An Assessment of Public Preferences for Newborn Screening Using Best–Worst Scaling

Beth A. Tarini, MD, MS[1,2], Norma-Jean Simon, MPH, MPA[3], Katherine Payne, PhD[4], Acham Gebremariam, MS[2], Angela Rose, MPH[2], and Lisa A. Prosser, PhD[2,5]

Objective To identify and quantify public preferences for attributes of newborn screening conditions.

Study design We conducted an online national survey of the public (n = 502) to evaluate preferences for attributes of candidate newborn screening conditions. Respondents were presented with hypothetical condition profiles that were defined using 10 attributes with 2-6 levels per attribute. Participants indicated whether they would recommend screening for a condition and which condition attributes were most and least important when making this decision (best–worst scaling). Difference scores were calculated and stratified by condition recommendation (recommend or not recommend for screening). Regression analyses were used to evaluate the effect of attributes on choice to screen or not screen.

Results The number of babies diagnosed was important to those who would recommend newborn screening for a profile, and age at which the treatment would start was important to those who would not recommend newborn screening. Cost was considered to be a key attribute, and treatment effectiveness and impact of making the diagnosis through newborn screening were of low importance for both groups.

Conclusion Public preferences identified through survey methods that provide an adequate baseline understanding of newborn screening can be used to inform newborn screening decisions. *(J Pediatr 2018;201:62-8).*

See editorial, p 8

Newborn screening is one of the most successful public health programs, having screened hundreds of millions of children for serious inherited disorders and saved many newborns from death and severe disability.[1] The decision about which conditions to add to a state newborn screening panel can be made through a number of processes, such as legislative mandate, department of public health regulations, or vote by a committee of medical and public health experts. The decisions also may be influenced by the Advisory Committee on Heritable Disorders in Newborns and Children, which provides evidence-based recommendations about which conditions should be added to a Recommended Uniform Screening Panel for all states.[2]

Public input in decisions about which disorders should be screened through newborn screening has been limited to public advocacy. This important, but select, view of public preferences[3] often is provided by parents of affected children or advocates for rare disease organizations. Missing is meaningful input from the general public, who finance the programs through tax dollars and whose children undergo this mandatory screening. The view of the general public is important because there are limited resources available to implement screening for the increasing numbers of candidate disorders. As a result, decisions must be made about which disorders should be given preference for screening based on a number of disorder characteristics (eg, number of children diagnosed, success of treatment). When available evidence must be placed in a value context, the public's preference regarding newborn screening conditions can provide useful additional information for newborn screening programs.

A number of US- and Canadian-based studies have attempted to examine the public's preferences about newborn screening.[4-7] However, these studies suffer from limitations that affect generalizability of their findings. First, these studies frequently queried public opinion about specific conditions—as opposed to the specific attributes, or characteristics (eg, age of onset, success of treatment), of conditions. Queries about conditions, rather than disease characteristics, may bias respondents based on their previous experience with the condition. Second, some studies were conducted by using focus groups and/or general surveys of select populations and did not have a broad and diverse representation of the general public.

Understanding which attributes, or characteristics, of candidate newborn screening conditions are important to the general public may be useful information for newborn screening program committees to consider alongside their current

From the [1]Department of Pediatrics, University of Iowa, Iowa City, IA; [2]Child Health Research and Evaluation Center, Department of Pediatrics and Communicable Diseases, University of Michigan, Ann Arbor, MI; [3]Division of Emergency Medicine, Lurie Children's Hospital, Chicago, IL; [4]Manchester Centre for Health Economics, The University of Manchester, Manchester, United Kingdom; and [5]Department of Health Management and Policy, School of Public Health, University of Michigan, Ann Arbor, MI

Funded by Agency for Healthcare Research & Quality (AHRQ) (R01HS020644). Study sponsors did not have a role in the study design; collection, analysis, and interpretation of the data; the writing of the report; or the decision to submit this manuscript for publication. The authors declare no conflicts of interest.

Portions of this study were presented at the Pediatric Academic Societies annual meeting, April 30-May 3, 2016, Baltimore, Maryland.

https://doi.org10.1016/j.jpeds.2018.05.035

evidence-based evaluation—especially when evidence is limited, as is often the case with rare disorders.[8,9] Given the paucity of information concerning US public preferences about candidate newborn screening conditions for this mandatory public health screening program, we aimed to identify and quantify public preferences for important condition attributes or characteristics.

