

Continuous Age- and Sex-Adjusted Reference Intervals of Urinary Markers for Cerebral Creatine Deficiency Syndromes: A Novel Approach to the Definition of Reference Intervals

Lars Mørkrid,^{1*} Alexander D. Rowe,¹ Katja B.P. Elgstoen,¹ Jess H. Olesen,² George Ruijter,³ Patricia L. Hall,⁴ Silvia Tortorelli,⁵ Andreas Schulze,⁶ Lianna Kyriakopoulou,⁷ Mirjam M.C. Wamelink,⁸ Jiddeke M. van de Kamp,⁹ Gajja S. Salomons,⁸ and Piero Rinaldo⁵

BACKGROUND: Urinary concentrations of creatine and guanidinoacetic acid divided by creatinine are informative markers for cerebral creatine deficiency syndromes (CDSs). The renal excretion of these substances varies substantially with age and sex, challenging the sensitivity and specificity of postanalytical interpretation.

METHODS: Results from 155 patients with CDS and 12 507 reference individuals were contributed by 5 diagnostic laboratories. They were binned into 104 adjacent age intervals and renormalized with Box–Cox transforms (Ξ). Estimates for central tendency (μ) and dispersion (σ) of Ξ were obtained for each bin. Polynomial regression analysis was used to establish the age dependence of both $\mu[\log(\text{age})]$ and $\sigma[\log(\text{age})]$. The regression residuals were then calculated as z -scores = $\{\Xi - \mu[\log(\text{age})]\} / \sigma[\log(\text{age})]$. The process was iterated until all z -scores outside Tukey fences ± 3.372 were identified and removed. Continuous percentile charts were then calculated and plotted by retransformation.

RESULTS: Statistically significant and biologically relevant subgroups of z -scores were identified. Significantly higher marker values were seen in females than males, necessitating separate reference intervals in both adolescents and adults. Comparison between our reconstructed reference percentiles and current standard age-matched reference intervals highlights an underlying risk of false-positive and false-negative events at certain ages.

CONCLUSIONS: Disease markers depending strongly on covariates such as age and sex require large numbers of reference individuals to establish peripheral percentiles with sufficient precision. This is feasible only through collaborative data sharing and the use of appropriate statistical methods. Broad application of this approach can be implemented through freely available Web-based software.

© 2015 American Association for Clinical Chemistry

Testing for creatine deficiency syndromes (CDSs)¹⁰ is required when patients manifest unexplained intellectual disability, delayed expressive speech and language development, autistic behavior, and seizures. The most common screening test for CDSs concurrently measures the urinary concentrations of creatinine (crn), creatine (CRE), and guanidinoacetic acid (GAA). Increased GAA, increased GAA/crn ratio, and decreased CRE/GAA ratio are informative markers of guanidinoacetate methyltransferase (GAMT) deficiency. Decreased GAA, decreased GAA/crn ratio, and increased CRE/GAA ratio define the biochemical phenotype of arginine:glycine amidinotransferase (AGAT) deficiency. Finally, increased CRE, increased CRE/crn ratio, and increased CRE/GAA ratio are usually but inconsistently found in males with the X-linked creatine transporter (CRT) defect (*I-4*). In symptomatic CRT female carriers, these 3 markers are at the high end of the respective reference

¹ Department of Medical Biochemistry, Oslo University Hospital, Rikshospitalet, Oslo, Norway; ² Rigshospitalet, University of Copenhagen, Copenhagen, Denmark; ³ Department of Clinical Genetics, Erasmus Medical Center, Rotterdam, The Netherlands; ⁴ Department of Human Genetics, Emory University, Atlanta, GA; ⁵ Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN; ⁶ Division of Clinical and Metabolic Genetics and ⁷ Division of Clinical Biochemistry, Hospital for Sick Children, University of Toronto, Toronto, Ontario, Canada; ⁸ Department of Clinical Chemistry and ⁹ Department of Clinical Genetics, VU University Medical Center, Amsterdam, The Netherlands.

* Address correspondence to this author at: Department of Medical Biochemistry, Oslo University Hospital, Rikshospitalet, Institute of Clinical Chemistry, PO Box 4950 Nydalen, Oslo, 0424, Norway. Fax +47-2-307-0902; e-mail lamo2@ous-hf.no.

Part of this work has been presented at the 2014 annual symposium of the Society for the Study of Inborn Errors of Metabolism (SSIEM), Innsbruck, Austria, September 2–5, 2014. Received November 17, 2014; accepted February 12, 2015.

Previously published online at DOI: 10.1373/clinchem.2014.235564

© 2015 American Association for Clinical Chemistry

¹⁰ Nonstandard abbreviations: CDS, cerebral creatine deficiency syndrome; crn, creatinine; CRE, creatine; GAA, guanidinoacetic acid; GAMT, guanidinoacetate methyltransferase; AGAT, arginine:glycine amidinotransferase; CRT, creatine transporter; CLIR, Collaborative Laboratory Integrated Reports.

