Practical Laboratory Medicine 9 (2017) 1–11

Contents lists available at ScienceDirect

Practical Laboratory Medicine

journal homepage: www.elsevier.com/locate/plabm

Generating method-specific Reference Ranges – A harmonious outcome?

Graham R. Lee*, Alison Griffin, Kieran Halton, Maria C. Fitzgibbon

Department of Clinical Biochemistry and Diagnostic Endocrinology, Mater Misericordiae University Hospital, Eccles Street, Dublin, D7, Ireland

ARTICLE INFO

Keywords:
Reference Ranges
Method-specific
Harmonisation

ABSTRACT

Objectives: When laboratory Reference Ranges (RR) do not reflect analytical methodology, result interpretation can cause misclassification of patients and inappropriate management. This can be mitigated by determining and implementing method-specific RRs, which was the main objective of this study.

Design and methods: Serum was obtained from healthy volunteers (Male + Female, n > 120) attending hospital health-check sessions during June and July 2011. Pseudo-anonymised aliquots were stored (at − 70 °C) prior t° analysis on Abbott ARCHITECT c16000 chemistry and *i*2000SR immunoassay analysers. Data were stratified by gender where appropriate. Outliers were excluded statistically (Tukey method) to generate non-parametric RRs (2.5th + 97.5th percentiles). RRs were compared to those quoted by Abbott and UK Pathology Harmony (PH) where possible. For 7 selected tests, RRs were verified using a data mining approach.

Results: For chemistry tests (n = 23), Upper or Lower Reference Limits (LRL or URL) were > 20% different from Abbott ranges in 25% of tests (11% from PH ranges) but in 38% for immunoassay tests (n = 13). RRs (mmol/L) for sodium (138 − 144), potassium (3.8–4.9) and chloride (102 − 110) were considerably narrower than PH ranges (133–146, 3.5–5.0 and 95–108, respectively). The gender difference for ferritin (M: 29–441, F: 8–193 ng/mL) was more pronounced than reported by Abbott (M: 22–275, F: 5–204 ng/mL). Verification studies showed good agreement for chemistry tests (mean [SD] difference = 0.4% [1.2%]) but less so for immunoassay tests (27% [29%]), particularly for TSH (LRL).

Conclusion: Where resource permits, we advocate using method-specific RRs in preference to other sources, particularly where method bias and lack of standardisation limits RR transferability and harmonisation.

1. Introduction

Reference Ranges (RRs) are intended to guide clinicians on the interpretation of a patient's test results, in the context of the patient's overall clinical assessment [1]. RRs are a familiar component of the result report and may have particular value when there is no previous patient data for comparison. However, their statistical derivation, inherent uncertainty and limitations are unlikely to be appreciated for routine test result evaluations, particularly by less experienced clinicians. For some analytes, e.g. Cholesterol and

Abbreviations: ALP, Alkaline Phosphatase; ALT, Alanine Aminotranfserase; AST, Aspartate Aminotransferase; Anti-TPO, Anti-Thyroid peroxidase; Anti-Tg, Anti-Thyroglobulin; CI, Confidence Interval; CK, Creatine Kinase; CRP, C Reactive Protein; FO, Far Out (Outliers); fT4, free Tetra-iodothyronine (thyroxine); fT3, free Tri-iodothyronine; LIS, Laboratory Information System; LRL, Lower Reference Limit; OS, Outside (Outliers); PH, Pathology Harmonisation; RR, Reference Range; TSH, Thyroid Stimulating Hormone; URL, Upper Reference Limit

* Corresponding author.
E-mail address: glee@mater.ie (G.R. Lee).

http://dx.doi.org/10.1016/j.plabm.2017.06.001
Received 7 November 2016; Received in revised form 19 May 2017; Accepted 30 June 2017
Available online 05 July 2017

HbA1C, RRs have been replaced by decision limits and therapeutic targets and so some of the laboratory's focus on reducing error is diverted to other extra-analytical and analytical stages of the Total Testing Pathway [2]. Where a patient has serial results available, RRs are believed to have less value in monitoring patient status [3]. Instead, an evaluation of the magnitude of any result change, as compared to a test's Reference Change Value (RCV), may provide a better means for monitoring patients and guiding management decisions [3]. RCVs are particularly useful for highlighting significant changes even when patient results remain within the RR and thereby may provide an earlier indication of changing patient status. Furthermore, use of RCVs may provide appropriate context particularly when results change from inside to outside the RR (and *vice versa*) but without significance (change < RCV).