Methods

An online survey was used to quantitatively measure the relative value of attributes for newborn screening candidate conditions using best–worst scaling. In this study, we used best–worst scaling, a preference measurement method from marketing economics and type of conjoint analysis.[10] Best–worst scaling provides the relative value of characteristics of a condition (eg, age of onset, success of treatment) by asking the respondent to choose the best and worst attribute from among a list. Best–worst scaling has been used in healthcare to identify characteristics that are most important to healthcare decision-making[11,12] and research policymaking[13] and to better understand the most and least important factors associated with health programs.[14] This study used a profile case approach as described by Flynn et al[12]. A set of attributes was identified to describe candidate conditions for newborn screening. Each attribute was associated with a set of levels to describe the possible states of that specific attribute (eg, the attribute for age of onset would include levels of infancy, childhood, and adulthood) (**Table I**). Respondents are asked to review a profile (hypothetical condition) and to select the most and the least important factors (attribute/level) in the profile.[15] By requesting 2 responses, more data about the respondent's preferences can be collected, providing greater insight into the respondent's decision-making process.[15]

Eligible participants were sampled randomly from GfK KnowledgePanel, a nationally representative online panel of US adults that has been widely used in other pediatric health surveys.[16-18] Eligible participants were noninstitutionalized adults age 18 years and older residing in the US and were surveyed from September to October 2015. This study was approved by

Table I. Condition attributes and levels

Attributes	Levels
1. Number of babies diagnosed	1. 1 in 100 000 2. 5 in 100 000 3. 10 in 100 000 4. 20 in 100 000
2. Chance that a positive newborn screening test result is wrong	1. 1% 2. 7% 3. 28% 4. 80%
3. Cost of confirming testing and diagnosis	1. $10 2. $100 3. $1000 4. $10 000
4. Likelihood of developing symptoms	1. Very unlikely to develop symptoms 2. Unlikely to develop symptoms 3. Likely to develop symptoms 4. Very likely to develop symptoms
5. Seriousness of symptoms without treatment	1. A child would be able to do all daily activities 2. A child would be able to do most daily activities 3. A child would be able to do some daily activities 4. A child would be able to do no daily activities
6. Age when symptoms develop and life expectancy without treatment	1. Symptoms develop in infancy; child expected to die in infancy 2. Symptoms develop in infancy; child expected to die in childhood 3. Symptoms develop in childhood; child expected to die in childhood 4. Symptoms develop in childhood; child expected to die in adolescence 5. Symptoms develop in childhood; child expected to die in adulthood 6. Symptoms develop in adulthood; child expected to die in adulthood
7. Start of treatment	1. Within a few weeks of birth 2. During the first year of life 3. During childhood 4. During adolescence
8. Success of treatment	1. Treatment will cure the disease 2. Treatment will prevent the disease from getting worse 3. Treatment will slow the worsening of the disease 4. No treatment is available, but care will be given to relieve symptoms
9. Side effects of treatment	1. Potential death 2. Serious side effects of treatment, but will not die 3. Mild side effects 4. None
10. Impact of diagnosis through newborn screening	1. Parents will know about the baby's disease sooner and treatment will be more successful 2. Parents will know about the baby's disease sooner, but this will have no change on the success of treatment

PRACTICE QUESTION: Imagine that, as a member of a committee, **you** can recommend screening for a disease through a state newborn screening program.

Would you recommend a newborn screening test for this disease?

□ Yes, I would recommend screening for the disease
□ No, I would not recommend screening for the disease

Please select what **you** think is the (1) most important and (2) least important characteristic that helped you decide whether or not to screen for this disease.

Least Important	Characteristic	Most Important
□	Number of babies diagnosed: 1 in 100 000	□
□	**7% (7 in 100)** chance that the positive newborn screening **test is wrong** and the **baby does not have the disease.**	□
□	Cost of testing and diagnosis: $10 000	□
□	Unlikely to develop symptoms	□
□	Without treatment, a child would be able to do **most** daily activities like taking care of themselves, school, sports, and chores	□
□	Symptoms develop in **childhood** (1–12 years) and expected to die in **childhood** (1–12 years) without treatment	□
□	Treatment starts during **childhood** (1–12 years)	□
□	Treatment **will cure** the disease	□
□	A child will **experience mild side effects** of treatment	□
□	With newborn screening, parents will know about the baby's disease sooner and **treatment will be more successful**	□

Figure 1. Sample best–worst scaling question.

the institutional review board at the University of Michigan Health System.

