Giacomo Montagna, Samuela Balestra, Federica D'Aurizio, Francesco Romanelli, Cinzia Benagli, Renato Tozzoli, Lorenz Risch, Luca Giovanella* and Mauro Imperiali

Establishing normal values of total testosterone in adult healthy men by the use of four immunometric methods and liquid chromatography-mass spectrometry

https://doi.org/10.1515/cclm-2017-1201
Received December 22, 2017; accepted April 23, 2018

Abstract

Background: The total testosterone (T) cutoffs clinically adopted to define late-onset hypogonadism (LOH) do not consider the differences that exist between different analytical platforms, nor do they consider the body mass index (BMI) or age of the patient. We aimed at providing method, age and BMI-specific normal values for total T in European healthy men.

Methods: A total of 351 eugonadal healthy men were recruited, and total T was measured with four automated immunometric assays (IMAs): ARCHITECT i1000SR (Abbott), UniCel DxI800 (Beckman Coulter), Cobas e601 (Roche), IMMULITE 2000 (Siemens) and liquid chromatography-tandem mass spectrometry (LC-MS/MS). Reference ranges (RRs) were calculated for each method.

Results: Passing and Bablok regression analysis and Bland-Altman plot showed an acceptable agreement between Abbott and LC-MS/MS, but a poor one between LC-MS/MS and the other IMAs. Age-specific T concentrations in non-obese (BMI <29.9 kg/m^2) men were greater than in all men. The total T normal range, in non-obese men aged 18–39 years, measured with LC-MS/MS was 9.038–41.310 nmol/L. RRs calculated with LC-MS/MS statistically differed from the ones calculated with all individual IMAs, except Abbott and among all IMAs. Statistically significant differences for both upper and lower reference limits between our RRs and the ones provided by the manufacturers were also noticed.

Conclusions: We calculated normal ranges in a non-obese cohort of European men, aged 18–39 years, with four commercially available IMAs and LC-MS/MS and found statistically significant differences according to the analytical method used. Method-specific reference values can increase the accuracy of LOH diagnosis and should be standardly used.

Keywords: age; body mass index; late-onset hypogonadism (LOH); reference ranges; testosterone.

Introduction

The understanding of the relevance of steroid hormones in health and disease has undergone substantial advancement [1]. Androgen deficiency contributes to many adverse conditions, including diabetes and prediabetic conditions, metabolic syndrome, reduced bone and muscle mass, impaired sexual function and fertility, poor sleep quality and decreased quality of life [2–9]. Over a 20-year period, testosterone (T) deficiency is involved in the development of approximately 1.3 million new cases of cardiovascular diseases, 1.1 million new cases of diabetes mellitus and over 600,000 osteoporosis-related fractures [10]. Longitudinal studies have documented a decline of T with aging starting from the early 40s [8, 11–14]. The average decline in serum T levels varies depending on the studies, ranging from 0.4% to 2% per year [8, 12, 15]. The clinical importance of this finding is still not completely understood [16, 17]. Differently from other endocrine glands, in which the classification and diagnosis of hypofunction are based on

***Corresponding author: Prof. Dr. med. Luca Giovanella**, PhD, Clinic for Nuclear Medicine and PET/CT Centre, Oncology Institute of Southern Switzerland, Bellinzona, Switzerland; Integrated Thyroid Centre, Ente Ospedaliero Cantonale, Bellinzona, Switzerland; and Clinic of Nuclear Medicine, University Hospital Zürich, Zürich, Switzerland, Phone: +41 91 811 86 72, Fax: +41 91 811 82 50, E-mail: luca.giovanella@eoc.ch
Giacomo Montagna: Department of Obstetrics and Gynecology, University Hospital Basel, Basel, Switzerland. http://orcid.org/0000-0003-3511-369X
Samuela Balestra, Cinzia Benagli and Mauro Imperiali: Department of Laboratory Medicine, Ente Ospedaliero Cantonale, Lugano, Switzerland
Federica D'Aurizio: Clinical Pathology Institute, University Hospital Udine, Udine, Italy
Francesco Romanelli: Department of Experimental Medicine, Sapienza University of Rome, Rome, Italy
Renato Tozzoli: Clinical Pathology Service, Regional Hospital, Pordenone, Italy
Lorenz Risch: Labormedizinisches Zentrum Dr. Risch, Liebefeld, Switzerland; Institute of Clinical Chemistry, University of Bern, Bern, Switzerland; and Private University of the Principality of Liechtenstein, Triesen, Liechtenstein

