

Revisiting the Wilson-Jungner criteria: how can supplemental criteria guide public health in the era of genetic screening?

Michael Petros, DrPH¹

Purpose: Advances in technology have made newborn screening for more than 50 inborn errors of metabolism possible using a dried blood sample. A framework is proposed that public health practitioners may use when considering candidate disorders for newborn screening panels.

Methods: The framework expands on the 10 Wilson-Jungner criteria with the addition of 11 criteria specific to newborn screening. A calculation, the “pNBS Decision Score,” is used to quantify results and rank candidate disorders.

Results: The pNBS Decision Scores that were calculated for phenylketonuria (OMIM# 261600), cystic fibrosis (OMIM# 219700), Pompe disease (OMIM# 232300), and severe combined immunodeficiency (OMIM# 102700) support their inclusion as newborn screening disorders. The pNBS Decision Score suggests that Krabbe disease (OMIM# 245200) is not a candidate disorder for inclusion at this time.

Conclusion: The proposed framework adds to the ability of policy makers to quantify an essential portion of the process for adding disorders to newborn screening panels. Other factors such as ethical, legal, and social issues, clinical utility, and advocacy are also part of the policy process. The framework is not intended to replace existing nomination processes but rather to enhance those processes by encouraging iterative review of newborn screening-specific criteria. The use of the framework will provide consistency across a portion of the decision process. The public health community should take the opportunity to revisit the screening determinants of the Wilson-Jungner criteria from a 21st century perspective. The results suggest that this framework provides the public health practitioner with a consistent process for making an evidence-based decision.

Genet Med 2012;14(1):129–134

Key Words: candidate disorders; expansion policy; framework; newborn screening; Wilson-Jungner criteria

INTRODUCTION

Newborn screening (NBS) is the process of testing newborn babies for treatable metabolic/genetic, endocrine, and hematologic diseases commonly referred to as inborn errors of metabolism. Many of these diseases are potentially fatal conditions that are not otherwise apparent at the time of birth. Although the incidence of some of these conditions is extremely rare in the newborn population, when one of these conditions is not found and treated, it can affect a newborn’s normal physical and mental development. In severe cases, the lack of or delay of treatment can lead to developmental disability, mental retardation, and/or premature death.

NBS is now widely recognized as a highly successful health promotion and disease prevention public health program. State health departments were among the first to recognize the importance of NBS to public health, and for more than 40 years, NBS has been an established public health function.^{1,2} NBS has become a world-wide endeavor of state, provincial, and national public health agencies. Evaluations of the impact of NBS have been published in the United States,³ Germany,⁴ and Australia.⁵

In the United States, advances in technology and the proactive stance of advocates have driven the expansion of NBS at the

state level. NBS began with relatively inexpensive screening for highly treatable conditions and has evolved into a more complex algorithm screening for rarer disorders, a number of which are more difficult to diagnose and treat. Over time, the state screening program has expanded its role in not only screening for more disorders but has also expanded its role in follow-up, diagnosis, and treatment.

Although the Secretary’s Advisory Committee on Heritable Disorders has made recommendations concerning candidate disorders for NBS panels at the national level and has recently issued a committee report,⁶ a number of states have moved forward on their own and have introduced legislation to add disorders to their respective panels. A recent effort to encourage expansion of NBS panels to include testing for lysosomal storage diseases has been undertaken by a number of advocacy and stakeholder groups in several states, including New York,⁷ Illinois,⁸ Missouri,⁹ and New Mexico.¹⁰ The questions that arise are “Can a process be developed that rationalizes the decisions of NBS expansion?” and “Would this process serve and enhance the mission of newborn screening?”

