Guidelines Review

Challenges to the Measurement of Oestradiol: Comments on an Endocrine Society Position Statement

Julie D Newman[1] and *David J Handelsman[2]
[1]Southern Cross Pathology Australia, Monash Medical Centre, Clayton Vic. 3168,
[2]ANZAC Research Institute, University of Sydney, Concord Hospital NSW 2139, Australia.
*For correspondence: Professor David Handelsman, djh@anzac.edu.au

Measurement of oestradiol (E_2) is increasingly important in clinical management for both diagnosis and monitoring of treatment for reproductive disorders as well as for research into hormone-dependent diseases. As the sole bioactive oestrogen in humans, E_2 is responsible for not only female reproductive development, function and treatment for fertility disorders, but also has increasingly recognised and wide-ranging effects on non-reproductive tissues including the pathogenesis of common disorders such as cardiovascular disease, hormone-dependent cancers and osteoporotic fracture. Recent research indicates that there are likely to be more clinical diagnostic and management insights fruitfully developed in the near future once improved E_2 assays are more widely available.

The US Endocrine Society recently published a position statement reviewing the current state of clinical E_2 assays highlighting their limitations, applications and prospects for future improvement.[1] We review and comment on the paper which is divided into reviewing methods of E_2 measurement and then applications to clinical practice, translational and epidemiological research. Finally, the position statement describes recommendations on standardising E_2 assays aiming to improve the presently available methods while awaiting the availability of the new generation of routine mass spectrometry(MS)-based steroid assays.

E_2 Assay methods

Oestrogen assays evolved during the 20th century from the whole animal bioassays that measure uterotrophic or vaginal cornification effects of any oestrogen (or pro-oestrogen). While laborious and costly, these oestrogen bioassays were crucial to Doisy's successful purification and identification of E_2 as the bioactive human oestrogen in the mid-1930s.[2] In the middle of the century, chemical methods to measure urinary oestrogen metabolites were developed.[3] Subsequently, oestrogen bioassays were refined to distinguish oestrogen agonist from antagonist effects as well as to eliminate animal use by developing oestrogen receptor(ER)-based assays, such as binding and, most recently, *in vitro* ER reporter gene bioassays.

Although these oestrogen bioassays have a continuing role in toxicology and pharmacological drug development to estimate oestrogenic bioactivity without regard to chemical structure,[4] they cannot measure the sole bioactive oestrogen, E_2, specifically. That became possible only with the advent of MS[5] and immunoassay[6] over 4–5 decades ago. The first generation of steroid immunoassays developed in the 1970s featured the essential triplet of validity criteria for steroid immunoassays, namely solvent extraction, chromatography and structurally authentic tracers.[7] An inherent limitation of E_2 immunoassays is their reliance on antibodies developed to the non-immunogenic steroid only after conjugation via bridge compound allowing for covalent binding to a larger immunogenic protein. This renders the E_2 antibodies 'blind' to the conjugation site of the steroid, ultimately limiting specificity of steroid immunoassays using that antibody, a defect most prominent at low steroid concentrations. Nevertheless, with the validity criteria satisfied, the original in-house E_2 immunoassays were of inestimable value in revealing much reproductive pathophysiology. This led to steeply growing demand for blood E_2 measurement in clinical and laboratory practice which, in turn, forced assay simplification in order to fit steroids into multiplex immunoassay platforms. Converting conventional immunoassays to fit into multiplex immunoassay platforms required abandoning extraction and chromatography as well as employing bulky non-authentic steroid tracers with convenient non-radioactive read-out signals. Although a few well-validated conventional E_2 immunoassays still exist, they remain too laborious to meet the high throughput demanded. In practice, only 'direct' (i.e. non-extraction) E_2 immunoassays are available in Australian clinical pathology laboratories. In recent years, although not a new finding, their limitations flowing from not satisfying the validity criteria for steroid immunoassays have been highlighted.

