



Development of a newborn screening tool based on bivariate normal limits: using psychosine and galactocerebrosidase determination on dried blood spots to predict Krabbe disease

Thomas J. Langan, MD¹, Joseph J. Orsini, PhD², Kabir Jalal, PhD³, Amy L. Barczykowski, MS³, Maria L. Escolar, MD⁴, Michele D. Poe, PhD⁴, Chad K. Biski, BS² and Randy L. Carter, PhD³

Purpose: Newborn screening for Krabbe disease (KD) originated in New York State in 2006 but has proven to have a high false positive rate and low positive predictive value. To improve accuracy of presymptomatic prediction, we propose a screening tool based on two biomarkers, psychosine and galactocerebrosidase enzyme activity (GalC).

Methods: We developed the tool using measures from dried blood spots of 166 normal newborns and tested it on dried blood spot measures from 15 newborns who later developed KD, 8 newborns identified as “high risk” by the New York screening protocol but were disease-free at follow-up, and 3 symptomatic children with onset before 4 years of age. The tool was developed from the (1–10⁻⁶)100% prediction region of the natural logarithms of psychosine and GalC measures, assuming bivariate normality, and their univariate normal limits.

Results: Krabbe disease was predicted correctly for every patient who developed symptoms in infancy or early childhood. None of the high-risk patients were incorrectly identified as having early KD.

Conclusion: Bivariate analysis of psychosine and GalC in newborn blood spots can accurately predict early Krabbe symptoms, control false positive rates, and permit presymptomatic treatment.

Genetics in Medicine (2019) 21:1644–1651; <https://doi.org/10.1038/s41436-018-0371-3>

Keywords: Krabbe disease; psychosine; galactocerebrosidase; newborn screening; biomarkers

INTRODUCTION

The “psychosine hypothesis” was advanced by Kunihiko Suzuki in 1972¹ to explain a central paradox about the pathogenesis of Krabbe leukodystrophy (OMIM 245200): Krabbe disease (KD) is a lysosomal disorder, but there is no accumulation of galactosyl-ceramide, the deficient enzyme’s presumably pathogenetic substrate.^{1,2} The amounts present are much lower in KD brains than in the brains of unaffected children.^{1,2} There is also a striking and rapid disappearance of the white matter-producing oligodendroglial cells early in the course of KD.²

Suzuki consequently hypothesized that galactosylsphingosine, or psychosine (PSY), accumulates in KD and has cytotoxic effects that explain the critical paradox about its unique pathological and biochemical character.² This elegant hypothesis remains the prevailing explanation of the pathogenesis of KD.

Only recently has the possibility emerged that PSY, in addition to its apparently key role in pathogenesis, is a diagnostically useful biomarker that may predict the symptoms of KD.^{3–5} Identification of a predictive biomarker could aid in resolution of the current newborn screening (NBS) controversy for KD.^{6,7} The largest experience with screening for KD is that of New York State, but after 10 years of screening almost 2 million infants, it was found that the screening algorithm had a high false positive rate resulting in an extremely low positive predictive value (PPV) of 1.4%.⁶ Wasserstein et al. concluded that based upon the New York experience, improvements in NBS protocols are needed before widespread adoption of Krabbe NBS.^{6,7} Other studies, however, have shown that incorporating PSY into NBS protocols may substantially reduce the false positive rate.^{3–5} Furthermore, KD screening has spread to several other states,⁸ reinforcing the urgent need for an improved newborn screening model.

¹Department of Neurology, School of Medicine and Biomedical Sciences, University at Buffalo, Buffalo, NY, USA; ²Newborn Screening Program, Wadsworth Center, New York State Department of Health, Albany, NY, USA; ³Department of Biostatistics, Population Health Observatory, School of Public Health and Health Professions, University at Buffalo, Buffalo, NY, USA; ⁴The Program for the Study of Neurodevelopment in Rare Disorders, Children’s Hospital Pittsburgh of UPMC, Pittsburgh, PA, USA. Correspondence: Thomas J. Langan (tjlangan@buffalo.edu)

Submitted 13 September 2018; accepted: 7 November 2018
Published online: 14 December 2018

Psychosine can be determined in dried blood spots, and it has been reported recently to be elevated in several newborn spots from infants who developed early infantile Krabbe disease (EIKD).^{3,5} However, equivocal values were noted in some high-risk and later-onset patients.⁵ PSY has not been established as a stand-alone test to predict KD symptoms without consideration of additional diagnostic tests.⁵ Krabbe NBS still requires enzyme determination and additional tests, comprising either the second tier of study of the original newborn spots analyzed for KD, or secondary additional diagnostic testing performed by physicians who later see the infants after a positive newborn test.^{4,5,8}

The first tier of Krabbe newborn screening is analysis of the activity of galactocerebrosidase (GalC), the enzyme considered to be deficient in all cases of KD.^{8,9} However, GalC measured in leukocytes, as determined during the New York KD screening experience, did not predict onset or type of symptoms.¹⁰ Even though improvements in the GalC leukocyte assay may be feasible,¹¹ there is no current evidence that GalC activity in leukocytes alone can predict symptoms with sufficient accuracy to obviate the significant false positive problem in Krabbe NBS.