Expansion of NBS may be opposed by some who are concerned that effective follow-up and treatment may not be

¹ University of Illinois, School of Public Health, Chicago, Illinois, USA. Correspondence: Michael Petros (Michael.P.Petros@gmail.com)

Submitted 12 April 2011; accepted 17 August 2011; advance online publication 7 October 2011. doi:10.1038/gim.0b013e31823331d0

Table 1 Wilson-Jungner criteria—phenylketonuria

Checklist A	Yes	No	Nonconsensus or inconclusive
The condition should be an important public health concern	X		
There should be a treatment for the condition	X		
Facilities for diagnosis and treatment should be available	X		
There should be a latent stage of the disease	X		
There should be a test or examination for the condition	X		
The test should be acceptable to the population	X		
The natural history of the disease should be adequately understood	X		
There should be an agreed policy on who to treat	X		
The total cost of finding a case should be economically balanced in relationship to medical expenditure as a whole	X		
Case finding should be a continuous process, not a "once and for all" project	X		
Subtotal, Section 1 ^a	10	0	0

^aScores in Checklist A and Checklist B are equally weighted.

available or that false-positive screening tests may cause more harm than good. Some of these disorders are extremely rare and have later onset forms that may not be detected during the newborn period. Additionally, the natural history for many of these diseases has not been well characterized, and effective treatment may not be available.^{11,12} A number of conditions may not lend themselves to inclusion following the classic screening protocol¹³ and as such, require discussion among stakeholder groups and the public health community before implementation.

The framework proposed in this study will contribute to an evidence-based process for selection of candidate disorders for state NBS panels. The framework will allow the public health practitioner and policy analyst to collect and review data, score and rank candidate disorders semiquantitatively, and prepare options or scenarios before the final decision making.

MATERIALS AND METHODS

In the mid-1960s, J.M.G. Wilson of the Ministry of Health of Great Britain and G. Jungner of Sahlgren's Hospital of Gothenburg Sweden compiled a set of criteria that would be used to determine the suitability for establishing screening programs. In this document, which is now considered to be a classic of public health literature, Wilson and Jungner reviewed the history of screening, defined key components of screening, and elucidated the principles that should guide screening efforts.

Table 2 Additional criteria—phenylketonuria

Checklist B	Yes	No	Nonconsensus or inconclusive
The test may be multiplexed or overlaid onto an existing structure or system	X		
The "diagnostic odyssey" for the patient/family may be reduced or eliminated	X		
Adverse outcome(s) are rare with a false-positive test	X		
Treatment costs may be covered by third parties (either private or public)	X		
Testing may be declined by parents/guardians	X		
Adequate pretesting information or counseling is available to parents/guardians	X		
Screening in the newborn period is critical for prompt diagnosis and treatment	X		
Public health infrastructure is in place to support all phases of the testing, diagnosis, and interventions	X		
If carriers are identified, genetic counseling is provided	X		
Treatment risks and the impact of a false-positive test are explained to parents/guardians	X		
The limitations of screening and risks of a false-negative test are explained to parents/guardians	X		
Subtotal, Section 2	11	0	0
TTL, overall (sum of Subsections 1 and 2)	21	0	0
Decision score	21.0		

Andermann et al.¹⁴ proposed revisiting the Wilson-Jungner criteria. In this publication, the authors referenced an earlier work by Pollitt¹⁵ that stated "The lack of even broad concordance at the level of national policy is extremely disturbing. Though all discussion is nominally founded on the ten principles laid down by Wilson and Jungner in 1968, there seems to be no generally accepted way of using these principles, or derived criteria, as objective decision tools." For example, when referring to the criteria "there should be an agreed upon policy who to treat," treatment policy may be likely established by each state NBS program as candidate disorders are added to the panels. In Illinois, this is accomplished by workgroups of clinical specialists who participate in the Genetic and Metabolic Diseases Advisory Committee. The workgroups, usually consisting of three or four members, review published literature and solicit advice from colleagues to reach consensus on a treatment protocol. This proposed protocol is presented to a subcommittee, for example, "The NBS Expansion Subcommittee"

Table 3 Wilson-Jungner criteria—Krabbe disease

Checklist A	Yes	No	Nonconsensus or inconclusive
The condition should be an important public health concern			X
There should be a treatment for the condition	X		
Facilities for diagnosis and treatment should be available	X		
There should be a latent stage of the disease	X		
There should be a test or examination for the condition	X		
The test should be acceptable to the population	X		
The natural history of the disease should be adequately understood			X
There should be an agreed policy on who to treat			X
The total cost of finding a case should be economically balanced in relationship to medical expenditure as a whole			X
Case-finding should be a continuous process, not a "once and for all" project	X		
Subtotal, Section 1	6	0	4

for recommendation to the overall Genetic and Metabolic Diseases Advisory Committee. Thus, treatment protocols are standardized across the population service area.