well-defined biochemical criteria, the diagnosis of late-onset hypogonadism (LOH) requires both clinical and biochemical criteria [18]. The etiology of LOH is complex and multifactorial; age and risk factors such as obesity, systemic disorders, drugs, environmental factors and lifestyle play a major role [8, 19–21]. Likely owing to many factors, including increased recognition and availability of therapies, the number of patients diagnosed and treated for LOH during the past two decades has increased significantly, with an exponential increase in prescribed T in many countries [22–26]. Because of the weak association of sexual symptoms and T levels in aging men, and the absence of pathognomonic signs for LOH, multiple symptoms combined with low T levels are required for the diagnosis of LOH [27]. However, the laboratory diagnosis of T deficiency is a challenge. A number of methods are available for total T testing, which include radioimmunoassay, immunometric assay (IMA) and extraction by liquid chromatography-mass spectrometry (LC-MS). Important biological and technical limitations among all these methods exist [28, 29]. A recent survey has shown that the most common method used among US laboratories is by IMA via high-throughput analyzers [30]. Ideally, valid T assay methods must demonstrate internal and external validity. Internal validity is gauged by the method providing essentially identical numerical values, that is, demonstrating high precision and minimal bias when calibrated against an independent reference method such as gas chromatography/mass spectrometry. External validity should provide consistent reference intervals when calibrated against a valid reference population [31], and therefore facilities are encouraged to validate their manufacturer's suggested reference range (RR) using a sample of local patients [28, 32]. Despite the efforts of the Endocrine Society and the Centers for Disease Control and Prevention in promoting T assay standardization [33], recent studies have shown that external validity is still often lacking; only a minority of laboratories in fact perform validation tests to establish their own age stratified RRs, and among those who do perform validation tests, sample sizes are limited, and there is little knowledge regarding the subjects' state of health [30]. In addition, there is no universally accepted lower limit for normal total T [34]. Currently, the T value below which substitution therapy is recommended ranges from 8 nmol/L (2.3 ng/mL) to 13.9 nmol/L (4.0 ng/mL) [27, 35, 36]. Wu et al. [27] suggested the presence of at least three sexual symptoms (decreased frequency of morning erections, decreased frequency of sexual thoughts and erectile dysfunction) preferably associated with the lower T threshold (<8.0 nmol/L, 2.3 ng/mL) as a diagnostic criterion to increase the probability of correctly diagnosing LOH. The authors also suggested that the estimation of free T, based on total T, sex hormone-binding globulin and serum albumin, may be useful in patients with borderline total T level ranging from 8 to 11 nmol/L (2.3–3.2 ng/mL) [27]. Nonetheless, considering the concerns about the variability and lack of accuracy among platform-based IMAs, the proposed T thresholds should not be translated from the literature into clinical practice without being adjusted to the immunoassay method used in local laboratories [37, 38]. The present study was undertaken to provide RRs for T in non-obese European healthy men using four largely available IMA platforms and LC-MS/MS.

Materials and methods

Subjects

We recruited 351 volunteer men aged >18 years with normal general and sexual health. All subjects underwent a standard medical history and physical examination conducted by a dedicated endocrinologist, which included recordings of prescribed and non-prescribed medications and dietary supplements, body weight, height, heart rate and blood pressure. Sexual health was assessed with a semistructured questionnaire based on a previously validated questionnaire and current literature [39, 40]. Exclusion criteria were history of testicular or androgenic pathology (gynecomastia, hypospadias, Kallmann or Klinelfelter syndrome, prostate cancer), male infertility, undertaking extreme exercise and training (i.e. agonistic athletes), chronic use of drugs (steroids, opiates, antihypertensive medications, antidepressants, antipsychotics and recreational drugs), use of any hormone treatment within the last year or any other serious medical illness; nightshift workers were excluded if the last nightshift was within 14 days.

Ethical approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of the Institutional and/or National Research Committee (Approval: CE 2819) and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. Written informed consent was obtained from all individual participants included in the study.

Blood collection and preparation

Blood samples were taken between 8:00 and 10:00 a.m. after a light breakfast to minimize the effects of circadian variation, fasting and dehydration. A total of 10 mL blood was collected by a single

Table 1: Total testosterone distribution (nmol/L) in a cohort of European healthy men measured by five analytical methods.

Total T, nmol/L	n	Min	1st Quartile	Median	95% CI	3rd Quartile	Max	IQR
Beckman	349	3.91	10.53	12.92	12.30–13.80	16.27	30.93	5.75
Immulite	347	4.06	11.12	13.90	12.30–14.60	17.28	37.40	6.17
Roche	345	3.7	12.50	15.70	15.10–18.80	20.13	40.20	7.63
Abbott	345	4.34	14.02	18.04	17.29–18.60	22.87	47.55	8.83
LC-MS	345	4.48	14.16	17.91	17.14–18.84	22.36	45.50	8.20

venipuncture into plain Vacutainer® (Becton Dickinson) tubes (without anticoagulants or separation gel). After standing at room temperature for 30 min, the blood sample was centrifuged for 15 min at 1800 *g*. The serum was carefully aliquoted into 2-mL screw-top microtubes for subsequent storage. At least six aliquots were collected from each study participant. Each microtube was labeled with a coded volunteer identification label and stored at −80 °C.

Testosterone measurement

The panel of 351 sera was used to measure total T concentrations over four widely available commercial automated analyzer systems; ARCHITECT i1000SR (Abbott), UniCel DxI800 (Beckman Coulter), Cobas e601 (Roche), IMMULITE 2000 (Siemens) and 1290 HPLC coupled to 6490 triple quadrupole equipped with standard ESI (Agilent Technologies). LC-MS/MS assay was performed in the Clinical Biochemistry and Pharmacology Laboratory, Department of Laboratory Medicine, Bellinzona, Switzerland. Analytical performance characteristics of the different platforms are reported in Table 1 in the Supplemental material. The analytes were separated and quantified in serum using Agilent Technologies 6490 triple quadrupole LC/MS system equipped with Jet-Stream ESI source. Samples were prepared accordingly to MassChrom® Steroids in Serum/Plasma LC-MS/MS kit instruction manual (Chromsystems GmbH). The kit uses stable isotopically labeled internal standards for each analyte to ensure reproducible and reliable quantification of the samples. The method is validated according to CLSI EP9-A3 [41] guidelines using the calibration Standard NIST 971 as reference material for total T. Intra-assay variability and interassay variability were checked for the three levels of quality controls obtaining values of 6.94% and 6.86% for 0.201 µg/L, 2.63% and 6.37% for 1.5 µg/L, and 5.27% and 5.05% for 7.85 µg/L.

