Clinica Chimica Acta 432 (2014) 68–71

Contents lists available at ScienceDirect

Clinica Chimica Acta

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

The harmonisation of growth hormone measurements: Taking the next steps

Gilbert E. Wieringa [a,*], Catharine M. Sturgeon [b], Peter J. Trainer [c]

[a] Department of Biochemistry, Bolton NHS Foundation Trust, Minerva Road, Farnworth, Bolton BL4 0JR, UK
[b] UK NEQAS [Edinburgh], Department of Laboratory Medicine, Royal Infirmary of Edinburgh, 51 Little France Crescent, Edinburgh EH16 4SA, UK
[c] Department of Endocrinology, The Christie NHS Foundation Trust, Wilmslow Road, Manchester M20 4BX, UK

ARTICLE INFO

Article history:
Received 6 September 2013
Received in revised form 8 January 2014
Accepted 8 January 2014
Available online 6 February 2014

Keywords:
Growth hormone
Method differences
Harmonisation collaboratives

ABSTRACT

For over 20 years differences in results of growth hormone (GH) measurement have been recognised as being significant enough to lead to misdiagnosis and inappropriate management of patients with GH-related disorders. Whilst issues of method standardisation, variable antibody specificity, use of different reporting units with different conversion factors, and interference from GH binding protein have been acknowledged as contributing to the discrepancies, inconsistent approaches to method harmonisation have hampered opportunities to enhance the evidence base for GH measurements. Amongst the first steps to be taken, international collaboratives recommended the universal adoption of the International Standard 98/547 and the reporting of results in mass units. Whilst inter-method variability may have improved over the last 10 years, clinically significant differences remain. A more recently recognised issue contributing to the discrepancies may be the differences in the matrix materials used by kit manufacturers to assign values to their calibrants. The establishment of an international harmonisation oversight group is recommended: its key roles to include identification of a commutable matrix reference material, assessing the clinical significance of assay interferents, the evaluation of liquid chromatography–mass spectrometry as a reference measurement procedure and the provision of acceptance criteria for the clinical application of GH methods.

1. Introduction

The measurement of growth hormone (GH) has been the cornerstone to diagnosis and management of growth hormone related disorders for many years. In the United Kingdom, amongst its criteria for offering GH replacement to adults with GH deficiency, the National Institute for Health and Clinical Excellence (NICE) specifies inclusion of a peak growth hormone response of less than 9 mU/L during an insulin tolerance test or a similar low result in another reliable test [1]. In 2002, a mean integrated 24 hour GH level less than 2.5 μg/L and/or suppression of GH below 1 μg/L in an oral glucose tolerance test (2 mU/L) have been recommended as criteria for excluding a diagnosis of acromegaly [2]. However, it has long been recognised that the variability in GH results produced by commercial kits and in-house methods challenges the appropriateness of diagnostic criteria and affects patient outcomes by prejudicing access to appropriate management. [3]. Ellis et al. in reporting performance from the UK National External Quality Assessment Service (UK NEQAS) scheme for GH noted that the most positively biased method could typically report values twice that of the most negatively biased [4]. Such differences were noted to cause 10% of laboratories to report the outcome of an insulin tolerance test as equivocal whilst 90% reported an adequate response. Arafat et al., in assessing the growth hormone suppression response during an oral glucose tolerance test, reported how results of one method were 2.3 times higher than those of a second and 6 times higher than those of a third [5]. In a retrospective study, Hauffa et al. reported that of 132 children who had been investigated for GH-related disorders, 36 would have been re-classified had the samples been measured by another method [6]. The differences are particularly relevant to the specialist endocrinologist to whom patients are referred with conflicting results from different centres which, given the low incidence of GH-related disorders (10 cases per million for isolated GH deficiency in children in the UK, 3 to 4 cases per million for acromegaly), is not an infrequent occurrence. The impact of the differences might not be so significant if method-specific diagnostic cut-offs were available. However, the evidence base for them also is poor, in part reflected by the low incidence of GH-related disorders but also affected by the limited collaboration between laboratories and their continuing willingness to transfer diagnostic cut-offs between methods despite the known biases [4,7]. These problems are compounded by the trend to apply international consensus criteria and guidelines to local practice without consideration, or even awareness, of the issues discussed here. (See Table 1.)

* Corresponding author. Tel.: +44 1204 390172.
E-mail address: gilbert.wieringa@boltonft.nhs.uk (G.E. Wieringa).

http://dx.doi.org/10.1016/j.cca.2014.01.014
0009-8981/

Table 1
GH method performance reported by UK NEQAS before and after the recommended adoption of GH IS 98/574 and reporting of GH results in mass units (μg/L).

	2005	2006	2007	2008	2009	2010	2011
No. participants	107	109	109	112	114	109	105
% labs. reporting GH in mass units (μg/L)	52%	68%	72%	79%	82%	85%	98%
GCV (%) across all methods	21%	22%	15%	9.1%	15.5%	15.2%	14.8%
VAR (%) for Siemens Immulite 2000 users	6.5%	6.1%	6.9%	7.2%	8.9%	7.3%	8.1%
No. participants using Siemens technology	69 (14/55)	92 (14/78)	94 (14/82)	99 (13/86)	99 (12/87)	94 (12/82)	87 (11/76)

GCV (%): Geometric coefficient of variation; VAR: Cumulative within method variability.

Because of the potential impact on patient care, the need to harmonise has been recognised for over 15 years [8–12]. Here we review the diversity of initiatives that have been undertaken, assess their contributions to a harmonised approach, and consider what further steps could be taken to improve methodologies for this heterogeneous polypeptide.

