Evaluation of serum and salivary cortisol responses to Synacthen stimulation

Dr Nadia El-Farhan

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Neurosciences & Mental Health Research Institute

School of Medicine, Cardiff University

December 2016

DECLARATION

This work has not been submitted in substance for any other degree or award at this or any other university or place of learning, nor is being submitted concurrently in candidature for any degree or other award.

Signed …… (candidate)

Date ……..……..……

STATEMENT 1

This thesis is being submitted in partial fulfilment of the requirements for the degree of PhD.

Signed ……..…… (candidate)

Date ……..……

STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated, and the thesis has not been edited by a third party beyond what is permitted by Cardiff University's Policy on the Use of Third Party Editors by Research Degree Students. Other sources are acknowledged by explicit references. The views expressed are my own.

Signed ……..…… (candidate)

Date ……..……

STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available online in the University's Open Access repository and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed ……..…….. (candidate)

Date ……

ACKNOWLEDGEMENTS

I would like to thank my supervisors, Dr Aled Rees and Dr Carol Evans for their support and encouragement throughout the past 6 years, for their quick return of corrections, their valuable pointers and for their seemingly endless positivity!

Laboratory analyses:

I am particularly grateful to Mr Alan Pickett who dealt with sample processing and storage within the lab, co-ordinated sample analysis by external labs, performed much of the local serum cortisol analysis and taught me how to run the CBG RIA! The laboratory aspect of this work would have been overwhelming without his help! I would also like to thank Dr Atilla Turkes for performing CBG and salivary cortisol analyses and Mrs Sarah Tennant for completing the salivary cortisol analysis following his retirement.

I am also grateful to Mr Dave Ducroq and his team at the WEQAS Reference Laboratory for performing the GC-MS cortisol analyses and teaching me how it's done and to Dr Alan Dodd for his help with serum cortisol analyses at UHW and WEQAS. I would also like to thank Dr Cath Bailey, Dr Kelly Mitchem, Dr Annie Armston, Dr Ann Bowron and Ms Nicky Morgan for their help coordinating and running serum cortisol assays at the participating external laboratories.

Patient recruitment and testing:

I would like to thank Ms Laila Jones for her help in recruiting patients and performing Synacthen tests and also Sister Nikki Davies and Sister Janet Lewis for their help with Synacthen tests. I would also like to thank Dr Andrew Lansdowne who helped with patient recruitment and Drs Stephen Riley and Lawrence Sunder Raj for allowing access to their renal and liver clinic patients, respectively. My thanks too, to Dr Kate Craig and the staff in the Clinical Research Facility, UHW for running such an efficient facility and for their help with patient care.

I am also deeply indebted to the many patients and volunteers who took part in this study, without whom it would not have been possible and to my many colleagues in the Biochemistry departments at UHW and RGH who have allowed me the space and time to get this done one piece at a time! In particular, I would like to thank Dr Cath Bailey who has, at times, single-handedly run the department, allowing me to "just get on with it"!

Finally, I would like to thank my parents for their constant love and support, my brother, Ramzi, for making sure it all came together at the end, my children, Haidar, Zaid, Muna and Hadi for their patience and for looking after each other when I wasn't there – with special thanks to Zaid, my Endnote adviser, and Muna, my lab buddy – and, of course, to Martin, whose quiet presence makes everything possible! I love you all!

SUMMARY

Several studies have demonstrated the assay-dependence of cortisol measurement and the potential impact this could have on the interpretation of the Synacthen test. Prompted by clinical concern that hypoadrenalism was being over-diagnosed, this work set out to determine assay-specific serum cortisol and salivary cortisol responses to Synacthen stimulation in healthy volunteers.

It begins with an overview of cortisol assays in serum and saliva from their inception 50 years ago to the present day; highlighting the limitations of serum total cortisol as an analyte and immunoassays for its measurement and exploring salivary free cortisol, calculated free cortisol and mass spectrometry as alternatives.

The study itself recruited a total of 206 healthy volunteers and patients to undergo serum and salivary Synacthen tests. Serum cortisol was measured by gas chromatography-mass spectrometry (GC-MS) and by 5 contemporary immunoassays, salivary cortisol was measured by liquid chromatography-tandem mass spectrometry (LC-MS/MS) and serum free cortisol was calculated using the Coolens' equation and the free cortisol index.

The results in serum confirmed the assay-dependence of immunoassay cortisol measurement, established valid reference ranges for each of the assays studied, proposed diagnostic cut-off concentrations and evaluated their performance in patients with suspected hypoadrenalism.

The salivary results were used to establish adult reference ranges specific to the in-house LC-MS/MS assay used and subsequently demonstrated comparable performance in patients with suspected hypoadrenalism to the serum Synacthen test.

Calculated serum free cortisol performed poorly compared to serum total and salivary cortisol and was subject to error because of difficulty measuring cortisol binding globulin (CBG).

Comparisons between the three measures in patients with altered serum protein concentrations showed that, of the three, only salivary cortisol had the potential to accurately diagnose hypoadrenalism using a single cut-off; although confirmation of this finding is needed in a larger patient group.

ABBREVIATIONS

11β-HSD2	11β Hydroxysteroid dehydrogenase
8-ANS	8-anilinonapthalene sulfonic acid
ACTH	Adrenocorticotrophin
ALTM	All laboratory trimmed mean
AVP	Arginine vasopressin
BSA	Bovine serum albumin
CBG	Cortisol binding globulin
CFC	Calculated free cortisol (SFC & FCI)
CRH	Corticotrophin releasing hormone
CV	Coefficient of variation
EQA	External quality assurance
FCI	Free cortisol index
GC-MS	Gas chromatography-mass spectrometry
HPA	Hypothalamic-pituitary-adrenal
HPLC	High pressure liquid chromatography
ID-GC/MS	Isotope dilution gas chromatography-mass spectrometry
iQC	Internal quality control
ITT	Insulin tolerance test
LC-MS/MS	Liquid chromatography-tandem mass spectrometry
LRL	Lower reference limit
OCP	Oestrogen-containing oral contraceptive pill
RIA	Radioimmunoassay
SFC	Calculated serum free cortisol (Coolens)
SST	Short Synacthen® test
UKNEQAS	United Kingdom National External Quality Assessment Scheme
WEQAS	Welsh External Quality Assessment Scheme

CHAPTER 1

INTRODUCTION

1.1 Cortisol and the hypothalamic-pituitary-adrenal axis

Cortisol is a steroid hormone produced by the zona fasciculata of the adrenal cortex. Its secretion is directly controlled by the release of adrenocorticotrophin (ACTH) from the anterior pituitary which, in turn, is regulated by corticotrophin-releasing hormone (CRH), and less directly, arginine vasopressin (AVP), secreted by the hypothalamus. Cortisol exerts negative feedback on both ACTH and CRH secretion whilst stress acts directly on the hypothalamus to stimulate CRH release (Figure 1.1).

Cortisol is lipophillic and is transported around the body bound to cortisol-binding globulin (CBG) and albumin (Mills 1962). Approximately 80% of total cortisol is bound to CBG, 10% is albumin-bound and a further 10% exists in a free, unbound state. It is this free fraction that is biologically active and mediates cortisol's metabolic effects.

Cortisol plays an essential role in maintaining health and well-being. It is involved in energy metabolism, with effects on the synthesis and breakdown of carbohydrates, protein and lipids, water and electrolyte balance, blood pressure control and glomerular filtration rate (Demers 2005; Stewart 2011). It plays a role in inflammation and in protecting against infection and maintaining immunity. There is evidence that cortisol has some central effects, as suggested by the emotional disturbances seen in both hypo- and hyper-adrenal states. Cortisol also plays a role in maintaining healthy pregnancy and in the initiation of birth (Kamel 2010).

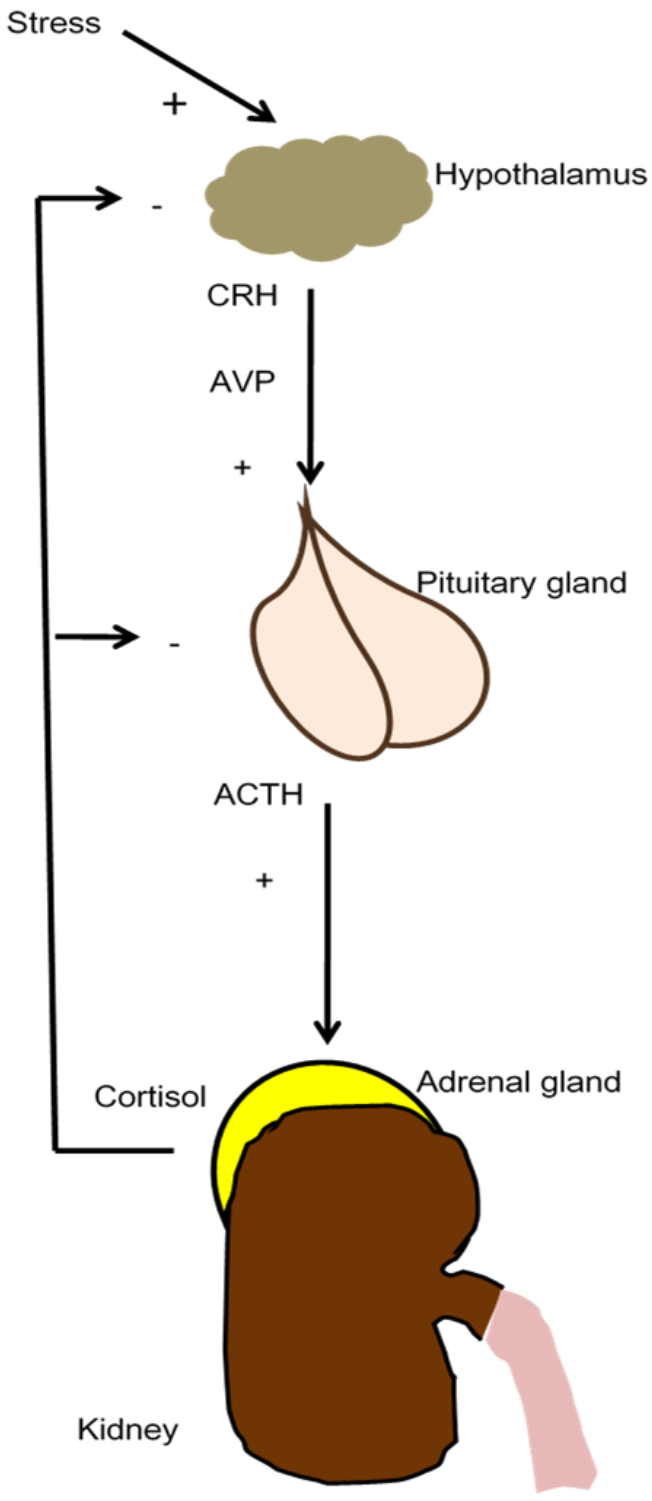


Figure 1.1 Regulation of the hypothalamic-pituitary-adrenal axis.

1.2 Measuring cortisol

Because of cortisol's central metabolic role, it is imperative that accurate methods are available for its measurement. This is particularly important since acute cortisol deficiency presents clinically as an Addisonian crisis, characterised by catastrophic dehydration and salt wasting, which is potentially fatal. Consequently, early diagnosis is essential. Milder degrees of adrenal insufficiency are difficult to detect clinically due to vague, non-specific symptoms such as fatigue, generalised weakness and lethargy. Thus, diagnosis relies entirely on biochemical testing and requires an alert clinician with a high index of suspicion. Similarly, the early symptoms of cortisol excess (Cushing's syndrome) which include weight gain, hypertension and impaired glucose tolerance are easily attributable to other causes, such as simple obesity or the metabolic syndrome, leading to delays in biochemical testing and diagnosis.

Establishing a diagnosis of adrenal disease requires more than simply measuring a random serum cortisol concentration, as this measurement alone, although widely used, rarely gives sufficient information about adrenal status (Grinspoon and Biller 1994; Bornstein et al. 2016). An early morning cortisol measurement is also rarely diagnostic, although it can contribute some diagnostically useful information; with a concentration below 140 nmol/L suggesting adrenal insufficiency (Kazlauskaite et al. 2008), a concentration above 450 - 550 nmol/L virtually excluding hypoadrenalism, and concentrations between these two extremes indicating the need for further testing. A low early morning cortisol combined with plasma ACTH measurement can indicate whether adrenal insufficiency is likely to be primary or secondary (Bornstein et al. 2016). Nevertheless, when adrenal disease is suspected, the most useful tests are those that stimulate or suppress the HPA axis – either in its entirety or partially – with subsequent measurement of cortisol to quantify the adrenal response.

1.2.1 Serum cortisol assays

Early plasma cortisol assays were restricted to research laboratories as they required several labour-intensive steps, given the low concentrations being measured and the presence in plasma of numerous other steroid hormones. In 1962 a simple fluorimetric assay that utilised the natural fluorescence of 11-hydroxycorticoids (cortisol and corticosterone) in plasma, and could turnaround six samples in an hour and a half, was described, marking the beginning of clinical cortisol assays (Mattingly 1962). However, fluorimetry was limited by its poor specificity for cortisol and low sample throughput and was replaced by radioimmunoassay (RIA) some 10 years later (Campuzano et al. 1973).

This first RIA used rabbit anti-cortisol antibodies raised against a cortisol-21-hemisuccinate-bovine serum albumin (BSA) conjugate which showed no cross-reactivity with corticosterone, 11-deoxycortisol, oestradiol, progesterone or testosterone. Samples were heated in boiling water for 5 minutes before being assayed to destroy endogenous CBG, and were then incubated with anti-cortisol antibody and radiolabelled cortisol for 40 hours at 5°C. Despite the prolonged incubation time, this assay was considered to be relatively quick, as it did not require pre-analytical sample extraction, and it became the method of choice for cortisol measurement for many years.

The main limitations of RIA included large variations in the specificity and affinity of different antibodies for cortisol and other steroid hormones and the requirement to release the steroid molecule from its carrier protein before it could be measured. This can be achieved in several ways, including heat treatment, which denatures CBG and other serum proteins, and typically involves incubating the sample at 60°C for 30 - 60 minutes (Holder 2006; Jung et al. 2011) or the addition of an agent, such as low pH, 8-anilinonaphthalene sulfonic acid (8-ANS) (Brock et al. 1978; McConway and

Chapman 1986) or a competing steroid e.g. Danazol (Haning et al. 1982), to displace cortisol from its binding proteins.

With time, cortisol RIAs underwent further development to enable pre-analytical sample extraction steps to be removed and enzyme and chemiluminescent detection to be introduced in place of radioactivity (Kominami et al. 1980; Lindstrom et al. 1982). The first automated cortisol immunoassay was described in 1992 (Bacarese-Hamilton et al. 1992) and these are now the method of choice for most clinical laboratories; although there is increasing interest in the definitive identification of cortisol offered by mass methods, particularly in more specialised laboratories (Owen et al. 2013a). Assays that measure bioactive, free cortisol are also available (Kirchhoff et al. 2011), but there has been little success in developing one that is easy to use, cost-effective and with sufficiently high throughput to meet the requirements of a clinical laboratory.

1.2.1.1 Immunoassays

Automated cortisol assays are solid phase heterogeneous competitive binding immunoassays. A sample of serum is added to a reaction cuvette and incubated with anti-cortisol antibodies attached to microparticles (the solid phase) and a labelled cortisol conjugate. Cortisol in the sample and the labelled cortisol conjugate compete for antibody binding sites on the microparticles (figure 1.2). Unbound cortisol from the sample and the labelled conjugate is washed away and a chemiluminescent signal, which is inversely proportional to the cortisol concentration in the sample, is produced and measured.

