

Open

Postanalytical tools improve performance of newborn screening by tandem mass spectrometry

Patricia L. Hall, PhD¹, Gregg Marquardt, MSS¹, David M.S. McHugh¹, Robert J. Currier, PhD², Hao Tang, PhD², Stephanie D. Stoway, BS¹ and Piero Rinaldo, MD, PhD¹

Purpose: The purpose of this study was to compare performance metrics of postanalytical interpretive tools of the Region 4 Stork collaborative project to the actual outcome based on cutoff values for amino acids and acylcarnitines selected by the California newborn screening program.

Methods: This study was a retrospective review of the outcome of 176,186 subjects born in California between 1 January and 30 June 2012. Raw data were uploaded to the Region 4 Stork Web portal as .csv files to calculate tool scores for 48 conditions simultaneously using a previously unpublished functionality, the tool runner. Scores for individual target conditions were deemed informative when equal or greater to the value representing the first percentile rank of known true-positive cases (17,099 cases in total).

Results: In the study period, the actual false-positive rate and positive predictive value were 0.26 and 10%, respectively. Utilization of the Region 4 Stork tools, simple interpretation rules, and second-tier tests could have achieved a false-positive rate as low as 0.02% and a positive predictive value >50% by replacing the cutoff system with Region 4 Stork tools as the primary method for postanalytical interpretation.

Conclusion: Region 4 Stork interpretive tools, second-tier tests, and other evidence-based interpretation rules could have reduced false-positive cases by up to 90% in California.

Genet Med advance online publication 29 May 2014

Key Words: cutoff values; newborn screening; postanalytical interpretive tools; second-tier test; tandem mass spectrometry

INTRODUCTION

Region 4 Stork (R4S) is a collaborative project for laboratory quality improvement of newborn screening by tandem mass spectrometry (MS/MS).¹ It began in 2004 and overlapped in time with the delineation of the recommended uniform screening panel.² Ten years later, R4S has engaged 1,050 users from 64 countries, and the true-positive database has exceeded 17,000 cases, with more than 1.2 million results. The key element of this project, which is now integrated into the Newborn Screening Translation Research Network (<https://www.nbstrn.org/>), is freely available, on-demand access to postanalytical tools designed to interpret analyte profiles of a single case.³ Tools generate a score reflecting the likelihood of a diagnosis based on condition-specific disease ranges of all informative analytes. Site-specific customization of these tools is available to correct for differences in analyte panels and sample preparation (derivatized versus underivatized method). In a recent study, R4S tools were credited for contributing to more effective diagnostic algorithms for very long-chain acyl-coenzyme A dehydrogenase deficiency.⁴

Despite more than 120,000 page views, in most instances, these tools have been utilized reactively after one or more analytes exceeded a chosen cutoff value, so their potential to prevent false-positive (FP) outcomes has not been investigated adequately. The impetus for this study comes from concerns

over poor specificity that have been voiced repeatedly.^{5–9} In R4S, the main barrier to a comparative assessment has been the necessity to process one case with one tool at a time. The introduction of two high-throughput portals has provided an opportunity for a comparison between the traditional interpretation model, based on analyte cutoff values, and one based on the scores of the R4S interpretive tools.

MATERIALS AND METHODS

Study cohort

The population study included 176,186 first specimens of subjects born in California between 1 January and 30 June 2012. The following exclusion criteria were applied to minimize skewing the count of cases with abnormal results based on cutoff-based interpretation: (i) birth weight <1,800 g; (ii) birth weight >1,800 g and age at collection <24 hours; and (iii) birth weight >1,800 g and age at collection >7 days. This study was approved by the Committee for the Protection of Human Subjects of the California Department of Public Health (protocol 13-10-1387).

Current practice in California is to assign profiles that are not completely normal to one of three categories: out of range, review, and presumptive positive. Out of range indicates a single result exceeding the analyte cutoff value but not meeting additional interpretation criteria (e.g., one or more ratios). These findings are considered of negligible clinical significance;

¹Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, Minnesota, USA; ²California Department of Public Health, Richmond, California, USA.
Correspondence: Piero Rinaldo (rinaldo@mayo.edu)

Submitted 28 February 2014; accepted 30 April 2014; advance online publication 29 May 2014. doi:10.1038/gim.2014.62

they are not reviewed individually, but a notification report is sent to the primary-care provider with no recommendation to pursue follow up and additional testing. A review category is triggered by multiple abnormalities according to the criteria described above. Further evaluation of the overall profile and demographic information is performed by an expert reviewer; in the vast majority of cases the resolution is the same of the out of range category, and a similar report is generated. Otherwise, a review case is escalated to become a presumptive positive, i.e., an abnormal result, and follow-up is initiated.