Isotope dilution gas chromatography-MS was the original steroid reference method for specificity and represented the gold standard method for E_2 measurement. However, this method was insensitive (relative to E_2 immunoassays) and laborious and required expensive equipment not widely available in chemical pathology laboratories, rendering it unsuited to routine, high-throughput clinical assays. Recent advances in technology coupling (ultra) HPLC with tandem, triple quadrupole MS (LC-MS/MS) in bench-top format now offer improved sensitivity, matching or exceeding the best steroid immunoassays, together with shorter run-times to allow for faster throughput while retaining reference levels specificity. While LC-MS/MS equipment remains expensive, this new technology offers crucial new biological insight into physiology and pathogenesis in the majority of clinical situations where E_2 concentrations are too low for valid measurement by direct E_2 immunoassays. As this limitation concerns children, adolescence during puberty, men at any age and women after menopause, in effect direct E_2 immunoassays are useful to provide reliable results for only eugonadal premenopausal women or patients with E_2-secreting tumours. The review[1] comments that, based on the ER's affinity for E_2, there are likely to be important biological E_2 effects manifest at sub-pg/mL circulating levels, which are beyond the limitations of E_2 immunoassays or even the still improving sensitivity of LC-MS/MS assays. Like all assays, LC-MS/MS steroid measurements require standardisation to minimise between-assay variability due to pre-assay factors such as sample handling and storage, calibration accuracy of secondary standards and traceability to certified reference materials, accurate analytical corrections for recovery and ion suppression as well as freedom from interference by structurally related chemicals. In addition, it has remained difficult to achieve high levels of between-laboratory accuracy and precision especially at low concentrations with all E_2 assays.

Blood E_2 assays must be sensitive, specific, accurate and precise over an unusually wide concentration range for a bioactive analyte. This ranges over three orders of magnitude from very high circulating E_2 concentrations (typically >1000 pg/mL or 3500 pmol/L) in women undergoing IVF hyperstimulation, ovulation induction, in pregnancy or bearing E_2-secreting tumours. At the other extreme, very low circulating E_2 concentrations in children and aromatase inhibitor-treated women (typically <5 pg/mL or 20 pmol/L) are mostly below the typical limits of quantification of immunoassays (30–100 pg/mL or 100–350 pmol/L) or LC-MS/MS (3–5 pg/mL or 10–15 pmol/L). For most patients however, including adolescents, men and postmenopausal women, circulating E_2 concentrations are often, but not always, undetectable or, at best, unreliable by immunoassays.[8]

In addition, E_2 is converted to conjugated and unconjugated metabolites while patients' serum may contain exogenous oestrogens, all of which may significantly interfere with E_2 assays, especially with methods that do not use extraction.

Measurement of E_2 in Clinical Practice, Translational and Epidemiological Research

The present clinical indications for E_2 measurement are severely limited, restricted mostly by the inadequacies of current direct E_2 immunoassays which are ubiquitous in clinical pathology laboratories. Potentially many more clinical applications of measurement of serum E_2 are likely with more sensitive and specific MS-based E_2 assays. An excellent contemporary review of E_2 methodology relative to current and future potential clinical applications of E_2 assays is available.[9]

Children and Puberty

Serum E_2 increases markedly during female puberty and measurement is likely to be useful for tracking progression of puberty. However, current direct E_2 immunoassays are unable to serve this purpose as prepubertal serum E_2 levels are undetectable and are only reliably quantifiable in late or completed female puberty when monitoring to guide management of pubertal delay is superfluous. Applications of serum E_2 measurement to male puberty, where circulating concentrations are lower, remain speculative without data from more sensitive E_2 assays.[8] Yet, while LC-MS/MS E_2 assays are at least an order of magnitude more sensitive, it remains to be determined whether they can quantify serum E_2 in all prepubertal children as ideally required to monitor early pubertal progression when delay is of most clinical relevance.

Adult Women

Most but not all current direct E_2 immunoassays can detect serum E_2 throughout the normal menstrual cycle. However, they lack sufficient sensitivity to define reference intervals across the menstrual cycle which is a prerequisite to support diagnosis and management of E_2 deficiency. By contrast, and contrary to an erroneous statement in the review, MS-based E_2 assays show lower circulating levels than direct E_2 immunoassays,[8] presumably due to superior specificity in avoiding interference from cross-reacting oestrogen metabolites.

In assisted reproductive technology services, serum E_2 assays are widely used for both ovulation induction and IVF hyperstimulation in conjunction with ultrasound monitoring. In these clinical situations, fast turnaround time is far more important than accuracy to determine within hours from blood collection whether to abort cycles (to avoid hyperstimulation in non-IVF ovulation induction) and/or when to administer the hCG ovulation trigger. Together with the extremely wide

working range (three orders of magnitude) required for a serum E_2 assay in these settings, these requirements are inimical to satisfying the high quality assay performance expected of serum E_2 measurement in other clinical practice and research. In such situations, direct E_2 immunoassays, or even simplified point-of-care tests derived from them, have major advantages over the slower turnaround times required for accurate E_2 assays.

