

Published in final edited form as:

*Ann Clin Biochem.* 2014 May ; 51(Pt 3): 368–378. doi:10.1177/0004563213506412.

## Salivary Testosterone Measurement by Liquid Chromatography Tandem Mass Spectrometry in Adult Males and Females

BG Keevil<sup>1</sup>, P MacDonald<sup>1</sup>, W Macdowall<sup>3</sup>, DM Lee<sup>2</sup>, FCW Wu<sup>2</sup>, and the NATSAL Team

<sup>1</sup>Department of Clinical Biochemistry, University Hospital of South Manchester, Southmoor Road, Manchester M13 9WL, , Manchester Centre for Endocrinology & Diabetes, Manchester Academic Health Science Centre, The University of Manchester, Manchester

<sup>2</sup>Andrology Research Unit, Old St Mary's Building, Hathersage Road, Manchester M13 9WL, , Manchester Centre for Endocrinology & Diabetes, Manchester Academic Health Science Centre, The University of Manchester, Manchester

<sup>3</sup>Centre for Sexual and Reproductive Health Research Department of Social and Environmental Health Research London School of Hygiene and Tropical Medicine London, United Kingdom

### Abstract

**Background**—Salivary testosterone (Sal-T) may be a useful surrogate of serum free testosterone. The study aims were to use a novel liquid chromatography tandem mass spectrometry (LC-MS/MS) assay to determine whether Sal-T concentrations accurately reflect Sal-T concentrations in both sexes and to investigate practical aspects of sample collection.

**Methods**—Saliva and serum samples were collected in 104 male and 91 female subjects. A more sensitive LC-MS/MS assay was developed to enable Sal-T quantitation in the low concentrations found in females. Saliva (200 µL) was extracted with 1 mL of methyl-*tert*-butyl ether following the addition of D5-testosterone. Quantitation was performed using a Waters TQ-S mass spectrometer.

**Results**—The assay achieved a lower limit of quantification of 5pmol/L, sufficiently sensitive to measure testosterone in female saliva. Sal-T showed a diurnal variation but samples taken at weekly and monthly intervals showed no significant differences. Sal-T was stable at ambient temperature for up to 5 days, after freeze-thawing and 3 years frozen storage. Reference intervals

---

Address all correspondence and requests for reprints to Brian G Keevil, Department of Clinical Biochemistry, University Hospital of South Manchester, Southmoor Road, Manchester M23 9LT, United Kingdom. brian.keevil@uhsm.nhs.uk.

**Contributorship:** FW, BK. co-designed the study;

BK, PM performed the study;

BK, FW, DL, PM wrote the draft paper;

FW, WM, BK revised the draft paper;

Declaration of conflicting interests

FCWW is a consultant for GSK, Bayer-Schering, Jenapharm, Ferring, TAP, Lilly-ICOS, Proctor & Gamble, Ardana Biosciences, Pierre Fabre Medicaments, Acrux DDS Pty Ltd, and has also chaired advisory board meetings and lectured on their behalf. FCWW has received lecture fees from Organon, Bayer-Schering, Lilly-ICOS, Ardana Biosciences, Pierre Fabre Medicaments. FCWW has received grant support (2009-2013) from Bayer Schering AG and the New England Research Institute, Watertown, MA, USA. All authors have nothing to disclose.

**Ethical approval**

This study was granted ethical approval from Oldham research Ethics Committee (09/H1011/18).

**Guarantor**

BK.

for Sal-T were 93-378 pmol/L in males and 5-46 pmol/L in females. Sal-T correlated significantly with serum calculated free-T in males ( $r=0.71$ ,  $P<0.001$ ) and in females ( $r=0.39$ ,  $P<0.001$ ).

**Conclusions**—These results confirm that testosterone can be reliably and accurately measured by LC-MS/MS in both adult male and female saliva samples. These results lay the foundation for further exploration of the clinical application of Sal-T as a reliable alternative to serum testosterone in the diagnosis and management of androgen disorders and assessment of androgen status in clinical research.

---

## Introduction

A self-administered non-invasive method to collect biological samples in which measurements of testosterone (T) levels can be accurately determined is clearly of enormous value to population research, population screening, patient diagnosis and treatment monitoring. As compared with the standard medium of blood (serum/plasma), the use of saliva will obviate the need for, and avoids the costs, stress and pain of venepuncture and attendance at a clinic or home visit by clinical personnel. Collection of saliva is easily acceptable, non-invasive, and requires minimal training, thereby facilitating collection of multiple samples to account for any natural within-individual biological variations. The advantages of salivary hormone collection have been recently reviewed (1), but there are also problems associated with saliva collection including blood contamination and adsorption of T onto saliva collection devices (2, 3).

Saliva contains only the free unbound fraction of T which freely diffuses across capillaries and salivary ducts and is unaffected by saliva flow rates, (4). T in saliva is thought to be closely related to the component of circulating T to which tissues androgen receptors are exposed. However, in contrast to serum testosterone (serum-T), salivary-T (Sal-T) is unaffected by variations in circulating sex hormone binding globulin (SHBG) and albumin. Measuring Sal-T concentrations may therefore provide an opportunity to directly assess 'tissue' testosterone concentrations which in turn may be a more accurate index of overall androgenicity and an alternative to serum free (free-T) or bio available-T. The latter are usually derived from mathematical formulae based on association constants of T with binding proteins, and their relationship to measured free-T and their clinical significance have been questioned (5, 6). Good agreement between Sal-T and free-T measurements by immunoassays has been shown in both eugonadal and hypogonadal men, suggesting that Sal-T is a reliable marker of testosterone bioavailability (7, 8).

Despite these obvious advantages, Sal-T measurements have not gained general acceptance in the clinical and research communities, largely due to uncertainties over the accuracy and reliability of the T assays (9, 10). These are related to difficulties in obtaining antibodies for the development of suitably sensitive and specific T immunoassays, poorly-optimized quantification, inadequate sample collection and preparation, limited sample size and, most critically, the lack of proper validation of salivary against currently accepted standard serum measurements. Recent studies have described the liquid chromatography tandem mass spectrometry (LC-MS/MS) analysis of Sal-T in males but the assays required either derivatization or large sample volumes to improve sensitivity (11, 12). More recently a

method has been described for measurement of Sal-T in males without derivatization that required only a small sample volume (13). To date, there have been no studies using LC-MS/MS to measure Sal-T in both males and females. We report here on the first study to systematically validate Sal-T measurement in both adult men and women. This validation work was carried out as part of the development work for Britain's Third National Survey of Sexual Attitudes and Lifestyles. The specific aims of the study were to: 1) develop a highly sensitive LC-MS/MS method to quantify T in saliva samples in adult men and women. 2) determine whether Sal-T concentration accurately reflect serum-T concentrations in both men and women and 3) investigate practical aspects of salivary sample collection, including stability in transport and storage.

