Indirect method for validating transference of reference intervals

Lykkeboe, Simon; Nielsen, Claus Gyrup; Christensen, Peter Astrup

Published in:
Clinical Chemistry and Laboratory Medicine

DOI (link to publication from Publisher):
10.1515/cclm-2017-0574

Publication date:
2018

Document Version
Publisher's PDF, also known as Version of record

Simon Lykkeboe, Claus Gyrup Nielsen and Peter Astrup Christensen*

Indirect method for validating transference of reference intervals

https://doi.org/10.1515/cclm-2017-0574
Received June 30, 2017; accepted August 23, 2017; previously published online October 14, 2017

Abstract

Background: Transference of reference intervals (RIs) from multicentre studies are often verified by use of a small number of samples from reference individuals or by the use of one serum sample (Serum X for NORIP RI). Despite recommended and appropriate methods, both have inconveniencies and drawbacks. Several attempts have been made to develop an indirect method, which uses historical data from the laboratory. These methods are retrospective relying on older test results. A near prospective method would be preferable for the laboratories introducing new methods or changing analytical platforms.
Methods: We performed a data mining experiment using results from our laboratory information system covering patients from a large geographic area. Request patterns for patients with assumed healthy characteristics were identified and used to extract laboratory results for calculation of new RI by an indirect method. Calculated RI and confidence intervals (CIs) were compared to transferred NORIP RI verified by NFKK Reference Serum X.
Results: We found that our indirect method and NFKK Reference Serum X in general produced similar results when verifying transference of RI. The method produces results for all stratifications. Only single stratifications and one analyte showed unexplained incongruences to the NORIP RI.
Conclusions: Our results suggest using request patterns as a surrogate measure for good health status. This allows for a data mining method for validation of RI or validating their transference, which is likely to be applicable in countries with similar healthcare and laboratory information system.

Keywords: data mining; indirect; NORIP; reference interval; transference; validation.

***Corresponding author: Peter Astrup Christensen,** Department of Clinical Biochemistry, Aalborg University Hospital, Hobrovej 18-22, 9000 Aalborg, Denmark, Phone: +45 97649000,
E-mail: Peter.christensen@rn.dk
Simon Lykkeboe and Claus Gyrup Nielsen: Department of Clinical Biochemistry, Aalborg University Hospital, Aalborg, Denmark

Introduction

The availability and use of reference intervals (RIs) from large multicentre studies supports mobility of patients between hospitals and regions. The individual clinical biochemistry laboratory also benefits from multicentre studies as they might transfer these RIs instead of establishing local RIs, a task beyond the scope of most laboratories. The common method according to (EP28-A3C) published by the Clinical and Laboratory Standards Institute (CLSI) and International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) recommends transference of RIs to be verified by ≥20 reference individuals from the laboratory's own healthy subject population [1–3]. The Nordic Reference Interval Project 2000 (NORIP) uses a different approach that relies on analysis of a single reference serum specimen NFKK Reference Serum X (Serum X). The Serum X results allow for assessing bias for the laboratory methods. If specified bias goals are met, NORIP RIs can be implemented [4]. Both the CLSI/IFCC and the Serum X methods have known issues. For the CLSI/IFCC method, the laboratory still has to identify reference individuals, and result evaluation makes it possible to oversee a clinical relevant bias due to power problems or accepting too wide RIs. The Serum X approach estimates bias at a single level only although bias often differs over the measuring range, which has to be somewhat subjectively accounted for, e.g. by inspecting data combined with general knowledge about methods in use. Theoretically, the CLSI/IFCC and the Serum X-based methods could lead to false acceptance or rejection of RI transference. This poses uncertainties when attempting RI transfer. Although difficult to develop, more objective and robust methods would be very preferable. Recently, via data mining methods, several comparisons have been made between the direct method for establishing RIs and the indirect method [5–9]. The indirect method makes use of patient data stored in laboratory information systems (LIS). A typical method assumes that individuals with good health have infrequent contact to healthcare. Infrequent contact is derived by counting the number of requested laboratory tests. These methods extract test results from individuals with a low number of test requests in preceding and following years. Unfortunately, the included

results are therefore not necessarily contemporary. This makes the methods less suitable for transference of RIs to present methods. A near prospective method would be preferable, for instance, when introducing new methods with RI supplied by the manufacturers or when changing analytical platforms. In the present study, we investigate a near prospective indirect method based on data mining an LIS for individuals with an apparent low indication for biochemistry investigations. The laboratory's healthy subject population was sampled from the outpatient population using a single selection criteria. Indirect RIs are calculated as 2.5th and 97.5th percentile with the non-parametric approach after proper selection of the suitable assumed healthy individuals. The obtained RIs are compared with NORIP RIs and Serum X verification.