Table 1. Age distributions of reference individuals provided by different sites.

Site	Age range										Total
	Months			Years							
	0-3	3-6	6-12	1-4	4-10	10-18	18-40	40-55	55-70	>70	
1	82	55	148	1063	1223	458	178	62	13	5	3287
2	502	158	211	558	424	272	40	24	7	3	2199
3	39	29	77	509	510	269	158	102	95	92	1880
4	288	106	176	585	642	363	231	89	37	2	2519
5	56	67	164	707	944	549	101	20	13	1	2622
Subtotal	967	415	776	3422	3743	1911	708	297	165	103	12507
6 (verification)	1	0	0	113	219	104	1	0	0	0	438
Total	968	415	776	3535	3962	2015	709	297	165	103	12945
Male prevalence, %	60	55	56	65	70	64	58	54	53	49	64

interval but rarely frankly increased. The normalization by crn attempts to correct for different degrees of urine concentration and relies on the assumption that the tubular excretion and reabsorption of these compounds is similar. However, preliminary evidence suggests that the synthesis of GAA within renal epithelial cells and subsequent transmembrane transport of CRE and GAA are not synchronized (5–8). Urinary crn concentration depends on many factors, among which age and sex are important covariates. After birth, mean urine crn declines for approximately the first month before rising steadily in response to increasing muscle mass and anabolism, reaching a maximum and gradually leveling off at approximately 18–20 years of age. This characteristic age-dependent pattern may be quite different from unrelated endogenous compounds that are excreted in urine. Because a disease marker in urine is often expressed as a ratio to the crn concentration, such a ratio will have its own specific age dependence, which needs to be corrected for when calculating reference and medical decision limits. Reference limits that are age dependent are generally calculated in separated age intervals. This procedure may result in abrupt discontinuities at the interval borders, which is a definite disadvantage. Such discontinuities are frequently seen in pediatric medicine (9–10), although they could be avoided with the use of different types of regression methods (11).

Cohorts containing at least 120 individuals are recommended to establish reference limits with 90% CIs for a variable that exhibits a unimodal and uniform distribution (12). When separated age bins are needed to describe the different age ranges properly, the total number of individuals should be 120 multiplied by the number of bins. This requirement can be somewhat lessened by the use of regression methods, which tie the different bins together, but the number requirement must be further

expanded if there are other covariates (e.g., sex) that must be corrected for. Because the cumulative prevalence of CDS is very low, routine results deemed to be uninformative can be used retrospectively for the calculation of reference limits. When several thousand results are available, the central part of the distribution constitutes a fairly adequate reference population after cases identified as outliers have been excluded.

The purpose of the present study was to establish continuous age-corrected reference limits for the CRE/crn and GAA/crn ratios. To achieve this with sufficient precision, a collaborative effort was initiated among 6 reference laboratories from different geographical sites. Combination of the data in 1 pool was based on finding sufficiently similar medians along the age axis between the different sites.

Methods

STUDY POPULATION AND ANALYTICAL METHODS

We collected reference data inclusive of age at collection and sex from 5 different laboratories (range of number of samples contributed was 1880–3287) that used similar analytical methods (see Supplemental Table 1, which accompanies the online version of this article at <http://www.clinchem.org/content/vol61/issue5>). These clinical specimens had been previously defined as uninformative for CDS and provide the basis for the statistical modeling. An additional laboratory contributed a smaller set ($n = 438$) of children with autistic spectrum disorder who were tested to rule out a possible CDS (Table 1). This group was processed separately for independent verification of the validity of the proposed model. True positive samples were diagnosed with either AGAT ($n = 3$), GAMT ($n = 32$), or CRT deficiency ($n = 120$). At least 84 of these cases have been reported previously in peer-

reviewed publications (13–14), and all were confirmed by combinations of molecular analysis, magnetic resonance spectroscopy, and enzymatic analysis. The specimens included in this study were collected before treatment. All laboratories have participated in external quality programs with satisfactory results.

STATISTICAL ANALYSIS

We partitioned the combined data into equally sized bins ($n = 120$) and used the Horn algorithm (15) within each bin to establish a simple polynomial relationship, preferably a first-order polynomial, i.e., a straight line, between the marker (X or a transform thereof, Ξ) and its corresponding z -score. By this approach, it is possible to calculate estimates for expected value μ and SD σ by means of first-order regression analysis, and any measured value can be converted to a z -score $[(\Xi - \mu)/\sigma]$ or its corresponding percentile in the reference population, analogous to the approach commonly used with bone density measurements or growth charts in pediatric practice. Both μ and σ typically depend on sex, age, and possibly other covariates but can be expressed as continuous functions of these factors, preferably by use of low-order polynomials where the coefficients are obtained by regression analysis.

