An audit of the measurement and reporting of male testosterone levels in UK clinical biochemistry laboratories

Mark Livingston, Paul Downie, Geoff Hackett, Rachel Marrington, Adrian Heald, Sudarshan Ramachandran

Abstract

Introduction: A number of guidance documents have been published in recent years for the diagnosis and management of hypogonadism (HG). Laboratory practice has a major role in supporting guidelines with accurate and precise serum total testosterone (TT) methods and standardised pre- and post-analytical protocols. Our study investigated whether laboratory practice currently supports the management guidelines for HG.

Methods: An internet-based questionnaire survey of senior laboratory biochemists (UK/Republic of Ireland) was conducted (April-May 2018). Questions reflected sampling, laboratory practice, reference ranges and reporting of results. The results were analysed in conjunction with data obtained from the UK National External Quality Assurance Service (UK NEQAS) on testosterone assay performance.

Results: Analyses of 96 laboratory surveys returned the following: 74 laboratories stated that the optimal sampling time was communicated to users; 81 laboratories used immunoassays; 76 laboratories included reference ranges for adult men (31 had dual/multiple age-related intervals). Wide variability in lower/upper limits was evident in the common immunoassays; the majority of reference ranges were from manufacturers (50.0%) or historical (18.8%). Action limits based on TT levels were used by 64 laboratories, but 63 did not report a borderline range as suggested by the guidelines. Protocols for cascading tests based on TT were evident in 58 laboratories, with 50 laboratories offering estimated free testosterone; interpretative comments were provided by 67 laboratories, but no references were made to the management guidelines. Data from UK NEQAS demonstrated considerable variation in testosterone assay performance.

Conclusions: Our survey has highlighted inconsistencies that could lead to HG (and other conditions requiring measurement of TT) not being managed appropriately. The results from this survey and from UK NEQAS reinforce the requirement for action to be considered regarding the standardisation of testosterone assays and harmonisation of laboratory practice.
1 | INTRODUCTION

Male hypogonadism (HG) is defined as a combination of low serum total testosterone (TT) levels and associated symptoms, such as reduced bone mineral density, muscular strength and cognition, increased fatigue and sexual dysfunction. 1-3 The prevalence is estimated at 6-12%, 4,5 with this figure as high as 40% in men with type 2 diabetes (T2DM). 6,7 The diagnosis is clinically important, as longitudinal studies demonstrate both HG and erectile dysfunction (ED), an associated symptom, to be independently associated with increased mortality. 8,9 The European Male Ageing Study (2599 men aged 40–79 years, 7% with T2DM) showed that the combination of symptoms associated with HG and TT <8 nmol/L (<230.5 ng/dL) was significantly associated with increased total and cardiovascular disease (CVD)-related mortality. 8,9

Accurate determination of testosterone levels in men will have a direct impact on decisions about the initiation of testosterone treatment (TTh). Snyder et al. 10 in the Testosterone Trial total cohort, showed significant benefits in sexual function, mood, depression, quality-of-life, physical performance, vitality, anaemia and bone mineral density following TTh. Shores et al. 11 investigated the effect of TTh on mortality in men aged >40 years with a TT ≤8.7 nmol/L (≤250.7 ng/dL) and found that mortality was significantly reduced in men with T2DM, but interestingly not in their non-diabetes counterparts. This finding was confirmed by two longitudinal studies in men with low TT levels and T2DM by Muraleedaran et al. 12 using a cut-off of 10.4 nmol/L (299.7 ng/dL) and Hackett et al. 13-15 using TT and calculated free testosterone (cFT) cut-offs of 12 nmol/L (345.8 ng/dL) and 0.25 nmol/L (7.2 ng/dL), respectively.

Some concern continues to exist regarding the cardiovascular safety following TTh. Thus, while most studies demonstrate either benefit or no increase in cardiovascular events, a few have reported higher CVD events in men on TTh; specifically, the retrospective co-visit study reported in 2013 by Vigen et al. 16 and Finkle et al. 17 who showed significant benefits in sexual function, mood, depression, quality-of-life, physical performance, vitality, anaemia and bone mineral density following TTh. Shores et al. 11 investigated the effect of TTh on mortality in men aged >40 years with a TT ≤8.7 nmol/L (≤250.7 ng/dL) and found that mortality was significantly reduced in men with T2DM, but interestingly not in their non-diabetes counterparts. This finding was confirmed by two longitudinal studies in men with low TT levels and T2DM by Muraleedaran et al. 12 using a cut-off of 10.4 nmol/L (299.7 ng/dL) and Hackett et al. 13-15 using TT and calculated free testosterone (cFT) cut-offs of 12 nmol/L (345.8 ng/dL) and 0.25 nmol/L (7.2 ng/dL), respectively.