Materials and methods

Data sets

In 2016, the laboratory performed around 8.5 million tests covering around 70% of the yearly requests for biochemical testing in the regional LIS database. As the LIS covers a large geographical area, all biochemical contacts to healthcare can be assumed to be registered in the LIS. Biochemical test are requested by hospitals, general practitioners and medical specialists. Approximately half of the overall test requests were on outpatients.

Analysis of request pattern

From a data set of outpatient results in the period of (01.04.2017–31.05.2017), all unique requests were identified. The time between these requests and the previous request on the patients was calculated and designated time to previous request (TTPR) (Figure 1). If no previous request was found we assigned a value of >24 months. For each month, we visualised the number of results and calculated the percentage of test results within their respective reference interval (Figure 2).

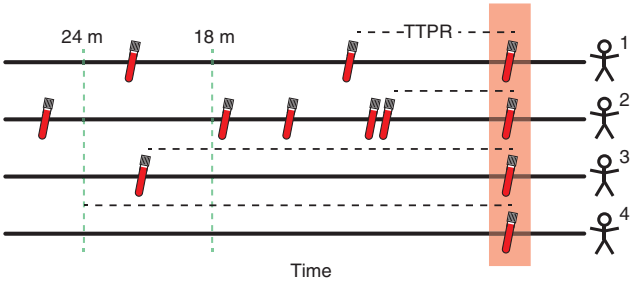

Figure 1: Diagram outlining the inclusion criteria and calculation of TTPR.
Four different possible request patterns are shown. Red box indicate sampling time. Tubes indicate requests for biochemical testing. Black dashed lines indicate calculated TTPR. Green dashed lines indicate TTPR 18 and 24 months, respectively.

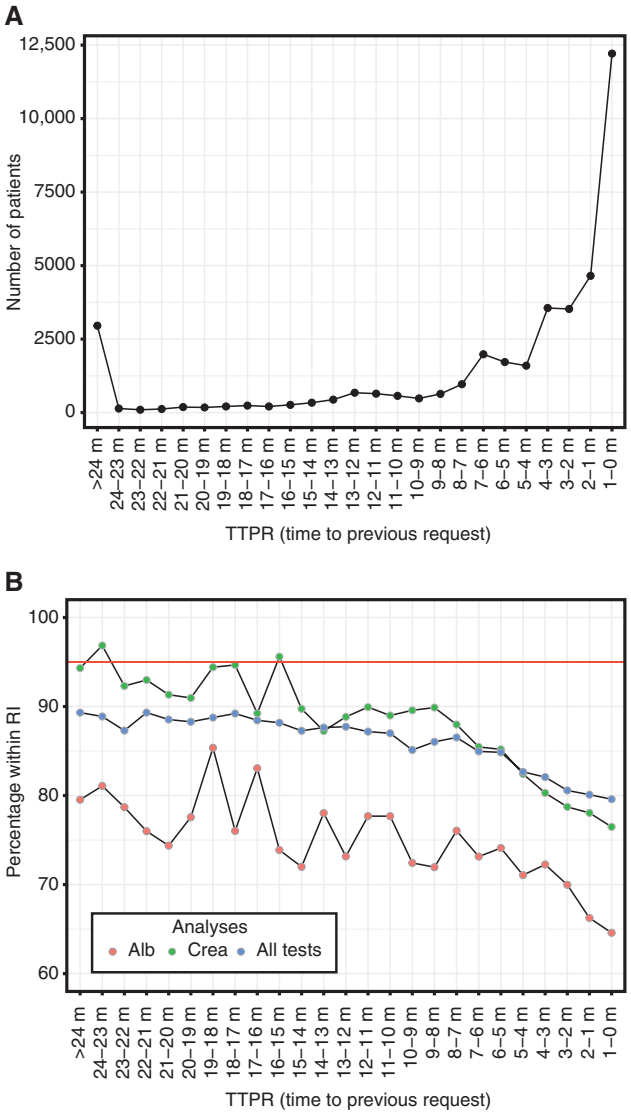

Figure 2: General analysis of request pattern.
(A) Number of patients identified in each interval of TTPR.
(B) Percentage of patient results in each interval of TTPR falling within the NORIP RIs. Blue is combined results for all tests. The red curve is albumin results and creatinine results are green.

Analysis of NFKK reference Serum X

Serum X was analysed according to the manufacturer's instructions, without investigating for correction [4, 10]. Briefly, 10 measurements of each analyte were performed in one analytical run, including controls before and after the analytical run. Results were analysed for bias to the certified or indicative values in the NORIP provided Excel sheet [11].