## Materials and Methods

A series of four studies were conducted to explore 1) analytical assay performance, 2) stability of T in saliva, 3) the correlations between Sal-T and serum-T and 4) individual variations in T, each of which are described in detail later. Saliva and blood samples were collected in tandem in all four studies and were stored, transported and analysed under standardized conditions (see below).

### Participants

Paired saliva and blood samples were collected from 104 men and 91 women recruited by advertisements from the general population (age range 16 – 74 years stratified to ~20 in each decade). Medical conditions and any medications (including hormonal contraceptives and hormone replace therapy) were recorded. Informed written consent was obtained from all participants.

### Collection and handling of saliva samples

Participants were asked to spit or drool directly into a 4 mL sealable polystyrene tube (Bibby Sterilin Ltd, Staffordshire, UK) and to provide at least 3 mL of saliva. Unstimulated saliva samples were used to avoid any assay interference. Several commercially available swab collection devices were evaluated but showed poor testosterone recovery due to non-specific binding (results not shown). We therefore used the passive drool technique throughout this study (14). In order to avoid blood and other oral contaminations that may interfere with the assay, volunteers were requested to:

Avoid dental work for 48 hours prior to sample collection

Avoid teeth brushing for 2h prior to sample collection

Avoid eating for at least 1h prior to sample collection

Rinse their mouth with water not less than 10 minutes and not more than 15 minutes prior to sample collection.

Each saliva sample was immediately put on ice until transported to the preparation laboratory. At the preparation laboratory, saliva specimens were checked for blood contamination visually. No specimens with visually evident blood contamination were analysed. Saliva samples were immediately frozen to -20°C, thawed the following day, and

then centrifuged at 1500 g for 15 minutes to precipitate mucopolysaccharides. The clear supernatant in each sample was either pooled for stability studies or divided into three equal aliquots for validation studies. All samples were identified by unique codes; no personal identifying information was available to laboratories.

### Collection of blood samples

Blood samples were collected immediately after (within 5 min) saliva collection. About 15 mL of venous blood was drawn from an antecubital vein to give 3 x 2.5mL aliquots of serum (in cryovials for storage at -80°C).

### Analytical assay performance

**Liquid chromatography tandem mass spectrometry**—Sal-T in males was initially measured using LC-MS/MS according to a previously published method (13), but due to the low concentrations of testosterone in the female range it was necessary to improve the assay sensitivity. This was achieved using a research grade LC-MS/MS instrument specifically to improve the sensitivity of testosterone measurement in female samples. Sample preparation for both males and females involved the same liquid–liquid extraction procedure by adding sample (200  $\mu$ L), D<sub>5</sub>-testosterone internal standard (10  $\mu$ L, 340 pmol/L) and methyl-*tert*-butyl ether (1 mL), vortex for 5 min and then freezing at -80 °C. After 1 h at this temperature, the organic layer was transferred and evaporated by heating and gentle N<sub>2</sub> gas flow. The residue was reconstituted with a 500 mL/L methanol mobile phase (80  $\mu$ L) and transferred to a 96-well microtiter plate. Extracted sample (35  $\mu$ L) was injected onto the analytical column. Liquid chromatography was performed with an ACQUITY™ Ultra Performance Liquid Chromatography system (Waters Corporation, Manchester, UK) and a C8 Kinetex column 3.0 × 100 mm 2.6  $\mu$ m (Phenomenex, Macclesfield, UK) maintained at 45 °C. The mass spectrometer was a Xevo TQ-S™ mass spectrometer (Waters Corporation, Manchester, UK) operated in positive ionization mode. Binary pump mixing of mobile phases A (2 mmol/L ammonium acetate and 1 mL/L formic acid in distilled water) and B (2 mmol/L ammonium acetate and 1 mL/L formic acid in methanol) produced a linear gradient that increased from 45% to 85% methanol for 4 min. The signal was optimized for the testosterone and D<sub>5</sub>-testosterone precursor ions (seen at  $m/z$  289.3 and  $m/z$  294.2) and the most abundant product ions (seen at  $m/z$  109.2 and  $m/z$  97.15 for testosterone and at  $m/z$  100.2 for D<sub>5</sub>-testosterone). Transitions were monitored in multiple reaction monitoring mode, with a dwell time of 0.15 s, the capillary was maintained at 0.8 kV and the desolvation temperature and gas flow were 650°C and 700 L/h respectively. The cone and collision energies were 30 V and 20 eV. The assay was validated according to the published Bioanalytical Method Validation guidelines issued by the U.S. Department of Health and Human Services Food and Drug Administration (15) which include acceptance criteria for linearity, precision, recovery and sample stability. Calibrator accuracy was assessed prior to the study by comparing calibrators, prepared from pure testosterone powder (Cat No; T1500, Sigma-Aldrich Company Ltd, Gillingham, UK) dissolved in methanol and then diluted into phosphate buffered saline PBS/bovine serum albumin as an artificial matrix, with routine LC-MS/MS testosterone calibrators. We have previously shown that measuring serum T samples using calibrators made in this manner has excellent agreement with the reference method (16). To evaluate linearity of the calibration curves, three curves were prepared and

analysed in separate batches. The ratios of analyte peak area to internal standard peak area were plotted against testosterone concentration in pmol/L. Calibration curves were judged linear if the coefficient of determination ( $r^2$ ) was better than 0.9900 as calculated by weighted linear regression. The lower limit of quantitation (LLOQ) was defined as the concentration for which 10 replicates of PBS-based samples prepared with low concentrations of testosterone gave a coefficient of variation (CV) of less than 20% and bias of less than 20%. Ion suppression is a matrix effect which occurs when compounds in a sample compete with the analyte for ionisation in the source. To investigate this we infused a 340 pmol/L solution of D<sub>5</sub> testosterone in 40% (v/v) methanol/water directly into the mass spectrometer to give a constant background signal. Extracted saliva samples (n=6) were injected simultaneously via the autosampler. A reduction in the background signal is observed when ion suppression is occurring and ion suppression is deemed significant if a reduction in signal of >10% is observed where the compound of interest elutes.

To determine the stability of the extracted saliva, 35 different samples were initially analysed on the LC-MS/MS. The 96-well microtitre plate was then resealed and analysed 24 h later after overnight storage at 4°C.

Serum-T was measured by LC-MS/MS using a previously validated method (17). Albumin was measured by the Bromocresol green (BCG) photometric method (Roche modular P analyser, Burgess Hill, UK) and SHBG was measured by a chemiluminescent enzyme immunoassay (Immulite 2000 automated analyser, Siemens, Newbury, UK). Calculated free-T concentrations were derived from law of mass action equations (Vermuelen) using serum-T, SHBG and albumin concentrations (18).

### Sample stability during transport and storage

Stability of T in saliva samples was investigated using LC-MS/MS, firstly at room temperature and at 4°C daily over 5 days, secondly after five freeze-thaw cycles over a period of 6 months and thirdly after storage at -20° for 1 week, 1 month, 2 months, and 6 months. Long-term sample stability of Sal-T at -20°C was also assessed using LC-MS/MS measurements in 19 men, from saliva samples collected and measured in 2009 and again approximately 3 years later.