MS/MS method

The California Department of Public Health screening program relies analytically on an underivatized method based on a commercial kit (StepOne Newborn Screening; Perkin Elmer, Waltham, MA). The analyte panel includes 13 amino acids and 38 acylcarnitines, integrated by 61 calculated ratios (23 for amino acids and 38 for acylcarnitines). The count of active cutoff values is 45, and 22% of them meet the R4S validation criteria.¹ Currently, the California Department of Public Health

does not routinely perform any of the following second-tier tests (2TTs): (i) homocysteine, methylmalonic acid, and methylcitric acid;^{10,11} (ii) allo-isoleucine;¹² and (iii) 3-hydroxyglutaric acid, glutaric acid, and ethylmalonic acid.

R4S data and tools

R4S is a custom-designed and coded application for the processing of laboratory data based on numerical results.¹⁻³ The application uses the Microsoft ASP.NET Framework version 3.5 and Microsoft SQL Server 2008 R2. As of 17 April 2014, the true-positive database includes 17,098 newborns, identified according to local protocols and/or professional guidelines,¹³ corresponding to 1,209,481 data points. Through an interface called tool builder, this multivariate pattern recognition software creates different types of postanalytical interpretive tools. These tools are designed to provide an evidence-based answer to one of three types of questions: (i) a yes or no situation (the one-condition tool: does a patient have or not have a specific condition?); (ii) a differential diagnosis between two conditions with similar biochemical phenotypes (the dual-scatter