Indirect RI determination

Patient results were extracted from the LIS using methods similar to previous published methods [6–8]. Included test results were all analysed at Aalborg University Hospital in the period of (26.03.2017–02.06.2017). Plasma samples from general practitioners were either collected in our outpatient clinic or delivered from the general practitioners by land transport held at 21 °C. The following inclusion criteria's were used:

- Each unique patient had only one request for biochemical testing in the regional LIS within 18 months retrospective of an included sample. This criterion ensured exclusion of patients undergoing yearly controls and thus focusing the included results on individuals with limited indication for biochemical testing;
- Furthermore, the data retrieved from the LIS database are all from outpatients consulting general practitioners and not from other medical specialists or hospital departments. These outpatients constitute the laboratory's subpopulation with the lowest percentage of diseased subjects.

Biochemical analysis

All tests were analysed in a routine clinical biochemistry laboratory on c502/c702/e602 Cobas 8000 clinical chemistry and immunochemistry modules (Roche Diagnostics A/S, Hvidovre, Denmark), except LDL-cholesterol, which was calculated with Friedewalds formula [12].

Calculation of RIs

Calculations, statistical evaluations, graphical representations and simulations were made in Rstudio (Version 1.0.136 with R Version 3.3.2). Tukey's fence was used for defining and removing outliers as described by Horn et al. [2, 13]. Tukey's fence is defined by 1st quartile–1.5 × (3rd quartile–1st quartile) and 3rd quartile + 1.5 × (3rd quartile–1st quartile). Data were visualised in histograms and Q-Q plots for evaluating the validity of outlier definition and possible transformation.

The results were evaluated in three groups:

- Stratifications with <60 results after outlier removal. The remaining results are evaluated according to % of results above NORIP upper limit (% AUL) and % of results under NORIP lower limit (% ULL). This group includes urea.
- Stratifications with ≥60 results after outlier removal and data assumes a normal distribution. The remaining results are evaluated for % AUL and % ULL. RI limits are calculated as well as a 90% confidence interval (CI).
- Stratifications with ≥60 results and data are not normal distributed. Results are transformed with natural logarithm and outliers are removed. The remaining results are evaluated for % AUL and % ULL and used to calculate a RI with CI. This group includes alanine transaminase; alkaline phosphatase; amylase, pancreatic; bilirubin; creatinekinase; γ-glutamyltransferase; and triglyceride.

The 95% RI limits were evaluated with non-parametric methods (2.5% and 97.5% percentiles) on the population/transformed population without outliers. The 90% CIs of the limits were calculated by bootstrap resampling with replacement as described in [14, 15]. Per distribution, 500 resamplings were made and RI limits were determined for all 500 resamplings. From these RI limits, the 5% and 95% percentile constitute the 90% CI of the RI [15].

Accept criteria for transfer of RIs

A validation of transference of the NORIP RI was considered accepted based on the following criteria:

- NORIP limit is within CI of indirect limit.
- Accept criteria based on desirable performance related to biological variation [16]. In total, between 0% and 6% of patient results without outliers are allowed outside the NORIP limits (sum of % ULL and % AUL). To make sure the distribution is not too biased, a maximum difference of 3% is allowed (difference between % ULL and % AUL). This allows for a maximum of 4.5% outside one RI limit, and this only in the case where 1.5% is outside the other RI limit.
- Accept criteria similar to validation and transference of RIs using small numbers of reference individuals, for sample sizes between 30 and 59. Ten percent outside the RIs is found acceptable as this leaves the probability of false rejection below 6.1%. If half or more of the stratifications for one analyte are verified, all stratifications are considered verified.

Ethics

The study was a technical and quality investigation in accordance with the guidelines of the Northern Denmark Regional Science and Ethics Committee.

Results

We performed a preliminary analysis of request patterns by identifying outpatients who consulted their general practitioners within a period of 2 months and calculated the TTPR (Figure 1). Investigating the number of patients in each TTPR interval shows that local maxima in the number of patients occur at 1, 3, 6 and 12 months, suggesting a significant amount of patients receiving regular scheduled follow-up testing (Figure 2A). The apparent absence of a local maximum later than 12 months suggests that after 12 months, regular scheduled follow-up consultations are rare events. Within each time interval, we calculated the percentage of patient results within the NORIP RIs (Figure 2B). With decreasing TTPR, we observe more patient results falling outside the RIs. The percentage of patient results within the RIs seems to be stable in the TTPR period >24–18 months, indicating a population with homogenous biochemistry. Based on these results, we chose 18 months as the cutoff for inclusion of patients for verification of RIs. Analysing the results from a single reference interval stratification in groups of TTPR was done for male creatinine results (Figure 3). This showed