After calculating z -scores corrected for the covariates, we removed statistical outliers with Tukey fences. A Tukey factor of 2 excluded samples with z -scores outside ± 3.372 . The measured values were sorted in ascending order: $X_i, i = 1, 2, \dots, n$, and each value was assigned a percentile value by the formula $p_i = (i - 3/8)/(120 + 1/4)$ (16). We found a corresponding z_i -score by use of the inverse standard normal distribution. To correct for skewness, all X_i values were then subjected to a Box–Cox power transformation:

$$Y = [B - C; \lambda] X = \begin{cases} (X^\lambda - 1)/\lambda & \text{if } \lambda \neq 0 \\ \ln(X) & \text{if } \lambda = 0 \end{cases} \quad (1)$$

and an optimal λ was sought to make the relation between Y_i values and z_i -scores as linear as possible. We use the abbreviated notation “[$B - C$]” when 2 subpopulations with different λ values are combined within the same bin [see Eqs. (2) and (3)]. Estimated expected value (μ) and SD (σ) of Box–Cox transformed marker values within each age bin were determined by the Horn method. Polynomial regression (order ≤ 3) on transformed parameters μ and σ vs log(age) (Briggs logarithm to base 10) established age-continuous curves for these estimates in separate age segments, which could be seamlessly connected at their borders. A flow chart of the statistical estimation procedure is shown in online Supplemental Fig. 1. We used ANOVAs with Student–Newman–Keuls post hoc tests to identify group differences on the basis of z -scores. We then applied the Gow-

ans criterion to statistically significant differences to evaluate biological significance (17). The latter considers a subgroup $|\text{bias}| < 0.25$ z -score units to be less biologically relevant and may thereby allow combination of subpopulations. In addition, this technique was supported by the application of a partition rule (18). To use combined reference limits, this rule requires that all proportions of the subgroup distributions cut off by the reference limits of the combined distribution should lie in the interval 0.9%–4.1%.

Nonparametric estimation of quantiles was performed by bootstrap resampling. This also allowed for calculation of 90% CIs on the estimates.

Results

ADJUSTMENT FOR AGE

After exclusion of outliers, 103 complete bins remained (see details below). It was not possible to find a unique Box–Cox λ for CRE/crn or GAA/crn that rendered the z -score plot rectilinear over the major part of its central course ($-1.96 < z\text{-score} < 1.96$). The curve always appeared kinked. This is illustrated in online Supplemental Fig. 2 for various age cohorts when $\lambda = 0.45$. However, it was noted that for values below a fixed threshold value = T , $\lambda = 0.10$ gave a rectilinear z -score plot in that interval (data points denoted by blue circles in online Supplemental Fig. 3). The T values were 30 and 8 $\mu\text{mol}/\text{mmol}$ for CRE/crn and GAA/crn, respectively. The z -score regression equation could then be written

$$z_i = a + b \times \zeta_i + e_i; i = 1, 2, \dots, k, \quad (2)$$

where $\zeta_i = [B - C; 0.1]X_i$, and e_i represents the regression residual.

Above T , $\lambda = 0.45$ resulted in a straight line relationship between z -scores and the Box–Cox transformed values. These data points are denoted by black circles in online Supplemental Fig. 3.

$$z_j = c + d \times \xi_j + \varepsilon_j; j = k + 1, k + 2, \dots, 120, \quad (3)$$

where $\xi_j = [B - C; 0.45]X_j$, and ε_j represents the regression residual.

The transformed values at threshold were $\tau = [B - C; 0.10]T$ and $t = [B - C; 0.45]T$, whereas T appeared to be largely invariant with respect to age, as shown in online Supplemental Fig. 2. The 2 optimal Box–Cox λ values 0.10 and 0.45 were chosen to give a maximal Pearson correlation coefficient r between z -score and transformed marker value within each of the 2 selected intervals. To have the same rectilinear relationship within a bin between z -scores and transformed values above and below T , a first-order transformation (linear shift) of the ζ_j values was applied:

$$\psi_i = \alpha + \beta \times \zeta_i. \quad (4)$$

To preserve continuity at the threshold, α is uniquely determined by β according the following equation:

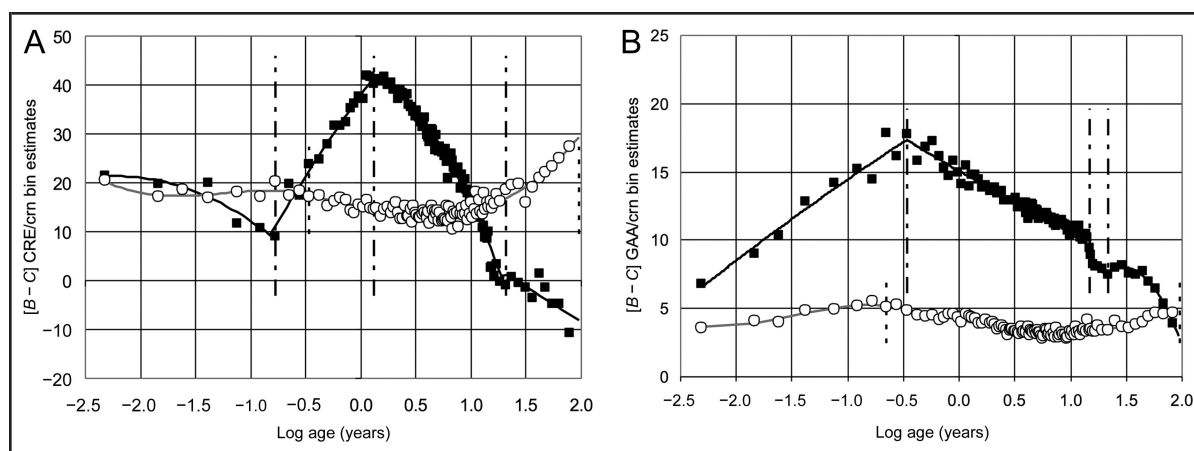


Fig. 1. Estimated parameters μ (■) and σ (○) as a function of logarithm (base 10) of age in years for Box-Cox transformed values of CRE/crn (A) and GAA/crn (B).

Polynomial approximations (solid line) were obtained by polynomial regression analysis in adjacent age intervals, delineated by vertical lines: dashed-dotted for μ and dotted for σ . For brevity, λ is omitted within the $[B - C]$ symbol.

$$\alpha + \beta \times \tau = t, \text{ by which } \alpha = t - \beta \times \tau. \quad (5)$$

From Eqs. (4) and (5),

$$\psi_i = t + \beta \times (\zeta_i - \tau). \quad (6)$$

The 2 variables are now connected by a single parameter β . In online Supplemental Fig. 3, ψ_i is symbolized by filled red circles. To let ψ have the same statistical properties as ξ (same expectation value and same SD), ψ_i and ξ_i should belong to the same linear regression line of z -score values. Because the upper part of the z -score curve (above the threshold) has the highest number of regression points, those coefficients were determined first, before adjusting the regression equation of the lower part to the upper part:

$$z_i = \hat{c} + \hat{d} \times \psi_i + \varepsilon_i = \hat{c} + \hat{d} \times (\alpha + \beta \times \zeta_i) + \varepsilon_i; \quad i = 1, 2, \dots, k, \quad (7)$$

where \hat{c} and \hat{d} are estimates already determined from regression analysis on Eq. (3). Equation (7) may be rearranged for the determination of β :

$$W_i = z_i - \hat{c} - \hat{d} \times t = \beta \times (\zeta_i - \tau) \times \hat{d} + \varepsilon_i = \beta \times \phi_i + \varepsilon_i. \quad (8)$$

The best estimate for β by the method of least squares is then:

$$\hat{\beta} = \frac{\sum_{i=1}^k W_i \times \phi_i}{\sum_{i=1}^k \phi_i^2}. \quad (9)$$

Beta as a function of log age is shown in online Supplemental Fig. 4, and the coefficients for the estimated poly-

nomials describing the central course are presented in online Supplemental Table 2. Having estimated a function β that combines transformed values below and above the threshold, ψ_i can be calculated from Eq. (6). The final, combined regression equation within a specific age bin is

$$z_i = A + B \times \Xi_i + \eta_i; \quad i = 1, 2, 3, \dots, n, \quad (10)$$

where $\Xi_i = \psi_i$ for $i \leq k$ and ξ_i elsewhere; η_i denotes the residuals.

Equation (10) can now be used to calculate estimates for expectation value:

$$\hat{\mu} = \hat{A}/\hat{B}, \quad (11)$$

and SD:

$$\hat{\sigma} = 1/\hat{B}, \quad (12)$$

within each bin. As with $\hat{\mu}$ and $\hat{\sigma}$, $\hat{\beta}$ will be a function of the covariate log age.

Having established expressions for β , the parameters A and B in Eq. (10) can be estimated for each age bin. The resulting estimates for μ and σ are depicted in Fig. 1 and with polynomial regression applied to adjacent age segments:

$$\begin{aligned} \mu \text{ or } \sigma = & D_0 + [D_1 \times \log(\text{age}_i)] + \{D_2 \\ & \times [\log(\text{age}_i)]^2\} + \{D_3 \times [\log(\text{age}_i)]^3\} + F_i, \end{aligned} \quad (13)$$

where F_i is the residual term. Estimates for the regression coefficients are shown in online Supplemental Table 3.