Because of the high false positive rate of newborn screening for KD, we previously investigated the potential for a bivariate analysis of PSY and GalC, as measured on the newborn screen dried blood spot specimens, to improve the false positive rate. Under the assumption that $\log_e(\text{GalC})$ and $\log_e(\text{PSY})$ have a bivariate normal distribution, a tool based on bivariate normal limits (BVNL) was developed from aggregate statistics previously reported in the literature. The tool showed great potential for highly sensitive and specific identification of EIKD cases.⁴ Lab-specific standardization of $\log_e(\text{GalC})$ broadened the successful application of the tool to data from multiple labs. Without the benefit of simultaneously measured GalC and PSY from a normative sample, however, it was not possible to test the bivariate normal assumption, and estimation of the distribution's parameters was necessarily ad hoc in nature. Furthermore, the questions of whether lab-specific standardization of $\log_e(\text{PSY})$ and/or time-specific standardization of either biomarker are necessary for general application remain open.

The goals of the current study are to finalize development and validation of the BVNL-based diagnostic tool for NYS based on simultaneous measurements of PSY and GalC from a normative sample and to continue investigating questions relevant to the more general applicability of the tool. The finalized tool could potentially prevent the anxiety experienced by parents whose infants screen positively but who are not destined to develop infantile KD, as well as the expense and potential discomfort of additional diagnostic testing of these infants.^{4,6,8} We report here on the development and accuracy of the finalized BVNL tool based on analysis of dried blood spots (DBS), including newborn spots that have been accumulated as part of the expansion of the World Wide Registry (WWR) for KD.^{12–14}

MATERIALS AND METHODS

Bloodspots and patient data

The WWR contains detailed genetic and phenotypic information regarding 198 KD patients.^{12–14} A de-identified version of the WWR data has been contributed to the Longitudinal Pediatric Data Resource (LPDR) of the Newborn Screening Translation Research Network (NBSTRN).^{15,16} This study has been approved by the University at Buffalo's IRB (IRB 030-385325). Each patient in the World Wide Registry has a signed consent form on file. Residual DBS were obtained, where available, from the newborn screening programs of states where registrants of the WWR were born. Fifteen of the newborn DBS were available for retrieval. Each of these DBS were analyzed for both PSY and GalC. These included the DBS of 12 infants who later developed EIKD, defined as onset of symptoms by age 6 months^{12,17} and 3 later-onset cases who developed symptoms prior to 30 months. We shall refer to these 15 cases studied as *early childhood KD*.

The analyses also included 174 DBS from New York State's newborn screening program. Of these, 166 were from normal newborns who were born in October 2016 and 8 were from newborns considered to be at high risk for KD, according to the current New York State protocol. These 8 infants were defined as high-risk due to low GalC enzyme levels on diagnostic testing consistent with the development of EIKD. However, they did not have genotypes that predicted the early onset of symptoms.^{6,9} All 8 remained symptom-free during the 6 to 24 month follow-up period.⁶ Three additional DBS were obtained from patients evaluated at the Program for the Study of Neurodevelopment in Rare Disorders at Children's Hospital of Pittsburgh of the University of Pittsburgh Medical Center (UPMC). These DBS were from symptomatic, early childhood, KD patients who developed symptoms before the age of 4 years.¹⁸

In summary, the samples analyzed in the current study consisted of 166 control newborn blood spots, 15 DBS from newborns who later developed KD, 8 high-risk New York cases, and 3 early childhood cases whose blood was collected early in their symptomatic course.

Additional normative samples comprised of either GalC or PSY measurements made at various times were available to address the need for time-specific standardization of either $\log_e(\text{GalC})$ or $\log_e(\text{PSY})$. Determinations of GalC enzyme activity and PSY were performed for each DBS as described previously using liquid chromatography and mass spectrometry methodologies.^{3,19,20}

All DBS analyses were conducted in the Wadsworth Laboratory of the New York State Department of Health using published protocols.⁵ Some cases had additional determinations at Thomas Jefferson University that were performed on blood drawn as part of subsequent diagnostic testing. The results of these analyses are presented in Table 1.