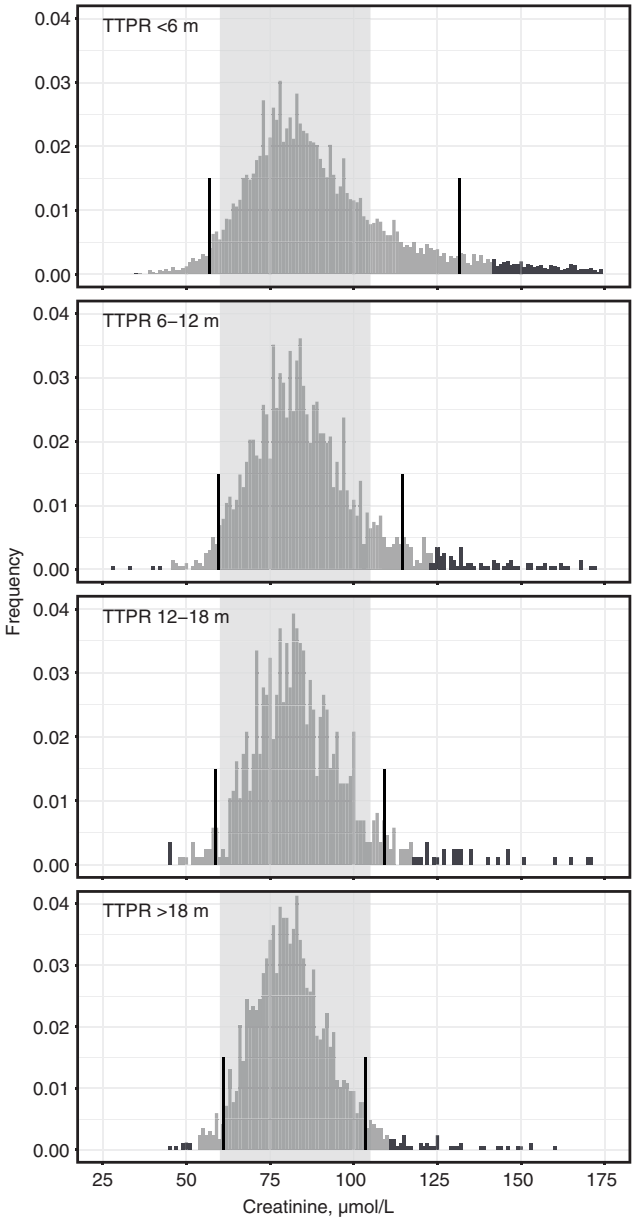

Figure 3: Histograms of creatinine results for males.
Light grey area covers NORIP RI. Black lines are new calculated RIs (95%). Black columns are outliers, grey columns are included in the calculation. Each panel outlines calculation in different TTPR intervals. Each TTPR panel contains >800 results.

that with increasing TTPR, the grouped results became more homogeneous, became increasingly normal, and contain less outliers, and the indirect determined limits approach the NORIP limits.

Using these inclusion criteria, the indirect method collected enough results for all tests in 9 weeks. These results enabled a possible validation of RIs or verification of transference for all RIs with all stratifications (Tables 1 and 2). Table 1 shows that only the analytes glucose, γ-glutamyltransferase, and one stratification of phosphate had an increased number of outliers. The remaining tests have between 0% and 3.5% outliers. The high number for glucose is most likely due to unknown fasting status. In general, the method gave results in concordance with Serum X verified transfer of NORIP RIs.

In Table 1 part I, RIs were verified by Serum X. Here 19 of 26 stratifications were verified by the indirect method, and all analytes have at least one stratification verified. Female stratifications of γ-glutamyltransferase have only upper limits verified. Similarly, urate in young females and males has only one limit verified. Cholesterol and LDL-cholesterol limits are not verified for the age stratification ≥50 as well as HDL-cholesterol for females.

The analytes in Table 1 part II were verified by Serum X but has zero or only single limits verified by the indirect method. For the potassium method, the manufacturer states the RI to 3.4–4.5 mmol/L (Roche, ISE indirect Na-K-Cl for Gen.2), which is a bit closer to the result of the indirect method. In relation to lactate dehydrogenase, the level increased by 20–25 U/L upon introduction of the transport service of whole blood at 21 °C. Moreover, as 90% of the included samples arrives from the general practitioners via the transport service, this effect is clearly seen here. Subtracting the transport introduced bias from the CIs brings the indirect method close to verifying the NORIP RIs. Upper limits of triglyceride and glucose are as expected difficult to verify as the method presented here does not take into account if the patient is fasting or not. Lower limits for glucose were though verified. For the iron method, the manufacturer states the RI to 5.83–34.5 µmol/L (Roche, Iron Gen.2), which for the lower limit also is closer to the indirect method.

In contrast to parts I and II, part III contains RIs not verified by Serum X. The indirect method agrees with this on all three tests. Albumin results are clearly under the NORIP RIs. We were already aware of this discrepancy and published this recently [17]. In our verification, bilirubin was found to have a very large imprecision in the lower analytical range (<10 µmol/L). This makes it impossible to evaluate the lower limit with our instruments. Similarly, it makes it impossible to transfer the RI by use of Serum X as the certified value is 8.97 µmol/L. However, for our method, the manufacturer states the RI to <21 µmol/L (Roche, Bilirubin total gen.3). Also for phosphate, the transport service of whole blood at 21 °C affected the analyte level. The level dropped 0.1 mmol/L upon introduction of the transport service. This decrease has also been reported previously in a study from a laboratory utilising a similar transport service [18]. Adding this transport

Table 1: Comparison of NORIP RI and RI determined by the indirect method.