### Comparison between salivary and serum T

Testosterone concentrations were measured in paired saliva and blood samples from male and female participants (n=93 and n=86, respectively). Sal-T concentrations compared with total and calculated free serum-T in men and women separately. These data were used for the determination of reference intervals (19).

### Intra-individual variability

Within-individual variations in Sal-T and serum-T were explored in paired samples of saliva and blood in 11 men and 12 women. To investigate circadian variation, three samples of saliva and blood were collected over 1 h (before 10.00h) at 30 min intervals. To investigate diurnal variation, two samples of saliva and blood were collected over 24 h (before 10.00 h and 22.00h) 12 h apart. To investigate week-to-week variation, saliva and blood samples

were collected (before 10.00h) once a week over 4 weeks. Lastly, to investigate month-to-month variation, saliva and blood samples were collected (before 10.00h), once a month over four months.

## Statistical analyses

Pearson's correlation coefficient was used to estimate the strength of association between the various paired measures of Sal-T and serum-T. For groups of more than two linked measurements taken from participants such as the weekly measurements (up to four), one-way repeated measures analysis of variance (ANOVA) was used to determine significant differences between the related means. Paired *t*-test or the Wilcoxon matched paired test where appropriate (depending on the distribution of the measurements) was used to determine the significance of the differences between saliva and serum-T concentrations, and other pairs of measurements on the samples. All statistical analyses were conducted using Intercooled STATA version 9.2 (StataCorp, College Station, TX, USA) and SPSS version 19.0 (SPSS Inc., Chicago, IL, USA). The sample size was sufficient to detect a statistically significant (5% significance level) correlation between salivary and blood testosterone if the underlying correlation is 0.7 with a power of more than 0.99.

This study was granted ethical approval from Oldham Research Ethics Committee (09/H1011/18). Informed written consent was obtained from all participants. All samples were identified by unique codes; no personal identifying information was available to laboratories.

## Results

### Analytical performance

Sal T analysis The LC-MS/MS method for measuring Sal-T in males has previously been shown to have good analytical performance (13). The extracted female saliva samples were measured using the more sensitive LC-MS/MS method described and produced clean chromatograms with no interfering peaks when using the testosterone quantifier transition  $m/z$  289.3>97.15 (qualifier 289.3>109.2) (Figure 1). Following injection of the extracted sample (35  $\mu$ L), testosterone and D<sub>5</sub>-testosterone co-eluted with clean, discrete and identifiable peaks at a retention time of 3.6 min. The total run time, injection to injection, was 6.5 min. Infusion experiments showed that there was no significant ion suppression present in any of the six chromatograms from extracted saliva when compared with an injected water sample. Standard curves were made by plotting testosterone concentrations on the x-axis and testosterone/D<sub>5</sub> testosterone peak area ratios on the y-axis. The curve was linear over the standard range and was reproducible between batches. The curves showed good correlation with the assigned standard values with an  $r^2$  value of 0.999. Overnight storage of the extracted saliva at 4°C in 35 different samples showed no decrease in measured testosterone values, indicating good sample stability.

The lower limit of quantification of the more sensitive assay was 5 pmol/L. Inter-assay CV(SD) bias was 12.9% (1.7) and 1.2%, 9.8%(2.5) and 0.4%; 4.5%(12.0) and 1.9% at concentrations of 12.9, 26.0 and 260 pmol/L. Intra assay CV(SD) bias was 9.5%(1.3) and

0.8%; 5.5% (1.6) and 12.6%; 2.1% (6.2) and 11.1%; at concentrations of 12.9, 26.0 and 260 pmol/L. Recovery was 104% (range 98.3 – 108.9).

### Serum T analysis

Serum-T concentrations analysed by LC-MS/MS (17) had a mean intra-assay CV of 4.0% (2.6 – 6.9) and mean inter-assay CV of 5.6% (3.2 – 9.8) with LLOQ of 0.3 nmol/L.

### Sample stability during transport and storage

One-way repeated measures ANOVA showed that Sal-T concentrations in different male saliva samples stored at both room temperature and at 4°C for a period of 5 days did not differ significantly ( $F = 0.88$ ,  $P = 0.50$  and  $F = 1.03$ ,  $p = 0.42$  respectively; [Figure 2(a) and (b)]. There was also no significant difference in Sal-T measurements over 5 freeze-thaw cycles ( $F = 1.29$ ,  $p = 0.31$ ). [Figure 2(c)] or after 3 years storage (Mean T before storage 217 pmol/L and after storage 202 pmol/L). Paired-sample t-tests showed that LC-MS/MS Sal-T ( $t = 2.88$ ,  $p = 0.01$ ) were significantly lower at the second (mean decrease = 15 pmol/L) as compared with the first measurement. However, the mean decrease in testosterone over the 3 year period was 4.7% (95% CI 0.4 – 9.0) which is within the analytical variation of the assay; this confirmed the stability of testosterone in saliva after 3 years storage at -80°C.

### Comparison between Sal-T and serum T

**Distribution of serum-T and sal-T in males and females**—Results of serum-T, serum calculated free-T and Sal-T measured by LC-MS in male and female participants are shown in Table 1. A small number of samples were missing or unsuitable or insufficient for analysis. Results of serum-T, serum calculated free-T and Sal-T in female participants excluding those on oral contraceptive (OCP) or hormone replacement therapy (HRT;  $n = 6$ ) are shown in Table 2.

Mean Sal-T was at least 10 times higher in males compared with females, mirroring the serum-T results. Reference intervals (2.5<sup>th</sup> to 97.5<sup>th</sup> percentile) for Sal-T measured by LC-MS/MS, constructed using non-parametric methods, were 5 - 46 pmol/L in females and 93 - 378 pmol/L in males. Sal-T results therefore showed good demarcation between the normal female and male reference intervals when measured by LC-MS/MS (Figure 3(a) and 3(b)). Two outliers were excluded from the analysis because there was insufficient sample to repeat the test. The range of SHBG results was greater in females both on and off OCP/HRT than in males. The results for all measurements were similar in females both on and off OCP/HRT.

**Correlation between Sal-T versus serum-T**—The comparison of serum T and free-T with Sal-T (LC-MS/MS) for both males and females is shown in Figure 4. In males, the correlation between Sal-T and serum free-T was higher ( $r = 0.71$ ,  $p < 0.001$ ) than the correlation between Sal-T and serum-T ( $r = 0.5$ ,  $p < 0.001$ ). In females there was a significant but weaker correlation between Sal-T and serum-T ( $r = 0.37$ ,  $p < 0.001$ ) and free-T ( $r = 0.39$ ,  $p < 0.001$ ) compared with males. The correlations found in females did not increase after excluding females taking oestrogen containing OCP or HRT or females with high serum SHBG concentration ( $>110$  nmol/L) (data not shown).