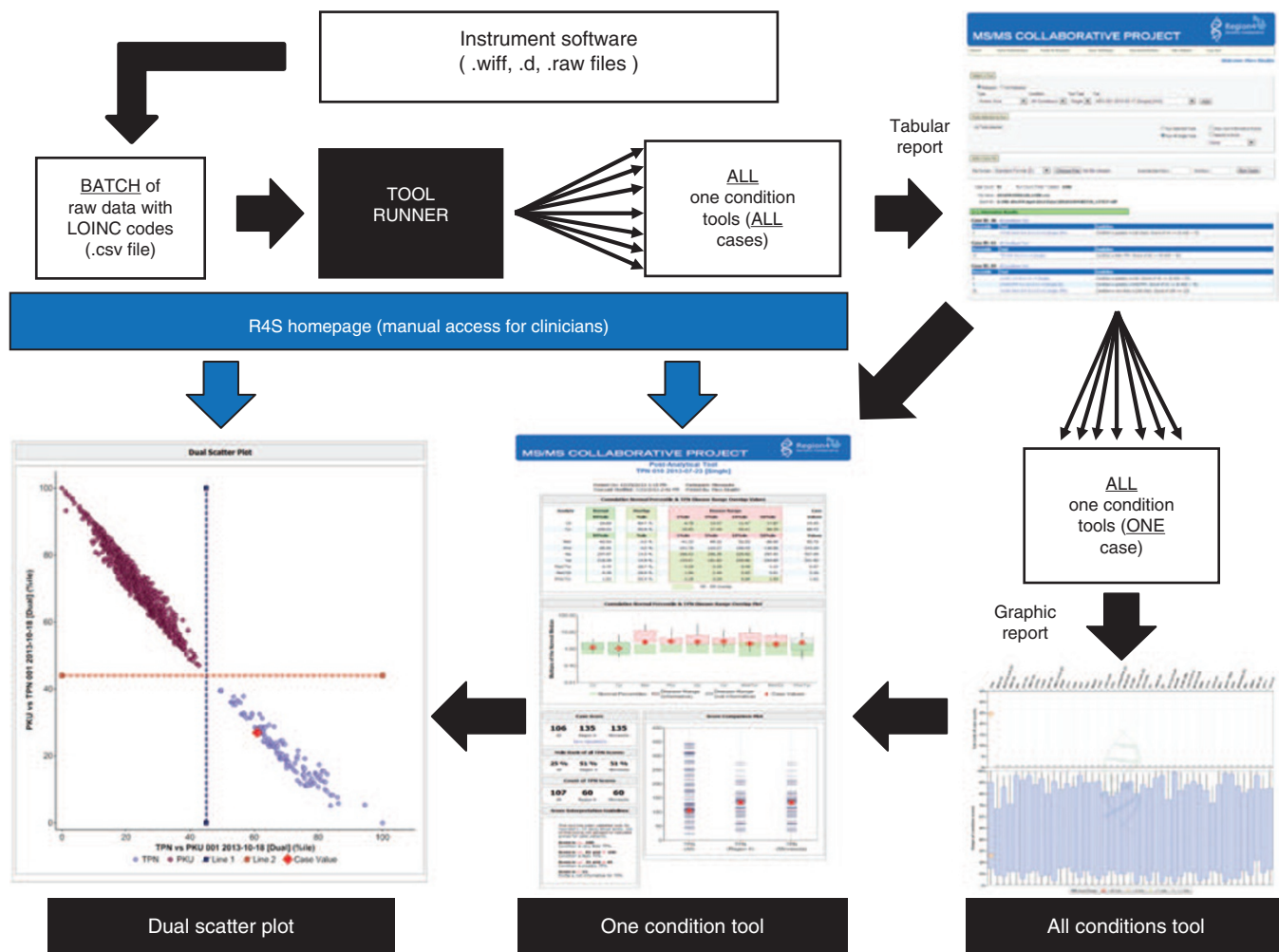


Figure 1 The process to upload data batches to the R4S tool runner. Details of the design, content, and clinical utilization of the one-condition tools and dual-scatter plots have been reported previously.³ A recorded educational presentation is also available at <http://www.mayomedicallaboratories.com/articles/hot-topic/2013/06-15-r4s-pt-3/index.html>. R4S, Region 4 Stork.

plot: e.g., differentiation between very long-chain acyl-CoA dehydrogenase deficiency and very long-chain acyl-CoA dehydrogenase carrier status); and (iii) to answer simultaneously the yes or no question for all conditions for which $n > 5$ and disease ranges have been established (the all-conditions tool). A defining characteristic of the R4S postanalytical tools is the evolution of clinical validation from the conventional static process, usually performed early during test development, to a constantly evolving, dynamic refinement of the disease ranges that continues to improve throughout the entire test life cycle process. The current count of active tools is 56 one-condition tools, 227 site-specific tools (13 of them created for California), and 24 dual-scatter plots. When a site-specific tool is released, it replaces the general tool for that particular condition when a data set is analyzed using either the tool runner or the all-conditions tool. Another available customization is to recalculate a score after switching from the cumulative reference percentiles to those uploaded by a user's own site.

Data uploading to R4S

The tool runner simultaneously calculates a score for all available tools for each case within a batch that is uploaded to the website as a comma-separated value (.csv) file void of any patient health information (Figure 1). Analyte results are paired with the corresponding logical observation identifiers names and code (LOINC; <http://www.loinc.org/>).¹⁴ LOINC is a universal code system for identifying laboratory and clinical observations maintained by the Regenstrief Institute at the University of Indiana. This system provides standardized terms for all kinds of observations and measurements that enable exchange and aggregation of electronic health data from many independent systems. A LOINC is unique to a combination of component, system (sample type), scale, and unit of measurement.

A 96-well plate includes ~90 patient samples and 9,000 results. After uploading of a .csv file of this size, processing time is typically less than 5 seconds before generating an interactive tabular report of any informative score. The all-conditions tool generates an interactive graphic report of all scores, including zeros, for one case at a time. Both types of reports are linked to each individual tool.

RESULTS

Current status and validation criteria of cutoff values

To establish a baseline for the comparison between performance outcomes, the size and adequacy of the cutoff panels of the California program were compared with those of all participating sites in R4S. Figure 2 shows the number of cutoff values plotted versus the proportion that meets the validation criteria described previously.¹ California is slightly below the median for both parameters.

Retrospective analysis of the California cohort

The outcome of this cohort is summarized in Table 1. Case resolution by tool score is based on a simultaneous, parallel evaluation of all conditions, not on a sequential/hierarchical manner.

Cases with multiple informative scores are resolved by application to every combination of two conditions of the differential diagnosis tool, the dual-scatter plot.³

In the true-negative group, 98.54% showed no informative scores. An additional 642 cases (0.39%) were settled as true-negative outcomes because recalculation of the score after switching to California's own reference percentiles rendered the score uninformative and/or produced a 0 percentile rank, i.e., a score lower than the lowest known true-positive case detected in California. The remaining 2,422 cases (1.46%) presented with at least one informative score. Because California relies on a network of 15 metabolic referral centers that would likely become aware of any case missed by newborn screening,¹⁵ the following analysis is based on the assumption that there were no additional false-negative events during the period under evaluation. Table 1 shows the relative proportions of these cases that could have been resolved conclusively as true negatives on the basis of a total parenteral nutrition score,¹⁶ a dual-scatter plot, or interpretation rules. For example, the cumulative experience in R4S with carnitine uptake defect ($n = 347$) and related maternal cases^{1,17,18} ($n = 168$) has shown that a borderline concentration of free carnitine could be disregarded when the sum of propionylcarnitine and palmitoylcarnitine is $>2 \mu\text{mol/l}$. Reflexing to a 2TT, particularly homocysteine measurement in a large number of cases with a concentration of methionine below the chosen cutoff value,¹⁰ would have resolved more than 800 cases. For the remaining cases, a more complete panel of analytes—in particular, the addition of argininosuccinic acid, glutamic acid, and glutamine, the latter two for improved interpretation of