Part	Test	Sex	Age	NORIP	n (n outlier)	% ULL	% AUL	Indirect	CI LL	CI UL	Verify	Limit within CI
I	Alanine transaminase, U/L	M	≥18	10–70	1680 (50)	0.2	2.2	13–69	11.8–13.2	67.1–73.3	X	UL
	Alanine transaminase, U/L	F	≥18	10–45	1909 (62)	2.3	1.6	10–43	9.3–10.1	40.1–44.2	X	LL
	Alkaline phosphatase, U/L		≥18	35–105	2651 (45)	1.5	4.5	36–114	35.7–37.3	111.5–116.6	X	
	Amylase, pancreatic, U/L		≥18	10–65	1266 (25)	0	0.2	13–53	12.6–13.7	51.2–54.2	X	
	Creatinekinase, U/L	M	18–49	50–400	194 (5)	2.6	4.8	45–507	42.3–59.7	360.4–645.7		LL, UL
	Creatinekinase, U/L	M	≥50	40–280	168 (1)	1.2	3	41–282	38.5–44.9	257.4–333.3	X	LL, UL
	Creatinekinase, U/L	F	≥18	35–210	348 (4)	2.3	3.8	32–247	28.8–36.5	202.9–264.9		LL, UL
	Creatinine, µmol/L	F	≥18	45–90	2099 (65)	1.6	0	46–84	45.4–47	83.3–85.3	X	
	Creatinine, µmol/L	M	≥18	60–105	1819 (48)	1.7	1.8	61–103	60.1–62.3	102.5–105.1	X	UL
	γ-glutamyltransferase, U/L	F	18–39	10–45	149 (8)	5.7	2.8	8–48	6.7–9.3	40.4–52.8		UL
	γ-glutamyltransferase, U/L	F	≥40	10–75	181 (4)	2.8	3.4	9–96	8.2–9.7	67.3–114.8		UL
	γ-glutamyltransferase, U/L	M	18–39	10–80	133 (6)	0	0	11–50	9.7–11.5	42.6–64.9	X	LL
	γ-glutamyltransferase, U/L	M	≥40	15–115	222 (15)	3.4	3.9	13–133	11.2–15.2	103–154.8		LL, UL
	Magnesium, mmol/L		≥18	0.71–0.94	65 (0)	1.5	0	0.69–0.92	0.69–0.75	0.91–0.92	X	LL
	Sodium, mmol/L		≥18	137–144	3696 (79)	0.7	3.7	137–145	137.2–137.4	144.6–144.9	X	LL[a]
	Urate, mmol/L	F	18–49	0.155–0.35	229 (3)	4.9	3.1	0.14–0.36	0.11–0.15	0.34–0.38		UL
	Urate, mmol/L	F	≥50	0.155–0.4	141 (3)	2.2	0	0.15–0.38	0.14–0.18	0.35–0.39	X	LL
	Urate, mmol/L	M	≥18	0.23–0.48	342 (5)	2.1	5.6	0.23–0.51	0.21–0.24	0.49–0.52		LL
	Cholesterol, mmol/L		18–29	2.9–6.1	304 (6)	1	0.7	3.0–6.0	2.9–3.07	5.53–6.1	X	LL, UL
	Cholesterol, mmol/L		30–49	3.3–6.9	855 (11)	1.9	2.3	3.3–6.9	3.16–3.44	6.81–7.06	X	LL, UL
	Cholesterol, mmol/L		≥50	3.9–7.8	1152 (22)	5.5	0.1	3.5–7.3	3.43–3.62	7.18–7.48		
	LDL-cholesterol, mmol/L		18–29	1.2–4.3	289 (6)	1.8	0	1.2–3.7	1.12–1.34	3.54–3.77	X	LL
	LDL-cholesterol, mmol/L		30–49	1.4–4.7	828 (10)	1.6	1	1.4–4.6	1.35–1.51	4.47–4.63	X	LL
	LDL-cholesterol, mmol/L		≥50	2–5.3	1113 (13)	10.3	0.5	1.3–4.8	1.14–1.39	4.72–4.98		
	HDL-cholesterol, mmol/L	F	≥18	1–2.7	1141 (17)	6.2	0	0.9–2.5	0.85–0.92	2.4–2.54		
	HDL-cholesterol, mmol/L	M	≥18	0.8–2.1	1146 (28)	4.6	1.4	0.8–2.1	0.72–0.76	2.01–2.1		LL[a], UL
II	Potassium, mmol/L		≥18	3.5–4.4	3753 (60)	2.3	8.2	3.5–4.6	3.44–3.47	4.61–4.65		LL[a]
	Lactate dehydrogenase, U/L		18–69	105–205	1204 (35)	0.1	20.9	132–242	127–135.3	237.8–246.3		
	Lactate dehydrogenase, U/L		≥70	115–255	128 (2)	0	3.2	140–257	124.7–150.7	245.7–264.7		UL
	Iron, µmol/L		≥18	9–34	1090 (23)	10	0	4–30	3.8–5.1	28.8–30.5		
	Glucose, mmol/L		≥18	4.2–6.3	231 (20)	0.9	13.7	4.3–7.1	4.19–4.55	6.83–7.33		LL
	Triglyceride, mmol/L		≥18	0.45–2.6	2278 (17)	1	9.5	0.5–3.8	0.49–0.52	3.65–4.24		
III	Albumin, g/L		18–39	36–48	1045 (12)	9	0	34–45	33–33.9	44.8–45.5		
	Albumin, g/L		40–69	36–45	1393 (18)	15.1	0	33–43	32.7–33.2	42.8–43.4		
	Albumin, g/L		≥70	34–45	266 (5)	22.2	0	30–42	29.2–30.3	40.6–42.1		
	Bilirubin, µmol/L		≥18	5–25	1713 (32)	14.5	0	3–19	3–3	18–19.1		
	Phosphate, mmol/L	F	≥18	0.76–1.41	203 (0)	13.8	0	0.59–1.25	0.49–0.64	1.22–1.31		
	Phosphate, mmol/L	M	18–49	0.71–1.53	85 (3)	14.6	0	0.51–1.18	0.5–0.57	1.1–1.27		
	Phosphate, mmol/L	M	≥50	0.71–1.23	64 (4)	16.7	0	0.56–1.06	0.56–0.62	1.03–1.07		
IV	Calcium, mmol/L		≥18	2.15–2.51	2101 (32)	0	3.8	2.22–2.54	2.21–2.23	2.53–2.54		