Some concern continues to exist regarding the cardiovascular safety following TTh. Thus, while most studies demonstrate either benefit or no increase in cardiovascular events, a few have reported higher CVD events in men on TTh; specifically, the retrospective co-visit study reported in 2013 by Vigen et al. 16 and Finkle et al. 17 who showed significant benefits in sexual function, mood, depression, quality-of-life, physical performance, vitality, anaemia and bone mineral density following TTh. Shores et al. 11 investigated the effect of TTh on mortality in men aged >40 years with a TT ≤8.7 nmol/L (≤250.7 ng/dL) and found that mortality was significantly reduced in men with T2DM, but interestingly not in their non-diabetes counterparts. This finding was confirmed by two longitudinal studies in men with low TT levels and T2DM by Muraleedaran et al. 12 using a cut-off of 10.4 nmol/L (299.7 ng/dL) and Hackett et al. 13-15 using TT and calculated free testosterone (cFT) cut-offs of 12 nmol/L (345.8 ng/dL) and 0.25 nmol/L (7.2 ng/dL), respectively.

Following a review of the existing evidence, guidelines for the diagnosis and treatment of men with HG were drawn up by the British Society for Sexual Medicine (BSSM) 18 stating that men with TT <8 nmol/L (<230.5 ng/dL) or cFT <0.180 nmol/L (<5.2 ng/dL) usually require TTh, while those with TT 8–12 nmol/L (230.5–345.8 ng/dL) may require TTh depending on the presence of symptoms associated with HG. Similar to the BSSM, 18 the Endocrine Society, 19 the American Association of Clinical Endocrinologists 20 and the American Urological Association 21 recommend screening testosterone levels in men with T2DM, metabolic syndrome and obesity, which will detect increased numbers as candidates for therapy.

Men with HG present to different specialties such as general practice, urology, endocrinology and diabetes. Thus, one option would be for the laboratory to assume a central role in the diagnosis and treatment guidance for HG. Treatment-based thresholds require the analytical methodology to be both accurate and precise. Further, it is important that laboratory pre-analytical protocols, assays and post-analytical advisory comments are standardised. It is important that clear interpretation is given and that relevant treatment guidelines are accessible and/or referenced in the laboratory report. Moreover, reference intervals based on population distributions are often misinterpreted as the “normal” range and can lead to confusion in the absence of any treatment guidelines. 22 A number of clinical guidelines have been published in recent years which address the matter of the diagnosis and treatment of hypogonadism in men. The fact that there are a number of guidance documents published (we counted eight published in recent years) may contribute to the level of uncertainty in this area. 23

The objective of this survey of laboratories in the United Kingdom (UK) and the Republic of Ireland (ROI) was to assess whether the laboratory measurement and reporting of TT levels were harmonised and supported the diagnostic and TTh guidelines for HG. To determine assay performance, we obtained quality assurance data on testosterone assays from the UK National External Quality Assurance Service (UK NEQAS).

2 | MATERIALS AND METHODS

2.1 | Survey questions, data gathering and analysis

This quantitative, questionnaire-based survey, conducted in the UK and ROI, was performed using an internet-based system (SurveyMonkey Inc, San Mateo, CA, USA). Questions pertaining to test requesting/sample collection, analytical methodology, reference ranges, action limits, laboratory reflex testing protocols and advisory comments reported with results were drawn up to assess current UK laboratory practice in clinical biochemistry. All the aspects covered were essential for optimal patient care in accordance with national and international guidelines on HG. The audit was supported by the UK NEQAS and the BSSM; it was registered as an audit with the local NHS Trust.
2.1.1 | Survey process

Questionnaires including the above nine questions were sent by email to clinical biochemists/chemical pathologists registered as such with the Royal College of Pathologists, Association of Clinical Biochemists or Institute of Biomedical Science working at a senior level (for clinical biochemists: Consultant, Principal or Senior level; for chemical pathologists: Consultant or Registrar level) across the UK and ROI in April 2018 and no geographic restrictions were applied when sending the surveys, thereby incurring no selection bias. A total of 551 email invitations were sent out. Reminders were sent until the survey was closed one month later (May 2018), this on the basis that 80% of responses are usually given within 7 days (https://www.surveymonkey.com/curiosity/time-to-respond/) and from observing that the number of responses (even with email reminders) had tailed off before 1 month.

Of all the emailed invitations, 106 (19.2%) participated; there were 20 duplicate responses from participants working in the same laboratory; hence, 10 of the questionnaires from the more senior healthcare professional within each of the laboratories were retained. This allowed 96 questionnaires from different laboratories available for analysis; of these, 88 (75.2%) were complete whilst the remaining 18 (15.4%) were partially filled. We split the type of hospital of those invited into university/teaching hospitals and district general hospitals and the same for those who responded to the survey; based on this, we performed a Chi-squared test for the comparison of two proportions. The proportion of either type of hospital for the invited participants compared with the proportion for the respondents was not found to be statistically different.

The 96 replies from the laboratories consisted of 94 from the UK (England: 77, Scotland: 10, Wales: 5, Northern Ireland: 2) with the remaining 2 from the ROI. Breakdown of the 77 English laboratory responses were as follows: North West England: 12; West Midlands: 12; South West England: 10; Greater London: 7; East Midlands: 6; East of England: 6; Yorkshire and Humber: 6; North East England: 4; and Greater Manchester: 2. Participants included: consultants: 63; principal grade scientists: 27; and senior grade scientists: 6. The respondents’ answers were exported into a Microsoft Excel (Microsoft Corporation; Redmond, WA, USA) spreadsheet for analysis. Results were pseudo-anonymised before analysis.