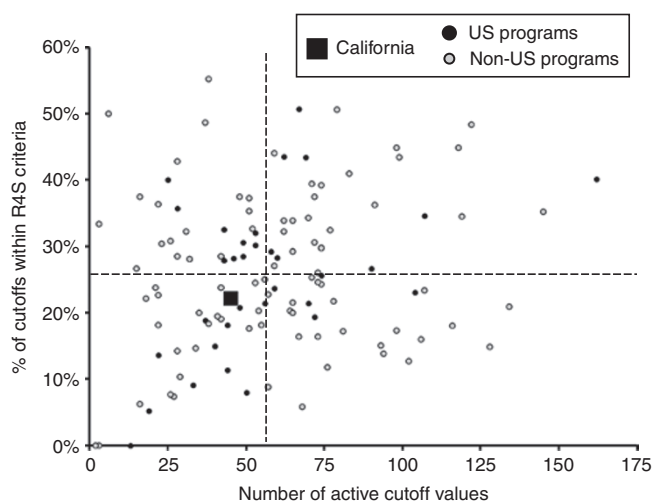


Figure 2 The number of active cutoff values ($n = 8,382$) used by R4S participating sites ($n = 151$). The y axis represents the proportion that meets the validation criteria. A given cutoff value is considered validated when it remains within the target range defined in R4S and also within the 25–75 percentile range of all cutoff values posted by all sites. High and low target ranges are defined by percentiles of the cumulative reference and disease ranges but could be replaced by override values in case of significant overlap.¹ Vertical and horizontal dotted lines indicate the median values (active cutoff values = 56; proportion within the validation criteria = 26%). R4S, Region 4 Stork.

Table 1 Comparison between cutoff-based system and R4S postanalytical interpretive tools in California

Screening outcome (actual)	Case count (actual)	No informative scores based on cumulative reference percentiles	No informative scores based on site-specific reference percentiles	Informative score for TPN	Carrier status assigned by dual-scatter plot	Informative scores overruled by interpretation rules	Informative scores likely to be overruled by 2TT	Informative scores likely to be overruled by added markers	Informative scores likely to be overruled by derivatization	Case count (residual)
Unaffected patients										
True negatives	166,118	163,696	642	343	3	237	801	271	125	0 ^a
Out of range cases ^b	8,276	6,024	258	899	1	27	414	615	38	0 ^a
Review cases ^c	1,285	261	90	764	1	9	99	57	4	0 ^a
False positives ^d	454	98	0	68	6	129	62	41	9	41 ^e
Subtotal	176,133	170,079	990	2,074	11	402	1,376	984	176	41
Percentage	—	96.56	0.56	1.18	0.01	0.23	0.78	0.56	0.10	0.02
Affected patients										
True positives	51	1 ^f	0	0	0	0	0	0	0	50 ^g
False negatives	2	1 ^h	0	0	0	0	0	0	0	1 ⁱ
Subtotal	53	2	—	—	—	—	—	—	—	51
Total	176,186	170,081	990	2,074	11	402	1,376	984	176	92

2TT, second-tier test; CoA, coenzyme A; R4S, Region 4 Stork; TPN, total parenteral nutrition.