n, total number of results; n outlier, number of outliers identified; % ULL, % of results (without outliers) under NORIP lower limit; % AUL, % of results (without outliers) above NORIP upper limit; Indirect, reference interval determined by indirect method; CI LL, confidence interval on lower limit; CI UL, confidence interval on upper limit; Verify, X denotes limits verified according to criteria 2; Limit within CI, LL denotes, NORIP lower limit within CI of indirect lower limit; UL denotes, NORIP upper limit within CI of indirect upper limit; LL[a] denotes, NORIP lower limit equals rounded indirect lower limit. RIs in parts I, II, and IV were verified by Serum X.

Table 2: Verification of transference of NORIP RI by a small number of results.

Test	Sex	Age	NORIP	n (n outlier)	% ULL	% AUL	Verify
Urea, mmol/L	F	18–49	2.6–6.4	54 (2)	5.8	0	X
Urea, mmol/L	F	≥50	3.1–7.9	46 (2)	0	0	X
Urea, mmol/L	M	18–49	3.2–8.1	55 (0)	7.3	0	X
Urea, mmol/L	M	≥50	3.5–8.1	31 (3)	17.9	3.6	(X)

n, total number of results; n outlier, number of outliers identified; % ULL, % of results (without outliers) under NORIP lower limit; % AUL, % of results (without outliers) above NORIP upper limit; Verify, X denotes limits verified according to criteria 3; (X) accepted as more than half of the stratifications are verified.

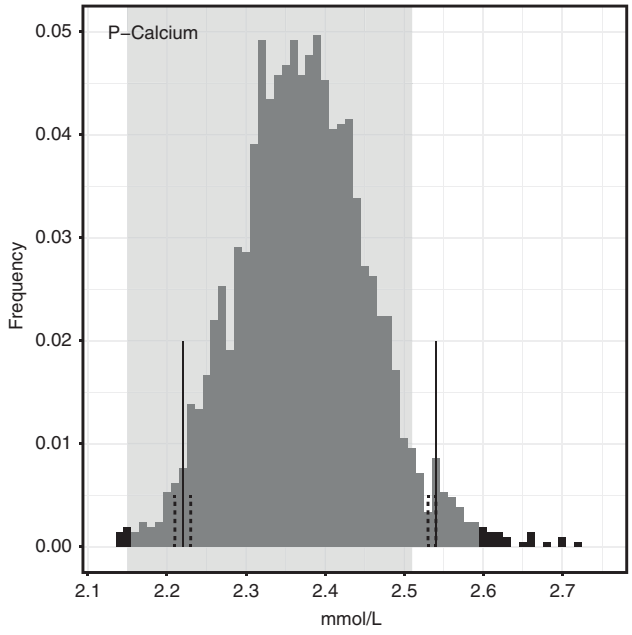

Figure 4: Histogram of calcium results.
Light grey area covers NORIP RI. Long lines are calculated RI (95%). Short dotted lines are 90% confidence interval for the reference limits. Black columns are outliers, grey columns are included in the calculation.

introduced bias to the CIs brings the indirect method close to verifying the NORIP RIs.