Previous work by Harris [4] showed that the usefulness of RRs in evaluating a patient's results is limited to a minority of analytes showing a large Intra- to Inter-individual Coefficient of Variation ratio (CV_I:CV_G > 1.4). For such analytes, serial results for a (stable) patient may potentially span the entire RR rather than only a limited part of it, and therefore the RR can be a useful and sensitive tool for assessing serial changes [4,5]. RRs also permit more meaningful result interpretation when they are relevant not only to the patient's demographics (e.g. race, gender) but also to the analytical methodology generating the test result [6]. The latter factor is particularly important when large inter-methodological bias exists. Careful consideration must therefore be given to defining, establishing and verifying RRs. For such purposes, the International Federation for Clinical Chemistry and Laboratory Medicine (IFCC) and the Clinical Laboratory Standards Institute (CLSI) offer guidance and specific recommendation for non-parametric RRs, derived from a minimum of 120 reference individuals [7]. Alternative approaches such as data mining of patient data are practically advantageous, less resource intensive and can provide large (and representative) data for determining RRs with a high level of confidence [1]. Where assay standardisation has assisted comparability of results between different analytical methods, a universal or harmonised RR may also be used preferentially [6]. The transferability of such RRs across healthcare systems is welcomed by clinicians, especially when movement of patients and clinicians within or between networked hospitals results in the same test being performed with different laboratory methods. This is the “pragmatic science” approach adopted by the United Kingdom (UK) Pathology Harmony (PH) group [8], whose aim is to develop a common set of RRs on the premise that minor methodological differences between laboratories are of little clinical significance. However, this approach is not currently viable for many immunoassay based tests e.g. troponins and tumour markers, whose assays are not standardized, with concomitant analytical bias between method manufacturers [9,10].

When appropriate RRs are however established, regardless of the approach, it is also important to periodically verify their use and applicability to the current measurement system [8,11]. This is in accord with the International Organisation for Standardisation's Requirements for Quality and Competence in Medical Laboratories (ISO15189:2012), which recommends verification following a change to the examination or pre-examination procedure [12].

In this study, we report on our own experiences and challenges in establishing method-specific RRs across our clinical biochemistry and diagnostic endocrinology test repertoire during a period of laboratory consolidation and verification of new analytical platforms.

2. Materials and methods

2.1. Reference Range subjects

This RR study was part of service evaluation, involving surplus sera from adult volunteers (non-fasted) who had previously attended hospital health check sessions in June and July 2011. Only sera from volunteers less than 18 years of age were excluded from analysis. There were no other exclusion criteria.

2.2. Pre-analytical: collection, archiving and preparation of serum samples

Blood from volunteers (10 mL) was initially collected into serum separator tubes (Becton Dickinson, Franklin Lakes, NJ, USA) and centrifuged at 10,000*g*, for 10 min at room temperature prior to analysis for “health check” purposes. Post-analysis, aliquots of residual serum (1.5 mL) were pseudo-anonymised (coded) and stored (at − 70 °C) until analysis for RR purposes. On the day of analysis, stored serum samples were thawed slowly while rotating at r.t.p. before transfer to 1.5 mL plastic tubes (Sarstedt, Numbrecht, Germany).

2.3. Analytical

All reagents and calibrators were obtained from Abbott Diagnostics (Abbott Laboratories, Lake Forest, IL, USA). Internal Quality Control (IQC) material was obtained from Technopath (Ballina, Ireland) and was run at 2 or 3 levels for each analyte. Running imprecision data (CVs) are shown in Table 1. Each test was also assessed through participation in relevant schemes run by the United Kingdom National External Quality Assurance Service (UKNEQAS, Birmingham, UK).

Serum samples (n > 120) were analysed over 5 days on Abbott ARCHITECT *c*16000 and *i*2000SR analysers for a comprehensive list of chemistry (n = 23) and immunoassay (n = 13) tests. All maintenance, calibration and IQC procedures were followed in accordance with the manufacturer's instructions.