In pregnancy, another setting of very high circulating E_2 concentrations but without requiring as fast turnaround as IVF and ovulation induction, whether serum E_2 monitoring has any role in identifying risks of pregnancy-related disorders (e.g. pre-eclampsia, hydatidiform mole) remains to be determined but the superior specificity of MS-based E_2 assays would be required to critically evaluate these opportunities given the wider range of cross-reacting steroids in pregnancy serum.

In postmenopausal women, the present commercial assays are unable to accurately measure serum E_2 as the prevailing concentrations are well below the limits of quantification of all direct E_2 immunoassays. Even the most sensitive LC-MS/MS assays are not yet able to measure serum E_2 in all older women, although MS-based assays are yet to be fully optimised for sensitivity. Clinicians would benefit from reliable serum E_2 assays which could reliably measure E_2 levels in postmenopausal women before and during E_2 replacement therapy in order to optimally titrate the lowest effective E_2 dose required to control symptoms. However, to date, only extraction-based E_2 assays (in-house conventional E_2 immunoassays or MS-based E_2 assays) can measure even the higher (but still low) circulating E_2 levels of E_2 treated women.

Other clinical circumstances where serum E_2 measurement is, or could be, useful in monitoring of treatment include (a) women with breast cancer treated with aromatase inhibitors where serum E_2 measurement is a measure of therapeutic efficacy; and (b) women undergoing medical castration for treatment of benign hormone-sensitive disorders (e.g. endometriosis, leiomyoma) where add-back E_2 dosage could be titrated against serum E_2 concentrations. However direct E_2 immunoassays cannot fulfil this role due to their insensitivity and inaccuracy at low serum E_2 concentrations and only MS-based assays are recommended in these situations.

Adult Men

Direct E_2 immunoassays are unable to measure the serum E_2 in men as all commercial assays lack specificity and usually sensitivity as well.[8] As usual, the problems of low sensitivity are compounded by non-specificity at the low prevailing ranges of circulating E_2.

As testosterone action is at least partially dependent on aromatisation, especially in the brain and bone,[10] sensitive and specific E_2 assays could provide a useful biomarker to predict risk of osteoporosis and fracture in men. However, the claim that there exists a threshold of circulating E_2 for maintenance of bone density remains to be well established by sufficiently sensitive and specific E_2 assays. Aromatisation of testosterone in bone may operate as a local rather than systemic regulatory mechanism, as it does in the brain.[11] Other androgen sensitive tissues such as muscle may have so far unrecognised dependence on aromatisation and oestrogen action, noting that muscle growth also occurs during female puberty without the androgen surge of male puberty and muscle lost after menopause is recovered by oestrogen replacement.[12]

A potential clinical role for measuring serum E_2 in diagnosis or monitoring treatment of men with hypogonadism remains to be established as existing direct E_2 immunoassays are neither sensitive nor accurate enough to fulfil these roles.[8] The utility of MS-based E_2 assays for men with hypogonadism remains to be determined when suitable assays are available.

In evaluation of gynaecomastia, it is highly likely that the nullifying effects of inaccurate serum E_2 measurements may have overlooked a hormonal basis for many cases of gynaecomastia.[13] While gynaecomastia is widely understood to reflect an androgen:oestrogen imbalance, authoritative reviews of clinical management of gynaecomastia do not recommend serum E_2 measurement.[14,15] It is likely the inability to measure E_2 accurately has hampered a better understanding of its pathogenesis as well as missing opportunities for effective treatment of what now remains considered idiopathic gynaecomastia. Whether this is substantiated by improved serum E_2 assays remains to be determined. In the interim, measurement of serum E_2 by direct immunoassay in men with gynaecomastia is likely to continue giving unreliable information and cannot be recommended.

The role of serum E_2 measurement in diagnosis and monitoring of treatment of men with prostate cancer requires further evaluation. Among men who undergo medical castration for advanced prostate cancer, the role of E_2 treatment to prevent symptoms and improve the adverse bone and cardiovascular outcomes may also depend on accurate measurement of serum E_2 which requires more sensitive and specific assays than current direct E_2 immunoassays.