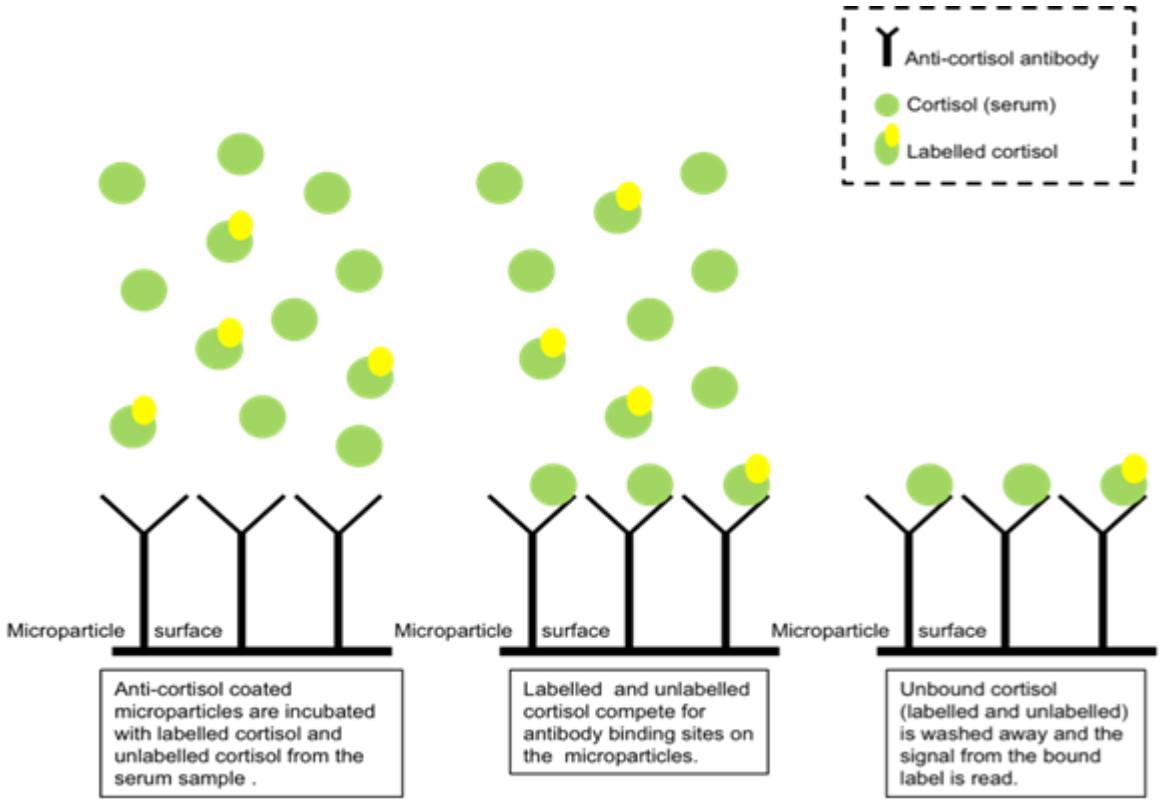


Figure 1.2 Principle of a solid phase competitive binding immunoassay

In common with other steroids, cortisol is not immunogenic. Thus, in order to obtain anti-cortisol antibodies a cortisol derivative is covalently bound to an immunogenic protein e.g. bovine serum albumin (BSA) and the complex is injected into an appropriate host animal (Campuzano et al. 1973). To enhance antibody specificity for cortisol a site that is remote from its specific functional groups should be chosen for attachment to the protein carrier, and extensive work has been carried out to find the optimum cortisol immunogen for creating specific antibodies. Nonetheless, the use of different antibodies, each with differing affinity and specificity for cortisol in the different cortisol immunoassays, will undoubtedly result in differences in cortisol measurement between them. Furthermore, as immunoassays are only able to identify the presence of an immunoreactive molecule, but not to confirm its nature, antibody cross-reactivity with other steroid hormones (table 1.1) will further contribute

to variability between the different commercial assays. Cortisol binding to its carrier protein can also interfere with immunoassay performance as antibody binding sites may be altered or hidden as a result. These factors, and others, which affect cortisol immunoassay performance are discussed in detail below.

Table 1.1 Manufacturer-reported interference from cross-reacting steroid hormones

Substance	Abbott Architect	Beckman Access	Centaur Advia	Roche E170	Immulite 2000
Aldosterone	ND	NR	0.3	NR	0.1
Allotetrahydrocortisol	NR	NR	6.5	165	ND
Corticosterone	0.9	2.08	5.3	5.8	1.2
Cortisone	2.7	8.06	31.1	0.30	1.0
11-deoxycorticosterone	ND	0.91	1.8	0.69	ND
11-deoxycortisol	1.9	17.8	23.3	4.1	1.6
21-deoxycortisol	NR	NR	8.1	45.4	NR
Dexamethasone	NR	0.04	0.2	0.08	ND
Fludrocortisone	36.6	NR	NR	NR	NR
6β-hydroxycortisol	0.2	NR	6.8	158	NR
17α-hydroxyprogesterone	0.6	5.33	1.2	1.5	0.2
Pregnenolone	ND	NR	0.5	NR	NR
Prednisolone	12.3	7.60	109	171	62
6-methyl-prednisolone	0.1	NR	26.2	389	22-25
Prednisone	0.6	3.05	34	0.28	6.1
Progesterone	ND	0.46	ND	0.35	ND
Spironolactone	ND	NR	ND	NR	ND
Testosterone	ND	NR	0.2	NR	NR

Samples were spiked with varying concentrations of the hormone or drug in question, starting at 10 ug/dL and increasing up to a maximum of 1000 ug/dL, and the percentage increase in cortisol concentration was reported. Thus, it can be seen that aldosterone does not interfere in the Siemens Immulite 2000 assay as there was no increase in serum cortisol concentration, even at 1000 ug/dL of Aldosterone. In contrast, 10 ug/L of prednisolone was sufficient to increase cortisol concentration by 171% with the Roche E170 assay. (NB. 1 ug/dL of cortisol = 27.6 nmol/L)
Interference is reported as a percentage.
ND = none detected; NR = not reported
Information extracted from each manufacturer's cortisol kit insert.

1.2.1.1.1 Assay variability and instability

Despite the limitations noted by early steroid RIA workers, the significance of differences between automated cortisol immunoassays was not widely considered until a study by Clark *et al* in 1998 (Clark et al. 1998). Cortisol concentration pre- and post- Synacthen stimulation of the adrenal gland was measured using four widely

available immunoassays: the TDX (Abbott Diagnostics), ACS 180 (Chiron Diagnostics), DELFIA (Pharmacia Wallac) and Coat-a-Count (Diagnostic Products Corp DPC). Median cortisol concentration 30 minutes post-Synacthen ranged from 707 nmol/L for the DELFIA assay to 866 nmol/L for the Coat-a-Count; with a lower limit (defined as the 5th percentile) which ranged from 510 nmol/L (DELFIA) to 626 nmol/L (TDX). This demonstrated the assay dependence of cortisol cut-offs and the risk that patients could be misdiagnosed with adrenal insufficiency simply by virtue of the cortisol assay used. Surprisingly, this work did not prompt widespread application of method-specific cortisol cut-offs for interpreting the Synacthen test and most laboratories continued to use the widely quoted, historical, 550 nmol/L cut-off (Chatha et al. 2010). But perhaps of even greater concern is the fact that the need for method-specific cut-offs did not become embedded in clinical practice and the diagnosis of adrenal insufficiency continued to rest on a single, poorly evidenced cortisol result.

Since that work a new generation of cortisol immunoassays has come into use and whilst some laboratories have continued to use their old validated cut-offs and reference ranges, others have attempted to adjust these values by direct comparison between their old and new assays. However, such adjustments have not been clinically validated and, in reality, they offer little improvement over the continued use of the old cut-off. More recent work has confirmed the ongoing need for method-specific cut-offs with the new assays (Klose et al. 2007), but for most of these, such cut-offs remain to be established. This is particularly important given the gravity of misdiagnosing adrenal insufficiency – both in terms of an Addisonian crisis in an undiagnosed patient and, conversely, unnecessary steroid replacement in an incorrectly diagnosed patient.

The long-term validity of any newly established cortisol cut-offs or reference ranges, however, is likely to require further investigation. A recent study designed to explore

the effect of gender and matrix on immunoassay cortisol measurements (Dodd et al. 2014) identified changes in assay performance over a three year period, despite no major changes in assay formulation over the same period (table 1.2). This change in cortisol measurement over time (assay drift or instability) is likely to be due to manufacturers' attempts to better align their assays with mass methods and, to some extent, to variability within antibody populations; but it means that any new cut-off is unlikely to remain valid for the entire lifetime of the assay. Thus, to accurately determine method-specific cortisol responses to the short Synacthen test, there is first a need for this response to be defined by the gold standard gas chromatography – mass spectrometry (GC-MS) method. This would provide a reference point from which current assay-specific characteristics could be identified and to which future assay shifts could be compared to ensure cut-offs are adjusted appropriately.

Table 1.2 Mean assay bias ratios in 2010 and 2013.

	Mean Bias Ratio 2010			**Mean Bias Ratio 2013**		
Assay	**Total**	**Males**	**Females**	**Total**	**Males**	**Females**
Roche E170	1.30	1.36	1.25	1.26	1.28	1.23
Siemens Centaur	1.06	1.08	1.05	1.17	1.19	1.14
Abbott Architect	1.02	1.04	1.00	0.94	0.95	0.93
Beckman Access	1.07	1.09	1.05	0.98	1.00	0.95

An individual sample's bias ratio was calculated by dividing its immunoassay cortisol concentration by the corresponding GC-MS concentration. A mean assay bias ratio was then calculated for each immunoassay.

Reproduced with permission from Dodd et al. The effect of serum matrix and gender on cortisol measurement by commonly used immunoassays. Ann Clin Biochem 2014;51:379-85. (Sage Publications Ltd)

The increasing divergence of cortisol assays is also evident when examining external quality assurance (EQA) data (figure 1.3) (UKNEQAS Annual review 2010-2013).

The majority of UK participants in the UKNEQAS cortisol scheme measure cortisol using one of the following immunoassays: Roche Elecsys (44%), Siemens ADVIA Centaur (20%), Abbott Architect (19%) and Beckman Access (10%). The scheme organisers calculate a B score, or specimen percentage bias, to demonstrate how far each laboratory's result deviates from the target concentration; which in this case is the all laboratory trimmed mean (ALTM)[1]. An arbitrary acceptable limit of performance of 10% above or below the target mean has been set. Figure 1.4 shows the percentage of laboratories failing to meet this target and demonstrates an increasing problem over the past 4 years. This is unlikely to be explained by worsening laboratory performance, but indicates that current cortisol immunoassays are so different to each other that attempts to combine their results to create a single, meaningful mean are impossible. UKNEQAS is, therefore, working towards defining a more suitable target concentration to replace the ALTM which is likely to utilise GC-MS or LC-MS/MS methodology.

[1] The all laboratory trimmed mean is the mean of all the results generated by participating laboratories, calculated after outlying results – typically the top and bottom 5 or 10% – have been excluded.

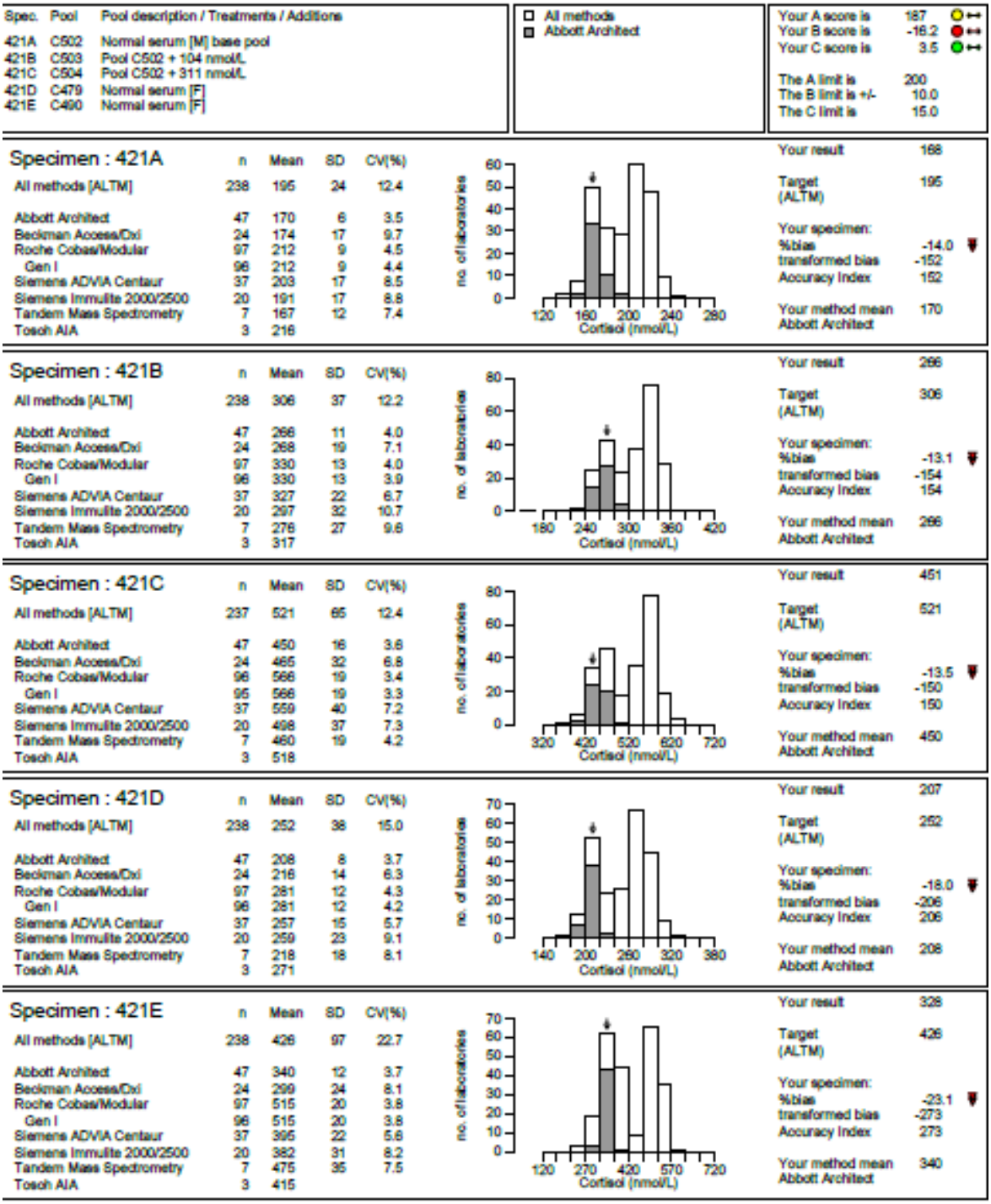


Figure 1.3 Cortisol EQA report (UKNEQAS). EQA samples are circulated to participating labs at monthly or fortnightly intervals and this summary report is produced to show the results. The methods being used are listed and a mean for each is provided. The ALTM is used as the target result for comparison despite reflecting each individual method mean poorly.