The survey instrument was developed based on information gathered from literature searches, expert consultation, and focus groups. Based on information gathered from all of these sources, 10 attributes were identified as being the most important to the consideration of candidate conditions for newborn screening. Each attribute was defined with 2-6 possible levels depending on the attribute (**Table I**). The survey instrument was then refined based on extensive cognitive testing with in-person interviews (n = 38).

The survey instrument contained an introductory section that provided a detailed introduction section on the rationale for newborn screening as well as the organization and delivery of newborn screening services, including how conditions are added to state-level newborn screening panels (**Appendix**; available at www.jpeds.com). Respondents were asked to imagine that they were a member of a committee responsible for recommending that a condition be screened for by a state-level newborn screening program. Respondents were then presented with 4 profiles of hypothetical newborn screening candidate conditions. Each profile contained a preset list of attributes and a selected value for each attribute that was generated using an efficient experimental design to identify a feasible list of profiles for inclusion.[10]

In each case, respondents were asked to make a recommendation for or against screening for the candidate condition given the specified profile and indicate which attribute was the most important and which attribute was the least important when making the decision to recommend or not recommend the condition (**Figure 1**). The final section also included questions on the respondent's confidence in their responses, experience with a family member with a newborn screening condition, and selected sociodemographics (household size, income). Respondent age, sex, race and ethnicity, and education were collected by the survey vendor.

Statistical Analyses

Difference scores for each attribute level were calculated by subtracting the number of times it was selected as the most

important characteristic minus the number of times it was selected as least important. Greater positive difference scores represent attribute levels that have been selected more often as "most important" to the respondent's decision whether or not to recommend screening. Conversely, more negative difference scores represent attribute levels that have been selected more often as "least important" to the respondent's decision whether or not to recommend screening. Attribute levels with difference scores close to zero are those that are inconsequential to respondents' decisions. We stratified these difference scores by whether or not the respondent recommended the condition.[19]

Regression Analysis. A generalized estimating equation logit model was used to analyze the best–worst questions and examine the relationships between the presence of different attribute levels and the respondent's recommendation in favor of or against each profile representing a newborn screening program. This approach adjusts for clustering at the respondent level. Effects coding was used for categorical attribute levels.[20] All analyses were conducted in STATA, version 13 (StataCorp, College Station, Texas).

Results

A total of 502 individuals completed the survey (response rate = 60%). More than 60% of the sample was older than 45 years; 83% of respondents were white, and 32% had at least a college education (**Table II**).

When presented with a profile representing a candidate newborn screening condition, respondents recommended screening for 56.8% of the profiles and recommended against screening for 29.6% of the profiles (respondents did not recommend for or against for 13.6% of the profiles). The importance of attributes differed by whether a respondent would recommend the condition for screening or not. Of note, we compared the demographic characteristics between respondents who answered all best–worse scaling questions and those who did not answer 1 or more questions. We found no statistical difference in the demographic characteristics of these 2 groups.

When respondents recommended screening for a condition profile, they indicated cost of testing and the number of babies diagnosed as the most important attributes behind this decision (**Figure 2**, A). Attributes least important to respondents' decision to recommend screening were age when symptoms start, success of treatment, and the impact of making the diagnosis through newborn screening (ie, whether or not earlier knowledge of disease would be likely to affect the success of treatment).

When respondents recommended against screening for a condition profile, they also indicated cost as the most important attribute behind this decision, followed by the age at which the child would start treatment (**Figure 2**, B). Attributes least important to respondents' decision to recommend against screening included the chance that the test was wrong and the success of the treatment.

Table II. Participant characteristics (n = 502)

Characteristics	Frequency (%)
Age, y	
18-29	76 (15.1)
30-44	117 (23.3)
45-59	151 (30.1)
60+	158 (31.5)
Sex	
Female	241 (48.0)
Male	261 (52.0)
Ethnicity	
Hispanic	56 (11.2)
Non-Hispanic	446 (88.8)
Race	
American Indian, Alaska Native	6 (1.2)
Asian	17 (3.4)
Black or African American	48 (9.6)
Native Hawaiian/ Pacific Islander	1 (0.2)
White	418 (83.3)
2 + races	12 (2.4)
Education	
Less than high school	43 (8.6)
High school	124 (24.7)
Some college	143 (28.5)
Bachelor's degree or higher	192 (38.2)
Household size	
1	99 (19.7)
2	201 (40.0)
3	76 (15.1)
4	67 (13.3)
5+	59 (11.8)
Household income	
<$25 000	71 (14.1)
$25 000-$49 999	111 (22.1)
$50 000-$74 999	81 (16.1)
$75 000-$99 999	78 (15.5)
≥$100 000	161 (32.1)

