467-021-25515-5
57. Posset R, Garcia-Cazorla A, Valayannopoulos V, et al. Age at disease onset and peak ammonium level rather than interventional variables predict the neurological outcome in urea cycle disorders. *J Inherit Metab Dis*. 2016;39(5):661-672. doi:10.1007/s10545-016-9938-9
58. Posset R, Gropman AL, Nagamani SCS, et al. Impact of diagnosis and therapy on cognitive function in urea cycle disorders. *Ann Neurol*. 2019;86(1):116-128. doi:10.1002/ana.25492
59. Rubio-Gozalbo ME, Haskovic M, Bosch AM, et al. The natural history of classic galactosemia: lessons from the GalNet registry. *Orphanet J Rare Dis*. 2019;14(1):86. doi:10.1186/s13023-019-1047-z
60. Boy N, Mengler K, Heringer-Seifert J, Hoffmann GF, Garbade SF, Kölker S. Impact of newborn screening and quality of therapy on the neurological outcome in glutaric aciduria type 1: a meta-analysis. *Genet Med*. 2021;23(1):13-21. doi:10.1038/s41436-020-00971-4
61. Couce ML, Castineiras DE, Boveda MD, et al. Evaluation and long-term follow-up of infants with inborn errors of metabolism identified in an expanded screening programme. *Mol Genet Metab*. 2011;104(4):470-475. doi:10.1016/j.ymgme.2011.09.021
62. Landau YE, Waisbren SE, Chan LM, Levy HL. Long-term outcome of expanded newborn screening at Boston children's hospital: benefits and challenges in defining true disease. *J Inherit Metab Dis*. 2017;40(2):209-218. doi:10.1007/s10545-016-0004-4
63. McKiernan PJ, Preece MA, Chakrapani A. Outcome of children with hereditary tyrosinaemia following newborn screening. *Arch Dis Child*. 2015;100(8):738-741. doi:10.1136/archdischild-2014-306886
64. O'Reilly D, Crushell E, Hughes J, et al. Maple syrup urine disease: clinical outcomes, metabolic control, and genotypes in a screened population after four decades of newborn bloodspot screening in the Republic of Ireland. *J Inherit Metab Dis*. 2021;44(3):639-655. doi:10.1002/jimd.12337
65. Porta F, Pagliardini V, Celestino I, et al. Neonatal screening for biotinidase deficiency: a 30-year single center experience. *Mol Genet Metab Rep*. 2017;13:80-82. doi:10.1016/j.ymgmr.2017.08.005
66. Simon E, Fingerhut R, Baumkötter J, Konstantopoulou V, Ratschmann R, Wendel U. Maple syrup urine disease: favourable effect of early diagnosis by newborn screening on the neonatal course of the disease. *J Inherit Metab Dis*. 2006;29(4):532-537. doi:10.1007/s10545-006-0315-y
67. Sperk A, Mueller M, Spiekerkoetter U. Outcome in six patients with mitochondrial trifunctional protein disorders identified by newborn screening. *Mol Genet Metab*. 2010;101(2–3):205-207. doi:10.1016/j.ymgme.2010.07.003
68. Wilcken B, Haas M, Joy P, et al. Expanded newborn screening: outcome in screened and unscreened patients at age 6 years. *Pediatrics*. 2009;124(2):e241-e248. doi:10.1542/peds.2008-0586
69. Wilcken B, Haas M, Joy P, et al. Outcome of neonatal screening for medium-chain acyl-CoA dehydrogenase deficiency in Australia: a cohort study. *Lancet*. 2007;369(9555):37-42. doi:10.1016/s0140-6736(07)60029-4
70. Kölker S, Garcia-Cazorla A, Valayannopoulos V, et al. The phenotypic spectrum of organic acidurias and urea cycle disorders. Part 1: the initial presentation. *J Inherit Metab Dis*. 2015;38(6):1041-1057. doi:10.1007/s10545-015-9839-3
71. Kölker S, Valayannopoulos V, Burlina AB, Sykut-Cegielska J, Wijburg FA, Teles EL, et al. The phenotypic spectrum of organic acidurias and urea cycle disorders. Part 2: the evolving clinical phenotype. J Inherit Metab Dis 2015;38(6):1059–74. doi: 10.1007/s10545-015-9840-x
72. Stepien KM, Geberhiwot T, Hendriksz CJ, Treacy EP. Challenges in diagnosing and managing adult patients with urea cycle disorders. *J Inherit Metab Dis*. 2019;42(6):1136-1146. doi:10.1002/jimd.12096
73. Torkzaban M, Haddad A, Baxter JK, Berghella V, Gahl WA, Al-Kouatly HB. Maternal ornithine transcarbamylase deficiency, a genetic condition associated with high maternal and neonatal mortality every clinician should know: a systematic review. *Am J Med Genet A*. 2019;179(10):2091-2100. doi:10.1002/ajmg.a.61329