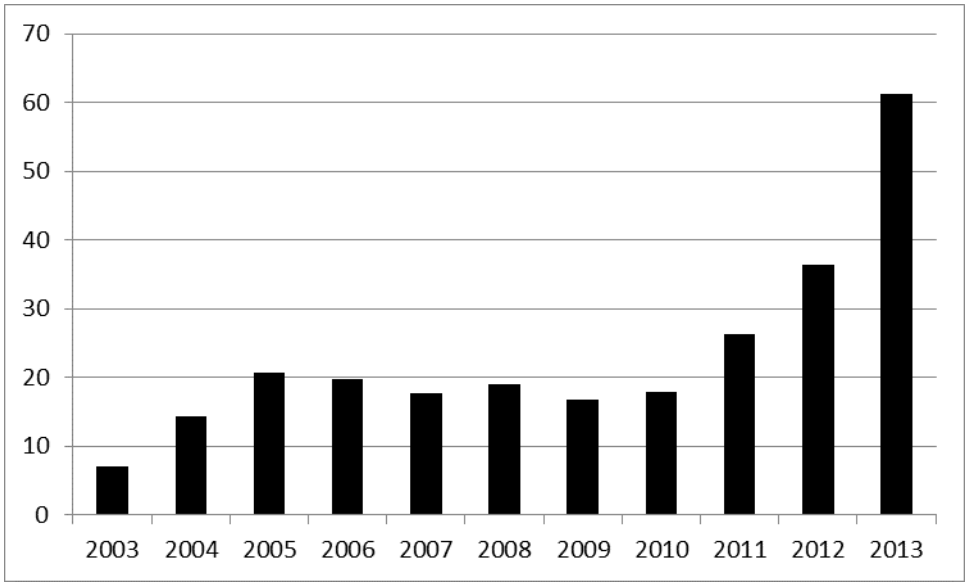


Figure 1.4 Trends in laboratories with bias scores outside acceptable limits. Each bar represents the percentage of laboratories participating in the UKNEQAS cortisol EQA scheme whose B scores are outside acceptable limits of performance.
The B score is the average %bias [(result-target)/(target) * 100%] of all 30 specimens distributed over a six month time window.

Given the ongoing assay changes, it is understandable that some laboratories and clinicians may be reluctant to adopt literature-based method-specific cut-offs. However, if assay differences continue to be overlooked when interpreting cortisol concentrations, particularly in the context of the Synacthen test, the clinical validity of these tests will be lost. In fact, this is already becoming apparent, with some authors recommending changes to sample timing in the Synacthen test to improve its specificity (Mansoor et al. 2007; Chitale et al. 2013).

1.2.1.1.2 Assay standardisation[2]

Poor standardisation, which is partly due to the different antibodies used by the different assays and partly explained by the lack of either a single reference material

[2]Assays should ideally be standardised against an approved reference material using an approved reference method in order that they produce interchangeable results. The International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) leads the way in defining appropriate reference methods and materials.

or method for cortisol, is another factor contributing to the variability of cortisol immunoassays. There are at least 6 certified reference materials for cortisol – ERM-DA192 and 193 (1985), ERM-DA451/IFCC (1999), SRM 921 (1993), SRM 971 (2011) and hydrocortisone USP reference standard, and information from manufacturers of current cortisol assays reveals traceability to several of these materials (personal correspondence, table 1.3). There are also several reference methods, the earliest of which used isotope dilution-gas chromatography-mass spectrometry (ID-GC-MS) (Breuer and Siekmann 1975), with more recent methods relying on liquid chromatography-tandem mass spectrometry (LC-MS/MS) for ease of use (Tai and Welch 2004); and a variety of these have been used to confirm the validity of current cortisol immunoassays (table 1.3).

Table 1.3 Immunoassay traceability.

Assay	Traceability
Beckman Access	Traceable to United States Pharmacopea reference material.
Abbott Architect	Assay is designed to have a slope of 1.0 +/- 0.1 and a correlation coefficient (r) of ≥ 0.95 for serum samples when compared to Liquid Chromatography Mass Spectrometry/Mass Spectrometry (LC-MS/MS).
Siemens Centaur	GC-MS
Roche E170	Standardised against another method which is standardised against ID-MS IFCC-451 Panel ID/GC/MS gives 89-111% recovery
DPC Immulite	Correlated to Coat-A-Count cortisol assay Traceable to an internal standard manufactured using qualified materials and measurement procedures

Information provided by assay manufacturers regarding the reference material and method to which each assay was traceable.

1.2.1.1.3 Binding proteins

Current assays measure total serum cortisol (bound and unbound fractions), and as a result are affected by changes in CBG and, to a lesser extent, albumin, concentration. This is a particular problem in patients with protein deficient states. The cortisol response to ACTH stimulation in nephrotic patients has been shown to be lower, by between 63 and 179 nmol/L, depending on the assay used, than that in healthy volunteers (Klose et al. 2005). This translates into a 50% failure rate with both the Immulite 2000 (Diagnostics Products Corp) and autoDelfia (Perkin-Elmer) assays, when a 500 nmol/L cut-off is used. Similar findings have been reported in patients with liver impairment and mean albumin concentrations of 21 g/L, with 46% of patients failing to achieve a total cortisol of 550 nmol/L post-ACTH stimulation (Vincent et al. 2009).

Acutely unwell patients are also at risk of misleading cortisol results due to low protein concentration; however, interpretation of cortisol measurements in this group is further complicated by increased synthesis of endogenous steroids by an activated HPA axis and a high prevalence of heterophilic antibodies (Cohen et al. 2006; Briegel et al. 2009)[38]. Assay differences also appear to be more significant in this patient group, particularly post-ACTH stimulation (Cohen et al. 2006; Briegel et al. 2009), leading some authors to suggest that cortisol immunoassays should not be used for measuring cortisol in the critically ill.

An excess of binding proteins, such as occurs in states of oestrogen excess e.g. pregnancy and treatment with oestrogen-containing oral contraceptives (OCP), can also cause difficulty in evaluating adrenal status/reserve through overestimation of

[3] *Heterophilic antibodies are human antibodies, typically produced against antigens from a specific animal, which are capable of binding to that animal's antibodies in an immunoassay; resulting in a false positive (or occasionally, negative) result. E.g. Human anti-mouse antibodies (HAMA).

cortisol concentration (Booth et al. 1961). This increase in total serum cortisol is artefactual and not a true reflection of bioavailable cortisol, and, in the context of Synacthen stimulation can suggest an adequate adrenal response in a cortisol-deficient patient.

1.2.1.1.4 Endogenous and exogenous steroids

Through their common precursor – cholesterol – and biosynthetic pathway, most steroid hormones share structural homology (figure 1.5), which makes the generation of entirely specific antibodies difficult (Holder 2006). Current cortisol immunoassays are hampered by varying degrees of antibody cross-reactivity with other steroids (table 1.1), both endogenous and exogenous (Playfair et al. 1974; Brossaud et al. 2009), and can be unreliable in certain clinical settings such as congenital adrenal hyperplasia (Curtis 2009), and in patients treated with synthetic glucocorticoids. To some extent, laboratories can use manufacturers' information about steroid cross-reactivity to identify interference; however, this information is far from exhaustive, and labs are not always aware of patients' medication. In fact, the recent implementation of an LC-MS/MS cortisol assay in a laboratory serving a large respiratory unit revealed the unreported presence of exogenous steroids in up to 50% of patients being tested (Owen et al. 2013a).

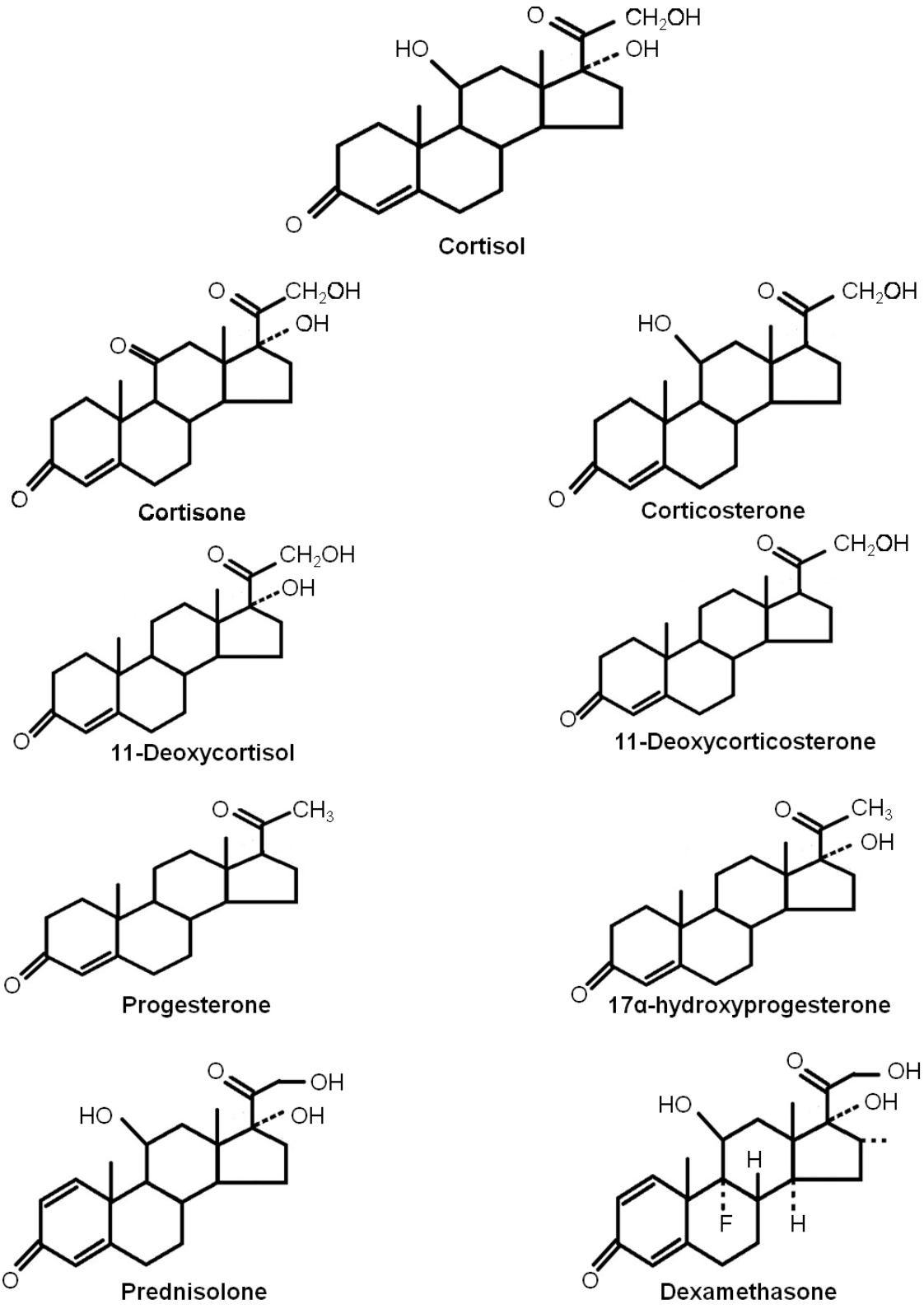


Figure 1.5 Structural homology of cortisol and other steroid molecules.

1.2.1.2 Mass methods

The first reference method for measuring cortisol used GC-MS and was described in 1975 (Breuer and Siekmann 1975). Cortisol was derivatised to its heptafluorobutyric ester to improve assay sensitivity and tritiated cortisol was added as an internal standard to correct for losses due to steroid adsorption to the chromatography column. Despite the high specificity and sensitivity of mass methods, GC-MS cortisol assays remain confined to reference and research laboratories as they are labour intensive and require expensive specialist equipment and highly-skilled laboratory staff.

In 2001, an LC-MS/MS method for measurement of multiple steroid hormones in patients with 11-β-hydroxylase and 21-hydroxylase deficiencies and Addison's disease was described (Kao et al. 2001). This method had the advantage over immunoassay of improved specificity, definitive identification of the substances being measured and the ability to perform multiple tests simultaneously. Advantages over GC-MS included higher throughput, elimination of the derivatisation step, lower cost and increasing availability of LC-MS/MS equipment in clinical laboratories. In 2004, an LC-MS/MS reference method was described (Tai and Welch 2004) as an alternative to GC-MS, and has since replaced the latter as the reference method of choice for routine method comparisons (Hawley et al. 2016).

As LC-MS/MS becomes more widely available in clinical laboratories, there has been a move towards using it for routine measurement of steroid hormones, either in isolation (Owen et al. 2013a) or as part of a panel of related molecules (Koal et al. 2012; Ruiter et al. 2012; Keevil 2013; Kyriakopoulou et al. 2013; Keefe et al. 2014; Lee et al. 2014). In fact, the UKNEQAS cortisol scheme demonstrates that several such assays are already in use (figure 1.3). However, if this interest in LC-MS/MS as an alternative to immunoassay continues, the cortisol response to Synacthen

stimulation and dexamethasone suppression will need to be defined to ensure correct interpretation in light of the increased specificity. Furthermore, unless better efforts at assay standardisation are made with LC-MS/MS; including traceability to a single reference method and material, it is likely to be affected by similar inter-assay differences to those which currently plague immunoassays. Particularly as such differences have already been observed with LC-MS/MS vitamin D and testosterone assays – the former due to the lack of common calibrators (Yates et al. 2008) and the latter due to factors thought to include differences in sample extraction techniques and the type of LC column and mass spectrometer used (Owen et al. 2013b).

The reproducibility and accuracy of LC-MS/MS assays can be affected by sample matrix. In particular, endogenous impurities e.g. salts in urine samples and co-eluting substances such as analyte and drug metabolites can cause ion suppression which is a reduction or increase in the efficiency of ion formation (Annesley 2003; Taylor 2005; Allende et al. 2014). Careful evaluation of this effect is, therefore, essential before introducing a new LC-MS/MS assay into clinical practice.

1.2.1.3 Free cortisol assays

Direct measurement of serum free cortisol has been possible since the 1950s, however the methods that have been described to date: equilibrium dialysis, ultrafiltration and gel filtration, are labour intensive and time consuming. Attempts to modify these assays for semi-routine, specialist services have been moderately successful, but there is, as yet, no prospect of a serum free cortisol assay for routine clinical use (MacMahon et al. 1983; Kirchhoff et al. 2011).

Equilibrium dialysis is particularly time consuming. It involves using a dialysis membrane to separate a small volume of plasma from a buffer solution and incubating the two for up to 24 hours (Vogeser et al. 2002). Free cortisol crosses the membrane until equilibrium is reached and can be measured in the dialysis buffer. However, the

accuracy of this method is affected by serum leakage across the dialysis membrane and by changes in cortisol binding equilibrium within the plasma sample which result from its dilution with dialysis buffer.

Ultrafiltration would seem to offer a more suitable method for adoption for clinical use due to its simplicity. Plasma, separated from a collection chamber by a dialysis membrane, is centrifuged for 10 minutes at 37°C and the cortisol concentration in the resulting filtrate is measured (MacMahon et al. 1983). However, substantial variations in filtrate yield have been reported; a finding which may be explained by the forced filtration disturbing cortisol binding equilibrium in the sample at the membrane surface. Precise temperature control can also be difficult to achieve, particularly in a routine laboratory setting, which somewhat limits wider application of this method.

Gel filtration relies on unbound cortisol moving down a packed glass column more slowly than its protein-bound counterpart (Burke 1969a). As the latter fraction passes through the column, the cortisol dissociates from its binding protein until the concentration of dissociated cortisol in the column equals the unbound fraction in the original sample and no further dissociation occurs. By replacing the sample with buffer at this point, the protein-bound cortisol elutes first and the concentration of serum free cortisol can be measured in the later fraction. This method requires large sample volumes to create a steady state in the gel column and, although analysis is possible at any temperature, it is not entirely clear how temperatures below 37°C affect cortisol-CBG binding equilibrium.