Table 2. Combined and sex-specific estimates of z-scores for Box-Cox transformed marker values. ^a						
Age, years	Both sexes	Female	Male	Sex difference	t	P
CRE/crn ratio						
<13.5						
n	10374	3679	6695			
Mean	0.012	0.102	-0.037	0.139	6.89	6.0 × 10 ⁻¹²
SD	0.991	0.960	1.005			
Less than 2.5th percentile ^b		2.0	2.8			
More than 97.5th percentile ^b		2.4	2.6			
>13.5						
n	2095	880	1215			
Mean	-0.035	0.244	-0.237	0.481	11.23	1.9 × 10 ⁻²⁸
SD	0.996	0.934	0.992			
Less than 2.5th percentile ^b		1.0	4.4			
More than 97.5th percentile ^b		3.0	1.8			
GAA/crn ratio						
<13.5						
n	10343	3664	6679			
Mean	0.006	0.147	-0.071	0.218	10.766	7.8 × 10 ⁻¹⁷
SD	0.988	0.987	0.981			
Less than 2.5th percentile ^b		2.2	2.7			
More than 97.5th percentile ^b		1.9	2.3			
>13.5						
n	2073	868	1205			
Mean	0.003	0.232	-0.161	0.393	8.540	2.9 × 10 ⁻¹⁷
SD	1.028	1.090	0.947			
Less than 2.5th percentile ^b		2.8	2.2			
More than 97.5th percentile ^b		5.8	1.9			

^a Values in bold exceed the bias criteria.
^b Portions in subpopulations lying outside the combined reference limits (%).

The following formula then converts every measured value X to a z-score:

$$z_i = \frac{\Xi_i - \mu[\log(\text{age}_i)]}{\sigma[\log(\text{age}_i)]} \quad i = 1, 2, \dots, N \quad (14)$$

The final z-scores were shown to follow a standard normal distribution, by which the grand mean and SD can be found as well as the corresponding subgroup estimates.

In the data set available for this study, the whole estimation process comprising Eqs. (2)–(14) was repeated twice for CRE/crn and 4 times for GAA/crn, leaving 2 (0.02%) low and 36 (0.29%) high outliers for CRE/crn and 35 (0.28%) low and 56 (0.45%) high outliers for GAA/crn. Hence, after outlier exclusion, the fi-

nal values for N in Eq. (14) were 12469 and 12416 for CRE/crn and GAA/crn, respectively.

ADJUSTMENT FOR SEX

Visual inspection of mean z-scores from each age bin showed an unequivocal sex difference both for CRE/crn and GAA/crn. Even in early childhood, females had higher values than males. The difference was statistically significant, was augmented in the course of adolescence (>13.5 years of age), and persisted throughout the adult period. The differences are shown in Table 2. Despite statistical significance, the Gowans criterion suggests that combined reference limits might be adequate for both sexes during childhood for both markers. This conclusion is also supported by another partitioning rule that also includes the effect of unequal SDs (18). However,

Table 3. Combined and site-specific estimates of z -score distribution for transformed values (age <13.5 years).^a

Ratio	All sites	Site 1	Site 2	Site 3	Site 4	Site 5	F^b	P^b	Site 6
CRE/crn									
n	10 374	2842	1989	1290	2258	1995			394
Mean	0.000	-0.027	0.058	0.121	-0.030	0.001	7.125	1.00×10^{-5}	-0.134
SD	1.000	0.940	1.048	1.211	0.893	0.943			0.719
Less than combined 2.5th percentile ^c	2.50	1.70	3.60	6.20	1.10	1.80			0.50
More than combined 97.5th percentile ^c	2.50	2.10	2.80	4.90	1.50	2.40			0.00
GAA/crn									
n	10 343	2831	1982	1289	2256	1985			392
Mean	0.000	-0.070	-0.065	0.292	-0.006	0.015	34.322	1.67×10^{-28}	-0.002
SD	1.000	0.954	0.998	1.064	0.941	0.996			0.814
Less than combined 2.5th percentile ^c	2.50	2.50	3.70	2.20	1.60	2.60			0.80
More than combined 97.5th percentile ^c	2.50	2.50	1.70	5.60	2.20	2.40			0.80

^a Values in bold indicate sites that require either separate reference limits or a site-specific adjustment.
^b Site group comparison (ANOVA).
^c Portions in subpopulations lying outside the combined reference limits (%).

both procedures point to the necessity of correction for sex in adolescents and adults.

Because of significant sex z -score differences with ages >13.5 years and the sex ratio varying significantly between sites in this age range (see Table 1), the effect of site was examined only for ages <13.5 years. The results are shown in Table 3. Despite the statistical significance, the Gowans criterion suggests that correction for level (mean differences) between sites should be unnecessary during the whole childhood for CRE/crn and for GAA/crn, except site 3. However, it is then obvious that it must be the difference in local SD that violates the partitioning rule for sites 3 and 6 and invalidates site-independent reference limits.

One may correct z -scores for sex and site differences with the following formula:

$$z\text{-score}_{\text{corrected}} = (z\text{-score} - \Delta\text{mean})/\text{SD}, \quad (15)$$

where Δmean denotes the deviation from the combined mean and SD is the locally obtained SD of z -scores. The grand mean and global SD are approximately equal to 0 and 1, respectively.

VERIFICATION

The algorithms described above were applied for independent verification to the smaller set ($n = 438$) of patients with autism spectrum disorder and uninformative results for CDS. The z -score plot for age-corrected $[B - C]$ CRE/crn appeared without outliers and was visually indistinguishable from a straight line (data not shown).