74. Moorthie S, Cameron L, Sagoo GS, Bonham JR, Burton H. Systematic review and meta-analysis to estimate the birth prevalence of five inherited metabolic diseases. *J Inherit Metab Dis*. 2014;37(6):889-898. doi:10.1007/s10545-014-9729-0
75. Spiekerkoetter U, Lindner M, Santer R, et al. Management and outcome in 75 individuals with long-chain fatty acid oxidation defects: results from a workshop. *J Inherit Metab Dis*. 2009;32(4):488-497. doi:10.1007/s10545-009-1125-9
76. Posset R, Kölker S, Gleich F, et al. Severity-adjusted evaluation of newborn screening on the metabolic disease course in individuals with cytosolic urea cycle disorders. *Mol Genet Metab*. 2020;131(4):390-397. doi:10.1016/j.ymgme.2020.10.013
77. Zaunseder E, Haupt S, Mütze U, Garbade SF, Kölker S, Heuveline V. Opportunities and challenges in machine learning-based newborn screening – a systematic literature review. *JIMD Rep*. 2022;63:250-261.

78. Tingley K, Coyle D, Graham ID, et al. Using a meta-narrative literature review and focus groups with key stakeholders to identify perceived challenges and solutions for generating robust evidence on the effectiveness of treatments for rare diseases. *Orphanet J Rare Dis*. 2018;13(1):104. doi:10.1186/s13023-018-0851-1
79. Pugliese M, Tingley K, Chow A, et al. Outcomes in pediatric studies of medium-chain acyl-coA dehydrogenase (MCAD) deficiency and phenylketonuria (PKU): a review. *Orphanet J Rare Dis*. 2020;15(1):12. doi:10.1186/s13023-019-1276-1
80. Potter BK, Hutton B, Clifford TJ, et al. Establishing core outcome sets for phenylketonuria (PKU) and medium-chain Acyl-CoA dehydrogenase (MCAD) deficiency in children: study protocol for systematic reviews and Delphi surveys. *Trials*. 2017;18(1):603. doi:10.1186/s13063-017-2327-3
81. Pugliese M, Tingley K, Chow A, et al. Core outcome sets for medium-chain acyl-CoA dehydrogenase deficiency and phenylketonuria. *Pediatrics*. 2021;148(2):e2020037747.
82. Strauss KA, Puffenberger EG, Robinson DL, Morton DH. Type I glutaric aciduria, part 1: natural history of 77 patients. *Am J Med Genet C Semin Med Genet*. 2003;121C(1):38-52. doi:10.1002/ajmg.c.20007
83. Bjugstad KB, Goodman SI, Freed CR. Age at symptom onset predicts severity of motor impairment and clinical outcome of glutaric acidemia type 1. *J Pediatr*. 2000;137(5):681-686. doi:10.1067/mpd.2000.108954
84. Kölker S, Christensen E, Leonard JV, et al. Guideline for the diagnosis and management of glutaryl-CoA dehydrogenase deficiency (glutaric aciduria type I). *J Inherit Metab Dis*. 2007;30(1):5-22. doi:10.1007/s10545-006-0451-4
85. Boy N, Mühlhausen C, Maier EM, et al. Proposed recommendations for diagnosing and managing individuals with glutaric aciduria type I: second revision. *J Inherit Metab Dis*. 2017;40(1):75-101. doi:10.1007/s10545-016-9999-9
86. Kölker S, Christensen E, Leonard JV, et al. Diagnosis and management of glutaric aciduria type I—revised recommendations. *J Inherit Metab Dis*. 2011;34(3):677-694. doi:10.1007/s10545-011-9289-5
87. Heringer J, Boy SP, Ensenauer R, et al. Use of guidelines improves the neurological outcome in glutaric aciduria type I. *Ann Neurol*. 2010;68(5):743-752. doi:10.1002/ana.22095
88. Kölker S, Garbade SF, Boy N, et al. Decline of acute encephalopathic crises in children with glutaryl-CoA dehydrogenase deficiency identified by newborn screening in Germany. *Pediatr Res*. 2007;62(3):357-363. doi:10.1203/PDR.0b013e318137a124
89. Strauss KA, Donnelly P, Wintermark M. Cerebral haemodynamics in patients with glutaryl-coenzyme A dehydrogenase deficiency. *Brain J Neurol*. 2010;133(Pt 1):76-92. doi:10.1093/brain/awp297
90. Strauss KA, Williams KB, Carson VJ, et al. Glutaric acidemia type 1: treatment and outcome of 168 patients over three decades. *Mol Genet Metab*. 2020;131(3):325-340. doi:10.1016/j.ymgme.2020.09.007
91. Märtner EMC, Thimm E, Guder P, et al. The biochemical subtype is a predictor for cognitive function in glutaric aciduria type 1: a national prospective follow-up study. *Sci Rep*. 2021;11(1):19300. doi:10.1038/s41598-021-98809-9