Routine measurement of serum free cortisol is further hampered by the lack of a single, robust reference range. Most studies have been small and have used in-house assays to measure cortisol, resulting in significant differences in the ranges quoted. Particularly low substrate concentrations and the lack of a stable quality control material have contributed to these differences. There have been few

comparisons between methods, although at least one study has shown good correlation between serum free cortisol measured by ultrafiltration and by equilibrium dialysis (Burke 1969b; Pretorius et al. 2011).

As a result of the difficulties in measuring serum free cortisol, interest has turned towards deriving calculated estimates as an alternative. In 1987, Coolens *et al* derived an equation for calculating serum free cortisol (SFC) taking into account the binding properties of both CBG and albumin (Coolens et al. 1987). Correlation between this estimate and actual serum free cortisol measured by ultrafiltration was found to be excellent. An alternative marker, the free cortisol index (FCI), defined as serum total cortisol divided by CBG, has also been shown to correlate well with serum free cortisol measured by gel filtration and with Coolens' estimate (le Roux et al. 2002). However, as is the case for measured serum free cortisol, neither the FCI nor SFC has been validated for use in patients with HPA axis disease and CBG measurement is, by no means, widely available or, indeed, well-suited to routine analysis.

The search for alternatives to measuring serum free cortisol has, therefore, explored cortisol measurement in other media, including urine and saliva; but it is the latter that shows more promise in the evaluation of patients with suspected adrenal insufficiency.

1.2.2 Salivary cortisol assays

Cortisol enters saliva by diffusion (Vining et al. 1983), independently of salivary flow rate. It is not significantly protein-bound and reflects changes in plasma cortisol quickly and reliably (Vining and McGinley 1987). Cortisol's circadian rhythm and early morning peak are detectable in saliva (Walker et al. 1978) and interest in measuring salivary cortisol to evaluate the HPA axis began in the 1960s (Riad-Fahmy et al. 1982; Vining et al. 1983). However, as cortisol concentrations in saliva are less than one

tenth those in serum, measurement was hampered by the lack of sensitive cortisol assays available at the time (Vining et al. 1983).

Salivary cortisol was first measured successfully by direct RIA in 1978 (Walker et al. 1978) and more recently automated serum cortisol immunoassays have been adapted to measure cortisol in saliva (Deutschbein et al. 2012; Sesay et al. 2013). However, the utility of immunoassays is limited by their poor specificity, particularly where there is significant antibody cross-reactivity with cortisone which is present in high concentration in saliva. Assay sensitivity can also present a problem as late night salivary cortisol concentrations often fall close to, or below, most assays' functional limit of detection.

In 2003, the first LC-MS/MS assay for measuring salivary cortisol was described and put forward as an alternative to RIA to reduce the impact of cross-reactivity (Jonsson et al. 2003). Efforts have since concentrated on adapting LC-MS/MS to provide a quick, high-throughput assay. However, salivary cortisol measurement as an alternative to serum is still limited by the lack of a single, validated reference range (Baid et al. 2007; Zerikly et al. 2010; Inder et al. 2012) and potential differences arising due to poorly standardised assays (Inder et al. 2012). Using different sample collection techniques has also been shown to contribute to the differences observed between salivary cortisol assays, as has sample contamination during collection.

Nonetheless, salivary cortisol is now widely used in clinical laboratories as a screening test for Cushing's syndrome (Nieman et al. 2008) and in monitoring hydrocortisone replacement (Wong et al. 2004) in patients with adrenal insufficiency. There is, also increasing interest in using it as an alternative to serum total cortisol in the interpretation of the cortisol response to Synacthen stimulation (Deutschbein et al. 2009; Perogamvros et al. 2010b), particularly in patients with altered protein

concentrations e.g. cirrhosis (Galbois et al. 2010); however a valid reference range does not yet exist.

Salivary cortisone has also been put forward as a useful marker of serum free cortisol (Perogamvros et al. 2010a; Raff and Singh 2012). Parotid tissue contains 11β-hydroxysteroid dehydrogenase type 2 (11β-HSD2) which converts cortisol to cortisone (Smith et al. 1996), which has also been shown to correlate well with serum free cortisol (Perogamvros et al. 2009; Perogamvros et al. 2010a). Further work is needed to establish suitable reference ranges and cut-offs.

1.3 Adrenal insufficiency (Oelkers 1996; Arlt and Allolio 2003)

Adrenal insufficiency was first described in 1855 by Thomas Addison as a syndrome characterised by wasting and hyperpigmentation. His description was of primary adrenal insufficiency caused primarily by direct destruction of the adrenal gland by tuberculosis; however, the clinical presentation remains the same whatever the underlying cause. The commonest cause of primary adrenal insufficiency in the developed world today is autoimmune adrenalitis; other common causes are listed in table 1.4.

Secondary adrenal insufficiency is caused by hypothalamic or pituitary disease preventing normal stimulation of the adrenal cortex. If ACTH deficiency is prolonged, the lack of stimulation of the adrenal gland results in cortical atrophy and down-regulation of adrenal ACTH receptors (Lebrethon et al. 1994). The end result is a clinical picture similar to that of primary adrenal insufficiency (table 1.5).

Table 1.4 Causes of adrenal insufficiency[4]

Primary:	Adrenal disease
	- Autoimmune adrenalitis – isolated or autoimmune polyglandular syndrome
	- Infection – TB, AIDS, fungus e.g. cryptococcosis, histoplasmosis
	- Infiltration – metastases, haemochromatosis, sarcoidsosis, amyloidosis
	- Iatrogenic – bilateral adrenalectomy, inadequate steroid replacement, cessation of high dose steroid therapy, drugs e.g. ketoconazole
	- Genetic – congenital adrenal hyperplasia, adrenoleukodystrophy
	- Bilateral adrenal haemorrhage
Secondary:	Pituitary
	- Tumours – adenomas, metastases
	- Infarction – postpartum Sheehan's syndrome
	- Infection – encephalitis
	- Iatrogenic – surgery, irradiation
	- Trauma
	- Infiltration e.g. TB, sarcoidosis, histiocytosis X
	Hypothalamus
	- Tumours – craniopharyngioma, optic glioma
	- Infection – meningitis, encephalitis, TB
	- Iatrogenic – surgery, irradiation
	- Trauma
	- Other – Anorexia nervosa

[4] Adapted from Arlt et al. 2003.

Table 1.5 Clinical features of adrenal insufficiency[5]

Symptoms:	Fatigue
	Generalised weakness
	Lack of energy
	Anorexia and weight loss
	Abdominal pain
	Myalgia and joint pain
	Dizziness
Signs:	Low blood pressure (and postural hypotension)
	Fever
	Pallor (Alabaster-skin, 2° adrenal insufficiency)
	Hyperpigmentation (1° adrenal insufficiency)
	Anaemia
Biochemistry:	Hyponatraemia
	Hyperkalaemia
	Raised urea
	Hypercalcaemia
	Hypoglycaemia
Acute adrenal crisis:	Hypovolaemic shock
	Acute abdominal pain
	Vomiting
	Fever

[5] Adapted from Arlt et al 2003.

1.4 Diagnosing adrenal insufficiency

The gold-standard test of HPA axis function is the insulin tolerance test (ITT), first described in 1963 (Landon et al. 1963). Insulin-induced hypoglycaemic stress stimulates the hypothalamus directly, resulting in CRH release and an increase in serum cortisol; thus demonstrating the integrity of the entire axis, from hypothalamus to adrenal gland. In 1969, Plumpton and Besser demonstrated that an adequate adrenal response to insulin-induced hypoglycaemia could accurately predict a cortisol rise in patients undergoing major surgery (Plumpton and Besser 1969). They noted marked inter-individual variability in the magnitude of the response, although the stress of surgery reliably provoked a larger and more prolonged response than that seen with the ITT. Following surgery, none of the patients with a normal cortisol response to insulin-induced hypoglycaemia developed features of adrenal insufficiency, confirming the ITT's accuracy as a test of HPA axis function.

However, this test is potentially dangerous and often unpleasant for patients. It is contra-indicated in several patient groups, including those with epilepsy and coronary artery disease and must be used with caution in children (Shah et al. 1992) and the elderly. It requires careful patient monitoring by experienced staff, with multiple timed blood collections to accurately identify a cortisol peak (Jones et al. 1994). Accordingly, it is rarely used as the first line test for investigating adrenal insufficiency (Reynolds et al. 2006).

Used even less is the Metyrapone test, which blocks conversion of 11-deoxycortisol to cortisol (Fiad et al. 1994; Grinspoon and Biller 1994). The resultant low cortisol exerts positive feedback on both the hypothalamus and pituitary, resulting in CRH and ACTH release. The ACTH stimulates the adrenal gland to produce cortisol, leading to an increase in cortisol precursors, and in particular, 11-deoxycortisol. This rise is taken to indicate an adequate pituitary response to cortisol deficiency.

As with the ITT, the Metyrapone test assesses the entire HPA axis, although the stimulus is cortisol deficiency rather than stress. Its main disadvantages lie in the lack of accurate assays to measure 11-deoxycortisol and the risk of inducing adrenal crisis in patients with adrenal insufficiency. It is not widely used in the UK and the short Synacthen test (SST) has now replaced the ITT as the first line for testing adrenal reserve (Reynolds et al. 2006).

The short Synacthen test was first described in 1965 as a safe and quick test of adrenal function (Wood et al. 1965). A supraphysiological dose of Synacthen, a synthetic peptide consisting of the first 24 amino acids of natural ACTH, is administered parenterally and stimulates the adrenal cortex directly, resulting in raised cortisol, unless there has been significant destruction or atrophy of the adrenal gland (Dorin et al. 2003). The cortisol response is measured 30 minutes after the Synacthen dose, giving a direct assessment of adrenal gland function but only an indirect measure of pituitary function (Borst et al. 1982).

Nevertheless, the cortisol response to Synacthen stimulation has been shown to correlate well with the response to insulin-induced hypoglycaemia in healthy volunteers and in patients with varying degrees of HPA axis impairment (Kehlet et al. 1976; Lindholm et al. 1978). The Synacthen test also correlated well with the stress of major surgery (Kehlet and Binder 1973), suggesting it could be used as a reliable alternative to the ITT in patients undergoing assessment of the HPA axis (Stewart et al. 1988).

Conversely, there have also been reports of patients with HPA axis dysfunction, confirmed by the Metyrapone test and clinical response to glucocorticoid replacement, whose cortisol response to Synacthen stimulation was “normal”, highlighting the potential danger of relying solely on the Synacthen test to diagnose HPA axis disease (Streeten et al. 1996).

These concerns about failing to detect HPA axis disease and adrenal insufficiency appear to have overshadowed any relating to overdiagnosis of adrenal insufficiency and consequent over-treatment, despite the potential harm from long-term glucocorticoid replacement. In a study comparing ITT and Synacthen test results in patients with hypothalamo-pituitary disorders, Stewart et al reported 10 discordant results from a total of 70 (Stewart et al. 1988). Nine of these failed the Synacthen test and subsequently passed the ITT, leading the authors to recommend using the Synacthen test as first line, followed by an ITT in those who failed. However, this is not routinely done.

1.5 Limitations of the short Synacthen test

Due to its widespread use and the serious nature of the disease it is used to detect, the SST has continued to come under increasing scrutiny over the past 30 years. It has been shown to have no role in the diagnosis of acute pituitary disease nor in the assessment of the HPA axis in the early weeks post pituitary surgery (Cunningham et al. 1983; Grinspoon and Biller 1994; Courtney et al. 2000; Dickstein 2001; Klose et al. 2005) because significant adrenal atrophy develops over weeks and, until it does, the risk of a false negative result is high (Dorin et al. 2003). Its role as the test of choice for investigating adrenal disease in other settings has also been brought into question by concerns over the most appropriate dose of Synacthen to use (Abdu et al. 1999), difficulties in determining the correct cortisol cut-off (Stewart et al. 1988; Clayton 1989; Fiad et al. 1994; Clark et al. 1998) and the impact of altered CBG levels on total cortisol measurement (Coolens et al. 1987; Dhillo et al. 2002; Moisey et al. 2006).

1.5.1 Synacthen dose

There has long been concern that the standard, 250 µg, dose of Synacthen is in such excess that it could elicit a cortisol response from a partially atrophied adrenal cortex and several authors have proposed a lower, 1 µg, dose as a more meaningful stimulus for assessing adrenal reserve (Borst et al. 1982; Crowley et al. 1991; Dickstein et al. 1991; Abdu et al. 1999). However, the low dose test may itself be subject to error due to the lack of commercial low dose Synacthen preparations and the consequent requirement to dilute a standard (1 mL) 250 µg Synacthen vial prior to administration (Thaler and Blevins 1998). Synacthen adsorption to the plastic syringes and other devices used for its administration can also pose a problem and in inexperienced hands the low dose SST is unlikely to be a reliable test of adrenal function.

1.5.2 Cortisol cut-off

Determining the most appropriate cortisol cut-off for use in the standard short Synacthen test has proved difficult due to disagreement over how it should be defined (Grinspoon and Biller 1994; Mayenknecht et al. 1998; Weintrob et al. 1998; Park et al. 1999). Some authors have suggested that the 30-minute cortisol cut-off value should be set to allow close correlation with the result of the insulin tolerance test (Hurel et al. 1996; Abdu et al. 1999), and, indeed, the widely quoted 550 nmol/L cut-off was used for precisely this reason. Others have defined a lower limit of normal using the 2.5th or 5th percentile of the response in a cohort of healthy volunteers (Clark et al. 1998; Klose et al. 2007). Still others have suggested using the 60-minute cortisol result to define the response (Mansoor et al. 2007) and it is now perfectly clear that, no matter how the cut-off is defined – and correlation with the insulin tolerance test would be ideal, although potentially difficult to achieve due to the risks associated with testing – a single cut-off is no longer applicable (Clark et al. 1998;

Klose et al. 2007). Furthermore, whilst correlation with the cortisol response to insulin-induced hypoglycaemia is desirable, this too is affected by changes to cortisol assays and will itself need to be redefined.

It is also time for the difficulties associated with using an absolute cortisol cut-off to be addressed. Those within the laboratory are well aware of the concept of assay precision and the slight differences between repeated analyses performed on the same sample. Clinicians are perhaps less aware of this concept; however, with manufacturer-quoted CVs of 3.8% to ≤ 10% for cortisol assays at the typical cut-off concentration of 500 – 550 nmol/L, continuing to use an absolute cut-off will mean some patients are misclassified purely because of assay imprecision.

1.5.3 Total versus free cortisol

As current cortisol immunoassays are only able to measure total serum cortisol (bound and unbound fractions), changes in CBG and, to a lesser extent, albumin concentration impact directly on cortisol measurement, making it an unreliable marker of adrenal function in patients with significantly altered protein concentrations (Coolens et al. 1987; Hamrahian et al. 2004; Ho et al. 2006; Vincent et al. 2009). Thus, adrenal status may be defined purely on the basis of a patient's clinical condition rather than true assessment of HPA axis function.

1.6 Assessing the adrenal glands in patients with acute disease

Cortisol production in the acutely unwell increases up to 6-fold due to a stress-induced increase in production of CRH, and consequently, ACTH, and reduced negative feedback from cortisol (Cooper and Stewart 2003). Loss of the diurnal pattern of cortisol secretion and a reduction in CBG concentration, leads to increased concentrations of circulating free corticosteroids. These changes are mediated by an increase in circulating cytokines, in particular tumour necrosis factor α, interleukin-1

and interleukin-6, which also act to increase cortisol at the tissue level by increasing glucocorticoid receptor affinity for cortisol and reducing peripheral cortisol metabolism (Chrousos 1995; Cooper and Stewart 2003).