The z -score mean and SD were -0.134 and 0.719 , respectively. The z -score plot for $[B - C]$ GAA/crn was also rectilinear, but contained 1 low and 1 high outlier (z -scores at -3.730 and 4.059 , respectively). The z -score mean and SD were -0.002 and 0.814 , respectively.

PERCENTILE CURVES

Conversely theoretical percentile curves can be constructed from Eq. (14) with concentration-specific dependent retransformation formulae. The results are illustrated in Fig. 2. The time course is not described by a single steady rise or fall throughout the age range, as in growth charts in children. Instead, there is a trough for CRE/crn and a peak for both CRE/crn and GAA/crn, and the variation in percentile form necessitates division into several age regions to describe them adequately with a low-powered polynomial (order cubic or less). This irregularity is greater for the median than for the dispersion. It is also interesting to note how patient values are located in relation to the 2.5%–97.5% percentile interval. The lack of patient data before 1 year of age must be viewed in the light of the difficulties of identifying characteristic clinical symptoms in that period. Nonparametric estimates seem also to correspond well with the parametric model (see online Supplemental Fig. 5, in which CIs are also represented).

Discussion

Reference limits on the basis of broad age bins may be misleading and lead to diagnostic misclassification. The

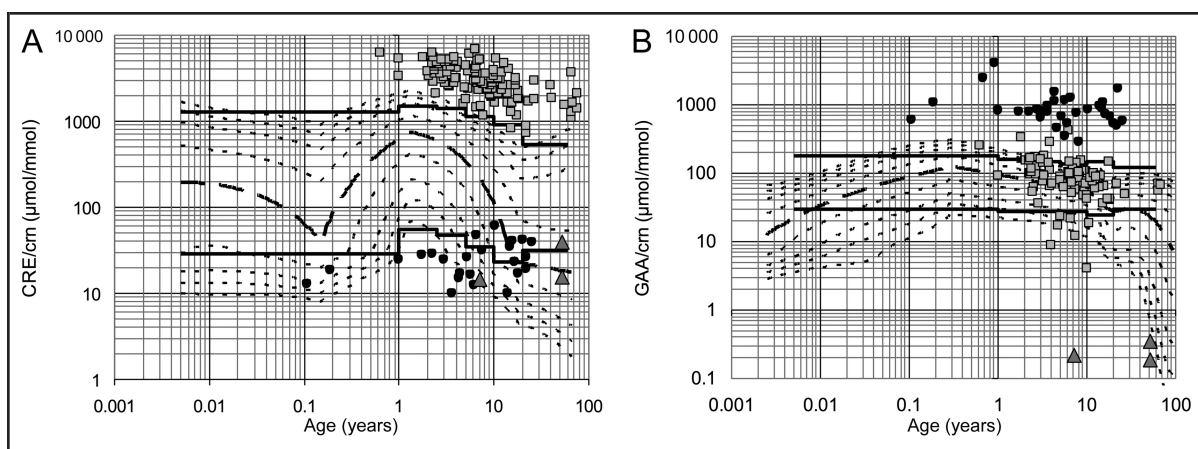


Fig. 2. Predicted percentile charts for CRE/crn (A) and GAA/crn (B). Dotted lines indicate, from bottom to top, 2.5th, 5th, 10th, 25th, 75th, 90th, 95th, and 97.5th percentiles.

The median is denoted by a dashed line. Horizontal solid lines are traditional cutoffs (20). □, 120 CRT patients (86 GAA/crn measurements); ●, 32 GAMT patients (24 CRE/crn measurements); ▲, 3 AGAT patients.

present study took advantage of a large number of reference individuals ($n = 12\,507$) from laboratory productions to circumvent the problems of obtaining samples from healthy children. This procedure tends to broaden the reference intervals and may therefore not be ideal. However when the prevalence of the actual disease is low (as here for AGAT, GAMT, and CRT), and when positively identified disease cases are removed together with regression outliers, the central part of the z -score distribution of the remnant cases should constitute a relevant reference population.

The total laboratory production was reasonably distributed in a uniform manner along the transformed age axis. Polynomial regression was used to obtain continuous curves for central tendency and dispersion for the transformed markers in discrete age regions. An analogous regression procedure has been proposed (19). Special care was taken to keep the polynomial order as low as possible and to avoid discontinuous jumps when crossing age limits. Examination of the reference material revealed a peculiar property of the CRE/crn and GAA/crn distributions: it was not possible to linearize the whole course of the z -score curve with a single Box–Cox transformation. Two different lambdas, $\lambda = 0.1$ and 0.45 , were needed, each in 1 of 2 separate sections of the curve, separated by a fixed threshold value of the marker. A linear transformation was then able to connect the 2 sections so that the major part of the z -score curve could be used in the estimation of its parameters μ and σ . To our knowledge, such a phenomenon has not been reported previously. The physiological basis for this behavior could be related to the production, reabsorption, and excretion of GAA and transport of CRE in the renal tubular cells.