Table 1 part IV contains the calcium RI. Serum X verified the transfer of the RI, whereas the indirect method documents a different RI (Figure 4). In particular, the lower limit differs with more than acceptable bias [19].

Table 2 shows urea. Serum X verified the transfer of this RI. The indirect method verifies three of four stratifications with ≤10% of the results outside the RI to transfer. The fourth stratification is considered verified as half or more than half of the stratifications are verified.

Discussion

Based on the current results, we find that the presented indirect method is very useful for verification of RI transference. Comparing our results to Serum X verification of transfer suggests that we can use the TTPR as a surrogate for health status. The obvious advantages by this method is that all stratifications can be quickly checked and not only one stratification as by the CLSI/IFCC method with a small number of reference individuals. The only criteria is that the algorithm collects preferably >60 results per stratification without outliers. The possible power problem by using the CLSI/IFCC method using a small number of reference individuals is also avoided. This method could verify the transfer of albumin RIs as only 9% of all samples are outside the RIs in one stratification despite a significant bias. The optimal acceptance criteria should approach the definition of an RI (2.5% outside the limits on each side). To allow some variation, is seem reasonable to use the desirable performance goal based on biological variation with minimum of 1.4% and a maximum of 4.4% outside each RI limit [16]. Modifying this slightly but in concordance with this, we choose 6% as a maximum (sum % ULL and % AUL) and a difference of 3% (difference % ULL and % AUL). By validation and transference of RIs using smaller numbers of reference individuals, it is accepted to have up to 10% outside the RIs. It is also accepted to have 0% outside one or both limits. This happens in situations where the proposed limits seem to be too wide, e.g. due to increased precision of the new method, or if the tested population is more homogenous than the original population [3]. Allowing 0% outside one or both limits of course increases the risk of accepting too wide limits, especially with the sample sizes in some tests in the present study. Serum X and the CLSI/IFCC method using a larger number of reference individuals do not overcome this problem, as both methods use accept criteria based on bias between means or to Serum X. Our indirect method is developed for validation of RIs to justify transference only. The intentions are not to change any limits. However, the indirect method does warrant new investigations of some test or test stratifications. Inspecting some of these cases shows the strengths of our approach. Pancreatic amylase RIs were calculated to 13–53 U/L. This is exactly similar to the RI supplied by the manufacturer (Roche, Pancreas-α-amylase EPS). This

would suggest the use of manufacturer RI instead of NORIP. Inspecting magnesium, urate and female stratifications of creatinine and γ-glutamyltransferase where either % ULL or AUL is 0 or only one limit is verified by CI, it is found that the absolute differences are minor. LDL-cholesterol (age ≥50) cannot be verified if cholesterol (age ≥50) is not verified as it is calculated based on cholesterol. However, it could warrant further investigation in the age ≥50 stratification, as it is very close to the 30–49 age stratification. For HDL-cholesterol (females), both limits determined by the indirect method are lower than the NORIP limits. This could be due to the well-known preanalytical factors that affect HDL-cholesterol concentrations in females, e.g. pregnancy and progestins [20]. The upper limits of young males γ-glutamyltransferase and iron are not verified; for this, we find no obvious reasons. Possibly all γ-glutamyltransferase stratifications could benefit from a longer sampling time (larger n, without outliers). In addition to the lower limit for iron supplied by the manufacturer, the very low lower limit found by the indirect method could also be due to inability to exclude some anaemic subjects.

In conclusion, we find our indirect approach useful, e.g. for verification of RIs from a manufacturer on a new method or for validation of the transference of RI between methods even after a short operation time. This is possible as it uses current test results and only historic request patterns without the addition of clinical information. We also find our indirect method fit for reviewing or validating current RIs [21, 22]. It avoids the obstacles by finding and obtaining plasma from reference individuals for the CLSI/IFCC method. The method is of limited use where information about the patients is essential, e.g. fasting in relation to lipid status. Using information from all stratifications makes it possible to perform a critical review of selected RIs. Rejection of RI transference would likely be due to true differences between populations or methods. However, some differences may arise from other factors, especially selection bias and disease prevalence. For instance, tests that are not typically ordered on healthy patients may be associated with investigation of likely or confirmed disease and therefore associated with higher disease prevalence. A primary limitation is the prerequisite of an LIS providing complete request history for each patient in the study period. Incomplete patient request history may result in the inclusion of unhealthy subjects compromising transfer evaluation. Ensuring complete patient request history is not trivial but is likely when the LIS covers a large geographical area as in this study. With this, we will encourage others to investigate their LIS and TTPR for finding their laboratory's subpopulation with the lowest percentage of diseased subjects for verification of RIs by indirect methods.