Collectively, these changes are important for regulating the body's response to insult; however, they can also be directly impaired by the insult. For example, CRH and ACTH release may be impaired in patients with head injuries, a deranged clotting process may lead to adrenal haemorrhage, significant exogenous corticosteroid use may suppress the HPA axis, certain medications can inhibit enzymes involved in cortisol biosynthesis and can increase hepatic metabolism of cortisol, and excessive cytokine levels can suppress the adrenal cortex and cause end-organ resistance to cortisol (Cooper and Stewart 2003). Thus, patients with an acute illness can develop a functional adrenal insufficiency i.e. a transient (for the most part) episode of subnormal corticosteroid production.

Current serum cortisol assays can add further confusion to this already complex interplay between acute disease and adrenal cortex, making true assessment of adrenal function virtually impossible. In a study by Hamrahian *et* al, which considered the impact of low serum protein concentrations (defined as an albumin concentration below 25 g/L) in the critically ill on their response to the standard SST, measurement of ACTH and serum free cortisol identified an activated HPA axis in all patients (Hamrahian et al. 2004). However, in 14 out of 36 patients, the serum total cortisol response was consistent with adrenal insufficiency. This led the investigators to suggest using serum free cortisol as a marker of the HPA axis in critically ill patients, although evidence to show that this measurement correlates with clinical outcome any better than serum total cortisol is currently lacking.

In a subsequent study (Cohen et al. 2006), in which a standard SST was performed in 9 critically ill patients admitted to the general intensive care unit of a metropolitan

hospital, the extent to which assay differences influence interpretation of the Synacthen test in this population was clearly demonstrated. Cortisol was analysed by HPLC and 3 immunoassays, with concordant SST results across all assays in only 4 out of the 9 patients, leading the authors to suggest that unrecognised interassay variability is likely to have contributed to confusion over the definition of relative adrenal insufficiency in intensive care patients. They also warned that clinical management of such patients was being determined by the cortisol immunoassay used by their local laboratory rather than their actual adrenal status, and recommended identifying other, more appropriate, methods for evaluating the HPA axis in these patients.

Similar findings were reported in an offshoot study of the multi-national CORTICUS (corticosteroid therapy of septic shock) study – whose main aim was to evaluate the role of hydrocortisone treatment in patients with septic shock (Briegel et al. 2009). Serum total cortisol measurements from 12 different local cortisol assays were compared with the Roche Elecsys immunoassay at a central laboratory and an LC-MS/MS reference method. The correlation coefficient for the overall concordance between the central laboratory immunoassay and local assays was 0.60; falling to 0.54 post-Synacthen stimulation. When local assays were compared to LC-MS/MS, the corresponding coefficients ranged from 0.43 to 0.97. Furthermore, although a comparison of cortisol measurements from an outpatient population showed excellent correlation between the Roche Elecsys assay and LC-MS/MS reference method, this fell when a critically ill population was used instead. Overall, differences between central and local laboratory immunoassays resulted in misclassification of approximately 25% of patients and LC-MS/MS was recommended as the assay of choice measuring cortisol in the acutely ill.

1.7 Thesis aims

This study was conceived following discussions with local Endocrinologists who perceived that an increasing number of patients were being diagnosed with hypoadrenalism following a failed Synacthen test, despite a lack of convincing clinical features. Differences between cortisol assays were well recognised, as was the need for assay-specific cut-offs. The cortisol cut-off in use in the laboratory at the time was derived from the literature and, at 550 nmol/L was not too dissimilar to the 5th percentile ranges of 510 – 626 nmol/L suggested by Clark et al (Clark et al. 1998). However, the laboratory's Siemens assay was one of several new generation assays, which also included the Abbott, Immulite, Beckman and Roche, and consequently had not been included in the study by Clark et al. There was, thus, no real evidence supporting the validity of the cortisol cut-off quoted by the lab.

This led to the hypothesis that the laboratory's cut-off was wrong and was leading to patients being incorrectly diagnosed with hypoadrenalism when, in fact, their HPA axis was intact. There was concern, however, that should a new cut-off be established locally, this might not be applicable across all assay platforms, given what was already known about immunoassay performance, meaning a similar situation could arise following any future platform changes.

Furthermore, given the effect of altered CBG concentration on serum total cortisol, it was postulated that a cut-off established in patients with CBG concentrations within the normal range could not be applied to patients with high or low serum CBG concentrations, and separate validated cut-offs would be needed for these groups. Conversely, a validated free cortisol cut-off – be it salivary or in serum – would be expected to apply to all patient groups, irrespective of serum CBG concentration.

The aims of this study were, therefore:

- To establish valid serum total cortisol cut-offs for use with the short Synacthen test in patients with normal and altered CBG concentrations.
- To investigate, using current assays, the effect of assay differences on these cut-offs.
- To explore salivary free cortisol and calculated serum free cortisol as alternatives to serum total cortisol in interpreting the short Synacthen test in patients with altered CBG concentrations.

Its primary objectives were to:

- Define the distribution of, and effect of gender on, serum total cortisol measured by GC-MS following Synacthen stimulation in healthy volunteers and establish a reference range and lower reference limit for this response.
- Establish the lower reference limit for the cortisol response to Synacthen in healthy volunteers for five widely-used commercially available cortisol immunoassays.
- Define the salivary cortisol and calculated serum free cortisol responses to Synacthen in healthy volunteers and establish a reference range and lower reference limit for each.
- Apply these newly defined cut-offs in groups of patients with suspected adrenal insufficiency and high and low serum CBG concentration, respectively to determine their validity.

CHAPTER 2

SUBJECTS AND METHODS

2.1 Study Subjects

2.1.1 Healthy volunteers

2.1.1.1 Recruitment

Volunteers aged between 20 and 66 were recruited from staff at the University Hospital of Wales (UHW) and Cardiff University and their friends. Posters (appendix 1) were displayed around the hospital and University, and interested parties were given further information and invited to participate. On completion of their SST, volunteers were given copies of the information sheet (appendix 2) to pass on to colleagues and friends.

2.1.1.2 Exclusion and inclusion criteria

Criteria for exclusion included pregnancy and breastfeeding, significant intercurrent disease, a history of thyroid or other autoimmune disease, previous sensitivity to Synacthen, asthma or an allergic disorder and treatment with corticosteroids. To be included in the study subjects had to be in self-proclaimed good health, free of illness on the day of testing and not taking any drug therapy. In women of childbearing potential, pregnancy was excluded by urinary pregnancy test before participation in the trial.

Of 172 subjects interested in participating a total of 7 were excluded due to asthma (3), topical steroid use (1), history of allergy to aspirin and walnuts (1) and autoimmune hypothyroidism (1); one was excluded due to difficulty collecting baseline blood (1). Following entry into the study, two female volunteers, in whom baseline cortisol was measured but the SST could not be completed due to difficulty obtaining a post-Synacthen blood sample, were subsequently excluded from analysis.

The study protocol (appendix 3) was approved by the South East Wales Research Ethics Committee, Cardiff University (study sponsors) and the UK Competent Authority (the Medicines and Healthcare Products Regulatory Agency). All subjects gave written informed consent before study commencement.

2.1.1.3 Subject characteristics

Of the 163 volunteers recruited, 60 were male and 103 female (88 premenopausal and 15 postmenopausal). At the time of testing, twenty-four of the premenopausal female participants were taking an oestrogen-containing oral contraceptive pill (OCP) and were excluded from the female healthy volunteer group. Seventy-nine women were included in the female group. The mean age (and range) of the male and female groups was 37.1 (22-62) years and 40.7 (20-66) years, respectively.

2.1.1.4 Female volunteers taking exogenous oestrogens

There were 24 premenopausal women taking an oestrogen-containing oral contraceptive pill (OCP) containing between 20 and 35 micrograms of ethinyloestradiol. The mean age of these subjects was 28.7 years (range 21 to 40). To avoid the possibility of differences in cortisol concentration due to menstrual cycle phase, Synacthen testing in this group was undertaken during the follicular phase of the menstrual cycle.

2.1.2 Patients with suspected adrenal insufficiency

Thirty patients with established or suspected adrenal insufficiency were recruited from Endocrine clinics at the UHW. Inclusion and exclusion criteria were as for healthy volunteers, with the exception that a history of autoimmune disease or steroid treatment did not preclude participation in the study.

Of the thirty patients recruited, three were known to have Addison's disease and were on hydrocortisone and fludrocortisone replacement therapy; three had other

autoimmune disease including type 1 diabetes and recurrent hypoglycaemic episodes (2) and autoimmune hypothyroidism and vitamin B12 deficiency (1); ten had a pituitary adenoma and five of these had undergone pituitary surgery; two had undergone intracranial surgery or radiotherapy for other reasons; one had undergone adrenal surgery; four were on long term steroid treatment and seven had non-specific symptoms but adrenal insufficiency was suspected clinically (table 2.1). This group comprised 13 men and 17 women with a mean age of 52.4 years (range 23 – 82).

Results from patients with known Addison's disease and, hence, undetectable serum cortisol were excluded from calculations of the mean in all arms of the study to avoid introducing negative bias to comparisons between mean concentrations in patients with suspected hypoadrenalism and healthy volunteers.

Table 2.1 Summary of the clinical history and medication of patients with suspected adrenal insufficiency

Patient	Age	Gender	Clinical Details	Medication
1	67	F	Addison's + Hypothyroidism	Hydrocortisone 15mg/10mg; Fludrocortisone 25mcg; T4 100mcg
2	63	M	Addison's, T2DM, Prostate Ca	Hydrocortisone 30mg/20mg; Fludrocortisone 100 mcg; Avandamet
3	65	F	Pituitary Macroadenoma	
4	82	F	Non-functioning Pituitary Macroadenoma (surgery 2004), T2DM, IHD	
5	57	F	Fatigue, Asthma, Recurrent steroids	Seretide, Atrovent, Salbutamol
6	61	M	Non-functioning Pituitary Adenoma	
7	74	M	Non-functioning Pituitary Adenoma, IHD, Psoriasis	
8	46	F	Symptoms of IBS	
9	54	M	Hypogonadotrophic hypogonadism, Gynaecomastia, Normal MRI	
10	62	M	Pituitary adenoma invading L cavernous sinus, Transsphenoidal resection Jan 09, T2DM	Hydrocortisone 10mg/10mg; T4, Metformin, Simvastatin, Sustanon
11	64	M	L adrenalectomy for subclinical Cushing's - Jan 09, T2DM, UC, IHD, Recently stopped high dose steroids	Hydrocortisone 10mg/10mg; Metformin; Insulin; Azathioprine; Irbesartan; Amlodipine; Simvastatin; Salbutamol; Omeprazole; Diltiazem; GTN; Aspirin
12	40	F	Addison's, Graves' in remission	Hydrocortisone 20mg/5mg/5mg; Fudrocortisone 100mcg
13	35	M	Craniopharyngioma, Partial Hypopituitarism due to surgery, T2DM	T4; Tostran; GH; Desmospray; Metformin; Ramipril; Bisoprolol; Aspirin; Simvastatin; GTN spray
14	39	F	Dizziness, postural low BP	
15	43	M	T1DM, recurrent severe hypos, weight loss	Lantus and Novorapid
16	50	M	Non-functioning Pituitary Adenoma; Surgery Feb 09; Hypogonadotrophic hypogonadism	
17	64	M	Crohn's Disease + ileal resections, on + off steroids for 2 years	
18	47	F	Autoimmune hypothyroidism, B12 deficiency, tiredness	
19	47	F	Non-functioning pituitary microadenoma, Primary hypothyroidism	Thyroxine
20	46	M	Low random cortisol and testosterone, Normal Pituitary scan	Gabapentin; Simvastatin; Raniditine; Allopurinol; Oxycontin; Amitriptyline; Lansoprazole; Colofac; Testosterone patch
21	81	F	Adrenal suppression secondary to recurrent steroid treatment, COPD, IHD, Cerebrovascular disease, osteoporosis	Prednisolone 10m; Lansoprazole
22	70	F	Non-functioning pituitary adenoma, transsphenoidal hypophysectomy, post-op transient DI, primary hypothyroidism	
23	36	M	T1DM, recurrent severe hypos, steroids	Hydrocortisone 10mg/10mg
24	55	M	Partial hypopituitarism secondary to cranial radiotherapy, GH deficiency, primary hypothyroidism, T2DM	GH; andropatch, Irbesartan; Simvastatin; Mebeverine; Allopurinol
25	23	F	Pituitary Microadenoma, PCOS	Cilest
26	43	F	Non-functioning pituitary adenoma, transsphenoidal surgery 2006, probably MEN1, partial anterior hypopituitarism	GH; Simvastatin; Amlodipine
27	46	F	?Addison's disease	
28	54	F	Non-specific symptoms - aches, pains, headaches, fatigue, low mood, forgetful - ? Hypopituitarism	Gabapentin; Solpadol; Fluoxetine
29	28	F	Iatrogenic hypoadrenalism, previous steroid treatment for sarcoid	Hydrocortisone 10mg/10mg
30	29	F	Fatigue, dizziness	

A pre-test likelihood of hypoadrenalism was derived for each patient from the available clinical information. This was an arbitrary score, designed to stratify patients into categories of low or high clinical risk of disease, to which the outcome of the Synacthen test could be compared. Presenting features were split into 2 categories – clinical history and medication – and a risk score, based on my clinical judgment, was assigned to each (table 2.2).

Within each category, a score of 1 was defined as "no convincing clinical risk factors for hypoadrenalism" and a score of 5 as "definite hypoadrenalism". Scores in between the two extremes were defined as "possible, but unlikely" (2), "possible" (3) and "probable" (4) hypoadrenalism; although no criteria were assigned this latter score.

The scores from the two categories were then added to give the pre-test likelihood of disease, with an overall score of 5 or more indicating a high probability of disease, as it would comprise at least one definite or probable criterion or two possible criteria, although one of these could be "possible, but unlikely" (2). An overall score of 10 was considered "definite hypoadrenalism" and a score of 2 or less indicated a low probability of disease.

Where more than one criterion within the same category was present these were added together, e.g. treatment with both oral and inhaled corticosteroids would result in a medication score of 5, to which the clinical score would be added. One major weakness of this scoring system, however, was that patients with an overall score of 3 or 4 were considered to have an indeterminate risk. This arose as a result of my concern not to underestimate risk.

Finally, agreement between this score and the actual cortisol response to Synacthen was evaluated and the percentage agreement established.

Table 2.2 Criteria for establishing the pre-test likelihood of hypoadrenalism.

Criteria	**Score:**
Clinical:	
Pre-existing Addison's disease or hypopituitarism	5
Unilateral adrenalectomy	3
Pituitary adenoma with evidence of partial pituitary insufficiency	3
Pituitary macroadenoma without evidence of partial pituitary insufficiency	2
Symptoms of hypoadrenalism (3 or more)	3
Other autoimmune disease	3
None of the above	1
Medication:	
Replacement hydrocortisone &/or fludrocortisone	5
Other oral corticosteroids	3
Inhaled corticosteroids	2
Other drugs affecting the HPA axis	1
No relevant medication	1

The criteria for determining the pre-test probability of adrenal disease were derived from the clinical information provided in table 2.1. Each criterion was given a score between 1 and 5, based on personal judgement. A score of 1 was defined as "no convincing clinical risk factors for hypoadrenalism" 2 "hypoadrenalism possible, but unlikely", 3 "hypoadrenalism possible", 4 "hypoadrenalism probable" and 5 "definite hypoadrenalism".

Patients were followed up 5 years after undergoing their Synacthen test, using information extracted from clinic letters and biochemistry test results stored on the electronic clinical portal, to confirm that the test outcome had been correct.