2.1.2 | Data from UK NEQAS (Birmingham quality)

Assay performance data were obtained from UK NEQAS for Steroid Hormones. This included results from the distributions in June 2018 (distribution 453) and January 2019 (distribution 460).

3 | RESULTS

Questions were analysed under the following themes: sample collection; analytical methodology; and communication of the result and clinical advice to the requesting clinician. Question 1 examines laboratory guidance on sample collection; Question 2 provides information on the methodology used to measure TT; and Questions 3–9 deal mainly with result reporting and clinical advice with the questions re-ordered to fit the following themes: (a) laboratory reference ranges, (b) borderline TT levels, (c) action limits based on TT levels and (d) interpretative comments added to the report.

3.1 | Sample collection

Question 1: Do you recommend that samples for serum testosterone from adult males are taken at a particular time of day?

If yes, what time, is this information provided to the user in advance or at the time of making the request?

96 laboratories responded to this question with 74 (77.1%) stating that optimal sampling time information was communicated to the users; of this group, 66 (89.2%) laboratories recommended phlebotomy before 11:00 hours, the remainder not specifying a time. Only 22 (29.7%) of the responding laboratories provided details of how the information was communicated: 8 via order communication systems, 10, via user handbook, 2 via both order communication systems and handbook, and 2 via user education, respectively.

3.2 | Analytical methods

Question 2: Which method do you use to measure male serum testosterone and are there any known interferences in your assay?

81 (84.4%) of the laboratories were using an immunoassay method to measure testosterone. The Roche immunoassay was the most popular (47) followed by Abbott (12), Siemens (10) and Beckman (8) immunoassay methods, with mass spectrometry utilised by two laboratories. A list of 70 interferences was provided by 44 (45.8%) of the laboratories with nandrolone (15), biotin (14), heterophilic antibodies (8), haemolysis/liquaemia/icterus (7) the most frequently cited.

3.3 | Result reporting and clinical advice

3.3.1 | Laboratory reference ranges

Question 3: What reference range for adult male serum testosterone do you quote? Please include details of any age-related ranges if you use them.

Question 4: Where does the serum testosterone reference range you use come from?

The lower and upper limits of the reference ranges obtained from responses to Question 3 were analysed separately, with greater focus on the lower limits (based on TT thresholds of <8 and 12 nmol/L in line with BSSM guidelines for TTh in men with adult
Of the 96 laboratories taking part in the survey, 76 quoted reference ranges for adult men in their survey report with 31 (32.3%) having dual/multiple age-related intervals (six of which quoted multiple ranges by decade from age 40 years onwards).

Figures 1A, B provide the lower and upper limits, respectively, of the TT reference ranges, arranged by the immunoassay methods (Roche, Abbott, Siemens and Beckman) in men aged >50 years (50 years was the commonest age demarcation when reporting age-related reference ranges). Wide variability in the lower limits was evident in all the common immunoassays: Figure 1A: Roche: 5.0–11.0 nmol/L (144.1–317.0 ng/dL), Abbott: 4.9–10.0 nmol/L (141.2–288.2 ng/dL), Siemens: 7.0–10.0 nmol/L (201.7–288.2 ng/dL), Beckman: 6.0–10.0 nmol/L (172.9–288.2 ng/dL) with many laboratories quoting figures below the 8 nmol/L (230.5 ng/dL) action limit threshold recommended by the BSSM guidelines (Roche: 25/41, Abbott: 2/10, Siemens: 1/9, Beckman: 4/5). A similar pattern was observed with the upper limits of the reference ranges, apart from the five laboratories using the Beckman immunoassay; Figure 1B: Roche: 25.0–40.0 nmol/L (720.5–1152.7 ng/dL), Abbott: 30.0–40.0 nmol/L (864.6–1152.7 ng/dL), Siemens: 25.0–40.0 nmol/L (720.5–1152.7 ng/dL), Beckman: 27.0–27.6 nmol/L (778.1–795.4 ng/dL). Interestingly, the TT reference ranges of the two laboratories using mass spectrometry varied: 6.3–26.5 nmol/L (181.6–763.7 ng/dL) and 7.1–31.1 nmol/L (204.6–896.3 ng/dL). The reference range across all methods included here ranged between 4.9 nmol/L (141.2 ng/dL) and 40 nmol/L (1152.7 ng/dL), which equates to more than an 8-fold difference between the two ends of the range.

Question 4 identifies the source(s) of the varying reference ranges seen above (Figure 2). The majority were obtained from the assay manufacturers (48) or were historical (18). Of the “In-house developed” group (n = 8), one laboratory commented that they were historical ranges adjusted for a new assay based on comparison data; a second laboratory stated that they were developed in conjunction with the literature; and a third informed that the range was developed in-house after Roche provided evidence that their male reference range was derived using blood samples collected between 08:00 and 13:00 hours.

Of the “Other” group (six laboratories), four laboratories further commented that the range was adapted from the literature. One laboratory stated that the figures used were those of the referral laboratory, whilst another stated that historical ranges from a previous referral laboratory were adjusted based on the comparison data when the assay was brought in-house.
3.4 | Borderline TT levels

**Question 6:** Do you have a range for serum testosterone where an adult male would be considered as having “borderline” hypogonadism? If yes, what is this range?