Statistical analyses

Normal probability plots of the natural logarithms of PSY and GalC observations in the normative sample ($n = 166$)

Table 1 GalC/PSY and age at onset for KD cases and high-risk NBS patients

Early childhood cases

Age at onset	Wadsworth measures		Jefferson measures		Standardized Wadsworth	
	GalC	PSY	GalC	PSY	GalC	PSY
4	0.17	109.5	N/A	N/A	-5.28	10.33
1	0.27	24.54	N/A	N/A	-4.52	7.6
4	0.27	52.04	N/A	N/A	-4.52	8.97
3	0.17	65.9	0.04	N/A	-5.28	9.4
0.5	0.17	38.96	0.1	N/A	-5.28	8.45
5	0.17	57.27	0.02	N/A	-5.28	9.15
5	0.17	8.25	0	N/A	-5.28	5.62
2	0.02	5.18	N/A	N/A	-8.78	4.77
5	0.07	5.4	0.05	N/A	-6.73	4.85
5	0.1	6.03	0.06	N/A	-6.15	5.05
4	0.69	8.08	0.06	N/A	-2.99	5.58
3	0.17	41.32	N/A	N/A	-5.28	8.55
29	0.22	7.37	0.17	N/A	-4.86	5.42
15.5	0.32	13	0	N/A	-4.24	6.44
11	0.12	40.39	0.06	N/A	-5.85	8.51
Mean (SD)	0.21 (.15)	32.22 (30.07)	0.06 (.05)		-5.35 (1.28)	7.25 (1.92)
High-risk NBS						
Wadsworth measures						
Age at onset	GALC	PSY	GALC	PSY	GALC	PSY
N/A	0.39	3.23	N/A	N/A	-3.92	3.92
N/A	0.48	3.64	N/A	N/A	-3.58	4.13
N/A	0.4	6.72	N/A	N/A	-3.88	5.25
N/A	0.58	2.47	N/A	N/A	-3.27	3.42
N/A	0.41	2.17	N/A	N/A	-3.84	3.19
N/A	0.36	1.28	N/A	N/A	-4.05	2.23
N/A	0.51	0.94	N/A	N/A	-3.48	1.68
N/A	0.44	1.26	N/A	N/A	-3.72	2.2
Mean (SD)	0.45 (.07)	2.71 (1.88)			-3.72 (.26)	3.25 (1.19)

While independent testing labs may disagree on GalC measurements, accounting for time variation when standardizing GalC shows early childhood cases exhibit lower standardized values than high-risk and normal patients.

GalC galactocerebrosidase, KD Krabbe disease, NBS newborn screening, PSY psychosine.

provided necessary (although, technically, not sufficient) evidence for the bivariate normality assumption. The points in both plots followed a straight-line pattern consistent with the assumptions that each has a normal distribution. Thus, development of the proposed NBS tool was based on multivariate normal distribution theory²¹ and formulas for $(1-p)100\%$ prediction regions.²² Prediction regions (ellipses) were calculated from univariately standardized $\log_e(\text{GalC})$ and $\log_e(\text{PSY})$ values of the 166 normal newborns. The prediction ellipse for $p = 10^{-6}$ was then used to develop a BVNL-based newborn screening tool as previously described.⁴

Tolerance for false positive rate (fp) dictated thresholds for the definitions of “low” GalC, “high” PSY, and our choice of p when developing a BVNL NBS tool. Thresholds of -2.9 and 2.9 for standardized $\log_e(\text{GalC})$ and $\log_e(\text{PSY})$, respectively, and $p = 10^{-6}$ were chosen in an attempt to control fp at approximately 10^{-7} and specificity at $(1-10^{-7})100\%$.

The $(1-10^{-7})100\%$ BVNL tool, so defined, was applied to appropriately standardized observations (see below) to test the 166 normal newborns, 15 KD cases from the WWR, 8 “high risk” but disease-free newborns, and 3 KD cases measured postsymptomatically for KD. The specificity of the tool for correctly identifying non-KD cases was calculated separately for the samples of 166 normal newborns and 8 “high-risk” non-KD cases as the percentage of each sample that tested negative for KD. The sensitivity of the tool was calculated separately from results of its application to the 12 EIKD and 3 later onset WWR patients as the percentage of each sample that tested positive. The sensitivity of the tool when testing postsymptomatic cases was similarly calculated from the results of its application to the three EIKD patients from the UPMC.

A newborn screening program that predicts KD for all newborns falling outside the $(1-10^{-7})100\%$ BVNL should

expect approximately 1 falsely predicted early childhood KD in every 10 million normal newborns. To determine whether this targeted *fp* and associated specificity were achieved, approximately, 100 million observations were generated from the standardized version of the following multivariate normal distribution:

$$\begin{pmatrix} \log_e(\text{GalC}) \\ \log_e(\text{PSY}) \end{pmatrix} \sim \text{MVN} \left[\begin{pmatrix} 0.67 \\ -1.00 \end{pmatrix}, \begin{pmatrix} 0.80 & -0.022 \\ -0.22 & 0.26 \end{pmatrix} \right],$$

where the mean vector and covariance matrix are sample estimates from the sample of 166 normal newborns measured simultaneously in October 2016. The prediction rule then was applied to the simulated observations. Treating generated points as observations from the actual standardized distribution of normal newborns, the false positive rate of the BVNL tool was calculated as the percentage of these 100 million simulated data points that fell inside the $(1-10^{-7})$ BVNL defined above. The specificity of the tool was then approximately estimated as 1.0 minus the calculated *fp*. The calculated *fp* was compared with the nominal $fp = 10^{-7}$ to illustrate that the targeted nominal rate is indeed approximately achieved.