Statistical analysis

Statistical analysis was performed using Analyse-it version 2.20 for Microsoft Excel (http://www.analyse-it.com). The analysis consisted of descriptive statistics to assess characteristics, where the data were presented as mean and standard deviation (SD) for continuous variables. Shapiro-Wilk test was used to assess the normality of the T distribution for the results of each assay. Because serum total T values were not normally distributed, and therefore better represented with the median, the statistical analyses were performed using non-parametric tests. Mann-Whitney U and Kruskal-Wallis tests were used for group comparison. Passing-Bablok regression and method were used to evaluate the relationship between different methods. Correlations were explored using the adjusted R^2-value. Agreement between the different assays was explored using the Bland-Altman plot. We used simple or multiple linear regression models to explore the association between the dependent variable T and the independent variables age, body mass index (BMI) and smoking. Age and BMI were analyzed as continuous variable and smoking as dichotomous variable. All tests were two-tailed, and a value of $p < 0.05$ was considered statistically significant. Finally, RRs were calculated based on guidance from the CLSI document EP28-A3C [42]. The lower (LRL) and the upper reference limits (URL) were established at 2.5th and 97.5th percentile, respectively, according to non-parametric method. Proportions' comparison and comparison of URL and LRL among the different analytical methods was done with χ^2-test.

Results

Subjects

Three hundred and fifty-one healthy men were recruited. The mean (SD) age of the all volunteers was 38.7 (11.8) years. One hundred and ninety-one were <40 years old (mean age 29.7 [5.5] years), the other 160 were >40 years old (mean age 49.5 [7.3] years). The mean (SD) BMI was 25 (3.7) kg/m^2, 332 had a BMI between 18.5 and 29.9 kg/m^2 (mean 24 [3] kg/m^2); the others (19) had a BMI ≥30 kg/m^2 (mean 33 [3] kg/m^2). Ninety-three volunteers (26%) were smokers. All of them had normal sexual history and no clinical signs of hypogonadism.

Measurement of total testosterone

Table 1 shows the distribution of total T measured by five methods. Analyses were performed on 349 (Beckman), 347 (Immulite) and 345 subjects (Roche, Abbott, LC-MS/MS), respectively, due to insufficient volume in few cases. Median total T values obtained with LC-MS/MS and Abbott were 17.91 and 18.04 nmol/L, respectively, whereas lower levels were obtained on Roche, Immulite and Beckman (15.70, 13.90 and 12.92 nmol/L) platforms, respectively). Figure 1 shows the results obtained with

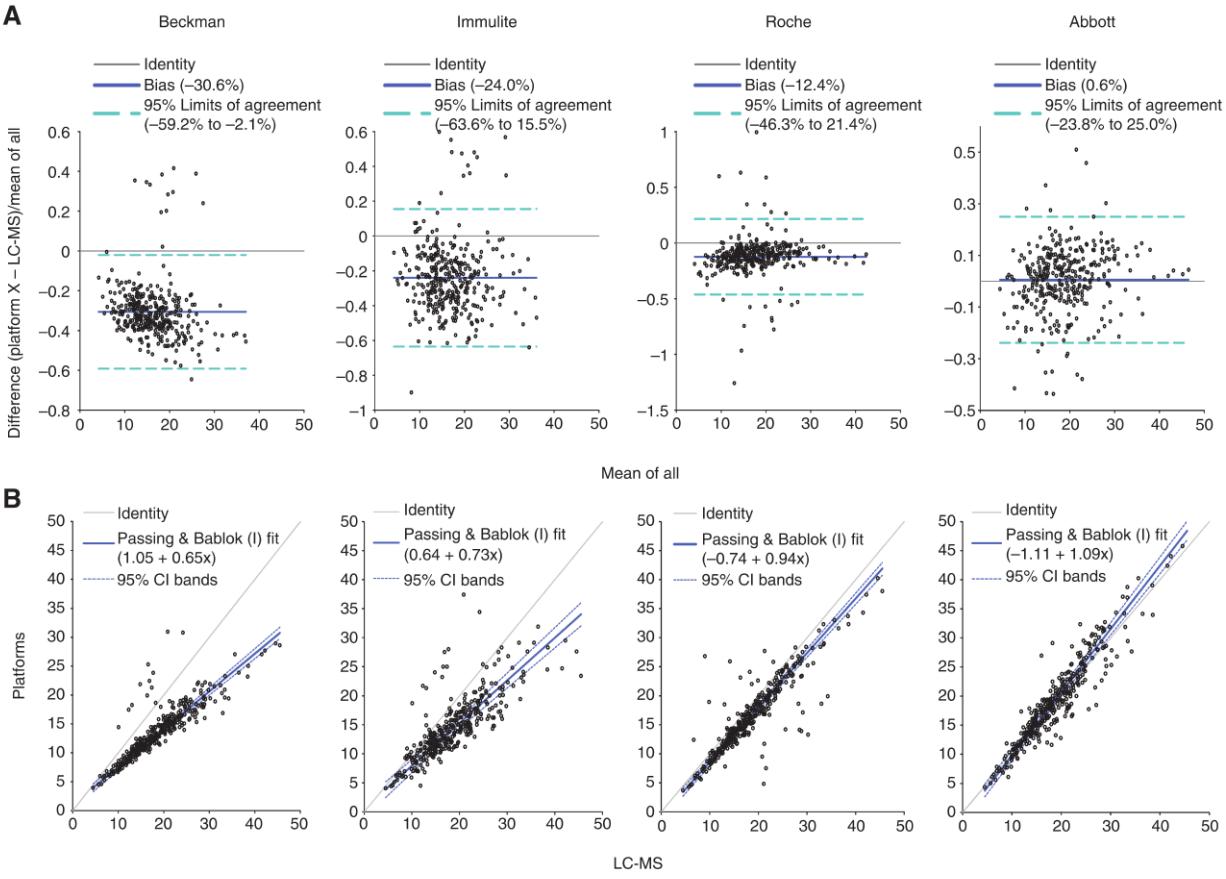

Figure 1: Interassay agreement and differences in serum testosterone levels between four commercial IMAs and LC-MS/MS. Bland-Altman plots (A): Y-axis depicts the % difference between values of the two measurements (IMA – LC-MS/MS/mean of all). Bias and limits of agreement are depicted as continuous and dashed line, respectively. The bias was –30.6% (95% CI –32.2 to –29.1) for Beckman, –24% (95% CI –26.2 to –21.9) for Immulite, –12.4% (95% CI –14.2 to –10.6) for Roche and 0.6% (95% CI –0.7 to 1.9) for Abbott. Passing-Bablok plots (B) for testosterone with four different IMAs and LC-MS/MS: Passing-Bablok equation and its confidence interval are depicted as continuous and dashed line, respectively. Constant and proportional bias for Beckman were 1.05 (95% CI 0.67–1.41) and 0.65 (95% CI 0.63–0.67); for Immulite 0.64 (95% CI –0.15 to 1.35) and 0.73 (95% CI 0.69–0.78); for Roche –0.74 (95% CI –1.17 to –0.34) and 0.94 (95% CI 0.91–0.96); for Abbott –1.11 (95% CI –1.67 to –0.47) and 1.09 (95% CI 1.05–1.13).