The technical capability to identify genes and screen for related disorders has progressed rapidly, whereas the ability of policy makers to assess the benefits and concerns of expanded genetic screening has lagged behind the technology curve.

Botkin¹⁶ provides an acronym, ACCE, which defines the approach to evaluation that is useful when contemplating expansion of population screening and subsequent selection of candidate disorders for the panel. ACCE refers to A: analytic validity, C: clinical validity, C: clinical utility, E: ethical, legal, and social implications.

From the NBS perspective, analytic validity refers to the ability of the test or analysis to correctly characterize a neonatal sample within an accepted matrix (a dried blood spot) under high-throughput laboratory conditions. Clinical validity refers to the ability of the test or analysis to correctly characterize the neonate as being affected or unaffected with the condition of interest. Clinical utility refers to the usefulness of the test results to benefit the neonate tested. Ethical, legal, and social implications involve application of public health law, acceptance of the process as the standard of care, and acceptability of the screening algorithm to the community.¹⁶ An example for consideration of the latter would be the instance where a screening test produces genetic information, which was not originally sought, consent not obtained, or sufficient prior information not provided to the patient, parent, or guardian.

The framework proposed in this study is based on the author's needs assessment from the perspective of a public health practitioner.¹⁷ The decision matrix, using the classic Wilson-Jungner criteria and other criteria developed specifically for NBS, provides an evidence-based, semiquantitative comparison of candidate disorders.

As part of the framework, a template and a decision matrix is proposed to assist the policy analyst in preparing recommendations for candidate disorders for expanded NBS. A semiquantitative calculation, the "pNBS Decision Score," is proposed to provide the policy analyst with a unitless number with which to compare attributes of candidate disorders under consideration for addition to NBS screening panels. Arithmetically, the pNBS Decision Score ("p" indicating "proposed" or "prototype") is expressed as: Decision Score = $([\text{Total Column "Yes"} - \text{Total Column "No"}] + [\text{Total Column "Non-Consensus or Inconclusive"}/2])$. **Tables 1** and **2** present the components of the decision template using phenylketonuria (PKU) as an example.

RESULTS

Using disorders considered as the standard of care for NBS, the pNBS Decision Score for PKU was calculated as 21.0 out of a possible 21.0 (**Tables 1** and **2**). Similarly, the proposed framework resulted in a pNBS Decision Score of 19.5 for severe combined immunodeficiency (see Table, Supplemental Digital Content 1, <http://links.lww.com/GIM/A207>, Wilson-Jungner Criteria for SCID; see Table, Supplemental Digital Content 2, <http://links.lww.com/GIM/A208>, Additional Criteria for SCID) and a pNBS Decision Score of 15.0 for Krabbe disease (**Tables 3** and **4**). The ability to use multiplex technologies such as tandem mass spectrometry has allowed testing for conditions such as short-chain acyl-CoA dehydrogenase deficiency (SCAD) where there is lack of knowledge about the natural history,² to be coupled with testing for other conditions with well-characterized natural histories such as medium-chain acyl-CoA dehydrogenase deficiency (MCAD). As such, the availability of using multiplex technologies is an important consideration and is included in Checklist B of this framework. For comparison, the pNBS Decision Scores for two fatty acid oxidation disorders, MCAD and SCAD, that are detectable by multiplex tandem mass spectrometry are presented (see Table, Supplemental Digital Content 3, <http://links.lww.com/GIM/A209>, Wilson-Jungner Criteria for MCAD; see Table, Supplemental Digital Content 4, <http://links.lww.com/GIM/A210>, Additional Criteria for MCAD; see Table, Supplemental Digital Content 5, <http://links.lww.com/GIM/A211>, Wilson-Jungner Criteria for SCAD; see Table, Supplemental Digital Content 6, <http://links.lww.com/GIM/A212>, Additional Criteria for SCAD). The major difference being the checked options for "The condition should be an important public health concern," "The natural history of the disease should be adequately understood," and "There should be an agreed policy on who to treat" for the reasons of low incidence of severe clinical cases and uncertainty about the extent variant forms of the disease have on health.