The *fp* calculated from the simulation results, the average prevalence reported in articles reviewed by Foss *et al.*,²³ and the sensitivity estimated in the current study were substituted into the formula for PPV²⁴ to obtain an estimate of the PPV of our proposed tool.

To test the need for temporally specific standardization, the data were categorized according to when the sample was analyzed at the Wadsworth lab. Group 1 consists of the 166 joint measurements of GalC and PSY, from October 2016. Group 2 consists of 30 joint measurements from spring 2017 (GalC in March and PSY in May). Group 3 consists of 1453 GalC measurements from November 2017. Lastly, group 4 consists of 57,796 GalC measurements from May 2009. To assess whether the mean of the distribution of these biomarkers varied from time to time, analysis of variance was performed on GalC measures across the four normative groups and on PSY measures across groups 1 and 2.

The 166 normal newborns' $\log_e(\text{GalC})$ and $\log_e(\text{PSY})$ measures and the simulated observations were standardized using group 1 means and standard deviations. Otherwise the statistics used to standardize observations were based on these

comparisons and the best measurement time match for each of our study samples. Specifically, the 15 KD cases, 8 high-risk NYS NBS, and UPMC patients' $\log_e(\text{GalC})$ measures were standardized using group 3 $\log_e(\text{GalC})$ statistics, and their $\log_e(\text{PSY})$ measures were standardized using the weighted mean and pooled standard deviation across the two normative samples of $\log_e(\text{PSY})$ in groups 1 and 2. Observations from the three UPMC patients' simulated observations were standardized using group 1 statistics.

RESULTS

Temporal change in GalC enzyme activity determined in dried blood spots

The time-specific samples are described in Table 2. $\log_e(\text{GalC})$ in normal newborns varied significantly across time-varying samples ($P < 0.0001$), while $\log_e(\text{PSY})$ did not ($P = 0.2634$).

BVNL classification of normal newborn bloodspots

Figure 1 illustrates results of the application of the proposed NBS tool to the 166 normal newborns from New York State. All 166 data points fall within normal limits; i.e., false positive rate = 0% and specificity = 100%.

BVNL results for Krabbe patients

Figure 1 also illustrates the results from the application of the BVNL NBS tool to measures from newborn DBS of the 15 newborns who developed early childhood KD. The BVNL tool correctly identified each of the 12 EIKD patients and each of the 3 later-onset early childhood patients. The ages of symptom onset varied from 1 to 29 months. Thus, the tool was 100% sensitive for early identification of early childhood KD.

All three postsymptomatic cases followed in Pittsburgh were correctly diagnosed by the BVNL tool (sensitivity = 100%; see Table 3). These children developed symptoms prior to 4 years of age.

BVNL results for high-risk New York cases

None of the eight high-risk cases fell outside the $(1-10^{-7})\%$ BVNL ellipse (Fig. 2) and thus none were incorrectly predicted to develop early childhood KD. All eight were below the univariate normal limit for $\log_e(\text{GalC})$ and five were above the univariate normal limit for $\log_e(\text{PSY})$. All remain free of Krabbe symptoms for over 2 years of observation.

Table 2 GalC and PSY measurement group comparison

Sample date	N	GalC		PSY		LN(GalC)		LN(PSY)	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
Oct 2016	166	2.83	2.43	0.42	0.20	0.67	0.90	-1.00	0.51
Mar 2017, May 2017	30	5.73	2.49	0.50	0.25	1.66	0.44	-0.88	0.73
Nov 2017	1453	5.15	3.66	N/A	N/A	1.45	0.61	N/A	N/A
May 2009	59,796	4.13	3.17	N/A	N/A	1.24	0.58	N/A	N/A

GalC galactocerebrosidase, PSY psychosine.