Table 2: Participants with a total T level <8 nmol/L, between 8–11 nmol/L and <11 nmol/L according to age groups.

	Beckman	Immulite	Roche	Abbott	LC-MS
Total samples	**349**	**348**	**347**	**345**	**345**
Samples <40 years	191	191	191	190	190
Samples with total T level <8 nmol/L (all)	**31 (18%)**	**23 (13%)**	**19 (11%)**	**15 (9%)**	**10 (6%)**
Samples with total T level T < 8 nmol/L (<40)	12 (6%)	11 (6%)	9 (5%)	8 (4%)	5 (3%)
Samples with total T level 8–11 nmol/L (all)	**73 (21%)**	**61 (18%)**	**36 (10%)**	**20 (6%)**	**27 (8%)**
Samples with total T level 8–11 nmol/L (<40)	40 (21%)	30 (16%)	16 (8%)	12 (6%)	17 (9%)
Samples with total T level < 11 nmol/L (all)	**105 (30%)**	**84 (24%)**	**55 (16%)**	**35 (10%)**	**37 (11%)**
Samples with total T level < 11 nmol/L (<40)	52 (27%)	41 (21%)	25 (13%)	20 (11%)	15 (8%)

Values measured in the -all age- group are displayed in bold.

immunometric methods compared with LC-MS/MS. The correlation between LC-MS/MS and the immunoassays was fairly good with all of them, R^2 being 0.92, 0.81, 0.78, 0.89 for Beckman, Immulite, Roche and Abbott, respectively.

Relationship between total testosterone age, BMI and smoking

A multivariate logistic regression analyses was performed to identify independent determinant of total T levels. With

all assays as well as with LC-MS, BMI was the only independent determinant ($p < 0.0001$).

Participants with low T values according to conventional clinical thresholds

Table 2 shows the number of participants with a total T level <8 nmol/L, between 8–11 nmol/L and <11 nmol/L, according to age groups (all ages and <40 years). When we compared the rate of participants with T < 8 nmol/L (Table 3) between each analytical platform, we found statistically significant differences between Beckman and Abbott and between Beckman and LC-MS/MS. Due to small sample size, we did not perform separate analyses according to age. Subsequently, we compared the rate of participants with T < 11 nmol/L (Table 4) between each analytical platform in both age groups and found statistical difference between most of the platforms. χ^2-Test showed no difference among the rate of T level <11 nmol/L among the all ages group vs. the <40 group.

Reference ranges calculation and comparison with manufacturer provided ones

To establish RRs, we first selected men 18–39 years, who were non-obese (BMI 18.5–29.9 kg/m^2); 180 volunteers were included in this analysis. Subsequently, we computed RRs for individuals of all ages with BMI 18.5–29.9 kg/m^2, 303 volunteers were included. Finally, RRs for all 351 subjects were calculated, regardless of obesity status and age (Table 5). We compared URLs and LRLs among all analytical methods in the entire participants' collective (all ages, all BMIs). URLs significantly differed among all methods except between LC-MS and Abbott, whereas LRLs differed between Abbott and all other IMAs and between LC-MS and all IMAS, except Abbott (Table 6). Among young (<40 years old) non-obese participants, we found statistically significant differences in the URLs among all IMAs and between LC-MS and all IMAs, except Abbott. LRLs statistically differed only between Abbott and Immulite and between LC-MS and all IMAs, except Abbott (Table 7).

We compared our RRs (LRL and URL) with the one provided by the manufacturers (Table 8). Unfortunately, not all manufacturers provide information about the age range of the population on which they calculate RRs, and none provide information about the BMI. When known (e.g. Abbott, Immulite and Roche), we compared RRs among the same age group. If not known (e.g. Beckman, LC-MS), comparison was carried out with the RRs calculated on all study participants (all age, all BMI). Chi-squared test showed statistically significant differences

Table 3: Comparison between the rate of all participants with T < 8 nmol/L among each analytical platform.

All ages	Beckman	Immulite	Roche	Abbott	LC-MS
Beckman		0.3711	0.1442	**0.0385**	**0.0032**
Immulite	0.3711		0.5667	0.5288	0.1185
Roche	0.1442	0.5667		0.5288	0.1185
Abbott	**0.0385**	0.5288	0.5288		0.3461
LC-MS	**0.0032**	0.1185	0.1185	0.3461	

Statistically significant p-values are displayed in bold.

Table 4: Comparison between the rate of participants with T < 11 nmol/L among each analytical platform in both age groups.

	Beckman	Immulite	Roche	Abbott	LC-MS
All ages					
Beckman		0.0774	**<0.0001**	**<0.0001**	**<0.0001**
Immulite	0.0774		**0.0063**	**<0.0001**	**<0.0001**
Roche	**<0.0001**	**0.0063**		**0.0257**	**0.0471**
Abbott	**<0.0001**	**<0.0001**	**0.0257**		0.8033
LC-MS	**<0.0001**	**<0.0001**	**0.0471**	0.8033	
< 40 years					
Beckman		0.1897	**0.0006**	**<0.0001**	**<0.0001**
Immulite	0.1897		**0.0304**	**0.0036**	**0.0002**
Roche	**0.0006**	**0.0304**		0.4384	0.0982
Abbott	**<0.0001**	**0.0036**	0.4384		0.3751
LC-MS	**<0.0001**	**0.0002**	0.0982	0.3751	

Statistically significant p-values are displayed in bold.