In regression analysis, increasing the number of babies diagnosed, decreasing the chance the test is wrong, and no treatment side effects experienced by the child were significantly associated with a recommendation in favor of newborn screening for the profiled condition (**Table III**; available at www.jpeds.com). However, the impact of newborn screening on treatment success was not associated with the recommendation for or against newborn screening for a profiled condition.

Discussion

In this study, we identified the relative importance of attributes of newborn screening conditions from a public perspective. In this study, we found that the public most frequently reported cost as the most important attribute for a candidate newborn screening condition, regardless of whether or not they would recommend screening for the condition. The age at which symptoms start, success of treatment, and impact of newborn screening on treatment success (ie, whether or not earlier knowledge of disease will affect success of treatment) were rated as the least important factors in deciding to recommend in favor of a screening for a condition. Unfortunately, there are a paucity of comparative data because most published studies on public attitudes about newborn screen-

A. Condition Profiles Recommended

Attribute	Attribute Level						Total
	1	2	3	4	5	6	
Cost of testing	69	79	70	60	-	-	278
No. of babies diagnosed	44	43	26	38	-	-	151
Chance test is wrong	16	18	36	21	-	-	91
Likelihood of symptoms	11	18	3	16	-	-	48
Start of treatment	2	10	11	5	-	-	28
Side effects of treatment	4	-5	9	13	-	-	21
Seriousness of symptoms	-17	-9	-17	-16	-	-	-59
Impact	-97	-79	-	-	-	-	-176
Success of treatment	-68	-38	-38	-36	-	-	-180
Age symptoms start	-34	-38	-26	-31	-47	-32	-208

B. Condition Profiles NOT Recommended

Attribute	Attribute Level						Total
	1	2	3	4	5	6	
Cost of testing	13	20	25	5	-	-	63
Start of treatment	15	6	12	13	-	-	46
Likelihood of symptoms	-1	0	8	13	-	-	20
No. of babies diagnosed	13	8	-8	6	-	-	19
Impact	-5	11	-	-	-	-	6
Side effects of treatment	-4	-7	1	13	-	-	3
Age symptoms start	-2	-2	-5	1	-3	5	-6
Seriousness of symptoms	-3	-8	5	-6	-	-	-12
Success of treatment	-14	-2	-12	-5	-	-	-33
Chance test is wrong	-12	-34	-33	-23	-	-	-102

Figure 2. Difference scores for attributes of newborn screen programs. **A,** Condition profiles recommended; **B,** condition profiles NOT recommended. Difference scores for each attribute level were calculated by subtracting the number of times it was selected as the most important characteristic minus the number of times it was selected as least important. Greater positive difference scores represent attribute levels that have been selected more often as "most important" to the respondent's decision whether to recommend screening. Conversely, more negative difference scores represent attribute levels that have been selected more often as "least important" to the respondent's decision whether to recommend screening. Attribute levels with difference scores close to zero are those that are inconsequential to respondents' decisions.

ing are conducted with small, select populations and focus on individual disorders.[21-25]

Our findings may indicate that the public prefers the knowledge provided by newborn screening regardless of the availability of effective treatment. This perspective challenges a historical tenet of newborn screening that conditions should only be screened for if there is an effective treatment for the condition. It is a perspective that has been raised among professionals in the newborn screening community and borne out in other studies of public preferences.[8,26] It is possible that our participants' preferences regarding effective treatment reflect this broader debate. Although a cure may not yet be available, the public may view that the knowledge gained by a diagnosis through newborn screening could help avoid a future diagnostic odyssey for the affected child, inform the parents' reproductive decisions, and allow the patient the opportunity to participate in early interventions and clinical trials to potentially slow disease progression.[8] Admittedly, this view challenges the current ethical framework of this mandatory screening program.[27]