Epidemiological Research

Epidemiological research studies describe the distribution of E_2 concentrations in human populations to determine how E_2 may influence disease risk and patient survival.

Serum E_2 measurements are crucial for epidemiological studies not only where E_2 has a primary role in breast carcinogenesis but also in many other hormone-sensitive diseases such as endometrial cancer, cardiovascular disease, bone fracture and cognitive function and brain injury, where sex steroids can represent either biomarkers of disease or even targets for hormone-based interventions. Highly sensitive, specific and accurate E_2 assays are crucial for such studies. Inaccurate E_2 assays will tend to nullify genuine associations and militate against arriving at sound, consistent findings between studies or adoption of valid diagnostic or therapeutic clinical guidelines based on inconsistent E_2 assays. It is therefore essential to foster the use of improved E_2 assays to allow progress in understanding the origins of numerous hormone-sensitive diseases and their improved clinical management.

Summary of Recommendations of the Position Statement

The recommendation of the US Endocrine Society review may be summarised as:

- all measurement should be traceable to a universally recognised E_2 standard;
- reference intervals should be generated based on age, gender, puberty, and menstrual cycle and menopausal status recognising this will be time-consuming and expensive;
- awareness by clinicians and laboratory staff of the unreliability of serum E_2 from direct E_2 immunoassays in all but premenopausal women;
- improvement is required for routine E_2 assays to allow accurate and precise measurement of serum E_2 levels between 0.2 and 2 pg/mL (~1–10 pmol/L).

The review authors acknowledge these requirements are difficult to achieve and costly in the changeover to MS-based assays as well as developing universal accuracy-based standardisation and possibly even genuine improvement in immunoassay-based methods. Additional barriers to such progress include the need to convince (a) physicians to insist on accuracy-based measurements; (b) journals to insist on accuracy-based measurements for all publications; and (c) government and third-party payers that a higher cost assay providing correct information is preferable to a cheaper assay providing incorrect information.

Conclusions

There are many potential benefits for patients of all ages from the sensitive, specific and accurate E_2 assays; however, only a small fraction are currently utilised clinically due to the limitations of presently available E_2 assays. At present, the commercially available direct E_2 immunoassays are capable of measuring serum E_2 reliably only in healthy eugonadal premenopausal women. In all other clinical settings, the direct E_2 immunoassays are likely to provide spurious results and misguide management. While MS methods would provide technically superior performance in terms of sensitivity and specificity, there remains a need to standardise MS methods as for all other clinically applicable analytes.

Surprisingly, given the manifest failures of current commercial direct E_2 immunoassays for most indications, the review authors believe harmonisation will improve E_2 assay performance. History suggests that worthy efforts at agreed-upon standardisation based on accuracy-based performance will be slow if ever adopted without some more direct drivers. Regrettably, the availability of reference methods and certified reference standards for decades has not improved performance of E_2 immunoassays[8,9] making it unlikely that this will happen in any near future. The authors regard accuracy-based standardisation as an alternative to the wider adoption of MS-based steroid assays in a recommendation that seems sentimentally geared towards salvage of direct E_2 immunoassays, despite their manifestly being not fit for purpose. Such direct immunoassays can only be brought into alignment with MS-based reference methods by adopting undesirable 'fudge factor' calibrations. However, it is beyond doubt that the limitations of direct E_2 immunoassays render them unfit for purpose in most clinical applications – indeed for all but premenopausal women – and thereby overlooks many opportunities to improve medical diagnosis and care for reproductive and hormone-sensitive disorders.

The availability of MS-based assays is currently limited by access to the equipment with its high initial purchase and running costs due to the time and skilled labour requirement to operate LC-MS/MS. These costs can be offset against the freedom from proprietary reagent costs.

Standardisation of E_2 assays has barely begun despite the availability of reference methods and certified reference standards for decades. A similar process of assay standardisation has been underway for seven years for testosterone since another US Endocrine Society position paper regarding testosterone measurement.[16] This culminated in a subsequent standardisation project[17] with the objective of establishing a basis for traceability of testosterone assay on the basis of accuracy in terms of well-established certified reference preparations to facilitate the goal that testosterone results would be comparable across methods, laboratories, time and location.