Table 5: Reference ranges of total testosterone (nmol/L) for each analytical assay, according to age and BMI (kg/m^2).

Reference ranges, nmol/L	BMI 18.5–29.9 and age <40 years	BMI 18.5–29.9, all ages	All BMI, all ages
Beckman	6.66–27.491	6.589–27.020	6.06–25.085
Immulite	6.847–29.100	6.372–28.340	6.022–28.083
Roche	7.36–36.38	7.28–33.73	6.01–31.81
Abbott	8.007–43.070	7.135–40.493	7.126–36.959
LC-MS	9.038–41.310	8.401–38.450	7.495–34.757

Table 6: Comparison of RRs in the entire participants' collective (all ages and all BMIs) among different IMAs and between IMAs and LC-MS/MS.

	Beckman	Immulite	Roche	Abbott	LC-MS
URL					
Beckman		**0.0039**	**<0.0001**	**<0.0001**	**<0.0001**
Immulite	**0.0039**		**0.0024**	**<0.0001**	**<0.0001**
Roche	**<0.0001**	**0.0024**		**0.0008**	**0.0040**
Abbott	**<0.0001**	**<0.0001**	**0.0008**		0.4335
LC-MS	**<0.0001**	**<0.0001**	**0.0040**	0.4335	
LRL					
Beckman		0.8107	0.9907	**0.0009**	**<0.0001**
Immulite	0.8107		0.9954	**0.0024**	**0.0001**
Roche	0.9907	0.9954		**0.0120**	**0.0072**
Abbott	**0.0009**	**0.0024**	**0.0120**		0.4613
LC-MS	**<0.0001**	**0.0001**	**0.0072**	0.4613	

Statistically significant p-values are displayed in bold.

Table 7: Comparison of RRs in the non-obese group of participants <40 years among different IMAs and between IMAs and LC-MS/MS.

	Beckman	Immulite	Roche	Abbott	LC-MS
URL					
Beckman		**0.0181**	**<0.0001**	**<0.0001**	**<0.0001**
Immulite	**0.0181**		**0.0024**	**<0.0001**	**<0.0001**
Roche	**<0.0001**	**0.0024**		**0.0043**	**0.0177**
Abbott	**<0.0001**	**<0.0001**	**0.0043**		0.4745
LC-MS	**<0.0001**	**<0.0001**	**0.0177**	0.4745	
LRL					
Beckman		0.4745	0.3112	0.0554	**0.0020**
Immulite	0.4745		0.4745	**0.0328**	**0.0008**
Roche	0.3112	0.4745		0.1549	**0.0185**
Abbott	0.0554	**0.0328**	0.1549		0.1526
LC-MS	**0.0020**	**0.0008**	**0.0185**	0.1526	

Statistically significant p-values are displayed in bold.

between our LRLs and the ones provided by Abbott ($p = 0.0138$) and Roche ($p = 0.019$), whereas our URLs only differed with the one provided by Immulite, for patients aged 20–49 years ($p = 0.0356$).

Discussion

Agreement among IMAs, between IMAs and LC-MS/MS and reference ranges

In our present study, we measured total T levels in a carefully selected cohort of 351 healthy men in central Europe, using four largely available commercial IMAs against LC-MS/MS as the standard method.

Our finding of high disagreement between measurements obtained with the four IMAs and LC-MS is consistent with results reported in previous studies [32, 43]. The patterns of individual sample bias differed among assays, with higher T concentrations measured with Abbott and lower concentrations measured with Roche, Immulite and Beckman, respectively. Intermethod variability was confirmed by the fact that outliers differed according to the analytical method.

Likewise, when we looked at independent determinants of total T, we found that, similar to previous reports [8, 37, 44], the magnitude of the age trend was modest (p non-significant with all methods) whereas the effect of increasing BMI was more substantial (total T was higher among non-obese men than in all patients, $p < 0.001$ with all methods).

We compared the URLs calculated with the different analytical methods among the entire study participants' collective, and we found statistically significant differences between all methods except between LC-MS/MS and Abbott. LRLs statistically differed between Abbott and all other IMAs and between LC-MS/MS and all IMAs except Abbott. Similarly, when we compared URLs among young (<40 years old) non-obese men, we found statistically significant differences among all IMAs and between LC-MS/MS and all IMAs except Abbott. Regarding LRLs, a statistically significant difference was noticed between Abbott and Immulite and between LC-MS/MS and all IMAs except for Abbott.

Table 8: Comparison between our RRs with the one provided by the manufacturers.

Manufacturer, age range, number of participants	Manufacturer's RRs	Our RRs	χ^2 p-value for URLs	χ^2 p-value for LRLs
Abbott, 21–49, n = 269	1.63–34	7.515–40.41	0.0798	**0.0138**
Abbott, >50, n = 70	4.41–35.38	6.25–39.63	0.5764	0.1543
Immulite, 20–49, n = 269	5.55–25.19	6.01–28.68	**0.0356**	0.5233
Immulite, >50, n = 70	4.47–26.61	5.69–30.35	0.5594	0.3156
LC-MS	8.675–34.7	7.49–34.75	1.0000	0.2678
Beckman, n = 349	6.07–27.1	6.06–25.08	1.0000	1.0000
Roche, 20–49, n = 200	8.65–29	6.01–32.63	0.3998	**0.0192**
Roche, >50, n = 67	6.68–25.7	5.33–32.90	0.3172	0.3100

Although, due to a lack of standardization with respect to choosing the subject population used to establish RRs, laboratories are strongly encouraged to establish their own RRs [28, 29], rather than using manufacturer-supplied ranges [37, 45], a recent study has shown that only a minority of laboratories (<10%) establish their own RRs [30]. When we compared our total T RRs with the ones provided by the manufacturers, we found statistically significant differences between our LRL and the ones provided by Abbott ($p = 0.0138$) and Roche ($p = 0.019$), and our URL differed from Immulite among patients aged 20–49 years ($p = 0.0356$).