92. Boy N, Heringer J, Brackmann R, et al. Extrastriatal changes in patients with late-onset glutaric aciduria type I highlight the risk of long-term neurotoxicity. *Orphanet J Rare Dis*. 2017;12(1):77. doi:10.1186/s13023-017-0612-6
93. Harting I, Boy N, Heringer J, et al. (1)H-MRS in glutaric aciduria type 1: impact of biochemical phenotype and age on the cerebral accumulation of neurotoxic metabolites. *J Inherit Metab Dis*. 2015;38(5):829-838. doi:10.1007/s10545-015-9826-8
94. Forny P, Hörster F, Ballhausen D, et al. Guidelines for the diagnosis and management of methylmalonic acidaemia and propionic acidaemia: first revision. *J Inherit Metab Dis*. 2021;44(3):566-592. doi:10.1002/jimd.12370
95. Hörster F, Tuncel AT, Gleich F, et al. Delineating the clinical spectrum of isolated methylmalonic acidurias: cblA and mut. *J Inherit Metab Dis*. 2021;44(1):193-214. doi:10.1002/jimd.12297
96. Molema F, Martinelli D, Hörster F, et al. Liver and/or kidney transplantation in amino and organic acid-related inborn errors of metabolism: an overview on European data. *J Inherit Metab Dis*. 2021;44(3):593-605. doi:10.1002/jimd.12318
97. Häberle J, Burlina A, Chakrapani A, et al. Suggested guidelines for the diagnosis and management of urea cycle disorders: first revision. *J Inherit Metab Dis*. 2019;42(6):1192-1230. doi:10.1002/jimd.12100
98. Morris AA, Kožich V, Santra S, et al. Guidelines for the diagnosis and management of cystathionine beta-synthase deficiency. *J Inherit Metab Dis*. 2017;40(1):49-74. doi:10.1007/s10545-016-9979-0
99. Huemer M, Diodato D, Martinelli D, et al. Phenotype, treatment practice and outcome in the cobalamin-dependent remethylation disorders and MTHFR deficiency: data from the E-HOD registry. *J Inherit Metab Dis*. 2019;42(2):333-352. doi:10.1002/jimd.12041
100. Huemer M, Diodato D, Schwahn B, et al. Guidelines for diagnosis and management of the cobalamin-related remethylation disorders cblC, cblD, cblE, cblF, cblG, cblJ and MTHFR deficiency. *J Inherit Metab Dis*. 2017;40(1):21-48. doi:10.1007/s10545-016-9991-4
101. Opladen T, Gleich F, Kozich V, et al. U-IMD: the first unified European registry for inherited metabolic diseases. *Orphanet J Rare Dis*. 2021;16(1):95. doi:10.1186/s13023-021-01726-3
102. Gramer G, Fang-Hoffmann J, Feyh P, et al. Newborn screening for vitamin B(12) deficiency in Germany—strategies, results, and public health implications. *J Pediatr*. 2020;216:165-72.e4. doi:10.1016/j.jpeds.2019.07.052
103. Gramer G, Fang-Hoffmann J, Feyh P, et al. High incidence of maternal vitamin B12 deficiency detected by newborn screening: first results from a study for the evaluation of 26 additional target disorders for the German newborn screening panel. *World J Pediatr*. 2018;14(5):470-481. doi:10.1007/s12519-018-0159-1
104. Röschinger W, Sonnenschein S, Schuhmann E, Nennstiel-Ratzel U, Roscher AA, Olgemöller B. Neue Zielerkrankungen im Neugeborenenscreening. *Monatsschr Kinderheilkd*. 2015;163(2):142-149. doi:10.1007/s00112-014-3297-0
105. Mütze U, Nennstiel U, Odenwald B, et al. Sudden neonatal death in individuals with medium-chain acyl-coenzyme A dehydrogenase deficiency: limit of newborn screening. *Eur J*

Pediatr. 2022. Online ahead of print. doi:10.1007/s00431-022-04421-y

106. Leonard JV, Vijayaraghavan S, Walter JH. The impact of screening for propionic and methylmalonic acidaemia. *Eur J Pediatr*. 2003;162(Suppl 1):S21-S24. doi:10.1007/s00431-003-1345-1
107. Dionisi-Vici C, Deodato F, Röschinger W, Rhead W, Wilcken B. 'Classical' organic acidurias, propionic aciduria, methylmalonic aciduria and isovaleric aciduria: long-term outcome and effects of expanded newborn screening using tandem mass spectrometry. *J Inherit Metab Dis*. 2006;29(2–3):383-389. doi:10.1007/s10545-006-0278-z
108. Grunert SC, Wendel U, Lindner M, et al. Clinical and neurocognitive outcome in symptomatic isovaleric acidemia. *Orphanet J Rare Dis*. 2012;7:9. doi:10.1186/1750-1172-7-9
109. Klose DA, Kolker S, Heinrich B, et al. Incidence and short-term outcome of children with symptomatic presentation of organic acid and fatty acid oxidation disorders in Germany. *Pediatrics*. 2002;110(6):1204-1211.