Of the 96 laboratories, 80 provided answers to this question with 63 (65%) not having a TT range for borderline HG as recommended by the guidelines, i.e., 8–12 nmol/L (230.5–345.8 ng/dL). The ranges quoted by the 17 (18%) remaining laboratories varied, with only two laboratories suggesting the 8–12 nmol/L borderline range for TT; the borderline ranges for TT levels varied between <4.6–14 nmol/L (129.7–403.5 ng/dL). It is worth highlighting that one laboratory considered a TT level <4.6 nmol/L (<132.6 ng/dL) in men aged >50 years to be borderline.

3.5 | Action limits based on TT levels

**Question 5:** Do you have action limits for adult male serum testosterone levels (clearly distinguishing normal/abnormal results) in addition to reference ranges? If yes, please provide.

**Question 7:** Do you have a protocol for further testing based on the serum testosterone level (eg, prolactin, Luteinising Hormone, Sex Hormone-Binding Globulin)? Other (please specify)/further details.

**Question 9:** Do your laboratory calculate a result for free testosterone? If yes, when is it calculated, which equation is used and what reference range/action limits do you quote?

Of the 96 laboratory responses analysed (16 laboratories did not answer question 5), 64 (66%) stated that action limits based on the TT level were used (flagging or telephoning of abnormal results; cascading of further tests (e.g., pituitary function tests, sex hormone-binding globulin (SHBG) and cFT or free androgen index (FAI) - either by protocol or clinical validator; suggesting referrals to endocrinologists; and referring to national/international guidelines), whilst 16 laboratories did not have any action limits. Specific action limits were quoted by 15 of the 64 laboratories having limits; these varied with TT cut-offs ranging from 4 to 12 nmol/L (230.5–345.8 ng/dL), with some laboratories citing multiple cut-off values.

Further tests were added at clinical validation (via laboratory protocol) by 58 (60.4%) of the laboratories based on the TT level. A further three laboratories stated that a computer rule/algorithm generated the add-on tests: two cascading cFT (albumin, SHBG) automatically on borderline TT results of 7–13 nmol/L (201.7–374.5 ng/dL) and 7–14 nmol/L (201.7–403.5 ng/dL); and another for TT <10 nmol/L (<288.2 ng/dL) added SHBG and a comment relating to cFT. Of the 58 laboratories adding tests during clinical validation, 19 (32.8%) provided more detailed actions that would be performed. Twenty-one (21.9%) laboratories stated they did not add further tests at clinical validation and 14 did not respond to the question. FAI or cFT were estimated by 50 laboratories depending on TT levels via various mathematical models; 30 laboratories offered neither of these, whilst 16 laboratories did not respond to the question. The most popular calculation was performed using the Vermeulen algorithm (34 laboratories). Wide variation in reference ranges/action limits was seen between the 28 laboratories which quoted relevant data regarding FAI and cFT levels (Table 1). Seventeen labs provided some information on when they utilised cFT; eight stated it would be performed if requested by the user or that it was added at clinical validation. One said they would add it if other tests proved inconclusive and a further three added it for borderline/low testosterone levels or in borderline clinical situations. Some provided TT limits when cFT was added: <10 nmol/L (288.2 ng/dL); 2; 7–14 nmol/L (201.7–403.5 ng/dL); 1; 8–11 nmol/L (230.5–317.0 ng/dL); 1; 8–12 nmol/L (230.5–345.8 ng/dL); 1; and 8–13 nmol/L (230.5–374.6 ng/dL); 1.

3.6 | Interpretative comments added to the report (Question 8)

**Question 8:** Do you add any interpretative comments on male serum testosterone results? If yes, please specify.

Interpretative comments were provided by 67 (69.8%) laboratories; 13 (13.5%) laboratories did not include any comments on reports whilst the remaining 16 (16.7%) did not provide a response to this question. Forty (59.7%) laboratories provided additional information on interpretative comments which are summarised in Table 2. A request for a repeat sample at 09:00 h was advised by 21 laboratories for TT levels below the reference range, whilst 19 laboratories reinforced the quoted reference range with a comment on the TT status. A further 13 comments pertained to cascading of extra endocrine tests, whilst 10 comments suggested a referral to an endocrinologist. Interestingly, nine comments were about diagnosing primary and secondary HG whilst no laboratory mentioned adult-onset HG. Moreover, no references were made to local or national patient management/treatment guidelines.
3.7 | Data from UK NEQAS (Birmingham Quality)

Table 3 shows the inter-assay performance for TT assays reported by Birmingham Quality for Steroid Hormones having distributed samples to 215 laboratories in December 2018 (distribution 459). Further more, data from the distribution 460 report (January 2019; shown in Figure 3) show the variation in assay performance for male testosterone. From this report, per cent method biases against the Tandem Mass Spectrometry (target) over various testosterone concentrations are given for the different assays/manufacturers. The variation is presented with positive and negative assay biases demonstrated, as well as changing bias over the concentration range.