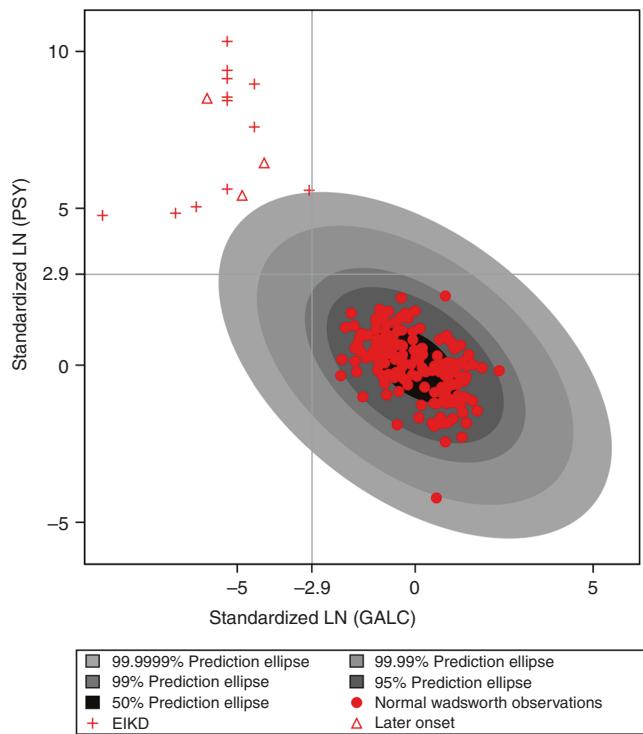


Fig. 1 Bivariate normal limits estimated from wadsworth normative sample univariate standardization of natural log of psychosine/GalC with wadsworth tested NBS from EIKD ($N = 12$), and later onset ($N = 3$) patients. Bivariate normal limits (BVNL) estimated from Wadsworth normative sample (circles; $N = 166$), with Wadsworth tested NBS from early-onset cases, including onset less than 6 months (EIKD/crosses; $N = 12$) and less than 30 months (later onset/triangles; $N = 3$). Points in the upper left quadrant and outside the ellipse represent positive screens for Krabbe disease (KD). EIKD early infantile Krabbe disease, GalC galactocerebrosidase, NBS newborn screening, PSY psychosine.

Specificity, sensitivity, false positive rate, and positive predictive value

Table 3 summarizes the data presented thus far and examines the accuracy of the BVNL tool for newborn KD screening. In terms of predicting symptoms from the newborn blood spots of infants destined to develop symptoms in early childhood, the tool demonstrated 100% sensitivity, as did psychosine alone (Fig. 1). Results provided from analysis of the 166 controls indicate 100% specificity, whereas the stand-alone PSY test had an unacceptable false positive rate (Fig. 2).

Results of BVNL classification of points in the simulation sample of 100 million observations generated from the estimated bivariate normal distribution for normal newborns are presented in Table 3. The estimated rate of false positives was 1.1 per 10 million normal newborns. (This translates to an estimated specificity of 99.9999%) Thus, the anticipated rate of false positives if each of the approximately 4 million normal newborns each year in the United States was screened, would then be only 1 every 2.5 years. The PPV would be 98.5%.

DISCUSSION

The results above establish the validity of bivariate analysis of PSY and GalC enzyme levels measured on newborn blood spots and show that this analysis can highly accurately predict the onset of symptoms of KD in infancy or early childhood. For children enrolled in the WWR identified as having high risk and for whom the exact age of onset is known, symptoms appeared between 1 and 29 months of age (Fig. 1). Early and accurate diagnosis of Krabbe symptoms presents a critical advance in management of this disorder because the optimal outcomes of hematopoietic stem cell transplantation result only if this treatment is applied presymptomatically.^{18,25} The efficacy of treatment will be enhanced by the earliest possible diagnosis.

The proposed BVNL tool has a PPV of 98.5% (Table 3). It would therefore virtually eliminate the false positive problem that has plagued KD newborn screening since its inception in New York State in 2006.^{4,6} This predictive potential also exceeds that of 80% reported previously as a potential outcome of newborn screening for lysosomal disorders.²⁶

Demonstrating that use of a newborn diagnostic tool based upon these two parameters is feasible will require a prospective study of a greater number of patients and normal newborns in states that screen for KD, because this disorder is quite rare. In the published New York State experience, only 5 cases were confirmed after screening over 2 million children.⁶ After screening almost 3 million newborns in NY State, an additional late infantile case has been identified and the false positive rate has improved, but remains a concern for this very rare disorder (J.J. Orsini, personal communication). Analysis of death certificates suggested a frequency of EIKD of 1 in 244,000 in the United States.²⁷ At the current time, only a few states are screening all newborns for Krabbe disease.⁸ Field testing of the BVNL tool may therefore be challenging, but based upon the current results, it is also timely.

High-risk cases such as those shown in Fig. 2 are extremely challenging for clinicians. Their low initial GalC enzyme levels provoke concern, but they do not appear to be at risk of EIKD even though they may develop later-onset variants.^{12,13,17,28} It has been estimated that there are over 140 distinct variants in the human gene encoding the GalC enzyme, with only a few predicting late-onset KD; genotype/phenotype correlation therefore remains elusive.^{17,29}

For these eight cases, only two were heterozygotes for the 30-kb deletion accompanied by a probable allele of late onset.^{9,29} The remainder had other putative late-onset alleles combined with variants of unknown significance. The BVNL tool's analyses of their newborn blood spots predicted correctly that they would remain free of symptoms of early childhood KD. Nevertheless, it is prudent to acknowledge the possibility that they may represent late-onset cases.^{6,12,13}

Figure 2 also illustrates an advantage of the BVNL tool over analysis of either PSY or GalC as univariate measures, or in a combined analysis lacking the elliptical portion of the BVNL. All of these cases were below the cutoff for GalC, so univariate

Table 3 Accuracy of bivariate normal limits (BVNL) for test groups

Test group	N	Within normal limits	Within high-risk region	% Correct	Accuracy measure
WWR EIKD	12	0	12	100	Sensitivity
WWR later onset	3	0	3	100	Sensitivity
UPMC later onset	3	0	3	100	Sensitivity
High-risk NYS NBS	8	8	0	100	Specificity
Simulated normals	100000000	99999989	11	99.999989	Specificity

EIKD early infantile Krabbe disease, NBS newborn screening, NYS New York State, UPMC University of Pittsburgh Medical Center, WWR World Wide Registry of Krabbe disease.