110. Pandor A, Eastham J, Beverley C, Chilcott J, Paisley S. Clinical effectiveness and cost-effectiveness of neonatal screening for inborn errors of metabolism using tandem mass spectrometry: a systematic review. *Health Technol Assess*. 2004;8(12):iii, 1-121.
111. Pollitt RJ, Leonard JV. Prospective surveillance study of medium chain acyl-CoA dehydrogenase deficiency in the UK. *Arch Dis Child*. 1998;79(2):116-119. doi:10.1136/adc.79.2.116
112. Gramer G, Haege G, Glahn EM, Hoffmann GF, Lindner M, Burgard P. Living with an inborn error of metabolism detected by newborn screening-parents' perspectives on child development and impact on family life. *J Inherit Metab Dis*. 2014;37(2):189-195. doi:10.1007/s10545-013-9639-6
113. Hewlett J, Waisbren SE. A review of the psychosocial effects of false-positive results on parents and current communication practices in newborn screening. *J Inherit Metab Dis*. 2006;29(5):677-682. doi:10.1007/s10545-006-0381-1
114. Waisbren SE, Albers S, Amato S, et al. Effect of expanded newborn screening for biochemical genetic disorders on child outcomes and parental stress. *JAMA*. 2003;290(19):2564-2572. doi:10.1001/jama.290.19.2564
115. Bessey A, Chilcott J, Pandor A, Paisley S. The cost-effectiveness of expanding the UK newborn bloodspot screening programme to include five additional inborn errors of metabolism. *Int J Neonatal Screen*. 2020;6(4):93. doi:10.3390/ijns6040093
116. Pfeil J, Listl S, Hoffmann GF, Kolker S, Lindner M, Burgard P. Newborn screening by tandem mass spectrometry for glutaric aciduria type 1: a cost-effectiveness analysis. *Orphanet J Rare Dis*. 2013;8:167. doi:10.1186/1750-1172-8-167
117. Mütze U, Roth A, Weigel JF, et al. Transition of young adults with phenylketonuria from pediatric to adult care. *J Inherit Metab Dis*. 2011;34(3):701-709. doi:10.1007/s10545-011-9284-x
118. Mütze U, Thiele AG, Baerwald C, Ceglarek U, Kiess W, Beblo S. Ten years of specialized adult care for phenylketonuria – a single-centre experience. *Orphanet J Rare Dis*. 2016;11:27. doi:10.1186/s13023-016-0410-6
119. Chen HF, Rose AM, Waisbren S, Ahmad A, Prosser LA. Newborn screening and treatment of phenylketonuria: projected health outcomes and cost-effectiveness. *Children (Basel)*. 2021;8(5):381. doi:10.3390/children8050381
120. Mütze U, Gleich F, Garbade SF, et al. Post-authorization safety study of Betaine anhydrous. *J Inherit Metab Dis*. 2022. Online ahead of print. doi:10.1002/jimd.12499
121. Lajic S, Karlsson L, Zetterström RH, Falhammar H, Nordenström A. The success of a screening program is largely dependent on close collaboration between the laboratory and the clinical follow-up of the patients. *Int J Neonatal Screen*. 2020;6(3):68. doi:10.3390/ijns6030068

How to cite this article: Mütze U, Mengler K, Boy N, et al. How longitudinal observational studies can guide screening strategy for rare diseases. *J Inherit Metab Dis*. 2022;1-13. doi:10.1002/jimd.12508