## Intra-individual variability

**Circhoral variation and repeatability**—Three paired serum and saliva samples obtained 30 minutes apart in the morning from 11 men and 12 women were analysed to assess the repeatability or test-retest reliability. One-way repeated measures ANOVA showed that serum-T measurements from both men ( $F = 1.65$ ,  $p = 0.23$ ) and women ( $F = 1.68$ ,  $p = 0.21$ ) did not differ significantly between the three time points. Similar results were seen for LC-MS/MS Sal-T repeat measurements (men:  $F = 0.40$ ,  $p = 0.59$ ; women:  $F = 0.99$ ,  $p = 0.35$ )

**Diurnal variation**—Variation in T concentrations between two samples of saliva and blood collected over 1 day ~12 hours apart to investigate diurnal variation in men and women is shown in Figure 5. Comparisons using paired-sample *t*-tests showed that LC-MS/MS Sal-T ( $t = 4.99$ ,  $p < 0.001$ ) but not serum-T were significantly higher in morning compared with evening samples for men. Mean morning serum T ( $t = 3.26$ ,  $p = 0.01$ ) and Sal-T were significantly higher than evening concentrations in women.

**Weekly variations**—Comparisons of paired serum and saliva using one-way repeated measures ANOVA, showed that serum-T ( $F = 1.03$ ,  $p = 0.40$ ) and LC-MS/MS Sal-T measurements ( $F = 0.74$ ,  $P = 0.54$ ) were not significantly different over the four consecutive weeks that samples were collected in men. The same was true for females: serum-T ( $F = 0.96$ ,  $P = 0.43$ ) and LC-MS/MS Sal-T ( $F = 1.90$ ,  $P = 0.19$ ) in samples collected in 4 consecutive weeks without reference to the menstrual cycle.

**Monthly variations**—Monthly samples of paired serum and saliva were obtained over four consecutive months to assess within-individual variations in T. Serum-T ( $F = 1.46$ ,  $P = 0.29$ ) and LC-MS/MS Sal-T measurements ( $F = 1.51$ ,  $P = 0.29$ ) were not significantly different over the 4 consecutive months for men. Over the same 4 month period, serum-T ( $F = 0.61$ ,  $P = 0.62$ ) and LC-MS/MS Sal-T ( $F = 2.64$ ,  $P = 0.13$ ) did not significantly differ in women.

## Discussion

Sal-T is an attractive alternative for measuring free-T directly in the body as it is thought to reflect free hormone available to target tissues and may therefore be a good index of bioactivity without resorting to the technically more challenging and expensive direct measurement of serum free-T or bioassays. However, testosterone in saliva, which is not bound to SHBG, will typically be <2-3% of serum total-T concentration and its accurate measurement presents its own set of challenges.

Commercially available salivary collection devices are popular for the measurement of cortisol and a variety of these devices were evaluated, but with only limited success. The use of commercially available swab collection devices was discarded because of poor testosterone recovery due to non-specific binding. We therefore used the tried and tested passive drool technique throughout this study (14).

The problem with LC-MS/MS sensitivity for the measurement of Sal-T in females was solved by using a more advanced mass spectrometer while still continuing the previous



sample preparation techniques. The Xevo TQ-S instrument has a larger sample cone aperture (0.8 mm diameter) which increases the gas and ion flow from the ionisation source. To reduce signal noise it also has a step wave focussing lens after the ion source to focus the larger ion cloud formed and also expel any neutral contamination ions accompanying the larger sampling orifice. The LC-MS/MS Sal-T assay was developed using strict validation criteria which gave us confidence in its performance. The assay was found to perform well for measurement of Sal-T across the whole female range with only 2% of female samples below the LLOQ. We have therefore developed, for the first time, a sufficiently sensitive method to accurately quantify T concentration in female saliva and generate results that could be used to construct reliable Sal-T reference intervals for females. Importantly, the clear demarcation between the female and male Sal-T reference intervals, mirrors the separation between the male and female serum-T reference intervals.

Sal-T measured by LC-MS/MS correlated significantly with serum-T across the full adult range of concentrations in males as well as females. The correlation between Sal-T and serum-T for males in this study was similar to that found in other studies (11, 12). However, it is noteworthy that the correlation between Sal-T and serum total or free-T in females, was somewhat lower than that in males. Excluding women on oestrogens and/or with high SHBG did not improve the correlation. It may be speculated that possible gender differences in the metabolism of T in the salivary gland and the much lower concentration of T may be potential explanations.

It should also be highlighted that Sal-T correlated best with free-T rather than serum total-T in both men and women. These results lend support to the view that Sal-T may provide a reliable read-out of tissue exposure to this hormone and a more accurate reflection of androgen status than circulating T concentrations in both males and females. It is important to emphasise however that clinical and biological data based on salivary T are scarce (1) and its place in clinical management and research needs to be established by further investigations. The application of these validated methods should now be explored in different patient and population groups to establish their clinical and research utility.

Our results also showed, for the first time, a diurnal variation in Sal-T concentrations in both males and females. Sal-T generally show the highest concentrations in the morning and lowest concentrations in the evening which agreed with other LC-MS/MS studies in male subjects (11, 12). As with serum measurements, therefore, saliva samples for T measurement should be collected in the morning, preferably before 10am. We have also demonstrated the lack of within-individual variation between samples at half-hourly, weekly and monthly intervals indicating that a single morning Sal-T measurement should therefore be sufficient and appropriate to provide an accurate representation of an individual's testosterone concentration. The exact timing of sample collection in the morning was not found to be critical because we did not demonstrate a significant difference in results when samples were taken 30 min apart. We did not investigate the effects of overnight fasting.

Under the defined conditions of this study, Sal-T samples were demonstrated to be stable for up to 5days at room temperature and for extended periods of time during storage at -20°C. These results are at slight variance with the 2 day stability found for Sal-T by Matsui *et al*

(12) but nevertheless suggest that posting saliva samples to a central laboratory should cause no problems in terms of sample deterioration. Stability of Sal-T was also demonstrated after storage for up to 3 years at -80 °C. This method has been successfully validated analytically and also clinically with the determination of intra-individual variability and reference intervals. Further work needs to be carried out to determine the usefulness of this assay in the investigation of hypogonadism in males and hyperandrogenism in females.

In conclusion, this study has shown that T can be reliably and accurately measured by LC-MS/MS in both adult male and female saliva samples and that Sal-T results correlated well with serum-T concentrations. These results lay the foundation for further exploration of the clinical application of Sal-T as a reliable alternative to serum-T in the diagnosis and management of androgen disorders and assessment of androgen status in clinical and population research.