^aNone of these cases would have been reassigned by the R4S tools to the presumptive positive category requiring follow-up evaluation. ^bA single result exceeding the analyte cutoff value but not meeting additional interpretation criteria. ^cMore than one result exceeding analyte cutoff values but not meeting additional interpretation criteria. See text for details. ^dAbnormal findings based on cutoff system that resulted in a referral to follow-up of a false-positive case. Amino acids (count): citrulline high (7), citrulline low (62), isoleucine-leucine (12), methionine high (22), methionine low (20), ornithine high (1), and phenylalanine high (12); acylcarnitines: free carnitine low (149), propionylcarnitine (16), butyryl-isobutyrylcarnitine (9), isovaleryl-2-methylbutyrylcarnitine (8), glutaryl-carnitine (60), hydroxy-isovaleryl-carnitine (15), octanoylcarnitine (14), tetradecenylcarnitine (8), and hexadecanoylcarnitine (2). ^eAbnormal findings based on R4S tools that still would have resulted in a referral to follow-up of a false-positive case. Amino acids (count): citrulline high (1), citrulline low (0), isoleucine-leucine (0), methionine high (0), methionine low (0), ornithine high (1), and phenylalanine high (1); acylcarnitines: free carnitine low (23), propionylcarnitine (0), butyryl-isobutyrylcarnitine (0), isovaleryl-2-methylbutyrylcarnitine (0), glutaryl-carnitine (0), hydroxy-isovaleryl-carnitine (10), octanoylcarnitine (1), malonylcarnitine (0), tetradecenylcarnitine (3), and hexadecanoylcarnitine (0). One true-positive case with carbamoyl phosphate synthetase deficiency not detected by the R4S tool for this condition. ^fList and count of conditions detected in true-positive cases. Amino acid disorders: carbamoyl phosphate synthetase deficiency (1), citrullinemia type 1 (3), maple syrup urine disease (1), ornithine transcarbamylase deficiency (2), disorder of biotin biosynthesis (1), phenylalanine hydroxylase deficiency (4), short-chain acyl-CoA dehydrogenase deficiency (6), maternal 3-methylcrotonyl-CoA carboxylase deficiency (1), organic acid disorders: 3-methylcrotonyl-CoA carboxylase deficiency (1), isobutyryl-CoA dehydrogenase deficiency (2), isovaleryl-CoA dehydrogenase deficiency (1), methylmalonic acidemia (2), and maternal vitamin B₁₂ deficiency (1). ^gOne false-negative case with ornithine transcarbamylase deficiency also not detected by the R4S tool for this condition. One false-negative case with ornithine transcarbamylase deficiency also not detected by the R4S tool for this condition.

cases with a low concentration of citrulline—could have prevented the generation of 20% of the residual scores.^{1,19} Another avenue of potential resolution could have been verification of an abnormal finding by using a derivatized method, particularly to mitigate the frequent flagging of malonylcarnitine due to either isobaric interference in the underivatized method or the choice of internal standard.²⁰ In summary, utilization of the tools and other interpretive options would not have reclassified any true-negative case to the presumptive positive category.

A total of 7,181 of 8,276 (87%) of combined out of range and review cases could have been prevented completely by combining lack of informative scores and total parenteral nutrition scores. 2TTs could have resolved 513 additional cases. **Table 1** shows the distribution of the remaining cases according to the same criteria applied to the true-negative cases.

A total of 454 cases initially reported as presumptive positives were later reclassified as FPs. Again, 66% of them could have been prevented by combining lack of informative scores, a score indicative of total parenteral nutrition, dual-scatter plots, and interpretation rules (**Figure 3**). Another 14% could have been resolved by a 2TT, with an estimated utilization rate in the entire cohort of 1 in 128 samples. With inclusion of additional

markers and verification by derivatization, the residual number of FPs could have been reduced below 10% of the initial count. In this group, there were some findings (e.g., citrulline >300 μmol/l; octanoylcarnitine >1 μmol/l; and glutarylcarnitine >5 μmol/l) that surprisingly normalized by the time of short-term follow-up. Low methionine concentration without the benefit of the 2TT for homocysteine caused 17 FP outcomes. More notably, 62 (11% of all FP cases) were flagged for low citrulline, with a FP rate of 0.035%. However, 28 of them could have been prevented by a noninformative score of the ornithine transcarbamylase deficiency tool. Another 30 could possibly have been prevented by a more complete analysis, especially by the inclusion of glutamic acid and glutamine, allowing the calculation of the respective ratios using citrulline as denominator.¹ Although it is not possible to calculate precise performance metrics under these circumstances, the same system proposed here has performed well prospectively in Minnesota over a 6-year period (2008–2013), leading to only 11 FP cases for low citrulline concentration among 431,748 newborns (0.0025%). There were five true-positive cases and one false-negative case over the same period. With one exception, a case confirmed to be affected with carbamoyl phosphate synthetase deficiency, all true-positive cases in the population study were correctly identified. The other false-negative case, affected with ornithine transcarbamylase deficiency, could have been detected by the site-specific tool for that condition, and it is possible that the resolution of both could have been aided by the measurement of glutamate and glutamine.