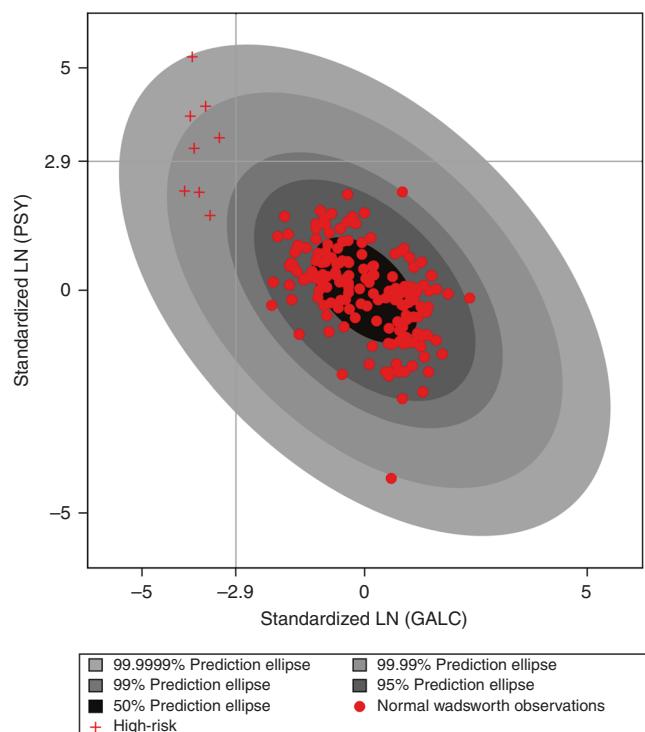


Fig. 2 Bivariate normal limits estimated from wadsworth normative sample univariate standardization of natural log of psychosine/GALC with wadsworth tested NBS from high-risk patients ($N = 8$). Bivariate normal limits (BVNL) estimated from Wadsworth normative sample (circles; $N = 166$), with Wadsworth tested NBS from high-risk patients (crosses; $N = 8$). Points in the upper left quadrant and outside the ellipse represent positive screens for Krabbe disease (KD). EIKD early infantile Krabbe disease, Ga/C galactocerebrosidase, NBS newborn screening, PSY psychosine.

analysis of the enzyme alone would cause false assignment or risk. But in addition, five of these eight high-risk cases were also above the upper limit for PSY (Fig. 2). Therefore, either separate or joint consideration of GalC or PSY relative to univariate thresholds of -2.9 and 2.9 , respectively, for these five cases would result in an incorrectly assigned risk of developing EIKD symptoms and place them at risk for receiving unneeded treatment. One could define more stringent univariate thresholds to solve this problem for these eight newborns but at the risk of identifying fewer true cases. The general point here is that the BVNL can be expected to

perform better (i.e., lower false positive rate given the same sensitivity) than simple application of univariate thresholds jointly. This is due to the fact that the BVNL takes advantage of additional information about the bivariate distribution of the biomarkers.

It should be emphasized that the tool can be adjusted to conform to different tolerances for false positives by appropriate choice of thresholds and p when defining the BVNL. Obviously, it is desirable that the tool have 100% sensitivity, so that no child who needs treatment will fail to receive it, and 0% false positive rate (100% specificity), so that no child who will not get KD will be subjected to unnecessary treatment. While this goal is not theoretically possible to attain in the entire population, our ($1 - 10^{-7}$) BVNL-based tool achieved it in the samples of 166 normal newborns and 18 early childhood KD cases in this study (See Table 3).

How might this tool be used in practice? Knowing that a child will develop KD in infancy or early childhood would result either in early treatment, or in extremely vigilant follow-up.^{5,6} At the current time, the decision for early treatment also considers other diagnostic information, such as genotype and the results of such ancillary tests as magnetic resonance imaging (MRI), spinal fluid protein, and evoked responses. But it is difficult to acquire this information in a timely fashion. Few of the children followed as part of the New York experience had the full recommended diagnostic panel in New York.⁶ There is clearly a need for fewer and more accurate tests and for better prediction from newborn blood spot data.

If the diagnostic paradigm after a positive newborn screen for KD could be simpler and more accurate, it would be reasonable to expect less parental anxiety and fewer painful medical procedures for the infants.^{6,7,30} In addition, improved and earlier prediction of KD symptoms would result in a substantial savings of health-care expenses. The cost of the New York State KD newborn screening program has been significant.³⁰ Screening for all lysosomal disorders in the United States was estimated to cost \$28–\$140 million annually with a false positive rate of at least 0.1%.²⁶ It would therefore be reasonable to expect that significant savings would result from reducing the significant false positive rate of Krabbe screening.