## Acknowledgements

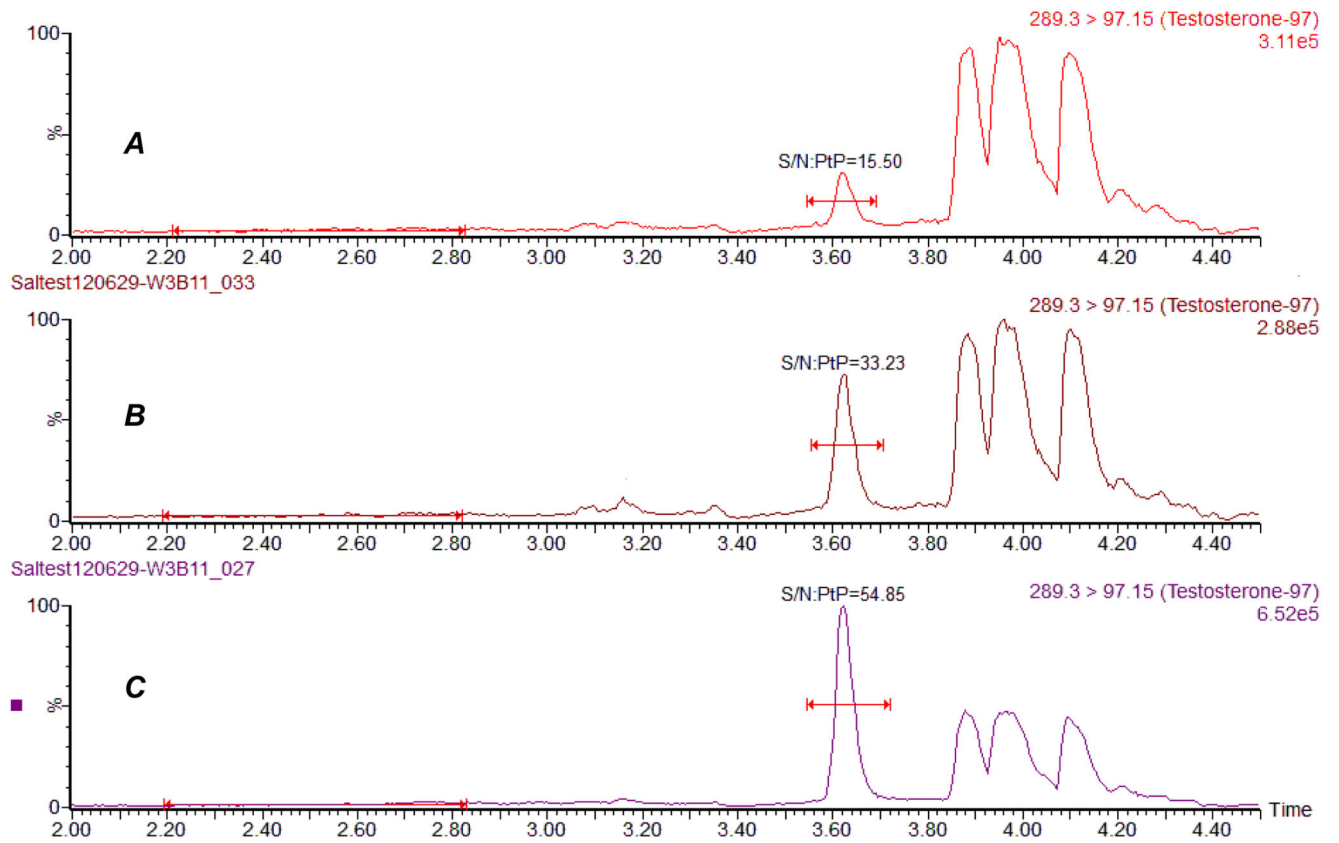
The authors wish to acknowledge the National Survey of Sexual Attitudes and Lifestyles (NATSAL) team for their scientific and operational support during this study. The NATSAL team comprises the following individuals, David Baird, John Bancroft, Soazig Clifton, Andrew Copas, Bob Erens, Anne M Johnson, Halina McIntyre, Catherine H. Mercer, Carly Mosely, Andrew Phelps, Pam Sonnenberg and Kaye Wellings.

**Funding:** This work was supported by the Wellcome Trust (084840).

## References

1. Gröschl M. Current Status of Salivary Hormone Analysis. *Clin Chem.* 2008; 54:1759–1769. [PubMed: 18757583]
2. Gröschl M, Rauch M. Influence of commercial collection devices for saliva on the reliability of salivary steroids analysis. *Steroids.* 2006; 71:1097–1100. [PubMed: 17070563]
3. Shirtcliff EA, Granger DA, Schwartz E, Curran MJ. Use of salivary biomarkers in biobehavioral research: cotton-based sample collection methods can interfere with salivary immunoassay results. *Psychoneuroendocrinology.* 2001; 26:165–73. [PubMed: 11087962]
4. Vining RF, McGinley RA, Symons RG. Hormones in saliva: mode of entry and consequent implications for clinical interpretation. *Clin Chem.* 1983; 29:1752–6. [PubMed: 6225566]
5. Hackbarth JS, Hoyne JB, Grebe SK, Singh RJ. Accuracy of calculated free testosterone differs between equations and depends on gender and SHBG concentration. *Steroids.* 2011; 76:48–55. [PubMed: 20816687]
6. Sartorius G, Ly LP, Sikaris K, McLachlan R, Handelsman DJ. Predictive accuracy and sources of variability in calculated free testosterone estimates. *Ann Clin Biochem.* 2009; 46:137–143. [PubMed: 19225026]
7. Morley JE, Perry HM, Patrick P, Dollbaum CM, Kells JM. Validation of salivary testosterone as a screening test for male hypogonadism. *The Aging Male.* 2006; 9:165–9. [PubMed: 17050116]
8. Arregger AL, Contreras LN, Tumilasci OR, Aquilano DR, Cardoso EML. Salivary testosterone: a reliable approach to the diagnosis of male hypogonadism. *Clin Endocr.* 2007; 67:656–62. [PubMed: 17953627]
9. Fulton A, Chan S, Coleman G. Effect of salivary proteins on binding curves of three radioimmunoassay kits: Amerlex-M progesterone, Amerlex cortisol, and Biodata testosterone. *Clin Chem.* 1989; 35:641–4. [PubMed: 2784724]
10. Dabbs JM Jr, Campbell BC, Gladue BA, et al. Reliability of salivary testosterone measurements: a multicenter evaluation. *Clin Chem.* 1995; 41:1581–4. [PubMed: 7586546]
11. Shibayama Y, Higashia T, Shimada K, et al. Simultaneous determination of salivary testosterone and dehydroepiandrosterone using LC–MS/MS: Method development and evaluation of