Table 2 shows a comparison between actual and estimated performance metrics. Notable improvements could have been achieved in positive predictive value and FP rate. This conclusion is not isolated: since its deployment, the tool runner has been applied 22,238 times by 30 programs (20 international), corresponding to 78 million calculated scores, on average 100,000 calculated scores per day. In terms of effectiveness, our findings are consistent with those that have been observed independently by other programs: a similarly designed comparative study applied to more than 180,000 Swedish newborns reduced the FP rate from 0.07 to 0.05% (Martin Engvall, personal communication). In another comparative study of 96,000 newborns in central Italy, the FP rate declined from 2.3 to 0.6% (Claudia Carducci, personal communication). Finally, the outcome of

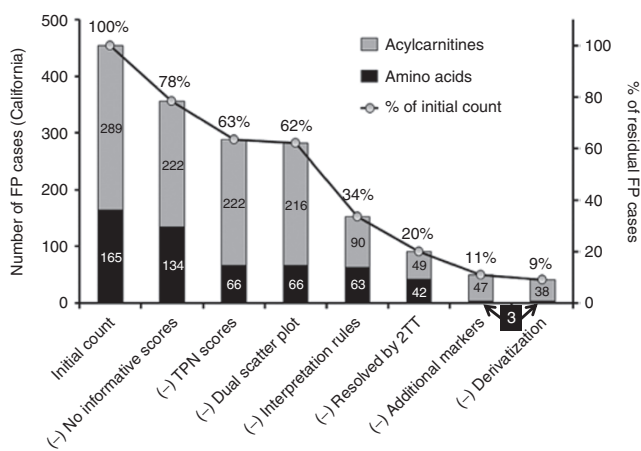


Figure 3 The potential impact of R4S tools, interpretation rules, 2TTs, and other potential interventions on the number of FP cases that were reported in California (January to June 2012). (-) Indicates subtraction of cases from the residual FP count on the basis of the criterion listed next to the symbol. 2TT, second-tier test; FP, false positive; R4S, Region 4 Stork.

Table 2 Comparison of performance metrics between cutoff-based systems (actual) and R4S postanalytical interpretive tools integrated with 2TTs, additional markers, and derivatization in Minnesota (actual) and California (estimated)

Performance metrics ^a	Minnesota	California		
	R4S tools interpretation rules and 2TTs (actual)	Cutoff values (actual)	R4S tools interpretation rules (estimated)	R4S tools, interpretation rules, 2TTs, and added markers, derivatization (estimated)
Sensitivity (%)	100	96	98	98
Positive predictive value (%)	69	10	25	55
Specificity (%)	99.98	99.74	99.91	99.98
False-positive rate (%)	0.02	0.26	0.09	0.02

2TT, second-tier test; R4S, Region 4 Stork.

^aSee ref. 5.

the R4S tools as the primary mode of prospective postanalytical interpretation in Minnesota for 2013 was as follows: among 71,207 newborns tested, 38 of 55 cases who were referred for follow-up because of abnormal MS/MS results were later confirmed to be true-positive cases, with a positive predictive value and FP rate of 69 and 0.024%, respectively. These figures are indeed very similar to those estimated by the retrospective analysis of the California data.

DISCUSSION

This report focuses on the second generation of R4S postanalytical interpretive tools and on how to use them effectively in a daily laboratory practice in which rapid, large-scale uploading and data processing are desirable. In a previous publication,³ we described how the postanalytical interpretive tools are produced using the tool builder functionality, but there was no mention of their use in a laboratory setting to underscore their clinical utility to a “user” rather than to a “producer” of laboratory results. Indeed, the one-condition tools and the dual-scatter plots can be described as clinician tools, applicable to aiding decisions about case resolution, test utilization, and therapeutic interventions. This was done deliberately to raise awareness of situations in which clinicians would have to reconcile the referral for follow-up of a case that in R4S would be considered not informative. Such a situation is less than ideal, suggesting that the application of the interpretive tools should occur at the earliest possible time, before a decision is made to proceed with further evaluation, including the request to submit a repeat specimen. This is the first report of the functionality that allows the simultaneous evaluation of all conditions with an active tool. Once an informative score has been generated, the all-conditions tool serves as an effective gateway to an unbiased set of differential diagnoses. This design allows a user to simultaneously and rapidly explore multiple one-condition tools, even those with a score of zero that was triggered by differentiator and outlier rules.³

Because additional conditions are likely to be added soon to the recommended uniform screening panel, improved performance is necessary to assure proper allocation of existing resources and avoidance of unnecessary distress to the patient's family.²¹ Most true-positive cases detected by MS/MS analysis of amino acids and acylcarnitines present with almost self-evident profiles that should not constitute a diagnostic challenge for the traditional cutoff-based system. The critical issue to be addressed is how to prevent the occurrence of FP cases driven by a rigid application of statistically defined cutoff values, especially when all potentially informative markers and ratios are not considered.