The age variation is substantial for both the estimated central tendency (μ) and dispersion (σ). The latter factor points to a heteroskedasticity that must be adequately corrected for in clinical practice. It is important to note the opposite course for CRE/crn and GAA/crn in the neonatal period, where the former decreases and the latter increases. Clinical awareness of these properties is essential to early diagnosis.

The present study has led to the recognition of age periods where either false-positive or, more concerning, false-negative events are likely to occur when applying published cutoff values for CRE/crn and GAA/crn, which are flat over broad age ranges (20–21). Review of available CRT cases illustrates the narrow demarcation between the upper reference limits and the lowest patient values after the first year of age. It remains to be determined where patient values are located in the neonatal period and during the first living year, and this will be critical information once an effective treatment becomes available for this disorder (22).

After correction for the continuous covariate age, z -scores may easily be used for group comparisons, e.g., between sexes and sites. It is important to check which statistically significant differences are biologically relevant with respect to reference intervals. Where differences in SD are also present, a partitioning rule (18) should be used. Females had significantly higher values of CRN/crn and GAA/crn than males for all ages. Nevertheless, it is known that symptomatic female carriers of the SLC6A8 defect do not show an increase of the CRE/crn ratio comparable to that in males. The mechanism behind this is not known, although it is reasonable to speculate that it could be attributed to differences in X-inactivation pattern in females. The reason for the sex

difference in childhood needs further study, although differences in muscular mass may contribute to this effect in adolescence and adulthood.

The site-specific differences in z -score mean and SD may be related to instrumentation, standardization of sampling, and population properties. The conversion of values to age-corrected z -scores allows correction for such differences as illustrated in Eq. (15). This also applies to sex effects.

It could be argued that the outliers in the present parametric regression model may actually represent affected cases. Even if that were the case, their exclusion would have been the desired outcome for a process designed to determine reference intervals. Further outcome evaluation of these cases could be undertaken, but its completeness will likely be hampered by the necessity to contact ordering physicians at referring institutions, and in some cases by the length of time passed since the testing was done.

Many biomarkers for inborn errors of metabolism are measured in urine and divided by creatinine. So far, we have not encountered a urinary component other than CRE/crn and GAA/crn (as well as the absolute concentrations of CRE and GAA) that exhibits the peculiar kinked appearance of the z -score plots observed in this study. Amino acids, uric acid, and other purines and pyrimidines apparently do not exhibit such behavior (L. Mørkrid, unpublished observations). Other disease markers excreted in the urine should also be thoroughly investigated in large populations for similar properties before they are factored into diagnostic algorithms. In addition to the collection of a sufficient number of reference results, the type of analysis described here is labor intensive and requires extensive statistical experience. For this reason, we have incorporated these complex calculations and the resulting adjustments for age and other covariates in postanalytical interpretive tools for CDS in a Web-based multivariate pattern recognition software

called Collaborative Laboratory Integrated Reports (CLIR) (<https://clir.mayo.edu>; access is password protected and requires user registration). CLIR is modeled after the Region 4 Stork worldwide collaborative project for laboratory quality improvement of newborn screening by tandem mass spectrometry (23–25). Like the newborn screening cluster of applications, our objective is to make them freely available to all interested users willing to share data and extend these applications to cover a broad spectrum of either clinical or research endeavors. This evidence-based approach could add substantial value to patient care by providing an environment conducive to standardization through normalization of z -scores, leading to enhanced interpretation of complex laboratory profiles that is driven by multisite evidence and objective peer comparison.

Author Contributions: All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.

Authors' Disclosures or Potential Conflicts of Interest: Upon manuscript submission, all authors completed the author disclosure form. Disclosures and/or potential conflicts of interest:

Employment or Leadership: P.L. Hall, Emory University; G.S. Salomons, Society of Inborn Errors of Metabolism.

Consultant or Advisory Role: None declared.

Stock Ownership: None declared.

Honoraria: None declared.

Research Funding: P. Rinaldo, the T. Denny Sanford professorship fund, Mayo Clinic.

Expert Testimony: None declared.

Patents: None declared.

Role of Sponsor: The funding organizations played no role in the design of study, choice of enrolled patients, review and interpretation of data, or preparation or approval of manuscript.