To address this problem, Travison et al. [37] recently attempted to harmonize total T RRs, using samples from four European and American epidemiological studies [the Framingham Heart study (FHS), the European Male Aging Study, the Osteoporotic Fractures in Men Study and the Sibling Study of Osteoporosis (SIBLOS)]. They found that the distribution of harmonized total T levels in young men, aged 19–39 years, with a normal BMI was similar among Europeans and Americans, and they concluded that the harmonized normal range in non-obese men was 264–916 ng/dL (9.2–31.8 nmol/L).

In our cohort, the normal range in non-obese men ($BMI \leq 29.9$ kg/m^2), aged 18–39 years, measured with LC-MS/MS was 9.038–41.310 nmol/L. The similarity of the LRLs (9.038 vs. 9.2 nmol/L) and the significant difference of the URLs (41.310 vs. 31.8 nmol/L) among our and Travison's study have many possible explanations. First, Travison et al. measured T concentrations in only 200 young non-obese men (from the FHS and the SIBLOS study) and then generated normalizing equations to calculate harmonized values; the introduction of the normalizing equations could have introduced a measurement bias. Second, there is a well-known lack of standardization among LC-MS/MS methods; it is possible that the technique we used differed from the one used by Travison et al. Last but not least, as significant geographical and racial differences in sex-steroid levels have been previously demonstrated [46, 47], a real difference in T concentration levels in our population is possible. Eisenhofer et al. [48] recently published total T RRs in 252 healthy men calculated with LC-MS/MS; they found an RR of 7.6–37.1 nmol/L (all age and BMI considered), which is comparable with ours (7.495–34.757 nmol/L) (Table 5). The similarity could be explained by an absence of geographic and racial differences (Germany vs. Switzerland). In addition, it is possible that the two chromatographic techniques used may have a certain degree of agreement, despite using different platforms.

Clinical impact of testosterone measurement

In recent years, T replacement therapy has become popular and marketing of T treatments and disease-awareness campaigns for "low-T" have been aggressive and effective [25, 49, 50]. It is estimated that by 2025, there will be approximately 6.5 million American men aged 30–80 years that will be diagnosed with androgen deficiency [51]. Increased cardiovascular events [52] and only modest benefit of T replacement therapy [53, 54] have been recently reported and the concern about misdiagnosis, overtreatment and long-term safety keep rising. A recent Food and Drug Administration report showed that among all men receiving a new T prescription, nearly a third had no evidence of a prior serum T measurement [55]. In our study, the rate of healthy individuals with low T (T <8 nmol/L) varied largely depending on the platform (Beckman 18%, Immulite 13%, Roche 11%, Abbott 9%, LC-MS 6%); nevertheless, statistically significant differences were found only between Beckman and Abbott and Beckman and LC-MS/MS. The rate of participants with a T level <11 nmol/L also varied according to the analytical method: 10% with Abbott, 11% with LC-MS, 16% with Roche, 24% with Immulite and 30% with Beckman. Consistent with the marginal role of age trend in total T concentrations, we did not find statistical differences in the rates of participants with $T < 11$ nmol/L between men aged <40 years and the entire participants collective.

The reason for such a high percent of healthy volunteers with a reduced T level could be due to a multitude of factors that have been proven to contribute to total T variation, including intraindividual variation [56, 57]. Certainly, the application of thresholds that do not consider the disagreement between different IMAs and between IMAs and LC-MS/MS introduces a bias. The two cutoffs of 8 and 11 nmol/L fall within the normal T range calculated for many assays used in this study. To prevent LOH overdiagnosis, RRs as well as T thresholds should be method specific and adjusted to the local population.

There are a number of limitations to this study. The results are based on cross-sectional data from questionnaires and a single T measurement, and clinical history was based on volunteers' statements. Testicular volume was not measured; therefore, no definitive statement to explain the difference with previously reported total T RRs could be made. On the other hand, this study is the first to provide total T RRs with four different IMAs and LC-MS/MS (performed in a single laboratory) in a large population of European healthy men.

Conclusions

Many questions about total T RRs remain currently a matter of debate, and there is a clear need to standardize T measuring. We have demonstrated that the agreement between most IMAs and LC-MS/MS is poor and that total T RRs differ among different analytical methods. While waiting for a standardization of IMA measuring, appropriate RRs and diagnostic thresholds, established in each laboratory and adapted to different analytical platforms, should be used in order to increase diagnosis accuracy and decrease the use of a potentially harmful drug.