- applicability for diagnosis and medication for late-onset hypogonadism. *J Chrom B*. 2009; 877:2615–2623.
12. Matsui F, Koh E, Yamamoto K, et al. Liquid chromatography-tandem mass spectrometry (LC-MS/MS) assay for simultaneous measurement of salivary testosterone and cortisol in healthy men for utilization in the diagnosis of late-onset hypogonadism in males. *Endocr J*. 2009; 56:1083–93. [PubMed: 19734692]
  13. Macdonald PR, Owen LJ, Wu FC, MacDowall W, Keevil BG. A Liquid Chromatography-Tandem Mass Spectrometry Method for Salivary Testosterone with Adult Male Reference Interval Determination. For the NATSAL team. *Clin Chem*. 2011; 57:774–775. [PubMed: 21300739]
  14. Granger DA, Kivlighan KT, Fortunato C, et al. Integration of salivary biomarkers into developmental and behaviorally-oriented research: Problems and solutions for collecting specimens. *Physiology & Behavior*. 2007; 92:583–590. [PubMed: 17572453]
  15. Food and Drug Administration, Department of Health and Health Services. Guidance for Industry: Bioanalytical Method Validation. 2001. <http://www.fda.gov/cder/Guidance/4252fml.pdf>
  16. Thienpont LM, Van Uytvanghe K, Blincko S, et al. State-of-the-art of serum testosterone measurement by isotope dilution-liquid chromatography-tandem mass spectrometry. *Clin Chem*. 2008; 54:1290–7. [PubMed: 18556330]
  17. Gallagher LM, Owen LJ, Keevil BG. Simultaneous determination of androstenedione and testosterone in human serum by liquid chromatography-tandem mass spectrometry. *Ann Clin Biochem*. 2007; 44:48–56. [PubMed: 17270092]
  18. 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; 84:3666–72. [PubMed: 10523012]
  19. Solberg HE, International Federation of Clinical Chemistry and International Committee for Standardization in Haematology. Approved recommendation (1987) on the theory of reference values. Part 5. Statistical treatment of collected reference values. Determination of reference limits. *J Clin Chem Clin Biochem*. 1987; 25:645–56.