2.1.3 Patients with low serum albumin

A total of 11 patients with either a recently diagnosed, untreated nephrotic syndrome (1) or established liver cirrhosis (10) were recruited from UHW renal and liver clinics. Inclusion and exclusion criteria were as for healthy volunteers except underlying renal or liver disease was a condition for inclusion. Patients were identified on the basis of a recent serum albumin concentration below the reference range of 35 – 50 g/L. Serum albumin concentration was repeated on the day of the Synacthen test and one of the 11 patients was subsequently excluded due to a serum albumin concentration within the normal range (38 g/L).

The final group comprised 7 men and 3 women with a mean age of 57.4 years (range 42 – 78) and mean albumin concentration 30.3 g/L (range 29 – 34).

2.2 Sample collection and processing

2.2.1 Short Synacthen test

The short Synacthen tests were carried out in the morning between 08.30 and 11.30 h. Subjects were not required to fast overnight, but were restricted from eating, drinking or smoking for the 30 minutes before the test. There were no restrictions on physical exercise prior to the test but participants were asked to rest in a sitting position for 15 minutes before the test began and then for the duration of the test. Women under the age of 40 were tested in days 1-7 of the follicular phase of their menstrual cycle, but in older women there were no particular timing requirements.

Once informed consent had been obtained, subjects were asked to collect a 5 mL saliva sample by passive drooling into a Universal container. An indwelling catheter was inserted into a superficial antecubital vein and 20 mL of blood were collected into plain tubes with no serum separator gel. A 250 µg bolus of synthetic $ACTH_{1-24}$ (Tetracosactide) (Synacthen, Alliance Pharmaceuticals Ltd, Wiltshire) was

administered intravenously and thirty minutes later a further 20 mL of blood were collected and subjects were asked to collect a second 5 mL saliva sample. This procedure was followed in healthy volunteers and patient groups.

2.2.2 Post collection sample handling

Blood samples were allowed to clot and following separation the serum was split into multiple aliquots; one of which was analysed immediately using the Siemens Centaur assay, to exclude adrenal insufficiency. The remainder were stored frozen, with the saliva samples, at -20°C. Subsequent sample processing was determined according to the group being studied. Thus, samples from healthy volunteers, who were recruited to allow the characteristics of baseline and post-Synacthen cortisol measurement to be identified, to examine the effect of assay on these characteristics and to define method-specific cortisol cut-offs post-Synacthen; were analysed by GC-MS and five automated immunoassays. Samples from patients with suspected hypoadrenalism, who were recruited to enable comparison between cortisol concentrations in a healthy population and in patients with a clinical indication for testing the adrenal axis, were analysed by a single automated immunoassay (Abbott Architect). Samples from patients with low serum albumin, who were recruited to explore the effect of altered protein concentration on cortisol measurement, were analysed by GC-MS and the Abbott Architect immunoassay.

Samples from healthy volunteers were removed for analysis in batches over a 10 week period (figure 2.1). This included repeat analysis on the Centaur to exclude sample degradation, whilst the remaining aliquots were sent to external laboratories for analysis according to the protocol (appendix 4). Samples from patients with established or potential adrenal disease were subsequently analysed in a single batch in the UHW Biochemistry laboratory using the Abbott Architect assay. Samples from

patients with low serum albumin were analysed in single batches by GC-MS and the Abbott Architect assay.

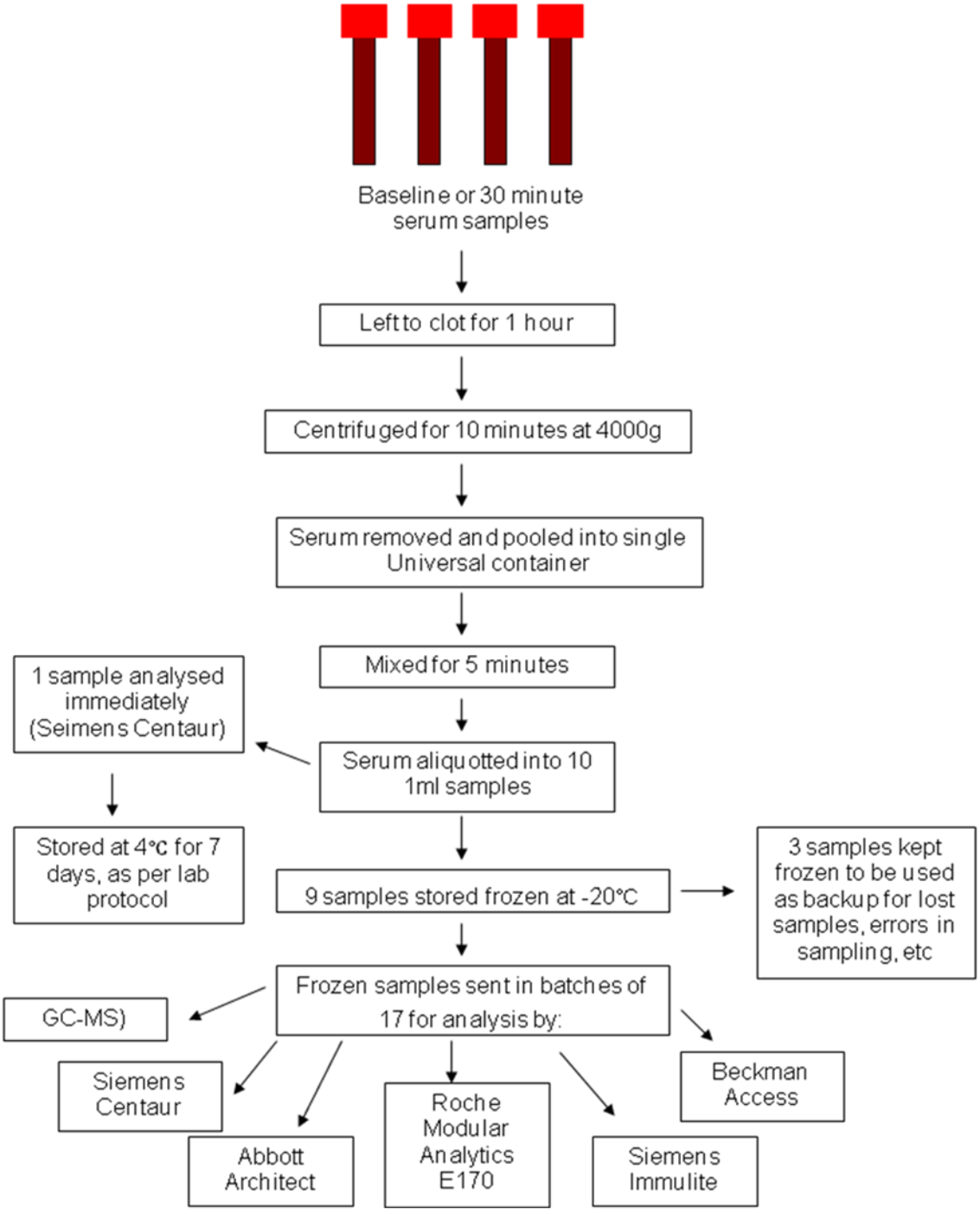


Figure 2.1 Description of processing and analysis of samples from healthy volunteers. Following storage, patient samples were analysed either by the Abbott Architect alone or by the Abbott Architect and GC-MS assays (see text).

2.2.3 Assay reformulation

During the course of the study, Siemens announced a minor reformulation of their Advia Centaur cortisol assay. This practice is not unusual and in the context of cortisol assays is likely to represent an attempt to align them more closely with an LC-MS/MS reference method. Because of the timing of the assay change it was possible to study its impact on cortisol measurement directly. Stored aliquots of serum from healthy volunteers were analysed in batches on the new Centaur assay for comparison to the Centaur results from the start of the study.

2.3 Analytical methods

2.3.1 Serum cortisol

Total cortisol from healthy volunteers, including women taking an OCP, was measured by the Welsh External Quality Assurance Scheme (WEQAS) Reference Laboratory using a modified version of their gas chromatography-mass spectrometry (GC-MS) reference method and by five automated immunoassays: the Advia Centaur (Siemens AG, Erlangen, Germany), Modular Analytics E170 (Roche, GmbH, Mannheim, Germany), Immulite 2000 (Siemens AG, Erlangen, Germany), Beckman Access (Beckman Coulter Inc., Brea, CA) and Abbott Architect (Abbott Laboratories, Illinois). Total cortisol in samples from patients with low serum albumin was measured by GC-MS and the Abbott Architect assay and from patients with suspected adrenal disease by the Abbott Architect assay only.

2.3.1.1 GC-MS cortisol[6]

The assay used was a modified version of the reference method developed and validated by the WEQAS Reference Laboratory, and differs in that it uses a

[6] Analysis performed by staff at the WEQAS reference laboratory

conventional 6-point calibration curve instead of the bracketed standard curve used in the reference method. It is described in some detail, below.

Samples and quality controls were thawed at room temperature and mixed. A 0.5 mL aliquot of each was dispensed in duplicate into 2.5 mL of 1% Sodium Chloride (NaCl) solution and deuterated cortisol was added as internal standard. Samples were then mixed and stored overnight at 4°C. The following day, cortisol was extracted with dichloromethane (DCM, 7 mL) and mixed vigorously. The tubes were centrifuged at 2500 rpm for 20 minutes at 10°C, and the upper aqueous layer discarded. The organic phase was decanted into labelled glass tubes, placed in a heating block and evaporated to dryness under a gentle stream of nitrogen at 30°C.

The dried extract was reconstituted in 200 µL of a chloroform/cyclohexane/methanol column solvent mix and further purified using LH-20 chromatography. LH-20 columns were constructed using a glass Pasteur pipette with a loose glass wool plug at the narrow end, and packed with LH-20 particles that had been soaked overnight at room temperature in the chloroform/cyclohexane/methanol (100:80:15, v/v) solvent mix. The columns were washed with 3 mL of column solvent and the reconstituted sample was applied. Following a final column wash with 1.5 mL of solvent mix, the cortisol was eluted into reaction vials containing 60 µg of 5α-androstane to reduce deterioration of cortisol and the samples were evaporated to dryness.

Derivatisation of cortisol to its heptafluorobutyric ester was initiated by reconstituting the dried extract with 25 µL each of acetone and heptafluorobutyric anhydride (HFBA) and mixing thoroughly. Samples were derivatised for 45 minutes, dried under nitrogen, reconstituted with 50 µL cyclohexane and thoroughly mixed before GC-MS analysis.

The GC-MS instrument was an Agilent 5973 mass spectrometer coupled with a 6890 gas chromatograph. The GC column was 15 m long with 0.32 mm internal diameter

and 0.25 µm film thickness (Agilent technologies). The derivatised cortisol was injected via a split/splitless ptv injection system at 75°C, ramped up to a final temperature of 270°C, held for 1 minute. Initial oven temperature was 100°C ramped to a final temperature of 290°C. Helium was used as the carrier gas at a constant flow with a column head pressure of 35 kPa (5 psi). The MS was operated under electron impact (70 eV) and selective-ion monitoring mode, monitoring at mass to charge ratios (*m/z)* of 489 and 491 for cortisol and d_2-cortisol respectively. Retention time of derivatised cortisol was 7.3 minutes. Data acquisition and quantitation of cortisol levels were achieved using Chemstation software.

High, medium and low quality control samples (average concentrations: 657 nmol/L, 363 nmol/L and 120 nmol/L respectively) were included in every run and processed identically to study samples.

2.3.1.2 Automated cortisol immunoassays[7]

For ease of reference the assays used will be referred to by manufacturer name throughout the text, while the two Siemens assays will be referred to as Centaur and Immulite respectively. All assays were solid-phase, competitive binding immunoassays with chemiluminescent detection, except the Roche which uses electrochemiluminescent detection. The Centaur and Abbott analyses were both undertaken at the University Hospital of Wales' Biochemistry laboratory which underwent a change of equipment during the study. The Roche, Immulite and Beckman analyses were undertaken at the Biochemistry laboratories at Prince Charles Hospital, Merthyr Tydfil, Bristol Royal Infirmary, Bristol and Southampton General Hospital, Southampton, respectively. All participating laboratories were enrolled in the UKNEQAS external quality assurance scheme at the time of analysis and their cortisol assays were performing well within their method groups (Table 2.3).

[7] Analysis performed by laboratory staff at the respective hospitals

Table 2.3 Cortisol assay EQA performance.

	Centaur	Abbott	Roche	Immulite	Beckman
UKNEQAS ABC Scores	A 64 B +5.1 C 6.1	A 93 B -8.5 C 8.0	A 65 B +5.7 C 6.6	A 82 B +0.4 C 10.6	A 133 B -9.9 C 10.6

UKNEQAS ABC scores for each of the cortisol assays included in the study.
A score = transformed accuracy score which includes a degree of difficulty factor to make it comparable across different analytes. For most analytes an A score below 100 indicates good performance.
B score = bias score which assesses the assay's bias (and its direction) relative to the target value.
C score = consistency of bias score i.e. are the magnitude and direction of the assay's bias consistent?
Data provided by each participating laboratory before analysis began and reproduced with kind permission of F MacKenzie, UKNEQAS.

Batches, comprising samples from a maximum of 17 subjects (i.e. a total of 34 samples, as each subject had a baseline and post-Synacthen sample) and three internal quality control (iQC) serum pools with GC-MS assigned target concentrations of 76, 528 and 696 nmol/L respectively, were sent to participating laboratories for analysis once a week over 10 weeks. Analysis of the three iQC pools over the 10-week period provided the data needed to determine the interassay (between batch) precision for each assay (Table 2.4).

Table 2.4 Assay precision.

	GC-MS Target [Cortisol] nmol/L			Mean [Cortisol]	CV%	Number of replicates
	76.1	**527.7**	**696.4**			
	Assay mean [Cortisol] nmol/L					
Centaur	121.3	646.4	891.1	NA	5.0 4.4 2.1	10 10 10
Roche	116.3	612.7	818.1	NA	9.8 6.1 6.6	10 10 10
Immulite	116.7	694.8	978.0	NA	10.2 7.7 7.2	9 9 9
Beckman	124.4	619.6	827.7	NA	9.9 5.0 2.5	10 10 9
Abbott	NA	NA	NA	96.5 549.2 840.1	10.4 5.4 6.8	135 131 146

Interassay CVs for the Centaur, Roche, Immulite and Beckman assays were established using pooled serum samples with GC-MS-assigned target concentrations, kindly provided by WEQAS. Interassay CV for the Abbott assay was determined from the mean and SD of multiple runs of 3 single serum samples with unknown cortisol concentration.

2.3.2 Serum CBG[8]

The serum CBG assay is a manual solid-phase, competitive binding radioimmunoassay from Biosource (Nivelles, Belgium). ^{125}I-labelled CBG competes with CBG in the sample or calibrator for binding sites on mouse anti-CBG antibodies which are immobilised onto the walls of specially designed polystyrene tubes by goat anti-mouse antibodies. The radioactive signal produced by bound labelled CBG is measured using a gamma counter, and is inversely proportional to the concentration of CBG in the sample. Prior to analysis, samples require a 25-fold dilution and this was done by adding 100 µL of sample to 2.4 mL of the dilution buffer provided and mixing well. Diluted samples, calibrators and controls were run in duplicate and 100

[8] Analysis performed by Dr Atilla Turkes, Mr Alan Pickett and myself.

µl of each was dispensed into the antibody-coated tubes. 100 µL of ^{125}I-labelled CBG was added to each tube, followed by 100 µL of anti-CBG anti-serum. The reaction mixture was incubated for 2 hours at room temperature, with continuous shaking at 400 rpm. The liquid content of each tube was aspirated and the tubes were washed with 2 mL of wash buffer and aspirated to dryness again. The tubes were counted in a gamma counter for 60 seconds.