Can the BVNL tool challenge or augment current NBS KD paradigms? Postanalytical platforms for lysosomal NBS have

been defined successfully.^{26,31} These involve an analysis of six to ten enzymes.²⁶ Advantages of the BVNL tool include lacking requirements to share patient profiles and reference data,²⁶ and free availability to anyone with access to the R statistical program.³² Because there is considerable variability between state newborn programs,^{33,34} it is conceivable that some states would not be able to provide six or ten enzyme results for multiplex analysis. The BVNL tool would, in contrast, rely on rapid second-tier analysis of newborn blood spots for only two biomarkers, while allowing state screening programs more flexibility/control over Krabbe screening decisions.

It is apparent that the successful application of a statistical tool that relies on determining GalC enzyme activities, such as the BVNL analysis presented here, will over time require up-to-date standardization of log₁₀(GalC) observations as noted in our previous study⁴ (Table 2). The experience of the New York State NBS program indicates that certain features of blood spot acquisition, such as age of the patient and birth weight, as well as changes in reagent lots, can cause variation in enzyme levels at NBS (unpublished data). The reasons for the temporal drift in the enzyme activity are unknown.

The New York State NBS program currently collects normative samples of GalC measures. In addition, while the current data suggests that infants who are destined to develop early childhood KD will have standardized values that fall in the high-risk region of the BVNL, no data exists yet about the course of these measures as the disease progresses.

The psychosine determinations used here for development of the BVNL tool did not reflect temporal variation. These were determined from normal newborn blood spots and performed at the Wadsworth Laboratory of the New York State Department of Health. However, it must be acknowledged that other laboratories have reported predictive PSY values that differ from those used here.^{4,5,26}

Broader use of the tool would consequently require restandardization of PSY. The cost of PSY determination may also vary among states.

The oldest patient who developed symptoms with high risk determined from BVNL newborn blood spot analysis was 29 months old at symptom onset, but we do not know the late age limit for symptom development that can be predicted. The data do indicate that none of the asymptomatic high-risk patients from NY State had high-risk BVNL status. Presumably these high-risk patients are at risk for emergence of symptoms in later childhood or even adulthood.^{6,7} Could periodic reapplication of the BVNL tool at intervals as they get older predict symptoms in these possible late-onset cases? Years of follow-up on these children may be necessary to answer this question.

Successful application of the BVNL tool would presumably also facilitate development and testing of experimental therapies, for example, those combining transplantation with gene therapy.³⁵

The results presented here support application of the BVNL analysis of PSY and GalC prospectively to second-tier analysis

of blood spots after a positive KD screen in states that are routinely screening for this illness. If the potential of the tool to predict symptoms is confirmed by its prospective use in more infants, this would be a significant step forward in fulfilling goals that are presumably shared by all who care for children who either have KD or are at risk of developing it. These goals include accurately diagnosing KD, ensuring the prompt treatment of children who need it, and protecting those not at imminent risk from the cost, morbidity, and potential mortality of unnecessary treatment.^{5–7,30}

ACKNOWLEDGEMENTS

Supported by the National Institute of Child Health and Human Development, R-21 HD087818-01 (to T.J.L., R.L.C., A.L.B., and K.J.)

DISCLOSURE

The authors declare no conflicts of interest.