Acknowledgments: The authors thank the staff at the Department of Clinical Biochemistry, Aalborg University Hospital, for valuable discussions and Professor Aase Handberg for critical reading of the manuscript.
Author contributions: All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.
Research funding: None declared.
Employment or leadership: None declared.
Honorarium: None declared.
Competing interests: The funding organisation(s) played no role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the report for publication.

References

1. Ceriotti F, Hinzmann R, Panteghini M. Reference intervals: the way forward. Ann Clin Biochem 2009;46:8–17.
2. Horn PS, Pesce AJ. Reference intervals: an update. Clin Chim Acta 2003;334:5–23.
3. CLSI. Defining, establishing, and verifying reference intervals in the clinical laboratory; approved guideline – third edition. CLSI document EP28 – A3c ed: Wayne, PA, USA: CLSI (Clinical Laboratory Standards Institute), 2010.
4. Rustad P, Felding P, Franzson L, Kairisto V, Lahti A, Martensson A, et al. The Nordic Reference Interval Project 2000: recommended reference intervals for 25 common biochemical properties. Scand J Clin Lab Invest 2004;64:271–84.
5. Bakan E, Polat H, Ozarda Y, Ozturk N, Baygutalp NK, Umudum FZ, et al. A reference interval study for common biochemical analytes in Eastern Turkey: a comparison of a reference population with laboratory data mining. Biochem Med (Zagreb) 2016;26:210–23.
6. Bock BJ, Dolan CT, Miller GC, Fitter WF, Hartsell BD, Crowson AN, et al. The data warehouse as a foundation for population-based reference intervals. Am J Clin Pathol 2003;120:662–70.
7. Grossi E, Colombo R, Cavuto S, Franzini C. The REALAB project: a new method for the formulation of reference intervals based on current data. Clin Chem 2005;51:1232–40.
8. Tozzoli R, Giavarina D, Villalta D, Soffiati G, Bizzaro N. Definition of reference limits for autoantibodies to thyroid peroxidase and thyroglobulin in a large population of outpatients using an indirect method based on current data. Arch Pathol Lab Med 2008;132:1924–8.
9. Katayev A, Balciza C, Seccombe DW. Establishing reference intervals for clinical laboratory test results: is there a better way? Am J Clin Pathol 2010;133:180–6.
10. Pedersen MM, Ornemark U, Rustad P, Steensland H, Loikkanen M, Olafsdottir E, et al. The Nordic Trueness Project 2002: use of reference measurement procedure values in a general clinical chemistry survey. Scand J Clin Lab Invest 2004;64:309–20.
11. Rustad P. Evaluation spreadsheet for X. Available at: http://nyenga.net/norip/X/x.htm. Accessed: 24 May 2017.

12. Friedewald WT, Levy RI, Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. Clin Chem 1972;18:499–502.
13. Horn PS, Feng L, Li Y, Pesce AJ. Effect of outliers and non-healthy individuals on reference interval estimation. Clin Chem 2001;47:2137–45.
14. Bjerner J, Theodorsson E, Hovig E, Kallner A. Non-parametric estimation of reference intervals in small non-Gaussian sample sets. Accredit Qual Assur 2009;14:185–92.
15. Solberg HE. The IFCC recommendation on estimation of reference intervals. The RefVal program. Clin Chem Lab Med 2004;42:710–4.
16. Fraser CG, Hyltoft Petersen P, Libeer JC, Ricos C. Proposals for setting generally applicable quality goals solely based on biology. Ann Clin Biochem 1997;34:8–12.
17. Christensen PA. Reference intervals for the P-Albumin bromocresol purple method. Scand J Clin Lab Invest 2017;77:472–6.
18. Henriksen LO, Faber NR, Moller MF, Nexo E, Hansen AB. Stability of 35 biochemical and immunological routine tests after 10 hours storage and transport of human whole blood at 21 degrees C. Scand J Clin Lab Invest 2014;74:603–10.
19. Ricos C. Desirable biological variation database specifications. Available at: http://www.westgard.com/biodatabase1.htm. Accessed: 24 May 2017.
20. Warnick GR, Nauck M, Rifai N. Evolution of methods for measurement of HDL-cholesterol: from ultracentrifugation to homogeneous assays. Clin Chem 2001;47:1579–96.
21. Haeckel R, Wosniok W, Arzideh F, Zierk J, Gurr E, Streichert T. Critical comments to a recent EFLM recommendation for the review of reference intervals. Clin Chem Lab Med 2017;55:341–7.
22. Henny J, Vassault A, Boursier G, Vukasovic I, Mesko Brguljan P, Lohmander M, et al. Recommendation for the review of biological reference intervals in medical laboratories. Clin Chem Lab Med 2016;54:1893–900.