4 | DISCUSSION

The aim of the current survey across the UK and ROI was to investigate whether laboratory practice supported the clinical guidance. We had 96 replies (94 laboratories in the UK and 2 from the ROI, although a proportion of questions were left blank by differing laboratories); all responses were included as originally planned in the data analysis with sufficient data available. All the questions led to responses showing considerable variation in practice. There has been the awareness of this issue for a number of years.

TABLE 1 | Reference ranges/action cut-off for free androgen index (FAI) and free testosterone (cFT) levels calculated by 28 participating laboratories

| Equation | Reference range(s) | Action cut-off | Labs reporting |
|----------|--------------------|----------------|----------------|
| FAI %    |                    | <24            | 1              |
|          |                    | <25            | 1              |
|          |                    | <30            | 1              |
|          |                    | <34            | 1              |
| cFT nmol/L (ng/dL) Vermeulen28 |                  | <0.200 (<5.76) | 3              |
|          |                    | <0.220 (<6.34) | 2              |
|          |                    | <0.225 (<6.49) | 2              |
|          |                    | <0.230 (<6.63) | 1              |
|          |                    | <0.245 (<7.07) | 1              |
|          |                    | <0.250 (<7.21) | 1              |

| 0.090–0.477a, 0.090–0.331b (2.60–13.76a), (2.60–9.55b) | 1 | |
| 0.150–0.570 (4.33–16.44) | 1 | |
| 0.160–0.470 (4.61–13.56) | 2 | |
| 0.174–0.729 (5.02–21.03) | 1 | |
| 0.198–0.619a, 0.163–0.473b (5.71–17.85a), (4.70–13.64b) | 3 | |
| 0.200–0.620a, 0.160–0.470b (5.77–17.88a), (4.61–13.56b) | 1 | |
| 0.200–0.620 (5.77–17.88) | 1 | |
| 0.215–0.760 (6.20–21.92) | 1 | |
| Age-related ranges used | 2 | |
| In-house range used (not given) | 1 | |

aAged <50 years; bAged 50+ years.

 TABLE 2 | Interpretative comments issued by the laboratories based on total testosterone (TT) levels. Of the 67 laboratories providing answers to question 8, 40 also provided additional comments grouped by theme as given in the table.

| Interpretative comments on TT levels grouped by theme | n |
|-----------------------------------------------------|---|
| Repeat on 9 am sample (some adding “if low on a pm sample”) | 21 |
| Low / borderline-low total testosterone comment | 20 |
| Comment when relevant tests were added on (various listed: FSH, LH, SHBG, cFT, PRL) | 13 |
| Primary and Secondary Hypogonadism comments | 13 |
| Endocrine referral comments | 10 |
| Low/estimated the cFT | 9 |
| Testicular failure, hypogonadotropic hypogonadism | 3 |
| Low/SHBG comment | 3 |
| Comment relating levels to function | 2 |
| Hypopituitary/pituitary tests added/suggested | 2 |
| Lower TT with age | 1 |
| Exclude other causes | 1 |

Abbreviations: cFT, calculated free testosterone; FSH, follicle-stimulating hormone; LH, luteinizing hormone; PRL, prolactin; SHBG, sex-hormone binding globulin.

Nevertheless, there continues to be a lack of standardisation of testosterone assay platforms. Now is surely the time to address this matter.
The diagnosis and management of HG are dependent on accurate and precise testosterone concentrations analysed by the laboratory. Male HG has a high prevalence (6-12%) and has been shown to be significantly associated with increased morbidity and mortality. Furthermore, TTh has been associated with improvement in symptoms\textsuperscript{10} and reduced mortality rates in men with T2DM\textsuperscript{11-14} where the prevalence is even higher than in the general population. Thus, it is essential that management is optimised and based on up-to-date evidence.

TT levels are known to vary during the day, but only 51 of 74 laboratories suggested a sampling time before 11:00 hours. Sampling at varying times could complicate the diagnosis of HG which is characterised by TT thresholds and symptoms and requirement for TTh; in view of the high prevalence of the condition, this could lead to large numbers of men being managed less effectively. In our view, harmonisation of sampling time is a straightforward issue to deal with, both by education and information given via the order communicating test requesting system.

Interestingly, most of the laboratories (84.4%) used immunoassays, with the Roche methodology being the most popular. A significant variation in reference intervals was evident; we had previously questioned whether reference intervals should be replaced by action limits in male HG.\textsuperscript{22} The move to action limits should be based on current evidence of benefit derived from the management of the pathology. In fields such as diabetes and dyslipidaemia, we do see report comments

| Method %bias vs Target for Testosterone [male] | Method %bias vs Target for Testosterone [male] |
|---------------------------------------------|---------------------------------------------|
| Method %bias vs Target for Testosterone [male]| Method %bias vs Target for Testosterone [male]|
| Method %bias vs Target for Testosterone [male]| Method %bias vs Target for Testosterone [male]|
| Method %bias vs Target for Testosterone [male]| Method %bias vs Target for Testosterone [male]|

**TABLE 3** Total testosterone assay performance from 215 laboratories (distribution 459 in December 2018) as reported by the UK National External Quality Assurance Service (UK NEQAS) for Steroid Hormones is summarised.