In addition to the samples, calibrators and controls, two plain plastic tubes containing 100 µl of ^{125}I-labelled CBG were used to determine the total radioactivity count and a "blank" coated tube was included to determine any non-specific binding (NSB)[9]. Included with the kit's calibrators was a "zero calibrator", also known as the "maximum binding" tube which contains both ^{125}I-labelled CBG and anti-CBG anti-serum but no competing CBG. This makes it possible to determine the maximum amount of binding that can occur between the radioactive CBG and its antibody in the absence of competition, and demonstrates that there is a sufficient amount of antibody in the reaction mixture.

Assay performance was deemed acceptable, with intra- and inter-assay CVs of 7.6% and 12.8% respectively at a concentration of 30 mg/L and 3.1% and 8.7% respectively at a concentration of 110 mg/L.

2.3.3 Salivary cortisol

Salivary cortisol was measured by Dr Atilla Turkes and Mrs Sarah Tenant from UHW Biochemistry using an in-house LC-MS/MS method. A 250 µL aliquot of saliva, containing 5 nmol/L deuterated cortisol was extracted with 2 mL of dichloromethane. The tubes were centrifuged for 5 min at 4000 rpm and the top aqueous layer was

[9] Non-specific binding is measured as the amount of radioactivity present in a control or "blank" coated tube which has been handled in an identical way to the sample tubes but contains only ^{125}I-labelled CBG and the reaction buffer, but no competing CBG and no anti-CBG antibodies. It is used to identify any direct, non-antibody mediated binding of the label to the walls of the reaction tube or to any components of the reaction buffer.

discarded. The solvent phase was evaporated under a gentle stream of nitrogen and the dried extract was reconstituted with 250 µl of mobile phase. A 20 µl volume of this extract was injected into the LC-MS/MS instrument for analysis.

The LC-MS/MS instrument was a Premier XE triple quadrupole tandem mass spectrometer (Micromass MS Technologies, Manchester, UK) with an Acquity ultra-performance liquid chromatography (UPLC) system comprising a binary pump and auto-sampler (Waters Ltd, California, USA). The LC column was a silica-based reverse-phase C18 (1.7µm, 2.1x50 mm) column (Waters Ltd) and the chromatographic mobile phases were composed of two solutions; (A) deionised water containing 2 mmol/L ammonium acetate and 0.1% v/v formic acid and (B) methanol containing 2 mmol/L ammonium acetate and 0.1% v/v formic acid. The mobile phase was delivered, after sample injection, at an initial flow rate of 0.40 mL/min. The retention time for cortisol and d_4-cortisol was 0.95 min and the analysis time for each sample was 4.5 min.

The MS/MS was operated with electrospray ionisation (ESI) source and Z-spray interface and selected reaction monitoring mode, monitoring at a mass to charge ratio (*m/z)* of 363.3 transitioning to 121.1 (363.3>121.2) for cortisol and 365.3 to 121.2 (365.3>121.2) for *d2*-cortisol. Data acquisition and quantitation of cortisol levels were achieved using MassLynx NT and QuanLynx (Waters Ltd.) software, respectively.

The intra- and inter-assay CVs were 5.6% and 6.0% respectively at a concentration of 1.2 nmol/L, 2.3% and 5.8% respectively at 5.4 nmol/L and 3.0% and 3.8% respectively at 15.1 nmol/L.

2.4 Statistics

Statistical analyses were performed using SPSS versions 16.0, 19.0 and 23.0 (SPSS Inc., Chicago, Illinois and IBM Corporation, New York). The Kolmogorov-Smirnov test was used to determine whether data were normally distributed and, for

consistency, given that not all data were normally distributed, all were log transformed before analysis. Calculations of the mean and reference ranges were carried out on the log transformed data and were transformed back to give meaningful results.

Comparisons between means were made using the paired and unpaired t-tests and Mann-Whitney U test, where data remained non-parametric following log transformation. In all cases, a p-value <0.05 indicated a significant difference. Specific details of the analyses carried out are included in the relevant chapters.

CHAPTER 3

DETERMINING THE METHOD-SPECIFIC SERUM CORTISOL RESPONSE TO SYNACTHEN

3.1 Introduction

The concept of using assay-specific cut-offs to interpret the cortisol response to Synacthen is not a new one (Clark et al. 1998; Klose et al. 2007). However, earlier studies have either been limited to a small number of immunoassays (Klose et al. 2007) or have used assays that are now obsolete (Clark et al. 1998). Thus, we do not have validated cortisol reference ranges or cut-offs for the majority of current immunoassays, despite evidence that we should. Differences in the characteristics of the cortisol response to Synacthen, including gender effect and distributional form, have also been identified but remain unexplained.

Clinically, there is growing concern that current cut-offs have resulted in overdiagnosis of adrenal insufficiency in patients with few, if any, definitive features of the disease. The need to review immunoassay measurement of cortisol in general and the cortisol response to Synacthen, specifically, is therefore clearly established from both a clinical and a laboratory perspective.

3.2 Aims

The work described in this chapter was undertaken to:

- Clarify the distributional pattern of serum cortisol measurements in men and women and identify any significant gender differences in concentration.
- Define the cortisol response to Synacthen in healthy volunteers using a modified version of the GC-MS reference method.
- Compare cortisol measurement by GC-MS to its measurement by five commercially available cortisol immunoassays to identify differences and quantify assay bias for each.
- Establish the lower reference limit for the cortisol response to Synacthen in the immunoassays studied.

- Explore the validity of an assay-specific lower reference limit for cortisol in a cohort of patients with suspected hypoadrenalism.

3.3 Subjects and sample analysis

3.3.1 Healthy volunteers

This group consisted of 60 male and 79 female volunteers, as described in chapter 2. Serum total cortisol in this group was measured by GC-MS and the Centaur, Abbot, Roche, Immulite and Beckman immunoassays.

3.3.2 Patients with suspected hypoadrenalism

This group consisted of 13 male and 17 female recruits, whose clinical details are summarised in table 2.1. Serum total cortisol in this group was measured by the Abbott immunoassay only.

3.4 Statistics

3.4.1 Samples from healthy volunteers

All results were log-transformed before calculating mean cortisol concentrations and 2.5^{th} to 97.5^{th} percentile ranges for each assay at each time point. A lower reference limit (LRL) for the cortisol response to Synacthen, equivalent to the 2.5^{th} percentile result, was calculated for each assay from the mean cortisol concentration at 30 minutes, using the equation "mean – 1.96*SD". For the purposes of this study, the LRL was used as the cut-off for differentiating between glucocorticoid sufficiency and deficiency.

An assay-specific bias ratio, indicating the magnitude and direction of each immunoassay's bias relative to GC-MS, was calculated for each sample (at each time point) by dividing the untransformed immunoassay result by the matched GC-MS result, and a mean value was calculated. Baseline and post-Synacthen results were

combined before the correlation between each immunoassay and GC-MS was studied using scatter plots and bias ratios.

Differences between method means and bias ratios were compared using the paired t-test. Gender differences were evaluated using the unpaired t-test, or the Mann-Witney U test in cases of non-parametric data. In all cases a p-value <0.05 was considered statistically significant.

3.4.2 Assay reformulation study

Mean cortisol concentration, overall assay bias and a lower reference limit for the cortisol response to Synacthen were calculated for the new Centaur assay as described above, and compared to the results from the old assay using the paired t-test.

3.4.3 Comparison to patients with established or potential adrenal disease

Mean cortisol concentration was calculated at baseline and post-Synacthen in the patient group and compared to the respective mean concentrations in healthy volunteers using the unpaired t-test. Patients whose baseline serum cortisol was undetectable were excluded from calculations of the mean.

3.5 Results

3.5.1 Defining cortisol characteristics using GC-MS

3.5.1.1 At baseline

Baseline cortisol was normally distributed in male but not in female volunteers, in whom concentrations were skewed to the right (figure 3.1). Further examination of the 8 women with the highest cortisol concentrations showed an over-representation of post-menopausal women; 37.5% (3/8) compared to 19% (15/79) overall. The mean age of this group was also significantly higher than that of the female group as a whole (43.5 vs 40.3 years, respectively). There was no significant gender

difference, however, (table 3.1) and on combining the male and female groups, cortisol concentrations were found to be independent of age (R^2 = 0.003, p-value = 0.26). Baseline cortisol concentration, measured by GC-MS, ranged between 92 and 603 nmol/L in males and 117 and 536 nmol/L in females.

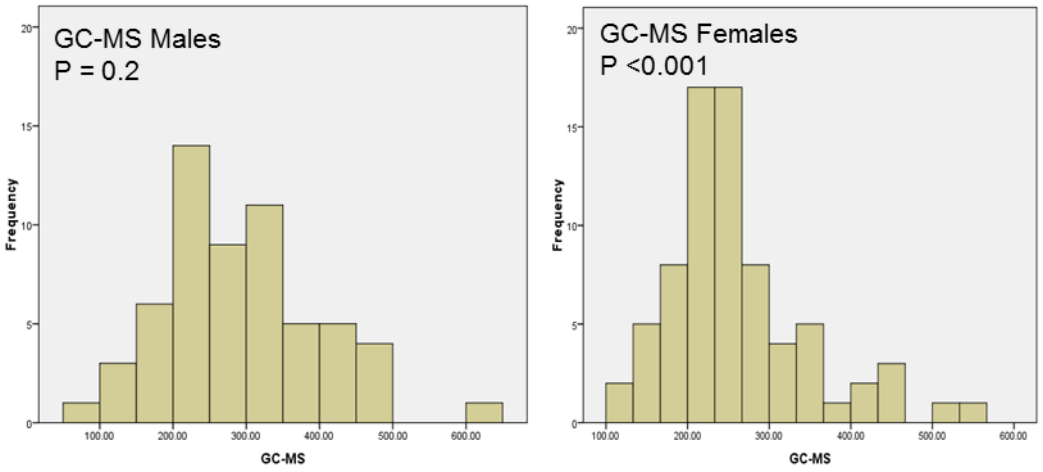


Figure 3.1 Distribution of baseline serum total cortisol concentration in males and females. The distribution was examined using the Kolmogorov-Smirnov test, with a P-value >0.05 indicating a normal distribution.

Table 3.1 Assay-specific baseline serum total cortisol concentration in healthy volunteers.

	Baseline mean [cortisol] (2.5th – 97.5th centile) nmol/L		
Assay	**Males n = 60**	**Females n = 79**	**P-value***
GC-MS	274 (131 - 575)	254 (139 - 463)	0.193
Centaur	298 (158 - 565)	257 (138 - 477)	0.023*
Abbott	289 (151 - 556)	247 (134 - 455)	0.018*
Roche	370 (182 - 750)	292 (147 - 581)	0.001*
Immulite	316 (165 - 604)	267 (144 - 495)	0.003*
Beckman	293 (160 - 538)	252 (143 - 444)	0.011*

* P-value <0.05 indicates a significant difference between gender means.

3.5.1.2 In response to Synacthen stimulation

After Synacthen stimulation, GC-MS cortisol remained normally distributed in male but not in female volunteers, in whom concentrations remained skewed to the right

(figure 3.2). There was no significant difference in cortisol concentration between genders (table 3.2) and concentrations remained unaffected by age (R^2 = 0.14; p = 0.11). Post-Synacthen cortisol concentrations ranged between 404 and 778 nmol/L in males and 388 and 824 nmol/L in females. As anticipated, stimulated mean cortisol concentration was significantly higher than mean baseline cortisol in both male and female subjects (p < 0.001; figure 3.3).

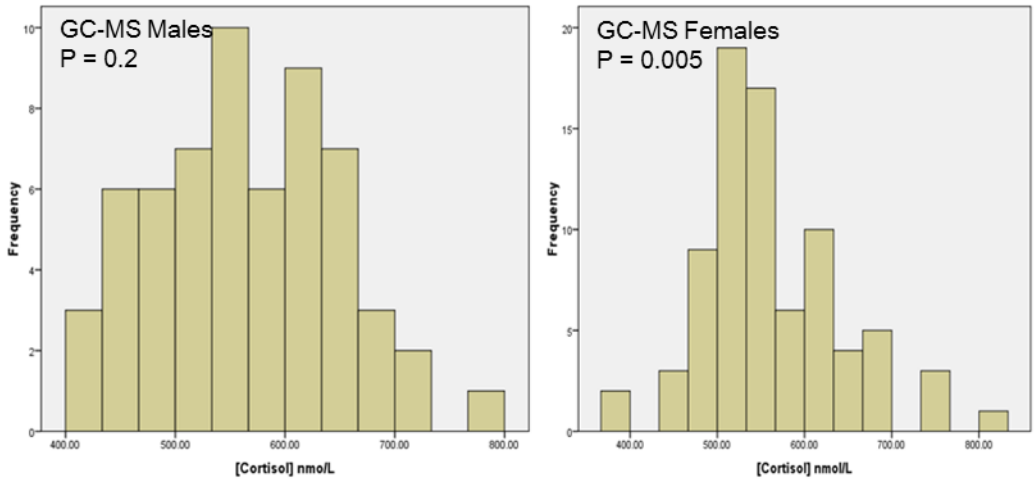


Figure 3.2 Distribution of post-Synacthen serum total cortisol concentration in males and females. The distribution was examined using the Kolmogorov-Smirnov test, with a P-value >0.05 indicating a normal distribution.

Table 3.2 Assay-specific post-Synacthen serum total cortisol concentration in healthy volunteers.

	Post-Synacthen mean [cortisol] (2.5th – 97.5th centile) nmol/L		
Assay	**Males n = 60**	**Females n = 79**	**P-value**
GC-MS	563 (418 - 757)	555 (421 - 731)	0.594
Centaur	599 (448 - 802)	578 (446 - 750)	0.138
Abbott	577 (430 - 773)	542 (416 - 707)	0.012*
Roche	772 (574 - 1039)	712 (524 - 967)	0.003*
Immulite	641 (469 - 874)	628 (478 - 826)	0.449
Beckman	625 (459 - 852)	594 (455 - 777)	0.045*

*P-value <0.05 indicates a significant difference between means.

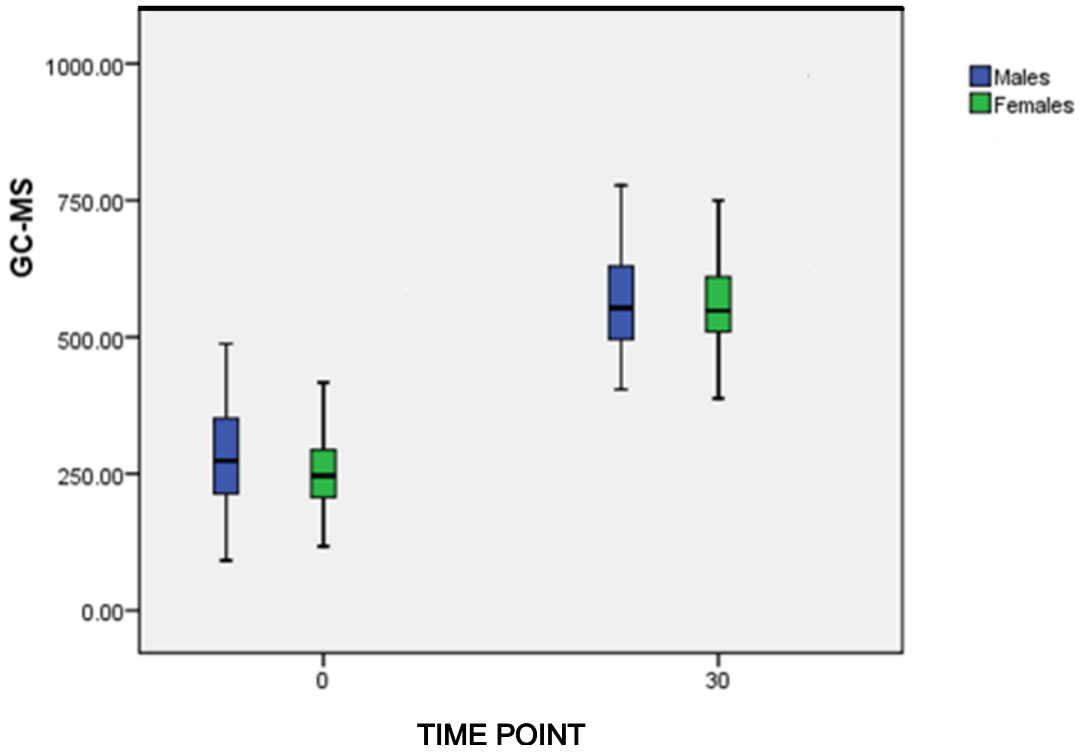


Figure 3.3 Mean GC-MS serum total cortisol concentration at 0 and 30 minutes by gender. There was no significant gender difference at either timepoint, but mean cortisol concentration was significantly lower at baseline (0) than post-Synacthen (30) ($p < 0.001$).