Important barriers to gathering data on the preferences of a wider public audience in the past have included easy access to a representative population and a methodology that allows for generalizable findings. Using best–worst scaling methodology in a diverse population, we have attempted to address these barriers. Although conditions that previously added to a state newborn screening panel and candidate conditions under consideration for addition to a panel are each unique, they can be compared using a set of common characteristics or attributes such as incidence rate and false positive rate. Comparing the range of values for different attributes allows for identification of condition types prized most by the public without the confounding cultural or personal biases associated with a specific condition. Although we conducted an extensive literature search, as well as expert review, of our attributes, it is possible that we did not include all relevant attributes. Of note, there is a limit to the number of attributes that can be used without complicating the analyses. Of note, regression results confirmed the most and least important attributes identified by importance scores. However, some of the middle-ranked attributes were not significant. The sample may not have sufficient sample size to identify significant differences for these attributes. An alternative explanation could be that these middle-ranked attributes were not significantly

different from each other and did not have a significant impact on the choice.

One limitation is that this study did not identify strong preferences for several attributes that were identified as important in other focus groups and previous literature. One explanation of this is that consideration of newborn screening programs is a complex cognitive task that requires lengthy reflection and deliberation. Given this possibility, future assessments of public preferences should explore the use of deliberative processes[28,29] that may better capture public preferences. Methods such as deliberative polling or consensus conferences represent one way to engage a larger, representative portion of the public.[30] Other examples of the use of deliberative processes in healthcare policymaking is the National Health Service Citizen Jury program in the United Kingdom, in which a purposive sample of the general population meets for several days to learn about and consider key public health policies.[31] Such intensive programs in which public representatives have an opportunity to learn about a complex program and its attendant benefits and harms may be a more effective strategy for deriving public input for newborn screening programs. Studies such as the one described here can provide important foundational data that can be used to guide the development of future deliberative projects in the US. The key drawback is that deliberative processes can be costly and resource intensive. However, we encourage newborn screening decision makers either at the state or federal level to consider using such deliberative approaches to elicit public feedback both for candidate conditions under consideration and also for input for ongoing newborn screening panels.

The study has additional limitations that also should be considered. We focused on a wide age range of the population, given that the entire population is responsible for the funding of newborn screening. As a result, we did not have enough power to examine important subgroups, such as pregnant women. Our population was slightly older, more educated, and had a higher income than the general US public. They also had a greater percentage of 2- and 5-person households. Therefore, it is likely that difficulties with comprehending newborn screening would be increased in a more representative population. Finally, we did not include visual aids to help explain the "number of babies diagnosed" attribute and respondents may not have had the math literacy to understand attributes presented numerically. Future studies should use risk graphics to better convey risk information to a wide audience.[32] Finally, although the majority of newborn screening costs are born by governmental and private insurance agencies, we chose to use personal cost as an attribute. Past research suggests that respondents are less sensitive to cost levels when they are borne by others.[33] Therefore, we sought to use personal cost as an indirect measure of respondents' sensitivity to the financing of newborn screening. ■

We thank Drs Alex Kemper (Nationwide Children's Hospital, Columbus, Ohio) and Scott Grosse (Centers for Disease Control and Prevention, Atlanta, Georgia) for thoughtful comments on preliminary findings; Caitlin Finan, MPH (NORC at the University of Chicago, Chicago, Illinois) for her contribution to developing the survey instrument; Michael Rozier, MPH (University of Michigan, Ann Arbor, Michigan) for expert research assistance; and the many study respondents who took the time to participate in focus groups, cognitive pretesting, and carefully respond to the study survey.

Submitted for publication Feb 20, 2018; last revision received Apr 18, 2018; accepted May 17, 2018

Reprint requests: Beth A. Tarini, MD, MS, Department of Pediatrics, University of Iowa, 200 Hawkins Dr, BT-1311, Iowa City, IA 52242. E-mail: beth-tarini@uiowa.edu