2. Why the discrepancies?

The reasons for the discrepancies in GH results are numerous and have been well described. They include:

- The use of calibrant materials whose values have been assigned in isolation of a defined international standard (IS). Until the availability of recombinant preparations, pituitary sourced IS 80/505 from the National Institute of Biological Standards and Control (NIBSC) with a nominally assigned bioactivity of 4.4 international units per ampoule was widely used [13]. Its advantage was that it reflected a physiological matrix that included GH moieties such as 22KDa and 20KDa monomers, dimers and hetero-isoforms thereof. Its main disadvantage was that the mass content was never defined although many users incorrectly used the approximate mass content (given only as a guide) as the assigned value. A further disadvantage was that an endless supply could not be maintained. As an IS it was widely used by kit manufacturers to then assign values to their (secondary) calibrators. The impact of individual practice in assigning those values was reflected in numerous reports highlighting the bias between methods albeit often in the face of correlation coefficients that could exceed 0.99. Since 1994, recombinant materials (IS 88/624 and its successor in 2001, IS 98/574) have been available [14]. IS 98/574 contains 22 kDa GH of >95% purity with a defined specific activity of 3.0 IU/mg so that assay results could be reported in mass, molar or activity units. In practice, results continued to be expressed in mass or activity units, the detraction from molar units reflecting the molecular heterogeneity of GH in physiological fluids. A key factor favouring adoption of IS 98/547 is that it meets the requirement of EU Directive 98/79/EC (in vitro diagnostics) for values of commercial calibrators to be traceable to higher-order reference materials or methods, if available [15]. However, the reference material validity is also dependent on its integration in a traceability chain (further discussed below). Despite its ready availability from 2001 onwards, its adoption was variable and was often dictated by individual country and/or industry preference. One consequence has been to perpetuate not only inter-method differences but also to generate intra-method differences when manufacturers tailor a method to the calibration demands of their users. Thus, Meazza et al. reported how results could vary by 2 fold depending on whether the calibration of an otherwise identical kit had been based on IS 98/574 or IS 80/505 [16].
- The use of a plethora of conversion units between μg/L and mU/L. Pokrajac et al. assessed the impact of using 3 different conversion factors found in articles in one edition of a leading endocrine journal [17]. Adopting reported factors of 2.0, 2.6 and 3.0 for converting mU/L to μg/L across 14 GH methods participating in UKNEQAS, 11%, 55% and 100%, respectively, of submitted results would have been consistent with acromegaly if applied as a patient's GH nadir in an oral glucose tolerance test. Use of a variety of factors can also extend to conflict amongst consensus statements when, for example, a conversion factor of 2.6 is used between μg/L and mU/L in the Growth Hormone Research Society's guideline for the diagnosis of GH deficiency [18] whilst a factor of 3.0 is used in the NICE guidance [1].
- Variable epitope specificity of antibodies used in commercial kits. Given the heterogeneity of GH, a key determinant of a reported value is the avidity and affinity of the method's antibody for the different GH moieties. The inter-individual physiological variation in the concentration of these moieties is thought to be considerable [19,20]. The clearest indication that antibody specificity impacts on variability has come from external quality assessment data during the transition from polyclonal (variably specific) to monoclonal (monospecific) antibodies in kit methods. From 1994 to 1998 the UK NEQAS-reported variability between methods increased from 17% to 30%, an increase which overlapped with the switch to the use of monoclonal antibodies by the scheme's participants.
- The interference of GH binding protein (GHBP) which can show considerable inter- and intra-individual variability and which may complex up to 50% of GH to cause falsely low results depending on the antibody epitope specificity in the method being used [21,22]. Reports of the significance of GHBP interference stem largely from a time when polyclonal antibodies were in use in competitive immunoassays [23,24]. Data on GHBP interference in modern monoclonal antibody excess non-competitive immunoassays are more limited although interference leading to 40% reduction in values at low GH levels has been reported for one such method [25]. Decreases of 9.7%, 10.6% and 14.8% in GH levels have also been reported for 3 currently popular methods when spiked with 10 μg/L of GHBP and with interference reportedly above 30% in the presence of larger concentrations [26].
- The interference of pegvisomant, a GH competitor receptor antagonist licensed for the treatment of acromegaly which, depending on assay principle, has been reported to show positive interference (when both antibodies in a sandwich immunometric assay react), negative interference (when one of the antibodies in a sandwich immunometric assay reacts) [27] or, most recently, no interference in an assay using a monoclonal antibody that does not recognise pegvisomant [28]. The degree of interference has been investigated for some methods but deserves follow-up to quantitate the effects within and beyond pharmaceutical levels for all commercially available methods.

3. Taking one step at a time towards harmonisation

In 2006 an International Growth Hormone Collaborative recommended that all manufacturers should adopt NIBSC's IS 98/574 and that mass units should be adopted as the reporting unit (1 mg corresponding to 3 international units somatropin), an initiative supported by 3 leading endocrine journals (*Clinical Endocrinology*, *Growth Hormone & IGF Research*, and *European Journal of Endocrinology*) that advised submitting authors in 2007 that only data reported in μg/L would be considered for publication [29]. In a second step, a 2011 consensus included recommendations that manufacturers should specify the degree of interference by GHBP in their methods, the identity and traceability to IS 98/574 of their calibrants, and the assay cross-reactivity characteristics [30]. The impact of the consensus statements on practice and