REFERENCES

- Miyatake T, Suzuki K. Globoid cell leukodystrophy: additional deficiency of psychosine galactosidase. *Biochem Biophys Res Commun*. 1972;48:539–543.
- Suzuki K. My encounters with Krabbe disease: a personal recollection of a 40-year journey with young colleagues. *J Neurosci Res*. 2016;94:965–972.
- Turgeon CT, Orsini JJ, Sanders KA, et al. Measurement of psychosine in dried blood spots—a possible improvement to newborn screening programs for Krabbe disease. *J Inher Metab Dis*. 2015;38:923–929.
- Carter RL, Wrabetz L, Jalal K, et al. Can psychosine and galactocerebrosidase activity predict early-infantile Krabbe's disease presymptomatically? *J Neurosci Res*. 2016;94:1084–1093.
- Escolar ML, Kiely BT, Shawgo E, et al. Psychosine, a marker for Krabbe phenotype and treatment effect. *Mol Genet Metab*. 2017;121:271–278.
- Wasserstein MP, Andriola M, Arnold G, et al. Clinical outcomes of children with abnormal newborn screening results for Krabbe disease in New York State. *Genet Med*. 2016;18:1235–1243.
- Dimmock DP. Should states adopt newborn screening for early infantile Krabbe disease? *Genet Med*. 2016;18:217–220.
- Orsini JJ, Saavedra-Matiz CA, Gelb MH, Caggana M. Newborn screening for Krabbe's disease. *J Neurosci Res*. 2016;94:1063–1075.
- Wenger DA, Luzi P, Rafi MA. Krabbe disease: are certain mutations disease-causing only when specific polymorphisms are present or when inherited in trans with specific second mutations? *Mol Genet Metab*. 2014;111:307–308.
- Jalal K, Carter R, Yan L, Barczykowski A, Duffner PK. Does galactocerebrosidase activity predict Krabbe phenotype? *Pediatr Neurol*. 2012;47:324–329.
- Liao HC, Spacil Z, Ghomashchi F, et al. Lymphocyte galactocerebrosidase activity by LC-MS/MS for post-newborn screening evaluation of Krabbe disease. *Clin Chem*. 2017;63:1363–1369.
- Duffner PK, Barczykowski A, Jalal K, Yan L, Kay DM, Carter RL. Early infantile Krabbe disease: results of the World-Wide Krabbe Registry. *Pediatr Neurol*. 2011;45:141–148.
- Duffner PK, Barczykowski A, Kay DM, et al. Later onset phenotypes of Krabbe disease: results of the world-wide registry. *Pediatr Neurol*. 2012;46:298–306.
- Langan TJ, Barczykowski AL, Dare J, Pannullo EC, Muscarella L, Carter RL. Evidence for improved survival in postsymptomatic stem cell-transplanted patients with Krabbe's disease. *J Neurosci Res*. 2016;94:1189–1194.
- Botkin JR, Lewis MH, Watson MS, et al. Parental permission for pilot newborn screening research: guidelines from the NBSTRN. *Pediatrics*. 2014;133:e410–417.
- Urv TK, Parisi MA. Newborn screening: beyond the spot. *Adv Exp Med Biol*. 2017;1031:323–346.
- Sakai N, Otomo T. Challenge of phenotype estimation for optimal treatment of Krabbe disease. *J Neurosci Res*. 2016;94:1025–1030.

18. Wright MD, Poe MD, DeRenzo A, Haldal S, Escolar ML. Developmental outcomes of cord blood transplantation for Krabbe disease: a 15-year study. *Neurology*. 2017;89:1365–1372.
19. Chuang WL, Pacheco J, Zhang XK, et al. Determination of psychosine concentration in dried blood spots from newborns that were identified via newborn screening to be at risk for Krabbe disease. *Clin Chim Acta*. 2013;419:73–76.
20. Orsini JJ, Kay DM, Saavedra-Matiz CA, et al. Newborn screening for Krabbe disease in New York State: the first eight years' experience. *Genet Med*. 2016;18:239–248.
21. Johnson R, Wichern D. Multivariate statistics, a practical approach. London: Chapman & Hall; 1988.
22. Chew V. Confidence, prediction, and tolerance regions for the multivariate normal distribution. *J Am Stat Assoc*. 1966;61:605–617.
23. Foss AH, Duffner PK, Carter RL. Lifetime risk estimators in epidemiological studies of Krabbe disease: review and Monte Carlo comparison. *Rare Dis*. 2013;1:e25212.
24. Altman DG, Bland JM. Diagnostic tests 2: predictive values. *BMJ*. 1994;309:102.
25. Escolar ML, Poe MD, Provenzale JM, et al. Transplantation of umbilical-cord blood in babies with infantile Krabbe's disease. *N Engl J Med*. 2005;352:2069–2081.
26. Minter Baerg MM, Stoway SD, Hart J, et al. Precision newborn screening for lysosomal disorders. *Genet Med*. 2017;20:847–854.
27. Barczykowski AL, Foss AH, Duffner PK, Yan L, Carter RL. Death rates in the U.S. due to Krabbe disease and related leukodystrophy and lysosomal storage diseases. *Am J Med Genet A*. 2012;158A:2835–2842.
28. Hossain MA, Otomo T, Saito S, et al. Late-onset Krabbe disease is predominant in Japan and its mutant precursor protein undergoes more effective processing than the infantile-onset form. *Gene*. 2014;534:144–154.
29. Saavedra-Matiz CA, Luzi P, Nichols M, Orsini JJ, Caggana M, Wenger DA. Expression of individual mutations and haplotypes in the galactocerebrosidase gene identified by the newborn screening program in New York State and in confirmed cases of Krabbe's disease. *J Neurosci Res*. 2016;94:1076–1083.
30. Lantos JD. Dangerous and expensive screening and treatment for rare childhood diseases: the case of Krabbe disease. *Dev Disabil Res Rev*. 2011;17:15–18.
31. 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;13:230–254.
32. Team RCR. A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2014.
33. Therrell BL, Adams J. Newborn screening in North America. *J Inherit Metab Dis*. 2007;30:447–465.
34. Mandl KD, Feit S, Larson C, Kohane IS. Newborn screening program practices in the United States: notification, research, and consent. *Pediatrics*. 2002;109:269–273.
35. Hu P, Li Y, Nikolaishvili-Feinberg N, et al. Hematopoietic stem cell transplantation and lentiviral vector-based gene therapy Krabbe's disease: present convictions and future prospects. *J Neurosci Res*. 2016;94:1152–1168.