| Method %bias vs Target for Testosterone [male] | Method %bias vs Target for Testosterone [male] |
|---------------------------------------------|---------------------------------------------|
| Method %bias vs Target for Testosterone [male]| Method %bias vs Target for Testosterone [male]|
| Method %bias vs Target for Testosterone [male]| Method %bias vs Target for Testosterone [male]|
| Method %bias vs Target for Testosterone [male]| Method %bias vs Target for Testosterone [male]|

**FIGURE 3** UK National External Quality Assurance Service (UK NEQAS) for steroid hormones method report for male testosterone (nmol/L) (Distribution number: 460; 29/01/2019; permission given to reproduce)
including evidence-based action limits for glycated haemoglobin A1c (HbA1c) and low-density lipoprotein cholesterol, respectively.

Further more, the TT reference intervals quoted varied significantly between different methods and even amongst users of the same method (as seen in Figures 1A,B). As stated previously in men with HG, the varying lower limit can have significant consequences with regards to the initiation of TTh. The upper limit could see an adjustment to treatment such as lowering of TTh dose (gels) or increasing the time between administrations (injectable) or even discontinuation. Treatment ambiguity often arises when TT levels are borderline. Unfortunately, a borderline range is not acknowledged by a majority of laboratories. Moreover, even when acknowledged, the TT ranges quoted showed significant variation. Once again, action limits with reference to the guidelines could overcome this clinical issue.

Disparity was revealed in the actions taken by the laboratories following the analysis of TT along with variable triggering levels. We feel that cFT or bioavailable testosterone is important in view of the free hormone hypothesis, especially when TT levels are between 8 and 12 nmol/L (230.5–345.8 ng/dL).26–29,30 Studies have demonstrated associations between free/bioavailable testosterone concentration and clinical conditions; higher bioavailable testosterone levels have been associated with a lower risk of the metabolic syndrome and of cardiovascular mortality.27 Furthermore, cFT or bioavailable testosterone may be more appropriate measures than TT; for example, their decrease with male ageing is steeper and associated with features of HG whilst a decrease in TT can be masked by SHBG levels rising with age.31,32 There was not a consistent approach to the measurement of SHBG which recently has been shown to be associated with symptoms of HG33 and mortality.34,35 These observations perhaps being due to lower FT/bioavailable testosterone levels.

Our survey indicates that harmonisation of laboratory function is necessary. Assay standardisation is another area that has to be addressed at the same time.36 In 2017, Cao et al37 distributed two male samples and studied assay performance compared with target values using reference measurement procedures operated by the Centers for Disease Control and Prevention (CDC) reference laboratory. Considerable bias existed for all the distributed samples, 13.94 nmol/L (401.7 ng/dL): −24.8% to 8.6%, 17.27 nmol/L (497.7 ng/dL): −22.1% to 6.8%.

Similarly, considerable inter-assay variation performance for TT is reported by UKNEQAS as seen in Table 3. The method-specific means showed variation that could lead to major variation in treatment decisions in a condition with a high prevalence such as HG. Further data from the UK NEQAS distribution 460 report (Figure 3) also showed concerning assay performance for male testosterone. From this report, per cent method biases against the Tandem Mass Spectrometry (target) across various testosterone concentrations are given for the different assays/manufacturers. The variation is clear to be seen with: (a) both positive and negative assay biases demonstrated, as well as changing bias over the concentration range for other assays; and (b) random variation shown in all assays with specimen biases reported at > ±25%, with some assays regularly having specimen biases beyond this level. Once again, this demonstrates the need for standardisation, which continues to be a problem despite previous calls for substantial improvements in automated TT immunoassay technologies or switch to mass spectrometry methods.25

We would like to propose a solution to the problem that our survey has identified; this would involve a discussion between UKNEQAS, the Association of Clinical Biochemists and/or the International Federation of Clinical Biochemistry and the assay manufacturers. The existing gap between clinical associations and laboratory medicine associations can be addressed with joint ownership of testing recommendations. All manufacturers should be encouraged to use international standard reference preparations better to standardise the measurement of TT assays to reduce the analytical variation of results. It is hoped that the programmes such as that of the CDC (https://www.cdc.gov/labstandards/pdf/hs/HoSt_Brochure.pdf) and the availability of reference materials will help assay manufacturers and laboratories to standardise testosterone methods. This is essential if action thresholds are to be recommended by guidelines and protocols. An interim (less robust) solution would be to use assay-specific treatment thresholds relating to method performance or to establish harmonised reference ranges/action limits for TT that can be applied across laboratories by cross-calibrating assays to a reference method and standard calibrator(s) as described by Travison et al24 in a healthy non-obese population of European and American men. Harmonisation of pre- and post-analytical laboratory function should also be addressed, including reference ranges and advice provided to clinicians by the laboratory, which should be audited. It must be ensured that this is all based on current evidence.