Over the past decade, the introduction and routine application of 2TTs to verify MS/MS results have encountered mixed reactions. To date, 44% of participants who have provided this information as part of their profile in R4S ($n = 84$) perform one or more 2TTs on site, 20% outsource them, and 36% do not use them. Although there is a steady trend toward greater utilization, persistent barriers include the need to set up additional MS/MS methods that also require chromatographic separation, lack of instrumentation, the perception of delayed reporting

and added cost to the laboratory, and last, but not least, the risk of overutilization. However, a workload of less than 1% should be manageable by most programs, especially when regionalization of services is considered to optimize resource utilization and quality of testing. Moreover, dedicated applications in R4S for each 2TT are or will soon be available where objective, large-scale comparison between true-positive and FP cases will further reduce the number of cases that still remain unresolved after the performance of a 2TT.

This study has provided new insight into the feasibility of newborn screening for a number of conditions with a biochemical phenotype based on low concentrations of amino acids. On the one hand, the verification of a low methionine level by a 2TT for homocysteine is extremely effective (no FPs and three confirmed cases in Minnesota since 2009), and a case could be made for the inclusion of remethylation disorders in the recommended uniform screening panel,²² considering the benefits of early identification and intervention.^{23,24} On the other hand, the California experience has shown that without 2TTs, targeting low methionine does cause many FP events, often not preventable by the R4S tool. The identification of proximal urea cycle disorders by low citrulline concentration has been regarded as a challenge in terms of sensitivity and specificity, also considering the growing spectrum of possible secondary targets.^{25,26} However, the observed FP rate in the Minnesota population has matched the average performance of all primary analytes that have caused FP outcomes over a 6-year period (0.0025%; range: 0.0002–0.0218%; $n = 24$) for the conditions detected by MS/MS. Therefore, once the availability and effectiveness of early intervention^{27,28} has been recognized as evidence of net benefit, the nomination of ornithine transcarbamylase deficiency and carbamoyl phosphate synthetase deficiency for inclusion in the recommended uniform screening panel should also be considered.

ACKNOWLEDGMENTS

This project was supported by a grant (U22MC03963) to the Region 4 Genetics Collaborative from the Health Resources and Service Administration of the Maternal and Child Health Bureau Cooperative Agreement, by contracts from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (contract no. HHSN275201000017C) and the Newborn Screening Translational Research Network (subcontract no. HHSN275200800001C 01), and by the T. Denny Sanford Professorship fund, Mayo Clinic. We recognize the expertise and dedication of the many metabolic specialists in California who were involved in the follow-up evaluation of presumptive positive cases and the 150 programs worldwide that have posted their analyte cutoff values on the R4S website. Access to R4S can be requested through NBSTRN (<http://www.nbstrn.org/>) or from the R4S login page (<http://www.clir-r4s.org/>).

DISCLOSURE

G.M., R.J.C., D.M.S.M., and P.R. are coinventors of a provisional patent application (Computer-based dynamic data analysis, US 61/722,761). The other authors declare no conflict of interest.

