summary and general discussion

based on ‘Testosterone assays: fitness for purpose’
Hong N. Bui, Marinus A. Blankenstein, Annemieke C. Heijboer

Annals of Clinical Biochemistry, provisionally accepted
During the last decade, the quality of commonly used serum total testosterone assays has been frequently questioned [1;2]. Many of these assays have been used despite their poor performance, in the absence of viable alternatives. Data obtained from these assays may lead to faulty conclusions. The appeal of the Endocrine Society for better assays has created more awareness of this problem [3] and was taken to heart by endorsing clinical organizations, clinical laboratories performing these assays, and testosterone assay manufacturers.

Serum total testosterone analysis
Since we wholeheartedly agreed with the appeal of the Endocrine Society, we developed a sensitive and accurate serum total testosterone method using ID-LC-MS/MS (Chapter 2). In the analytical validation, our method showed a comparable performance to a registered reference method gas chromatography-mass spectrometry (ID-GC-MS). Moreover, we performed a clinical validation and established that our ID-LC-MS/MS method was able to distinguish testosterone concentrations between women in whom menopause was induced by removal of the ovaries and women who went into menopause naturally. Because the choice of internal standard is important, we tested two isotopically labeled internal standards in the validation process and found that these internal standards show equal performance. After this thorough analytical and clinical validation we concluded that we developed a state-of-the-art method for the reliable assessment of testosterone.

The appeal of the Endocrine Society also led to the development and marketing of 2\textsuperscript{nd} generation assays by two manufacturers that show improved performance compared to their predecessors. The accuracy of the first and second generation immunoassays was investigated in Chapter 3. We used our ID-LC-MS/MS method to test the commercially available testosterone immunoassays that are most commonly used in the Netherlands, anno 2012. We found the performance of these assays in the normal male range to be acceptable. However, at low testosterone concentrations, typically found in women, there is a considerable range in quality among different immunoassays. In the immunoassay’s procedure there is a step incorporated to release testosterone from its endogenous binding protein, but its execution may not always be fully effective. We therefore also studied the effect of a manual liquid extraction before analysis. We concluded that manual sample extraction may boost the performance of most of the immunoassays, but may also lead to a less accurate calibration.

Serum total testosterone in clinical practice
Diagnoses of testosterone related-pathologies of patients with low testosterone concentrations, i.e. women, hypogonadal men, and children, will benefit most from the improved testosterone assays and we now have the tools to re-evaluate testosterone testing in these populations. Because of conflicting data in literature, we also
investigated the daily dynamics of testosterone parameters across the menstrual cycle in Chapter 4. Our data indicates that even though testosterone concentrations were statistically significantly elevated around ovulation, the biologic variation in individual women is independent on the menstrual cycle and of such a magnitude, that there is no need to test testosterone in a particular phase of the menstrual cycle and the reference ranges can be applied irrespective of the day of sampling. Based on the intra-individual variation we do recommend, however, that assessment of the androgen status of women is based on multiple measurements. With these data, we also re-evaluated the reference ranges for testosterone in women.

In continuation of Chapter 4, we investigated testosterone derived parameters, i.e. free testosterone and free androgen index (FAI), and the diagnostic value in polycystic ovary syndrome (PCOS) (Chapter 5). We found that the parameters that take sex hormone-binding globulin into account have a better discriminative power between healthy and hyperandrogenic women with PCOS. In addition, we conducted both studies (Chapters 4 and 5) also by a 2nd generation testosterone immunoassay. Compared to our ID-LC-MS/MS, this immunoassay showed a persistent positive bias that seems to be related to the SHBG levels. Nevertheless, there was a close analytical and clinical agreement between this immunoassay and ID-LC-MS/MS. Based this data, we were able to estimate the biological variation of testosterone in women and determine performance criteria for testosterone assays.

In hypogonadal men, either due different pathological mechanism or induced through androgen deprivation therapy in prostate cancer, testosterone concentrations are low. Especially in androgen deprivation therapy, testosterone is lowered to castration levels, i.e. < 1.7 nmol/L, and monitoring of therapy is necessary. Thus far, urologists do not request testosterone measurements consequently because this was of limited use due to poor testosterone assays [4]. With the ID-LC-MS/MS, we evaluated testosterone concentrations in patients using luteinizing hormone-releasing hormone agonist-therapy (LHRH agonist) and surgically castrated patients (Chapter 6). Remarkably, medically castrated men had lower testosterone levels than men who had their prostate removed. The clinical relevance remains unclear and should be further investigated. Also, the concentrations that we reported raise the discussion whether the cut-off value for testosterone concentrations in castrated men is appropriate or should be adjusted downwards.

The third group of patients in which an accurate serum testosterone method is required are children, this was not investigated as part of this thesis. Children are a vulnerable group and research is limited due to ethical considerations. A sample collected by a non-invasive acquisition procedure as an alternative for serum could take away these limitations. Saliva collection, for instance, is practical and essentially non-invasive. We therefore developed an ID-LC-MS/MS method to measure testosterone in saliva (Chapter 7). Apart from setting salivary testosterone reference ranges in men, we
The method for the assessment of salivary testosterone is also applicable in children. The extreme concentrations, i.e. supra-physiological levels after injection and subsequently relatively low levels towards the end of the inter-injection period, may affect the wellbeing of the patient. We anticipate that the method for the assessment of salivary testosterone is also applicable in children.

**GENERAL DISCUSSION**

**Serum total testosterone analysis**

The ID-GC-MS is considered the gold standard for serum total testosterone measurement provided it is executed properly; however, the cumbersome procedure renders this technique unsuitable for routine use. ID-LC-MS/MS is theoretically less labor intensive. In light of the still compromised performance of immunoassays, does this mean that all clinical laboratories should switch to ID-LC-MS/MS immediately? Until now, the introduction of LC-MS/MS into clinical routine laboratories is still limited. Several reports state that ID-LC-MS/MS is a highly accurate technique for testosterone measurements. However, its accuracy strongly depends on the instrument in combination with the procedure used. If extensive sample preparation is needed in order to achieve this high performance, throughput is usually compromised to such an extent that the method is not suitable for routine use. Meanwhile, one should keep in mind that the implementation into routine practice requires financial investment and expertise, both of which aspects could be problematic, especially in small laboratories without a back-up instrument. Even though the new generation MS/MS instruments are more stable and user-friendly, downtimes of a week for troubleshooting and maintenance are not exceptional.

A great advantage of ID-LC-MS/MS is its potential to measure multiple steroids within one run. Testosterone-related disorders in newborns, i.e. ambiguous genitalia and congenital adrenal hyperplasia (CAH), are generally disorders in which the concentration of multiple steroids deviates from normal ranges. In addition, the volume of blood that can be obtained is small. Therefore, an ID-LC-MS/MS method that can determine multiple steroids in one run for routine use is particularly useful in this population [5]. The current trend towards more sensitive state of the art instruments could enable measurement of a steroid profile in saliva in the future. This could benefit these patients, however, reference values in saliva need to be established.

In summary, no fast and easy total testosterone measurement in serum with high accuracy, precision, and sensitivity is available today. It is however critical to select an assay that has sufficient discriminative power to answer the clinical question at hand, or, in other words, 'fit for purpose'. The performance of an assay relies on the whole...