References

- Valongo C, Cardoso ML, Domingues P, Almeida L, Verhoeven N, Salomons G, et al. Age related reference values for urine creatine and guanidinoacetic acid concentration in children and adolescents by gas chromatography-mass spectrometry. *Clin Chim Acta* 2004;348:155–61.
- Arias A, Ormazabal A, Moreno J, González B, Vilaseca MA, García-Villoria J, et al. Methods for the diagnosis of creatine deficiency syndromes: a comparative study. *J Neurosci Methods* 2006;156:305–9.
- Mercimek-Mahmutoglu S, Stöckler-Ipsiroglu S, Salomons GS. Creatine deficiency syndromes. In: Pagon RA, Adam MP, Ardinger HH, Bird TD, Dolan CR, Fong CT, Smith RJH, Stephens K, editors. *GeneReviews*. Seattle, WA: University of Washington; 1993–2014.
- van de Kamp JM, Mancini GM, Salomons GS. X-linked creatine transporter deficiency: clinical aspects and pathophysiology. *J Inher Metab Dis* 2014;37:715–33.
- Sims EA, Seldin DW. Reabsorption of creatine and guanidinoacetic acid by the renal tubules. *Am J Physiol* 1949;157:14–20.
- Kadono K, Yamaguchi T, Tetsutani T, Yasunaga K. Clinical test of renal guanidinoacetic acid metabolism by oral citrulline and creatine loading. *Clin Chim Acta* 1992; 209:9–17.
- Edison EE, Brosnan ME, Meyer C, Brosnan JT. Creatine synthesis: production of guanidinoacetate by the rat and human kidney in vivo. *Am J Physiol Renal Physiol* 2007;293:F1799–804.
- Reznichenko A, Sinkeler SJ, Snieder H, van den Born J, de Borst MH, Damman J et al. SLC22A2 is associated with tubular creatinine secretion and bias of estimated GFR in renal transplantation. *Physiol Genomics* 2013; 45:201–9.
- Hilsted L, Rustad P, Aksjøglæde L, Sørensen K, Juul A. Recommended Nordic paediatric reference intervals for 21 common biochemical properties. *Scand J Clin Lab Invest* 2013;73:1–9.
- Colantonio DA, Kyriakopoulou L, Chan MK, Daly CH, Brinc D, Venner AA, et al. Closing the gap in pediatric laboratory reference intervals: a CALIPER database of 40 biochemical markers in a healthy and multiethnic population of children. *Clin Chem* 2012;58:854–68.
- Koenker R, Hallock KF. Quantile regression. *J Econom Persp* 2001;15:143–56.
- Solberg HE. The IFCC recommendation on estimation of reference intervals. The RefVal program. *Clin Chem Lab Med* 2004;42:710–4.
- van de Kamp JM, Betsalel OT, Mercimek-Mahmutoglu S, Abulhoul L, Grünewald S et al. Phenotype and genotype in 101 males with X-linked creatine transporter deficiency. *J Med Genet* 2013;50:463–72.
- Comeaux MS, Wang J, Wang G, Kleppe S, Zhang VW, Schmitt ES et al. Biochemical, molecular, and clinical

-
- diagnoses of patients with cerebral creatine deficiency syndromes. *Mol Gen Metab* 2013;109:260–8.
15. Horn PS, Feng L, Li Y, Pesce AJ. Effect of outliers and nonhealthy individuals on reference interval estimation. *Clin Chem* 2001;47:2137–45.
 16. Altman DG. *Practical statistics for medical research*. London: Chapman & Hall; 2006.
 17. Gowans EM, Pedersen PH, Blaaberg O, Hørdér M. Analytical goals for the acceptance of common reference intervals for laboratories throughout a geographical area. *Scand J Clin Invest* 1988;48:757–64.
 18. Lahti A, Petersen PH, Boyd JC, Rustad P, Laake P, Solberg HE. Partitioning of nongaussian-distributed biochemical reference data into subgroups. *Clin Chem* 2004;50:891–900.
 19. Cole TJ. Using the LMS method to measure skewness in the NCHC and Dutch national height standards. *Ann Hum Biol*. 1989;16:407–19.
 20. Haas D, Gan-Schreier H, Langhans CD, Anninos A, Haege G, Burgard P et al. Diagnosis and therapeutic monitoring of inborn errors of creatine metabolism and transport using liquid chromatography-tandem mass spectrometry in urine, plasma and CSF. *Gene* 2014; 538:188–94.
 21. Joncquel-Chevalier Curt M, Cheillan D, Briand G, Salomons GS, Mention-Muliez K, Dobbelaere D et al. Creatine and guanidinoacetate reference values in a French population. *Mol Genet Metab* 2013;110:263–7.
 22. Dunbar M, Jaggamantri S, Sargent M, Stockler-Ipsiroglu S, van Karnebeek CD. Treatment of X-linked creatine transporter (SLC6A8) deficiency: systematic review of the literature and three new cases. *Mol Genet Metab* 2014;112:259–74.
 23. McHugh DMS, Cameron CA, Abdenur JE, Abdulrahman M, Adair O, Al Nuaimi SA 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; 13:230–54.
 24. Marquardt G, Currier R, McHugh DM, Gavrilov D, Magera MJ, et al. Enhanced interpretation of newborn screening results without analyte cutoff values. *Genet Med* 2012;14:648–54.
 25. Hall PL, Marquardt G, McHugh DM, Currier RJ, Tang H, Stoway SD, Rinaldo P. Post-analytical tools improve performance of newborn screening by tandem mass spectrometry. *Genet Med* 2014;12:889–95.