We accept that a weakness of the survey is that a number of laboratories did not respond. However, the coverage of the survey in terms of laboratories contacted was comprehensive with representation from all regions of the UK and some replies from the Republic of Ireland. Moreover, the proportion of university/teaching hospitals and district general hospitals represented by the survey participants was not different from that found in the total number invited to participate. We cannot fully rule out bias from respondents in this survey, potentially from respondents who are more pro-active. However, this may suggest that the laboratories included are providing a more responsive service; if this is the case, the actual picture in the UK may be even worse than reflected in our findings.

All surveys, including this one, are limited by their narrow scope with the authors’ questions potentially exposing bias in their views; as such, respondents are required to answer the questions asked rather than what may be important. The current authors tried to ask questions requiring fact-based answers rather than opinions.

5 | CONCLUSION

Our survey has highlighted inconsistencies between laboratories that could lead to HG (and other conditions requiring the measurement of TT) being underdiagnosed and not managed optimally. The
results from this survey and from UK NEQAS reinforce the requirement for action to be considered regarding the standardisation of testosterone assays and harmonisation of laboratory function. If work can continue to improve testosterone assay standardisation and deal with the factors that introduce variation into the measurement of serum testosterone, this situation could change in a relatively short period of time.

CONFLICT OF INTEREST
Dr Heald, Professor Ramachandran and Professor Livingston have received educational grants to attend meetings and honoraria for serving as a speaker for Besins Healthcare Ltd. Professor Hackett has received honoraria for acting as a speaker for Bayer plc and research grants from Bayer plc. Professor Hackett has spoken at national and international meetings on testosterone and PDE5I inhibitor treatments in men and sat on the committee of the European Society for Sexual Medicine. Dr Downie and Dr Marrington have nothing to declare.

ORCID
Mark Livingston https://orcid.org/0000-0001-6878-0769
Geoff Hackett https://orcid.org/0000-0003-2073-3001
Adrian Heald https://orcid.org/0000-0002-9537-4050