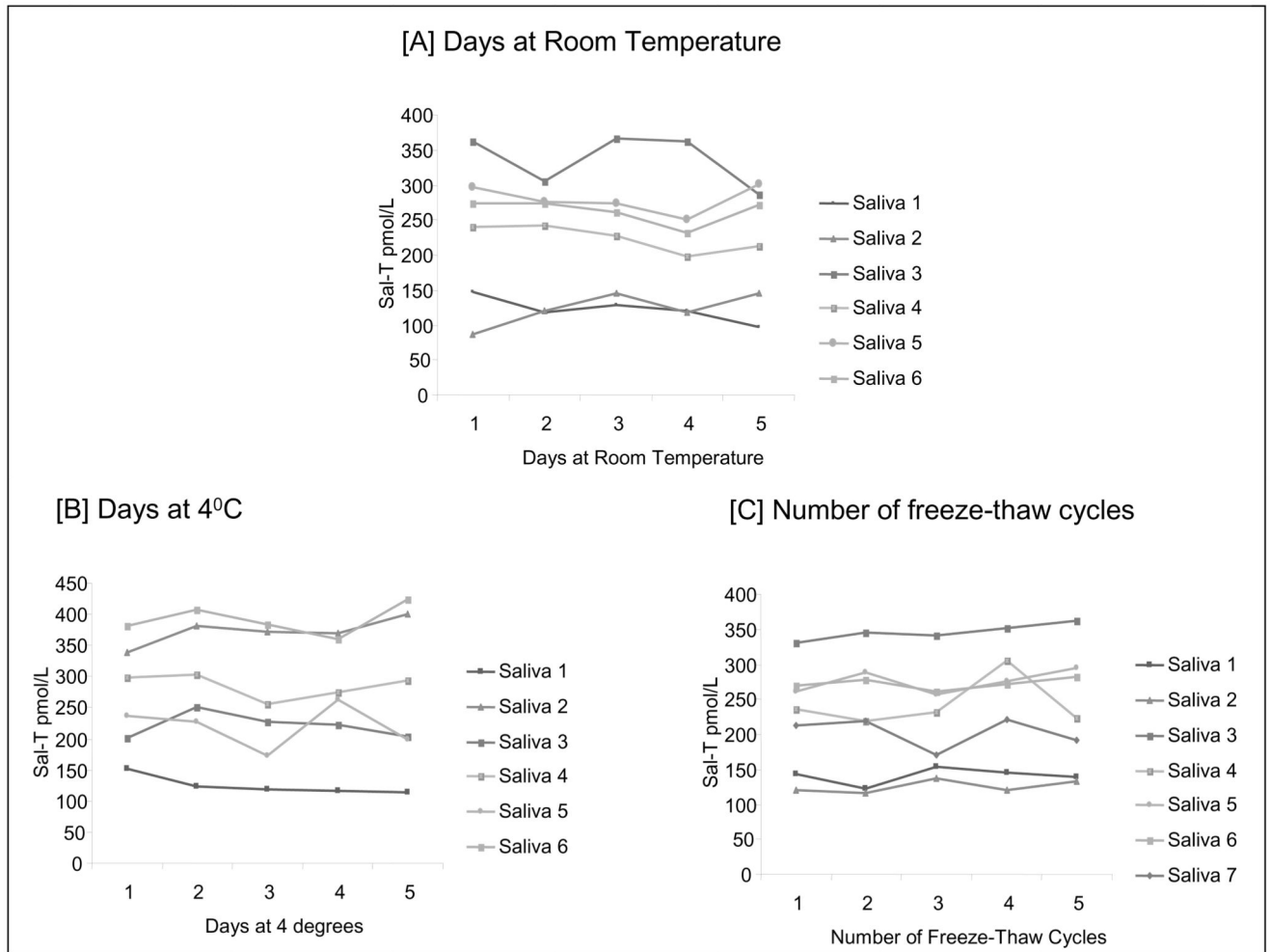


**Figure 1.**

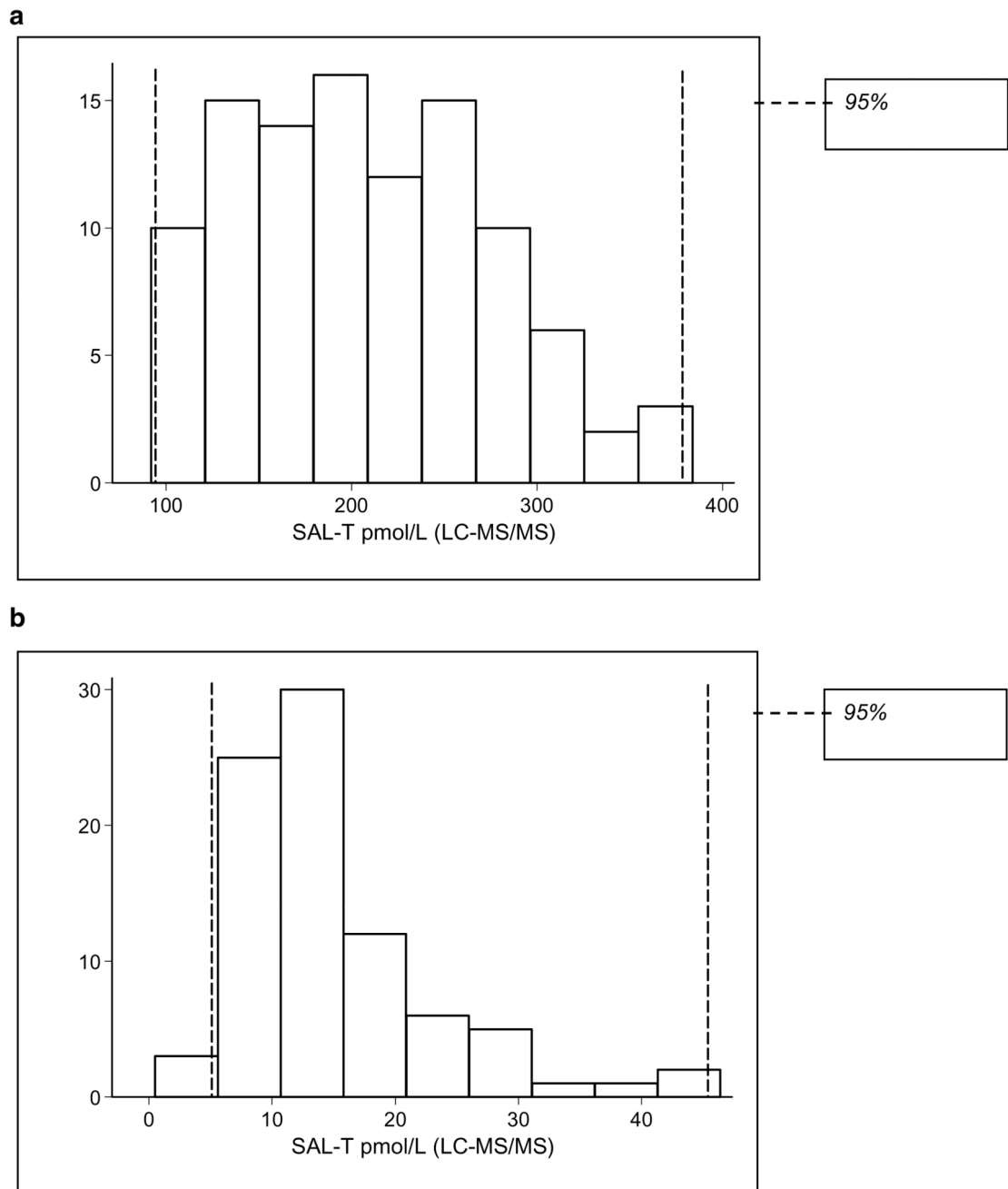
Chromatograms of extracted saliva samples showing absence of interfering peaks. [a].

Female sample containing 9 pmol/L testosterone, [b]. Female sample containing 30 pmol/L

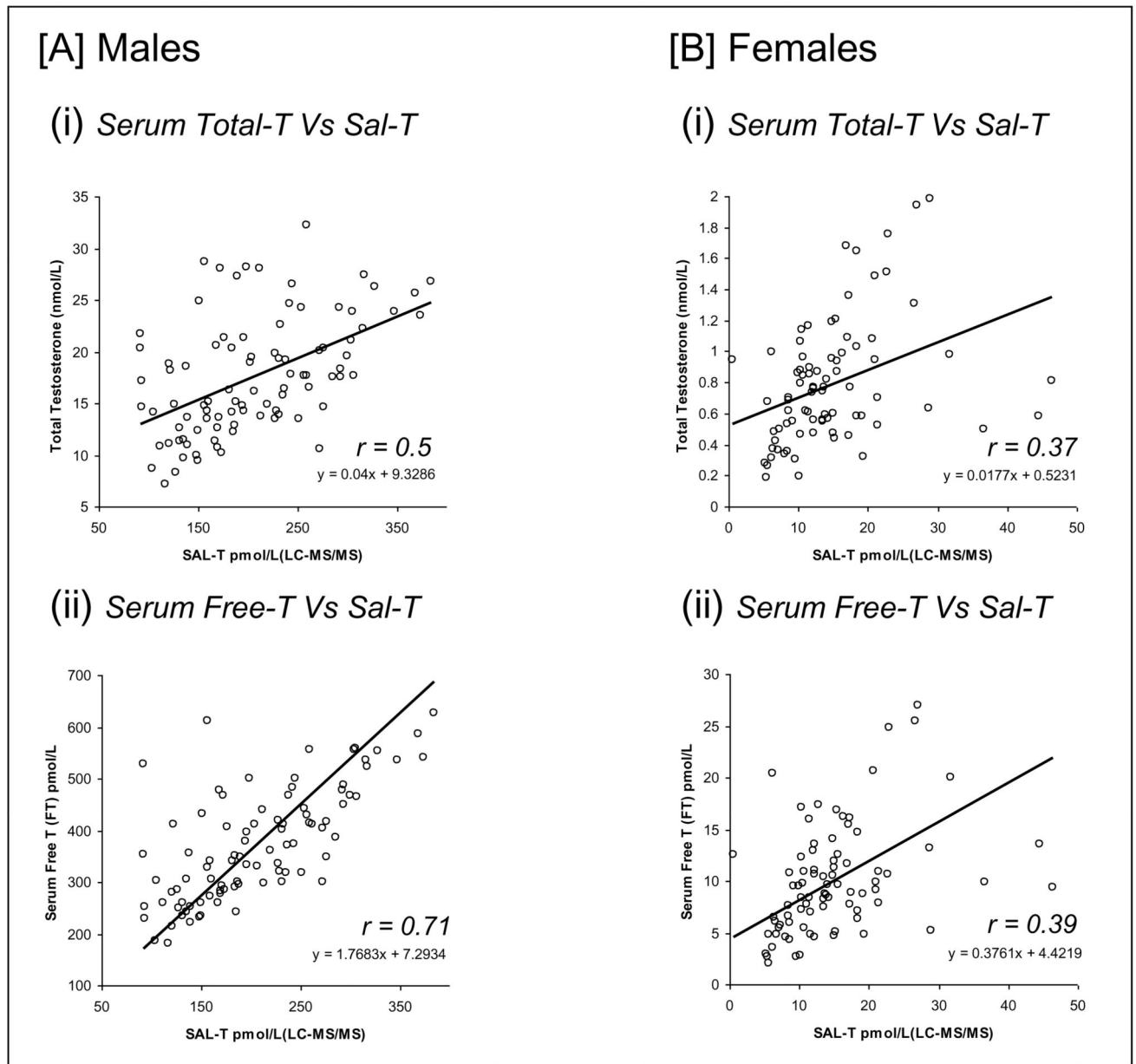
testosterone. [c]. Male sample containing 100 pmol/L



**Figure 2.** Stability of salivary testosterone at [a] room temperature, [b] 4°C and [c] after multiple freeze thaw cycles.