Table 4 Additional criteria—Krabbe disease

Checklist B	Yes	No	Nonconsensus or inconclusive
The test may be multiplexed or overlaid onto an existing structure or system	X		
The “diagnostic odyssey” for the patient/family may be reduced or eliminated	X		
Adverse outcome(s) are rare with a false-positive test			X
Treatment costs may be covered by third parties (either private or public)			X
Testing may be declined by parents/guardians			X
Adequate pretesting information or counseling is available to parents/guardians	X		
Screening in the newborn period is critical for prompt diagnosis and treatment	X		
Public health infrastructure is in place to support all phases of the testing, diagnosis, and interventions		X	
If carriers are identified, genetic counseling is provided			X
Treatment risks and the impact of a false-positive test are explained to parents/guardians	X		
The limitations of screening and risks of a false-negative test are explained to parents/guardians	X		
Subtotal, Section 2	6	1	4
TTL, overall (sum of Subsections 1 and 2)	12	1	8
Decision score	15.0		

Table 5 Summary of pNBS decision scores for selected disorders

Disease	pNBS decision score (maximum = 21.0)	pNBS decision score, relative to phenylketonuria
Phenylketonuria	21.0	NA
Medium-chain acyl-CoA dehydrogenase deficiency (MCAD)	20.0	95%
Cystic fibrosis	19.5	93%
Severe combined immunodeficiency	19.5	93%
Short-chain acyl-CoA dehydrogenase deficiency (SCAD)	18.5	88%
Pompe disease	17.0	81%
Krabbe disease	15.0	71%

NA, not applicable.

Calculations for the proposed “pNBS Decision Score,” presented in **Table 5**, have been completed for PKU, MCAD, cystic fibrosis, severe combined immunodeficiency, SCAD, Pompe disease, and Krabbe disease using the proposed framework.¹⁸ The results have been as predicted and correlate well with the current positions of the American College of Medical Genetics and the Secretary’s Advisory Committee on Heritable Disorders in Newborns and Children.¹⁹ Also included in **Table 5** are percentages based on calculated pNBS Decision Scores relative to the calculated pNBS Decision Score for PKU. In doing so, PKU or another selected condition may be used as an “internal standard” to “normalize” the assessments. A normalized calculation using a relative percentage may be more consistent across assessments by different individuals or programs.

DISCUSSION

Referring to Steiner’s commentary²⁰ on the NYS Krabbe screening program, the observation was made that “... it may be best to mandate newborn screening for disorders only after careful study and deliberation, to ensure the lowest risk and greatest potential for a favorable outcome. The slow, deliberate, thoughtful approach to adding disorders to newborn screening panels may be distasteful to advocacy groups who see children continuing to suffer from these disorders during the process, but nevertheless, the end product of a carefully designed screening program with low risk for harm and high likelihood of success may justify the process.”

Calonge et al.⁶ published an important committee report on an evidence-based review process from nomination of a NBS disorder through the inclusion of a disorder as part of the recommended uniform panel. This process uses 6 key questions covering improved outcomes, case definitions, test/screening methodology, clinical validity of the screening algorithm, clinical utility of the screening algorithm, and cost-effectiveness.

This committee report provides guidance on how evidence is evaluated and which factors are essential for developing a recommendation for inclusion of a nominated disorder into a screening panel. Further, this committee report presents a decision matrix with which to categorize the recommendation to include, defer, or not include a nominated disorder into a screening panel.

The framework proposed in this study differs somewhat from that published by Calonge et al.⁶ in that this framework’s focus is based primarily on decisions made at the state level rather than at the national level. The scoring system, using the pNBS Decision Score, provides a semiquantitative comparison of a candidate disorder to specific disorders presently included in screening panels. Recent decisions by some states to include disorders in their screening panels that have been previously determined to be “not recommended” for inclusion by the Secretary’s Committee indicate that other factors are at work at the local level as driving forces for inclusion of candidate disorders.