It is timely that a recent editorial in *The Journal of Clinical Endocrinology and Metabolism (JCEM)* has announced that the journal has upgraded its submission requirements for publications of studies of sex steroid measurements.[7] From

1st January 2015 *JCEM* will allow only MS-based assays for reports using sex steroid assays as important endpoints, together with enough detail to allow results to be reproduced together with standard quality control, specificity and reproducibility. Requirements for MS-based assays will likely also be extended to adrenal steroids and Vitamin D in the future. This editorial is an important step forward for clinical research which will eventually impact on clinical practice via pathology laboratories.

In particular, developments in the last decade of MS-based methods may provide E_2 assays with greater sensitivity, specificity and accuracy, albeit at a higher cost and assay time than direct immunoassay. Provision of a universal traceable E_2 standard for all E_2 methods will aid in accurate measurement especially of low E_2 concentrations in children, women and men. It will also allow clinicians to make valid clinical decisions if patient results are obtained using different methods at different laboratories. Development of age- and gender-specific as well as biologically specific reference intervals will also be important. Improved E_2 assays will also introduce more accurate data into epidemiological studies which will ultimately benefit clinical decision making. Hopefully, these important endeavours will improve the standard of care of patients with reproductive or endocrine disorders, hormone-dependent cancers and osteoporosis in the near future.

Competing Interests: None declared.

References

1. Rosner W, Hankinson SE, Sluss PM, Vesper HW, Wierman ME. Challenges to the measurement of estradiol: an endocrine society position statement. J Clin Endocrinol Metab 2013;98:1376-87.
2. MacCorquodale DW, Thayer SA, Doisy EA. The isolation of the principal estrogenic substance of liquor folliculi. J Biol Chem 1936;115:435-48.
3. Brown JB. A chemical method for the determination of oestriol, oestrone and oestradiol in human urine. Biochem J 1955;60:185-93.
4. Owens JW, Ashby J; Organisation for Economic Co-operation and Development. Critical review and evaluation of the uterotrophic bioassay for the identification of possible estrogen agonists and antagonists: in support of the validation of the OECD uterotrophic protocols for the laboratory rodent. Crit Rev Toxicol 2002;32:445-520.
5. Shackleton C. Clinical steroid mass spectrometry: a 45-year history culminating in HPLC-MS/MS becoming an essential tool for patient diagnosis. J Steroid Biochem Mol Biol 2010;121:481-90.
6. Abraham GE. Solid-phase radioimmunoassay of estradiol-17 beta. J Clin Endocrinol Metab 1969;29:866-70.
7. Handelsman DJ, Wartofsky L. Requirement for mass spectrometry sex steroid assays in the Journal of Clinical Endocrinology and Metabolism. J Clin Endocrinol Metab 2013;98:3971-3.
8. Handelsman DJ, Newman JD, Jimenez M, McLachlan R, Sartorius G, Jones GR. Performance of direct estradiol immunoassays with human male serum samples. Clin Chem 2014;60:510-7.
9. Middle JG, Kane JW. Oestradiol assays: fitness for purpose? Ann Clin Biochem 2009;46:441-56.
10. Finkelstein JS, Lee H, Burnett-Bowie SA, Pallais JC, Yu EW, Borges LF, et al. Gonadal steroids and body composition, strength, and sexual function in men. N Engl J Med 2013;369:1011-22.
11. Handelsman DJ. Mechanisms of action of testosterone—unraveling a Gordian knot. N Engl J Med 2013;369:1058-9.
12. Greising SM, Baltgalvis KA, Lowe DA, Warren GL. Hormone therapy and skeletal muscle strength: a meta-analysis. J Gerontol A Biol Sci Med Sci 2009;64:1071-81.
13. Stratakis CA. An aroma of complexity: how the unique genetics of aromatase (CYP19A1) explain diverse phenotypes from hens and hyenas to human gynecomastia, and testicular and other tumors. J Clin Endocrinol Metab 2013;98:4676-81.
14. Braunstein GD. Clinical practice. Gynecomastia. N Engl J Med 2007;357:1229-37.
15. Ma NS, Geffner ME. Gynecomastia in prepubertal and pubertal men. Curr Opin Pediatr 2008;20:465-70.
16. 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.
17. Vesper HW, Botelho JC. Standardization of testosterone measurements in humans. J Steroid Biochem Mol Biol 2010;121:513-9.