REFERENCES

1. Khera M, Adaikan G, Buvat J, et al. Diagnosis and treatment of testosterone deficiency: recommendations from the Fourth International Consultation for Sexual Medicine (ICSM 2015). J Sex Med. 2016;13:1787-1804.
2. Livingston M, Kalansooriya A, Hartland AJ, Ramachandran S, Heald A. Serum testosterone levels in male hypogonadism: why and when to check—a review. Int J Clin Pract. 2017;71:e12995.
3. Dohle G, Arver S, Bettocchi C, Jones TH, Kliesch S, Punab M. EAU Guidelines on Male Hypogonadism. European Association of Urology. 2016. https://uroweb.org/wp-content/uploads/EAU-Guidelines-Male-Hypogonadism-2016-1.pdf. Accessed July 19, 2019.
4. Harman SM, Metter EJ, Tobin JD, Pearson J, Blackman MR. Longitudinal effects of aging on serum total and free testosterone levels in healthy men. Baltimore Longitudinal Study of Aging. J Clin Endocrinol Metab. 2001;86:724-731.
5. Araujo AB, Esche GR, Kupelian V, et al. Prevalence of symptomatic androgen deficiency in men. J Clin Endocrinol Metab. 2007;92:4241-4247.
6. Kapoor D, Aldred H, Clark S, Channer KS, Jones TH. Clinical and biochemical assessment of hypogonadism in men with type 2 diabetes: correlations with bioavailable testosterone and visceral adiposity. Diabetes Care. 2007;30:911-917.
7. Hackett G, Cole N, Deshpande A, Popple M, Kennedy D, Wilkinson P. Biochemical hypogonadism and type 2 diabetes in primary care. Br J Diabetes Vasc Dis. 2009;9:226-231.
8. Pye SR, Huhtaniemi IT, Finn JD, et al. Late-onset hypogonadism and mortality in aging men. J Clin Endocrinol Metab. 2014;99:1357-1366.
9. Dong JY, Zhang YH, Qin YQ. Erectile dysfunction and risk of cardiovascular disease meta-analysis of prospective cohort studies. J Am Coll Cardiol. 2011;58:1378-1385.
10. Snyder PJ, Bhasin S, Cunningham GR, et al. Effects of testosterone treatment in older men. N Engl J Med. 2016;374:611-624.
11. Shores MM, Smith NL, Forsberg CW, Anawalt BD, Matsumoto AM. Testosterone treatment and mortality in men with low testosterone levels. J Clin Endocrinol Metab. 2012;97:2050-2058.
12. Muraliedaran V, Marsh H, Kapoor D, Channer KS, Jones TH. Testosterone deficiency is associated with increased risk of mortality and testosterone replacement improves survival in men with type 2 diabetes. Eur J Endocrinol. 2013;169:725-733.
13. Hackett G, Heald AH, Sinclair A, Jones PW, Strange RC, Ramachandran S. Serum testosterone, testosterone replacement therapy and all-cause mortality in men with type 2 diabetes: retrospective consideration of the impact of PDE5 inhibitors and statins. Int J Clin Pract. 2016;70:244-253.
14. Hackett G, Jones PW, Strange RC, Ramachandran S, Statin, testosterone and phosphodiesterase 5-inhibitor treatments and age related mortality in diabetes. World J Diabetes. 2017;8:104-111.
15. Hackett GI, Cole N, Mulay A, Strange RC, Ramachandran S. Long-term Testosterone Therapy in Type 2 Diabetes is Associated with reduced Mortality without improvement in conventional cardiovascular risk factors. BJU Int. 2019;123:519-529.
16. Vigen R, O’Donnell CI, Barón AE, et al. Association of testosterone therapy with mortality, myocardial infarction, and stroke in men with low testosterone levels. JAMA. 2013;310:1829-1836.
17. Finkle WD, Greenland S, Ridgeway GK, et al. Increased risk of non-fatal myocardial infarction following testosterone therapy prescription in men. PLoS One. 2014;9:e85805.
18. Hackett G, Kirby M, Edwards D, et al. British Society for Sexual Medicine guidelines on adult testosterone deficiency, with statements for UK practice. J Sex Med. 2017;14:1504-1523.
19. Seftel AD, Kathrins M, Niederberger C. Critical Update of the 2010 Endocrine Society Clinical Practice Guidelines for Male Hypogonadism: a systematic analysis. Mayo Clin Proc. 2015;90:1104-1115.
20. Handelsman Y, Bloomgarden ZT, Grunberger G, et al. American association of clinical endocrinologists and american college of endocrinology - clinical practice guidelines for developing a diabetes mellitus comprehensive care plan - 2015. Endocr Pract. 2015;21:1-87.
21. Mulhall JP, Trost LW, Brannigan RE, et al. Evaluation and management of testosterone deficiency: AUA guideline. J Urol. 2018;200:423-432.
22. Ramachandran S, Konig CS, Hackett G, Livingston M, Strange RC. Managing clinical heterogeneity: an argument for benefit based action limits. J Med Diag Therapy. 2018;1:034701.
23. Salter CA, Mulhall JP. Guideline of guidelines: testosterone therapy for testosterone deficiency. BJU Int. 2019;124:722-729.
24. Travison TG, Vesper HW, Orwell E, et al. Harmonized reference ranges for circulating testosterone levels in men of four cohort studies in the United States and Europe. J Clin Endocrinol Metab. 2017;102:1161-1173.
25. Sikaris K, McLachlan RL, Kazlauskas R, de Kretser D, Holden CA, Handelsman DJ. Reproductive hormone reference intervals for healthy fertile young men: evaluation of automated platform assays. J Clin Endocrinol Metab. 2005;90:5928-5936.
26. Antonio L, Wu FCW, O’Neill TW, et al. Low free testosterone is associated with hypogonadal signs and symptoms in men with normal total testosterone. J Clin Endocrinol Metab. 2016;101:2647-2657.
27. Lin J-W, Lee J-K, Wu C-K, et al. Metabolic syndrome, testosterone, and cardiovascular mortality in men. J Sex Med. 2011;8:2350-2360.
28. Vermeulen A, Verdonck L, Kaufman JM. A critical evaluation of simple methods for the estimation of free testosterone in serum. J Clin Endocrinol Metab. 1999;84:3666-3672.
29. Nanjee MN, Wheeler MJ. Plasma free testosterone—is an index sufficient? Ann Clin Biochem. 1985;22:387-390.
30. Morgentaler A, Zitzmann M, Trush AM, et al. Fundamental concepts regarding testosterone deficiency and treatment: international expert consensus resolutions. Mayo Clin Proc. 2016;91:881-896.
31. Morley JE, Perry HM. Androgen treatment of male hypogonadism in older males. J Steroid Biochem Mol Biol. 2003;85:367-373.
32. Van den Beld AW, de Jong FH, Grobbee DE, Pols HA, Lamberts SW. Measures of bioavailable serum testosterone and estradiol and their relationships with muscle strength, bone density, and body composition in elderly men. J Clin Endocrinol Metab. 2000;85:3276-3282.

33. Rastrelli G, Corona G, Cipriani S, Mannucci E, Maggi M. Sex hormone-binding globulin is associated with androgen deficiency features independently of total testosterone. Clin Endocrinol. 2018;88:556-564.

34. Tint AN, Hoermann R, Wong H, et al. Association of sex hormone-binding globulin and free testosterone with mortality in men with type 2 diabetes mellitus. Eur J Endocrinol. 2016;174:59-68.

35. Ramachandran S, Strange RC, Fryer AA, Saad F, Hackett G. The association of sex hormone binding globulin with mortality is mediated by age and testosterone in men with type 2 diabetes. Andrology. 2018;6:846-853.

36. Ramachandran S, Livingston M, Hackett G, Strange RC. Should laboratory assay quality be described in study publications and clinical guidelines? A focus on testosterone assay performance. Edorium J Biochem. 2019;3:100007B01SR2019.

37. Cao Z, Botelho JC, Rej R, Vesper H. Accuracy-based proficiency testing for testosterone measurements with immunoassays and liquid chromatography-mass spectrometry. Clin Chim Acta. 2017;469:31-36.

How to cite this article: Livingston M, Downie P, Hackett G, Marrington R, Heald A, Ramachandran S. An audit of the measurement and reporting of male testosterone levels in UK clinical biochemistry laboratories. Int J Clin Pract. 2020;74:e13607. https://doi.org/10.1111/ijcp.13607