Another important consideration presented in the report by Calonge et al.⁶ is that of evidence and potential knowledge gaps.

This leads to the question: Where will the evidence be obtained? Considering the rarity of many disorders, population screening, either through pilot testing or full-scale screening, may be the only source of evidence that would be available for decisions and recommendations at the national level. By using a relative scoring process as suggested in this study, state screening programs would be the sources of evidence for the NBS community and for the Secretary's Committee.

Although the recommendations of the Secretary's Committee and the various state advisory committees provide needed expertise in the areas of clinical efficacy, technology, and operations and other traditional considerations of public health, Watson²¹ notes "consumer demand will continue to force public health professionals to reassess their programs with a focus on the more subjective criteria by which conditions are evaluated." Expanded criteria and a means to rank or score such criteria, such as presented in this framework, will be useful at the state or regional levels when evaluating candidate disorders.

As stated earlier, scoring, using the pNBS Decision Score, provides a means to rank candidate disorders relative to disorders already included in decision panels. This provides two benefits. First, this process fosters a review of current disorders and thus provides a means to revisit and gain perspective on the existing test panel (at the state level). Second, this process allows reviewers or policy makers to gain insight on issues specific for the candidate disorders and subsequently develop plans either to investigate alternate sources of evidence or by recommending a plan of action to implement pilot testing or other action that would provide the necessary evidence. This evidence should be shared with the NBS community. In the future as experience is gained, a weighted scoring system might prove to be useful. For example, the 11 criteria listed in Checklist B may be weighted more heavily than the criteria in Checklist A, as the Checklist B criteria are specific to NBS. In this way, credence is still given to the 10 Wilson-Jungner Criteria of Checklist A, but the weighted criteria in Checklist B may provide for a more enlightened score and subsequent ranking.

The proposed framework may have limited usefulness in states that have already passed laws requiring the addition of specific disorders to NBS panels, but other states contemplating expansion will find this framework useful for selecting candidate disorders based on their population needs and within their program limitations. This proposed framework will be a particularly useful tool for outcome evaluation and needs assessment after several years of data are accumulated for lysosomal storage diseases.

The pNBS Decision Score described in this study is proposed as component of a toolkit that public health policy analysts and practitioners at the state level may use to evaluate candidate disorders both before implementation and after pilot testing or other reliable data are available.

Limitations and procedure notes

A primary assumption of the method is that the policy analyst should be familiar with population screening programs and has foundational knowledge of the NBS system in place at the

time of the analysis. A second assumption is that the process of adding disorders is supported by evidence-based proposals. A third assumption is that the policy analyst will revisit the framework and modify or otherwise revise the process to include additional data as it becomes available. Paralleling this assumption is that NBS programs will publish or otherwise make available data obtained during and after the addition of candidate disorders to their respective screening panels.

The policy analyst or public health practitioner must also exercise consistency when evaluating candidate disorders. The proposed "pNBS Decision Score" is a unitless, relative number and could be used by a single analyst using a consistently applied process. Decision score thresholds may be established. However, thresholds should not be set as to arbitrarily eliminate a candidate disorder from consideration. The pNBS Decision Score described in this study is best used as a relative ranking tool, not as an absolute measure of "worthiness" when evaluating a candidate disorder.

Sources of data must be reliable and decisions must be made with foundational knowledge of population dynamics. Therefore, the policy analyst must be prepared to address public health issues as part of an agenda heavily weighted with data.

The proposed framework encourages participation of multiple reviewers, working as specialists for the initial review process then working together as a committee to achieve consensus. It is suggested that reviewers should be drawn from dissimilar backgrounds or perspectives to gain a rich source of expertise such as clinical genetics or pediatrics, laboratory and associated screening operations, public health law/ethics, and policy development. Recently, the framework's checklists were shared with a large Midwestern NBS program. The director of this program is using the framework as part of that agency's process for consideration of Pompe disease as part of their NBS panel. The results of the application of the pNBS Decision Score framework will be shared with that state's NBS advisory committee. Another Midwestern state NBS program has placed this framework on their fall 2011 meeting agenda. This framework may find acceptance as an evaluation tool among members of the Health Resources and Services Administration regional NBS workgroups as they consider candidate disorders. Members of the workgroups could use this framework to score and rank candidate disorders from their states' perspectives, share that information, and discuss to reach a regional consensus. This may open the possibility of sharing resources among regional members in the cases where expertise for evaluating and treating cases of extremely rare disorders crosses state boundaries.