References

1. Petros M. Revisiting the Wilson-Jungner criteria: how can supplemental criteria guide public health in the era of genetic screening? Genet Med 2012;14:129-34.
2. Kemper AR, Green NS, Calonge N, Lam WK, Comeau AM, Goldenberg AJ, et al. Decision-making process for conditions nominated to the recommended uniform screening panel: statement of the US Department of Health and Human Services Secretary's Advisory Committee on Heritable Disorders in Newborns and Children. Genet Med 2014;16:183-7.
3. Paul DB. Patient advocacy in newborn screening: continuities and discontinuities. Am J Med Genet C Semin Med Genet 2008;148C:8-14.
4. Miller FA, Hayeems RZ, Bombard Y, Cressman C, Barg CJ, Carroll JC, et al. Public perceptions of the benefits and risks of newborn screening. Pediatrics 2015;136:e413-23.
5. Bunnell ME, Tarini BA, Petros M, Goldenberg AJ, Arjunan A, Wicklund C. Biobank participant support of newborn screening for disorders with variable treatment and intervention options. J Community Genet 2016;7:291-302.
6. Lipstein EA, Nabi E, Perrin JM, Luff D, Browning MF, Kuhlthau KA. Parents' decision-making in newborn screening: opinions, choices, and information needs. Pediatrics 2010;126:696-704.
7. Hayeems RZ, Miller FA, Bombard Y, Avard D, Carroll J, Wilson B, et al. Expectations and values about expanded newborn screening: a public engagement study. Health Expect 2015;18:419-29.
8. Alexander D, van Dyck PC. A vision of the future of newborn screening. Pediatrics 2006;117:S350-4.
9. Bailey DB Jr. Newborn screening for fragile X syndrome. Ment Retard Dev Disabil Res Rev 2004;10:3-10.
10. Louviere JJ, Flynn TN, Marley AA. Best-worst scaling: theory, methods and applications. Cambridge: Cambridge University Press; 2015.
11. Potoglou D, Burge P, Flynn T, Netten A, Malley J, Forder J, et al. Best-worst scaling vs discrete choice experiments: an empirical comparison using social care data. Soc Sci Med 2011;72:1717-27.
12. Flynn TN, Louviere JJ, Peters TJ, Coast J. Best-worst scaling: what it can do for health care research and how to do it. J Health Econ 2007;26:171-89.
13. Pollitt A, Potoglou D, Patil S, Burge P, Guthrie S, King S, et al. Understanding the relative valuation of research impact: a best-worst scaling experiment of the general public and biomedical and health researchers. BMJ Open 2016;6:e010916.
14. Cheung KL, Wijnen BF, Hollin IL, Janssen EM, Bridges JF, Evers SM, et al. Using best-worst scaling to investigate preferences in health care. Pharmacoeconomics 2016;34:1195-209.
15. Bridges J, Hauber A, Marshall D, Lloyd A, Prosser L, Regier D, et al. A checklist for conjoint analysis applications in health: report of the ISPOR Conjoint Analysis Good Research Practices Taskforce. Value Health 2011;14:403-13.
16. Gollust SE, Niederdeppe J, Barry CL. Framing the consequences of childhood obesity to increase public support for obesity prevention policy. Am J Public Health 2013;103:e96-102.
17. Pynnonen MA, Handelsman JA, King EF, Singer DC, Davis MM, Lesperance MM. Parent perception of newborn hearing screening: results of a US National Survey. JAMA Otolaryngol Head Neck Surg 2016;142:538-43.

18. McFarlane M, Brookmeyer K, Friedman A, Habel M, Kachur R, Hogben M. GYT: Get Yourself Tested campaign awareness: associations with sexually transmitted disease/HIV testing and communication behaviors among youth. Sex Transm Dis 2015;42:619-24.
19. Flynn TN, Louviere JJ, Peters TJ, Coast J. Best–worst scaling: what it can do for health care research and how to do it. J Health Econ 2007;26:171-89.
20. Bech M, Gyrd-Hansen D. Effects coding in discrete choice experiments. Health Econ 2005;14:1079-83.
21. Campbell E, Ross LF. Parental attitudes regarding newborn screening of PKU and DMD. Am J Med Genet A 2003;120:209-14.
22. Helton JL, Harmon RJ, Robinson N, Accurso FJ. Parental attitudes toward newborn screening for cystic fibrosis. Pediatr Pulmonol 1991;11:23-8.
23. Lang CW, Stark AP, Acharya K, Ross LF. Maternal knowledge and attitudes about newborn screening for sickle cell disease and cystic fibrosis. Am J Med Genet A 2009;149:2424-9.
24. Rothwell E, Anderson RA, Swoboda KJ, Stark L, Botkin JR. Public attitudes regarding a pilot study of newborn screening for spinal muscular atrophy. Am J Med Genet A 2013;161:679-86.
25. Skinner D, Sparkman KL, Bailey DB Jr. Screening for fragile X syndrome: parent attitudes and perspectives. Genet Med 2003;5:378.
26. Etchegary H, Dicks E, Green J, Hodgkinson K, Pullman D, Parfrey P. Interest in newborn genetic testing: a survey of prospective parents and the general public. Genet Test Mol Biomarkers 2012;16:353-8.
27. Ross LF. Mandatory versus voluntary consent for newborn screening? Kennedy Inst Ethics J 2010;20:299-328.
28. Carman KL, Mallery C, Maurer M, Wang G, Garfinkel S, Yang M, et al. Effectiveness of public deliberation methods for gathering input on issues in healthcare: results from a randomized trial. Soc Sci Med 2015;133:11-20.
29. Wang G, Gold M, Siegel J, Sofaer S, Yang M, Mallery C, et al. Deliberation: obtaining informed input from a diverse public. J Health Care Poor Underserved 2015;26:223-42.
30. Abelson J, Forest P-G, Eyles J, Smith P, Martin E, Gauvin F-P. Deliberations about deliberative methods: issues in the design and evaluation of public participation processes. Soc Sci Med 2003;57:239-51.
31. Maer L. Citizens' juries. Parliament and Constitution Centre. http://researchbriefings.files.parliament.uk/documents/SN04546/SN04546.pdf. Accessed April 18, 2018. 2007.
32. Harrison M, Rigby D, Vass C, Flynn T, Louviere J, Payne K. Risk as an attribute in discrete choice experiments: a systematic review of the literature. Patient 2014;7:151-70.
33. Mitchell RC, Carson RT. Using surveys to value public goods. Washington (DC): Resources for the Future; 1989.