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 organization(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. Vesper HW, Botelho JC, Wang Y. Challenges and improvements in testosterone and estradiol testing. Asian J Androl 2014;16:178–84.
2. Harman SM. Testosterone in older men after the Institute of Medicine Report: where do we go from here? Climacteric 2005;8:124–35.
3. Kaufman JM, Vermeulen A. The decline of androgen levels in elderly men and its clinical and therapeutic implications. Endocr Rev 2005;26:833–76.
4. Stellato RK, Feldman HA, Hamdy O, Horton ES, McKinlay JB. Testosterone, sex hormone-binding globulin, and the development of type 2 diabetes in middle-aged men: prospective results from the Massachusetts male aging study. Diabetes Care 2000;23:490–4.
5. Wittert G. The relationship between sleep disorders and testosterone in men. Asian J Androl 2014;16:262–5.
6. Wang C, Jackson G, Jones TH, Matsumoto AM, Nehra A, Perelman MA, et al. Low testosterone associated with obesity and the metabolic syndrome contributes to sexual dysfunction and cardiovascular disease risk in men with type 2 diabetes. Diabetes Care 2011;34:1669–75.
7. Oskui PM, French WJ, Herring MJ, Mayeda GS, Burstein S, Kloner RA. Testosterone and the cardiovascular system: a comprehensive review of the clinical literature. J Am Heart Assoc 2013;2:e000272.
8. Wu FC, Tajar A, Pye SR, Silman AJ, Finn JD, O'Neill TW, et al. Hypothalamic-pituitary-testicular axis disruptions in older men are differentially linked to age and modifiable risk factors: the European Male Aging Study. J Clin Endocrinol Metab 2008;93:2737–45.
9. Lombardo F, Sgro P, Salacone P, Gilio B, Gandini L, Dondero F, et al. Androgens and fertility. J Endocrinol Invest 2005;28(3 Suppl):51–5.
10. Moskovic DJ, Araujo AB, Lipshultz LI, Khera M. The 20-year public health impact and direct cost of testosterone deficiency in U.S. men. J Sex Med 2013;10:562–9.
11. Harman SM, Metter EJ, Tobin JD, Pearson J, Blackman MR, Baltimore Longitudinal Study of Aging. Longitudinal effects of aging on serum total and free testosterone levels in healthy men. Baltimore Longitudinal Study of Aging. J Clin Endocrinol Metab 2001;86:724–31.
12. Feldman HA, Longcope C, Derby CA, Johannes CB, Araujo AB, Coviello AD, et al. Age trends in the level of serum testosterone and other hormones in middle-aged men: longitudinal results from the Massachusetts male aging study. J Clin Endocrinol Metab 2002;87:589–98.
13. Travison TG, Araujo AB, Kupelian V, O'Donnell AB, McKinlay JB. The relative contributions of aging, health, and lifestyle factors to serum testosterone decline in men. J Clin Endocrinol Metab 2007;92:549–55.
14. Corona G, Monami M, Rastrelli G, Aversa A, Tishova Y, Saad F, et al. Testosterone and metabolic syndrome: a meta-analysis study. J Sex Med 2011;8:272–83.
15. Muller M, den Tonkelaar I, Thijssen JH, Grobbee DE, van der Schouw YT. Endogenous sex hormones in men aged 40–80 years. Eur J Endocrinol 2003;149:583–9.
16. Wu FC. Commentary: guideline for male testosterone therapy: a European perspective. J Clin Endocrinol Metab 2007;92:418–9.
17. Shames D, Gassman A, Handelsman H. Commentary: guideline for male testosterone therapy: a regulatory perspective. J Clin Endocrinol Metab 2007;92:414–5.
18. Foresta C, Calogero AE, Lombardo F, Lenzi A, Ferlin A. Late-onset hypogonadism: beyond testosterone. Asian J Androl 2015;17:236–8.
19. Neaves WB, Johnson L, Porter JC, Parker CR, Jr., Petty CS. Leydig cell numbers, daily sperm production, and serum gonadotropin levels in aging men. J Clin Endocrinol Metab 1984;59:756–63.
20. Landry D, Cloutier F, Martin LJ. Implications of leptin in neuroendocrine regulation of male reproduction. Reprod Biol 2013;13:1–14.
21. Mah PM, Wittert GA. Obesity and testicular function. Mol Cell Endocrinol 2010;316:180–6.
22. Spitzer M, Huang G, Basaria S, Travison TG, Bhasin S. Risks and benefits of testosterone therapy in older men. Nat Rev Endocrinol 2013;9:414–24.
23. Baillargeon J, Urban RJ, Ottenbacher KJ, Pierson KS, Goodwin JS. Trends in androgen prescribing in the United States, 2001 to 2011. JAMA Intern Med 2013;173:1465–6.
24. Handelsman DJ. Global trends in testosterone prescribing, 2000–2011: expanding the spectrum of prescription drug misuse. Med J Aust 2013;199:548–51.
25. Nguyen CP, Hirsch MS, Moeny D, Kaul S, Mohamoud M, Joffe HV. Testosterone and "Age-Related Hypogonadism" – FDA Concerns. N Engl J Med 2015;373:689–91.
26. Busnelli A, Somigliana E, Vercellini P. 'Forever Young'-Testosterone replacement therapy: a blockbuster drug despite flabby evidence and broken promises. Hum Reprod 2017;32:719–24.
27. Wu FC, Tajar A, Beynon JM, Pye SR, Silman AJ, Finn JD, et al. Identification of late-onset hypogonadism in middle-aged and elderly men. N Engl J Med 2010;363:123–35.