**Figure 3.**

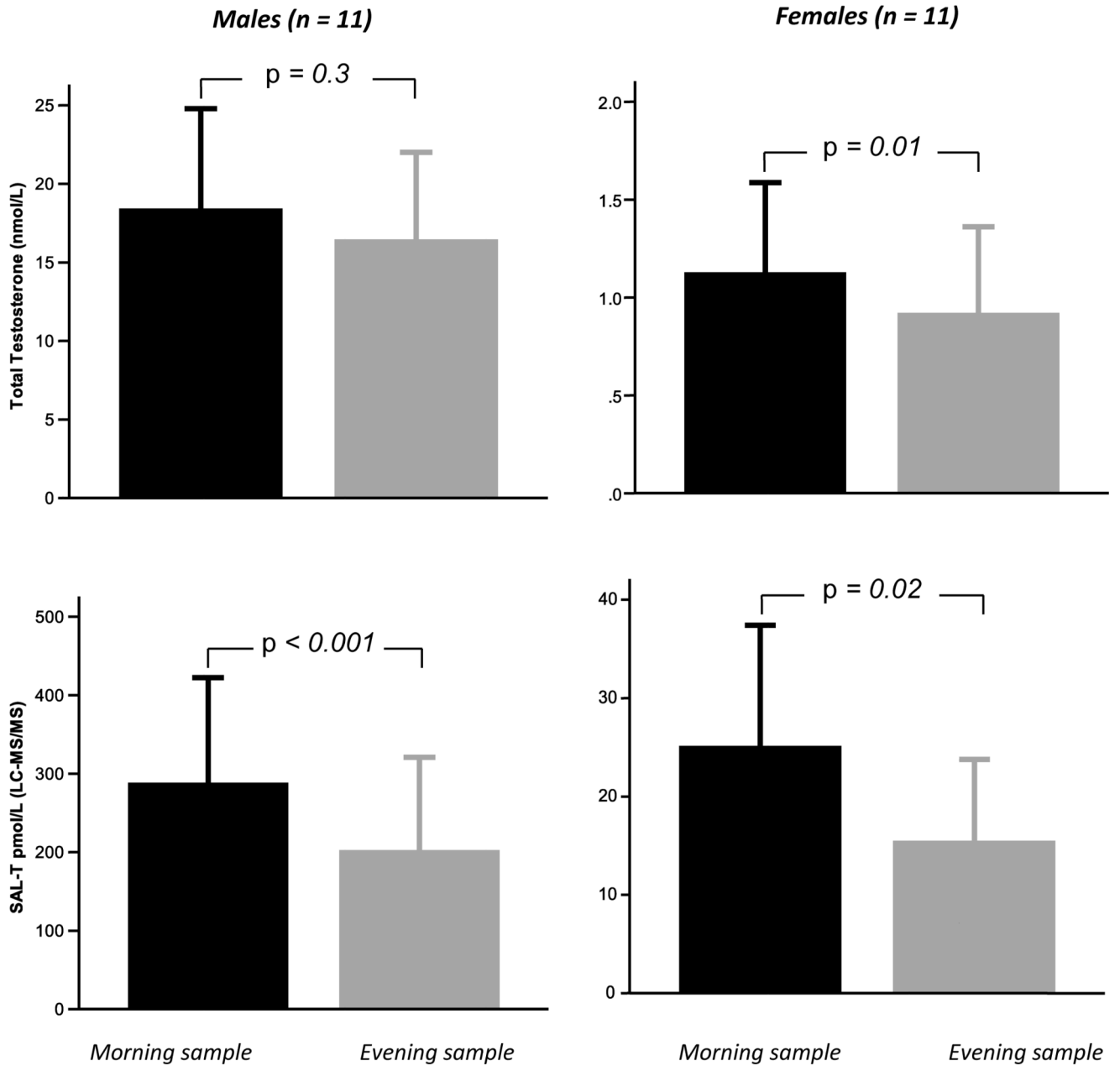
**a.** 95% reference interval for male salivary testosterone measured by LC-MS/MS (n=104). The reference interval was 93 - 378 pmol/L (2.5<sup>th</sup> to 97.5<sup>th</sup> percentile). **b.** 95% reference interval for female salivary testosterone measured by LC-MS/MS (n=84). The reference interval was 5 - 46 pmol/L (2.5<sup>th</sup> to 97.5<sup>th</sup> percentile)



**Figure 4.**

Comparison of serum total-T (i) and serum free-T (ii) with salivary testosterone on both males (n=93) [a] and females (n=86) [b].

Note: Solid line = linear regression fit, dashed lines = 95% confidence intervals



**Figure 5.**

Comparison of serum total-T and Sal-T from samples collected in the morning (before 10.00h) and evening (after 18.00h). Note: Bars represent mean values and whiskers one standard deviation, P values refer to paired *t*-test.



**Table 1**

Results of serum T, serum calculated free-T and Sal-T measured by LC-MS/MS in male and female participants.

Males (mean age 38.7 years, range 17.3 - 66.6 years)											
Hormone	n	Mean	SD	Median	IQR	Percentiles				Range	
						2.5%	5%	95%	97.5%	Min	Max
Serum T (nmol/L)	94	18.0	7.5	16.9	7.8	8.4	9.7	28.1	30.9	7.2	65.7
Serum free T (pmol/L)	93	387.3	212.4	350.0	173.1	190.6	219.2	566.4	621.4	181.0	2128.0
Sal T (pmol/L)	104	221.2	150.7	200.5	103.3	92.6	104.3	343.0	377.8	92.0	1577.0
SHBG (nmol/L)	93	29.9	11.1	28.0	13.5	12.0	14.7	52.3	53.7	11.0	72.0
Females (mean age 39.7 years, range 16.0 - 63.0 years)											
Hormone	n	Mean	SD	Median	IQR	Percentiles				Range	
						2.5%	5%	95%	97.5%	Min	Max
Serum T (nmol/L)	91	0.8	0.5	0.7	0.5	0.2	0.3	1.7	2.0	0.2	3.3
Serum free T (pmol/L)	90	10.0	6.0	8.8	6.5	2.7	2.9	22.5	26.6	2.1	38.0
Sal T (pmol/L)	86	15.9	12.3	13.5	8.5	5.3	5.6	34.9	46.1	<5.0	99.8
SHBG (nmol/L)	91	66.2	54.9	50.0	29.0	20.2	24.6	170.6	285.4	19.0	360.0

**Table 2**

Results of serum-T, serum calculated free-T and Sal-T measured by LC-MS/MS in female participants excluding those on oral contraceptive or hormone replacement therapy (n=6).

Hormone	n	Mean	SD	Median	IQR	Percentiles				Range	
						2.5%	5%	95%	97.5%	Min	Max
Serum T (nmol/L)	85	0.8	0.5	0.7	0.5	0.2	0.3	1.7	2.0	0.2	3.3
Serum free T (pmol/L)	84	10.3	6.0	9.2	6.3	2.7	3.8	23.8	26.8	2.6	38.0
Salivary T (LC-MS) (pmol/L)	84	16.0	12.1	13.7	8.3	5.5	6.2	31.6	46.2	<5.0	99.8
SHBG (nmol/L)	85	61.6	48.4	50.00	26.5	23.2	25.0	127.7	235.3	19.0	360.0