Appendix

Survey Introduction

State newborn screening programs exist in every state. They make sure babies are screened for certain inherited diseases before they leave the hospital. These screening programs can help identify diseases that a baby could develop within a few days, months, or years even though the baby might appear healthy at birth. Early detection and treatment of these diseases can prevent brain damage, physical disabilities, costly medical care, and death.

State newborn screening programs are run by state public health departments and are funded by taxpayer dollars. Every state has its own newborn screening program, and each program can decide for which diseases in newborns it will screen. If new screening tests become available to detect a disease, a state newborn screening program might consider screening newborns for that disease.

When we refer to **disease** throughout the survey, we are not asking you to think about any specific disease, but just diseases in general. Many diseases screened for by state newborn screening programs are **inherited** or are **genetic**.

Again, we are interested in your opinions about what types of diseases you think states should screen babies for at birth.

Table III. Attribute levels and likelihood that respondent's will recommend screening for a newborn screening profile: generalized estimating equation logit model

Attribute-level descriptions	Coef. (SE)	*P* value
Intercept	**0.657 (0.084)**	**.000**
No. of babies diagnosed (continuous)	**0.043 (0.020)**	**.034**
Chance test is wrong (continuous)	–0.001 (0.021)	.951
Cost of testing (continuous)	0.016 (0.021)	.452
Likelihood of symptoms (continuous)	–0.023 (0.019)	.219
Seriousness of symptoms: able to do all daily activities	–0.076 (0.078)	.327
Seriousness of symptoms: able to do most daily activities	–0.101 (0.071)	.156
Seriousness of symptoms: able to do some daily activities	0.125 (0.081)	.123
Seriousness of symptoms: able to do no daily activities	0.052 (0.068)	.446
Expected to die: infancy	–0.066 (0.096)	.494
Expected to die: childhood	0.126 (0.065)	.052
Expected to die: adolescence or early adulthood	–0.156 (0.083)	.058
Expected to die: adulthood	0.097 (0.076)	.201
Treatment starts: during the first year of life	–0.020 (0.065)	.752
Treatment starts: during childhood	–0.134 (0.074)	.071
Treatment starts: during adolescence	**0.154 (0.079)**	**.049**
Success of treatment: will cure the disease	0.005 (0.080)	.955
Success of treatment: will not cure the disease, prevents progression	–0.104 (0.072)	.149
Success of treatment: will not cure the disease, slows progression	–0.081 (0.079)	.309
Success of treatment: no treatment is available	0.180 (0.120)	.134
Side effects of treatment: may die	0.087 (0.084)	.304
Side effects of treatment: serious side effects	0.070 (0.086)	.418
Side effects of treatment: mild side effects	0.093 (0.075)	.211
Side effects of treatment: none	**–0.250 (0.106)**	**.018**
Impact: will know about disease sooner; treatment more successful	–0.044 (0.056)	.437
Impact: will know about disease sooner; no change of treatment success	0.044 (0.056)	.437

Values in bold indicate statistically significant associations.