28. Rosner W, Auchus RJ, Azziz R, Sluss PM, Raff H. Position statement: utility, limitations, and pitfalls in measuring testosterone: an endocrine society position statement. J Clin Endocrinol Metab 2007;92:405–13.
29. Paduch DA, Brannigan RE, Fuchs EF, Kim ED, Marmar JL, Sandlow JI. The laboratory diagnosis of testosterone deficiency. Urology 2014;83:980–8.
30. Le M, Flores D, May D, Gourley E, Nangia AK. Current practices of measuring and reference range reporting of free and total testosterone in the United States. J Urol 2016;195:1556–61.
31. Sikaris K, McLachlan RI, 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;90:5928–36.
32. Wang C, Catlin DH, Demers LM, Starcevic B, Swerdloff RS. Measurement of total serum testosterone in adult men: comparison of current laboratory methods versus liquid chromatography-tandem mass spectrometry. J Clin Endocrinol Metab 2004;89:534–43.
33. Rosner W, Vesper H, Endocrine Society, American Association for Clinical Chemistry, American Association of Clinical Endocrinologists, Androgen Excess/PCOS Society, et al. Toward excellence in testosterone testing: a consensus statement. J Clin Endocrinol Metab 2010;95:4542–8.
34. Wang C, Nieschlag E, Swerdloff RS, Behre H, Hellstrom WJ, Gooren LJ, et al. ISA, ISSAM, EAU, EAA and ASA recommendations: investigation, treatment and monitoring of late-onset hypogonadism in males. Aging Male 2009;12:5–12.
35. Nieschlag E, Swerdloff R, Behre HM, Gooren LJ, Kaufman JM, Legros JJ, et al. Investigation, treatment, and monitoring of late-onset hypogonadism in males: ISA, ISSAM, and EAU recommendations. J Androl 2006;27:135–7.
36. Bhasin S, Basaria S. Diagnosis and treatment of hypogonadism in men. Best Pract Res Clin Endocrinol Metab 2011;25:251–70.
37. Travison TG, Vesper HW, Orwoll E, Wu F, Kaufman JM, Wang Y, 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;102:1161–73.
38. Huhtaniemi IT, Tajar A, Lee DM, O'Neill TW, Finn JD, Bartfai G, et al. Comparison of serum testosterone and estradiol measurements in 3174 European men using platform immunoassay and mass spectrometry; relevance for the diagnostics in aging men. Eur J Endocrinol 2012;166:983–91.
39. Rosen RC, Riley A, Wagner G, Osterloh IH, Kirkpatrick J, Mishra A. The international index of erectile function (IIEF): a multidimensional scale for assessment of erectile dysfunction. Urology 1997;49:822–30.
40. Travison TG, Morley JE, Araujo AB, O'Donnell AB, McKinlay JB. The relationship between libido and testosterone levels in aging men. J Clin Endocrinol Metab 2006;91:2509–13.
41. Institute CaLS. Measurement Procedure Comparison and Bias Estimation Using Patient samples; approved guidelines, 3rd ed. ISBN 1-56238-888-6 (Electronic); CLSI Document EP9-A3, 2013. Available at: https://clsi.org/media/1435/ep09a3_sample.pdf.
42. Clinical and Laboratory Standards Institute (CLSI). Defining E, and Verifying Reference Intervals in the Clinical Laboratory. Approved guideline-third Edition. CLSI Document EP28-A3C. Horowitz, GL: CLSI, 2010. Available at: https://clsi.org/media/1421/ep28a3c_sample.pdf.
43. Taieb J, Mathian B, Millot F, Patricot MC, Mathieu E, Queyrel N, et al. Testosterone measured by 10 immunoassays and by isotope-dilution gas chromatography-mass spectrometry in sera from 116 men, women, and children. Clin Chem 2003;49:1381–95.
44. Mohr BA, Bhasin S, Link CL, O'Donnell AB, McKinlay JB. The effect of changes in adiposity on testosterone levels in older men: longitudinal results from the Massachusetts Male Aging Study. Eur J Endocrinol 2006;155:443–52.
45. Matsumoto AM, Bremner WJ. Serum testosterone assays–accuracy matters. J Clin Endocrinol Metab 2004;89:520–4.
46. Orwoll ES, Nielson CM, Labrie F, Barrett-Connor E, Cauley JA, Cummings SR, et al. Evidence for geographical and racial variation in serum sex steroid levels in older men. J Clin Endocrinol Metab 2010;95:E151–60.
47. Vesper HW, Wang Y, Vidal M, Botelho JC, Caudill SP. Serum total testosterone concentrations in the US Household Population from the NHANES 2011–2012 study population. Clin Chem 2015;61:1495–504.
48. Eisenhofer G, Peitzsch M, Kaden D, Langton K, Pamporaki C, Masjkur J, et al. Reference intervals for plasma concentrations of adrenal steroids measured by LC-MS/MS: impact of gender, age, oral contraceptives, body mass index and blood pressure status. Clin Chim Acta 2017;470:115–24.
49. Huhtaniemi I. Late-onset hypogonadism: current concepts and controversies of pathogenesis, diagnosis and treatment. Asian J Androl 2014;16:192–202.
50. Layton JB, Li D, Meier CR, Sharpless JL, Stürmer T, Jick SS, et al. Testosterone lab testing and initiation in the United Kingdom and the United States, 2000 to 2011. J Clin Endocrinol Metab 2014;99:835–42.
51. Araujo AB, Esche GR, Kupelian V, O'Donnell AB, Travison TG, Williams RE, et al. Prevalence of symptomatic androgen deficiency in men. J Clin Endocrinol Metab 2007;92:4241–7.
52. Elsherbiny A, Tricomi M, Bhatt D, Dandapantula HK. State-of-the-art: a review of cardiovascular effects of testosterone replacement therapy in adult males. Curr Cardiol Rep 2017;19:35.
53. Snyder PJ, Bhasin S, Cunningham GR, Matsumoto AM, Stephens-Shields AJ, Cauley JA, et al. Effects of testosterone treatment in older men. N Engl J Med 2016;374:611–24.
54. Orwoll ES. Establishing a framework–does testosterone supplementation help older men? N Engl J Med 2016;374:682–3.
55. FDA briefing document for the joint meeting for Bone, Reproductive, and Urologic Drugs Advisory Committee (BRUDAC) and Drug Safety and Risk Management Advisory Committee (DSARMAC). Silver Sring, MD: FDA Advisory Committee, 17 Sep 2014. Available at: http://www.dfda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/ReproductiveHealthDrugsAdvisoryCommittee/UCM412536.pdf.
56. Cunningham GR, Toma SM. Clinical review: why is androgen replacement in males controversial? J Clin Endocrinol Metab 2011;96:38–52.
57. Swerdloff RS, Wang C, Cunningham G, Dobs A, Iranmanesh A, Matsumoto AM, et al. Long-term pharmacokinetics of transdermal testosterone gel in hypogonadal men. J Clin Endocrinol Metab 2000;85:4500–10.

Supplemental Material: The online version of this article offers supplementary material (https://doi.org/10.1515/cclm-2017-1201).