REFERENCES

- 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.
- Watson MS, Mann MY, Lloyd-Puryear MA, Rinaldo P, Howell RR. Newborn screening: toward a uniform screening panel and system [executive summary]. *Genet Med* 2006;8(suppl):1S–11S.
- Marquardt G, Currier R, McHugh DM, et al. Enhanced interpretation of newborn screening results without analyte cutoff values. *Genet Med* 2012;14:648–655.
- Merritt JL 2nd, Vedal S, Abdenur JE, et al. Infants suspected to have very-long chain acyl-CoA dehydrogenase deficiency from newborn screening. *Mol Genet Metab* 2014;111:484–492.
- Rinaldo P, Zafari S, Tortorelli S, Matern D. Making the case for objective performance metrics in newborn screening by tandem mass spectrometry. *Ment Retard Dev Disabil Res Rev* 2006;12:255–261.
- Gurian EA, Kinnamon DD, Henry JJ, Waisbren SE. Expanded newborn screening for biochemical disorders: the effect of a false-positive result. *Pediatrics* 2006;117:1915–1921.
- Tarini BA, Christakis DA, Welch HG. State newborn screening in the tandem mass spectrometry era: more tests, more false-positive results. *Pediatrics* 2006;118:448–456.
- Tarini BA, Clark SJ, Pilli S, et al. False-positive newborn screening result and future health care use in a state Medicaid cohort. *Pediatrics* 2011;128:715–722.
- Weisfeld-Adams JD, Morrissey MA, Kirmse BM, et al. Newborn screening and early biochemical follow-up in combined methylmalonic aciduria and homocystinuria, cblC type, and utility of methionine as a secondary screening analyte. *Mol Genet Metab* 2010;99:116–123.
- Tortorelli S, Turgeon CT, Lim JS, et al. Two-tier approach to the newborn screening of methylenetetrahydrofolate reductase deficiency and other remethylation disorders with tandem mass spectrometry. *J Pediatr* 2010;157:271–275.
- Turgeon CT, Magera MJ, Cuthbert CD, et al. Determination of total homocysteine, methylmalonic acid, and 2-methylcitric acid in dried blood spots by tandem mass spectrometry. *Clin Chem* 2010;56:1686–1695.
- Oglesbee D, Sanders KA, Lacey JM, et al. Second-tier test for quantification of alloisoleucine and branched-chain amino acids in dried blood spots to improve newborn screening for maple syrup urine disease (MSUD). *Clin Chem* 2008;54:542–549.
- American College of Medical Genetics and Genomics. ACMG act sheets and confirmatory testing. <http://www.ncbi.nlm.nih.gov/books/NBK55832/>. Accessed 25 February 2014.
- Downs SM, van Dyck PC, Rinaldo P, et al. Improving newborn screening laboratory test ordering and result reporting using health information exchange. *J Am Med Inform Assoc* 2010;17:13–18.
- Feuchtbaum L, Dowray S, Lorey F. The context and approach for the California newborn screening short- and long-term follow-up data system: preliminary findings. *Genet Med* 2010;12(suppl 12):S242–S250.
- Chace DH, De Jesús VR, Lim TH, Hannon WH, Clark RH, Spitzer AR. Detection of TPN contamination of dried blood spots used in newborn and metabolic screening and its impact on quantitative measurement of amino acids. *Clin Chim Acta* 2011;412:1385–1390.
- Schimmenti LA, Crombez EA, Schwahn BC, et al. Expanded newborn screening identifies maternal primary carnitine deficiency. *Mol Genet Metab* 2007;90:441–445.
- Lee NC, Tang NL, Chien YH, et al. Diagnoses of newborns and mothers with carnitine uptake defects through newborn screening. *Mol Genet Metab* 2010;100:46–50.
- Trinh MU, Blake J, Harrison JR, et al. Quantification of glutamine in dried blood spots and plasma by tandem mass spectrometry for the biochemical diagnosis and monitoring of ornithine transcarbamylase deficiency. *Clin Chem* 2003;49:681–684.
- Chace DH, Lim T, Hansen CR, Adam BW, Hannon WH. Quantification of malonylcarnitine in dried blood spots by use of MS/MS varies by stable isotope internal standard composition. *Clin Chim Acta* 2009;402:14–18.
- Waisbren SE, Albers S, Amato S, et al. Effect of expanded newborn screening for biochemical genetic disorders on child outcomes and parental stress. *JAMA* 2003;290:2564–2572.
- Kemper AR, Green NS, Calonge N, 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;16:183–187.
- Strauss KA, Morton DH, Puffenberger EG, et al. Prevention of brain disease from severe 5,10-methylenetetrahydrofolate reductase deficiency. *Mol Genet Metab* 2007;91:165–175.
- Schiff M, Benoist JF, Tilea B, Royer N, Giraudier S, Ogier de Baulny H. Isolated remethylation disorders: do our treatments benefit patients? *J Inher Metab Dis* 2011;34:137–145.
- Atkuri KR, Cowan TM, Kwan T, et al. Inherited disorders affecting mitochondrial function are associated with glutathione deficiency and hypocitrullinemia. *Proc Natl Acad Sci USA* 2009;106:3941–3945.
- de Sain-van der Velden MG, Rinaldo P, Elvers B, et al. The proline/citrulline ratio as a biomarker for OAT deficiency in early infancy. *JIMD Rep* 2012;6:95–99.
- Campeau PM, Pivalizza PJ, Miller G, et al. Early orthotopic liver transplantation in urea cycle defects: follow up of a developmental outcome study. *Mol Genet Metab* 2010;100(suppl 1):S84–S87.
- Diaz GA, Krivitzky LS, Mokhtarani M, et al. Ammonia control and neurocognitive outcome among urea cycle disorder patients treated with glycerol phenylbutyrate. *Hepatology* 2013;57:2171–2179.



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/3.0/>