The framework presented in this study is offered for consideration as one of the steps used to evaluate candidate NBS disorders at the state and/or regional level. The use of this framework, along with consideration of the recommendations put forth from the Secretary's Advisory Committee, will enhance the process used for adding disorders to NBS test panels by fostering a close examination of the factors outside the clinical arena such as cost, ethics, screening operations, and case management. Watson²¹ stated "... criteria are important in

that they describe the relevant questions that are to be asked.” By asking the relevant questions, NBS practitioners and policy makers will be able to develop and continuously refine a system for identifying and selecting NBS panel disorders and adding those disorders to NBS panels that would best serve the newborn population and families.

SUPPLEMENTARY MATERIAL

Supplementary material is linked to the online version of the paper at <http://www.nature.com/gim>

DISCLOSURE

The author declares no conflict of interest.

REFERENCES

1. MacCready RA, Hussey MG. Newborn phenylketonuria detection program in Massachusetts. *Am J Public Health Nations Health* 1964;54:2075–2081.
2. Levy HL. Newborn screening conditions: What we know, what we do not know, and how we will know it. *Genet Med* 2010;12(suppl 12):S213–S214.
3. Newborn screening: toward a uniform panel and system. *Genet Med* 2006;8(suppl 1):1S–252S.
4. Schulze A, Lindner M, Kohlmüller D, Olgemöller 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.
5. 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;369:37–42.
6. Calonge N, Green NS, Rinaldo P, et al. Committee report: Method for evaluating conditions nominated for population-based screening of newborns and children. *Genet Med* 2010;12:153–159.
7. New York Public Health Law Sec. 69-1.2 as amended, 2006.
8. Illinois Senate Bill 1566, signed into law, 2007.
9. Missouri House Bill 716, signed into law, 2009.
10. New Mexico House Bill 201, signed into law, 2010.
11. Moyer VA, Calonge N, Teutsch SM, Botkin JR; United States Preventive Services Task Force. Expanding newborn screening: process, policy, and priorities. *Hastings Cent Rep* 2008;38:32–39.
12. Baily MA, Murray TH. Ethics, evidence, and cost in newborn screening. *Hastings Cent Rep* 2008;38:23–31.
13. Wilson JM, Jungner YG. [Principles and practice of mass screening for disease]. *Bol Oficina Sanit Panam* 1968;65:281–393.
14. 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;86:317–319.
15. Pollitt RJ. International perspectives on newborn screening. *J Inherit Metab Dis* 2006;29:390–396.
16. Botkin JR. Assessing the new criteria for newborn screening. *Health Matrix Clevel* 2009;19:163–186.
17. Petros M, Lenihan P, Jinks D, Swartzman DS. A cost analysis and planning model for newborn screening program expansion. Presented at the 4th National Conference on Genomics and Public Health, Bethesda, MD, 2010.
18. Petros M. A cost analysis and decision framework for state newborn screening program expansion [doctoral thesis]. School of Public Health, University of Illinois: Chicago, 2011.
19. Advisory Committee on Heritable Disorders in Newborns and Children. External evidence review workgroup reports 2008–2010. <http://www.hrsa.gov/heritabledisorderscommittee/reports/ExternalEvidence.htm>. Accessed 6 April 2011.
20. Steiner RD. Commentary on “Newborn screening for Krabbe disease: the New York state model” and “the long-term outcomes of presymptomatic infants transplanted for Krabbe disease. A report of the workshop held July 11–12, 2008, Holiday Valley, New York.” *Genet Med* 2009;11:411–413.
21. Watson MS. Current status of newborn screening: decision-making about the conditions to include in screening programs. *Ment Retard Dev Disabil Res Rev* 2006;12:230–235.