3.5.2 The effect of assay on cortisol characteristics

3.5.2.1 At baseline

The normal distribution of baseline cortisol concentrations in males persisted for all immunoassays apart from the Immulite and, in line with GC-MS results, baseline cortisol in females was not normally distributed with any immunoassay. In contrast to GC-MS, all immunoassays showed a statistically significant gender difference in mean baseline cortisol concentration, with significantly higher concentrations in male subjects (table 3.1). The difference between GC-MS mean cortisol concentration and immunoassay cortisol was statistically significant in males for all assays and in females for all but the Abbott and Beckman assays (figure 3.4).

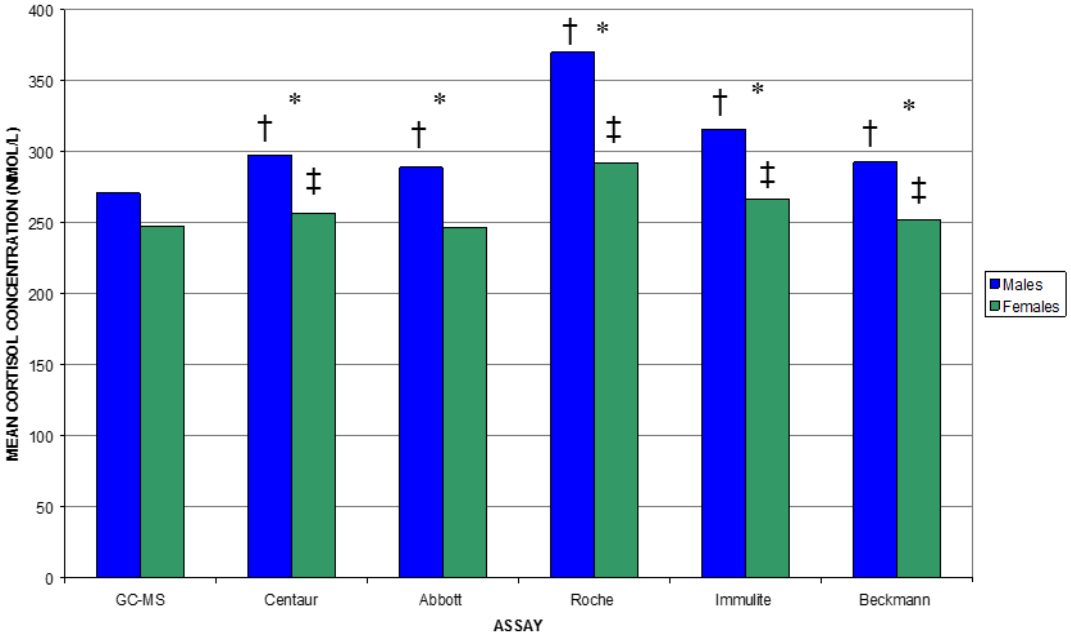


Figure 3.4 Mean assay-specific baseline serum total cortisol concentrations in males and females**.** * denotes statistically significant (P<0.05) gender differences within an assay, † denotes a statistically significant difference (P<0.05) from GC-MS cortisol in males; ‡ denotes a statistically significant difference (P<0.05) from GC-MS cortisol in females.

3.5.2.2 In response to Synacthen stimulation

When measured by immunoassay, post-Synacthen cortisol remained normally distributed in males and, in contrast to its measurement by GC-MS, became normally distributed in females with all immunoassays apart from the Centaur. The gender difference identified in baseline cortisol persisted in post-stimulation cortisol concentrations with the Abbott, Roche and Beckman assays, with a significantly higher cortisol response in males (table 3.2). For all assays, the mean cortisol concentration differed significantly from GC-MS cortisol in both male and female subjects, although the clinical significance of this difference is arguable for the Centaur and Abbott assays (figure 3.5).

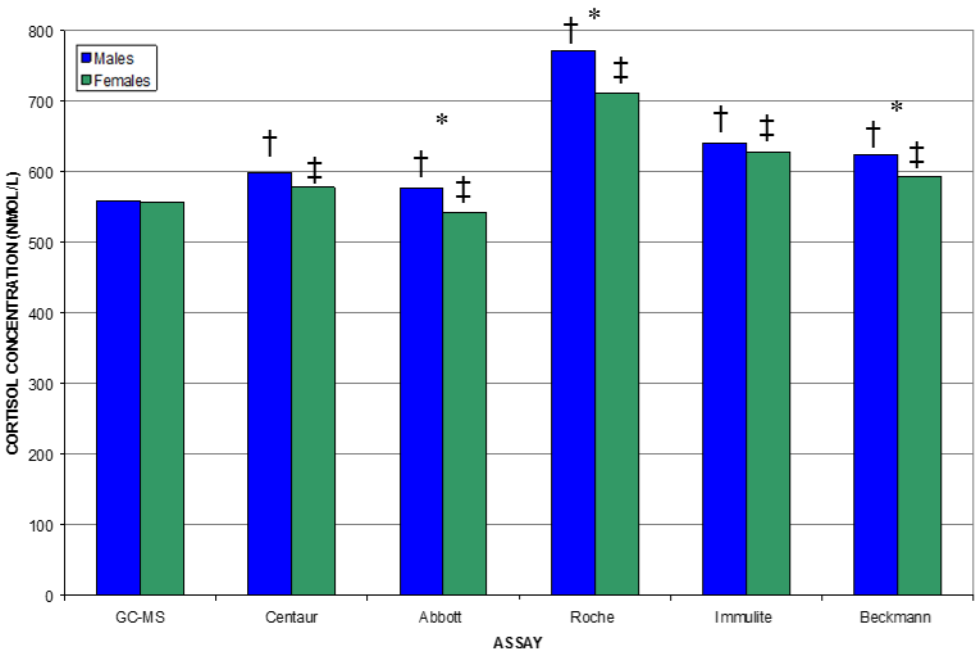


Figure 3.5 Mean assay-specific post-Synacthen serum total cortisol concentrations in males and females**.** * denotes statistically significant (P<0.05) gender differences within an assay, † denotes a statistically significant difference (P<0.05) from GC-MS cortisol in males; ‡ denotes a statistically significant difference (P<0.05) from GC-MS cortisol in females.

3.5.3 Correlation between immunoassay and GC-MS

3.5.3.1 Correlation plots

There was good correlation between immunoassay and GC-MS cortisol measurements for both baseline and post-Synacthen samples. The data were, therefore, combined to create a single correlation plot for each assay, covering a range of cortisol concentrations between 92 and 824 nmol/L (figure 3.6).

The correlation plots show there was a slight gender effect on the relationship, with samples from male subjects showing slightly higher positive bias than samples from female subjects for all assays apart from the Architect for which there was virtually no bias for male samples, and a small negative bias for samples from females. Bias also appear to be influenced to some extent by cortisol concentration, with the line of best

fit showing overall negative bias at higher concentrations with the Abbott assay and marginally less bias for the Centaur, Immulite and Beckman assays. The Roche assay showed positive bias across the entire concentration range.

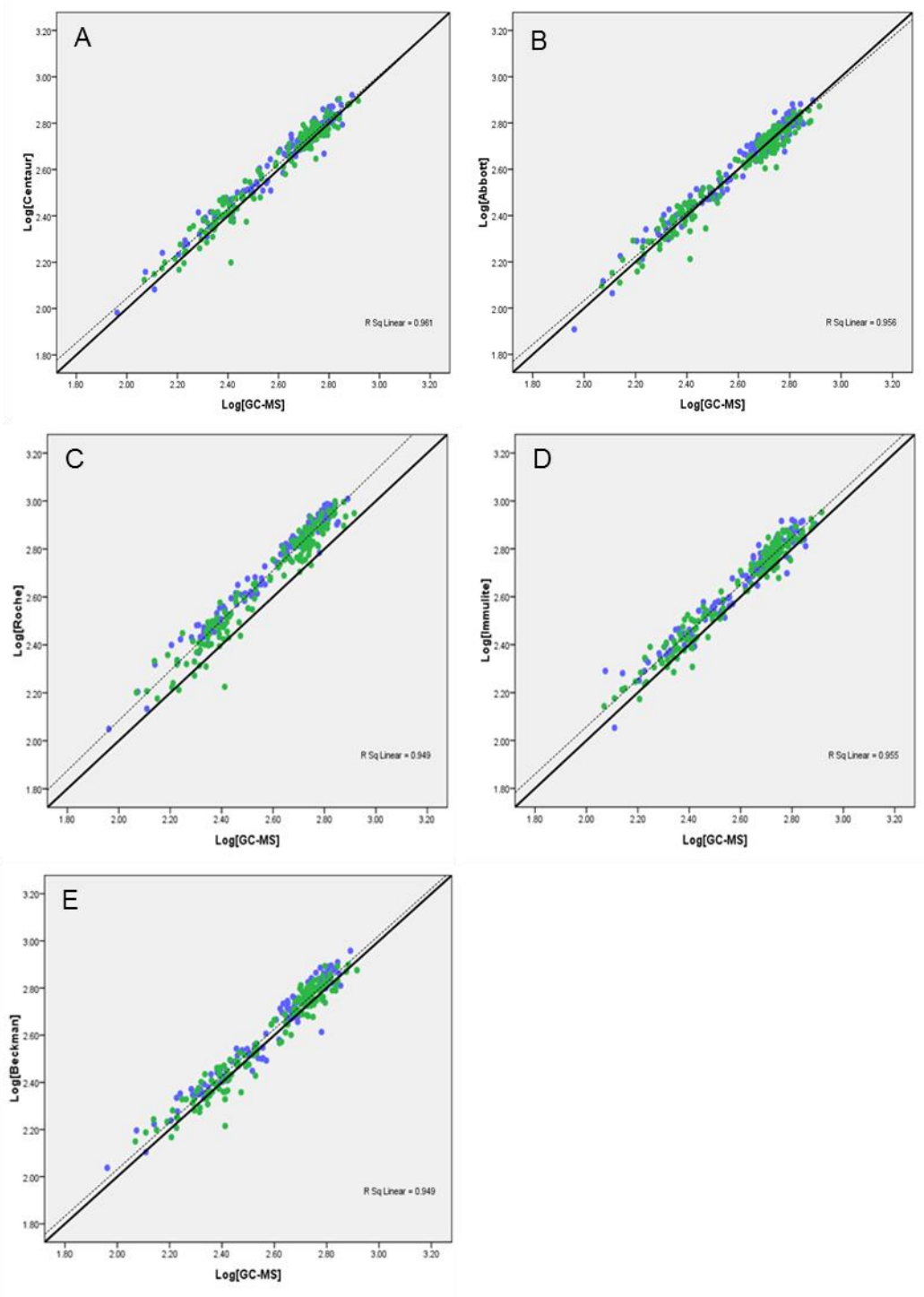


Figure 3.6 Correlation between log-transformed immunoassay and GC-MS cortisol in healthy volunteers. A) Centaur, B) Abbott, C) Roche, D) Immulite and (E) Beckman assays. The solid line represents equivalence between methods; the dotted line is the line of best fit. ● = males; ● = females.

3.5.3.2 Bias ratios

A single mean bias ratio was calculated from all male and female samples at baseline and post-Synacthen stimulation. The overall bias ratio was positive for all assays and greatest in the Roche and Immulite assays, reaching almost 30% for the Roche (table 3.3). Overall bias for the Centaur, Abbott and Beckman assays was less marked, ranging from 2 to 7%.

Table 3.3 Mean assay bias ratios for male and female subjects combined.

	Mean Bias Ratio		
Assay	**Overall**	**0 Minute**	**30 Minute**
Centaur	1.06	1.07	1.05
Abbott	1.02	1.03	1.00
Roche	1.30	1.27	1.33
Immulite	1.13	1.12	1.14
Beckman	1.07	1.04	1.09

Overall mean bias ratio was calculated using both male and female baseline (0 minute) and post-Synacthen (30 minute) cortisol concentrations.

Assay bias was affected by cortisol concentration, with greater bias noted at lower concentrations (baseline measurements) for the Centaur and Abbott assays, and the opposite effect (greater bias at higher concentrations) for the Roche, Immulite and Beckman assays (table 3.3). This concentration-dependent bias difference was statistically significant for all assays other than the Centaur, although its clinical significance is less clear as it ranged from as little as 2% for the Immulite assay to a maximum of 6% for the Roche.

Assay bias was also affected by gender, with greater positive bias in males than in females (table 3.4). This difference was statistically significant for all assays; despite its small magnitude (3-4%) in all apart from the Roche. For the Roche, the difference between genders was 11% which is likely to be of clinical, as well as statistical, significance.

Table 3.4 Mean assay bias ratio by gender.

	Overall Mean Assay Bias Ratio		
Assay	**Males**	**Females**	**P-value**
Centaur	1.08	1.05	P=0.008
Abbott	1.04	1.00	P<0.001
Roche	1.36	1.25	P<0.001
Immulite	1.15	1.11	P=0.011
Beckman	1.09	1.05	P=0.001

Gender-specific assay bias ratios were calculated using baseline (0 minute) and post-Synacthen (30 minute) cortisol concentrations combined.

3.5.4 Assay-specific lower reference limits for the cortisol response to Synacthen

The lower reference limit for the cortisol response to Synacthen was defined by GC-MS as 418 nmol/L in males and 421 nmol/L in females. A gender-specific lower reference limit was also determined for each immunoassay (Table 3.5).

Table 3.5 Assay-specific lower reference limits for post-Synacthen cortisol by gender:

Assay	Males	Females
GC-MS	418	421
Centaur	448	446
Abbott	430	416
Roche	574	524
Immulite	469	478
Beckman	459	455

Lower reference limit of cortisol was calculated from log-transformed data as mean – 1.96*SD.

Since mean cortisol concentration post-Synacthen was not gender-dependent for the GC-MS, Centaur and Immulite (2000) assays (table 3.2) a single lower limit was calculated by combining the males and females and deriving a single mean and lower reference range. This was not possible for the Roche, Abbott and Beckman assays for which mean cortisol concentration post-Synacthen was gender dependent.

However, the difference between the lower reference limit for cortisol in males and females for the Abbott and Beckman assays (14 nmol/L and 4 nmol/L, respectively) fell within their precision limits (CVs of 5.4% and 5.0% respectively at 549 and 620 nmol/L (table 2.4)), indicating that gender-specific reference ranges would not be

necessary for these assays in clinical practice. Thus, a single lower limit was derived using the higher of the two gender limits (in both cases, the male limit). For the Roche assay, gender-specific lower limits were retained, given the 50 nmol/L difference between the male and female lower limits (table 3.6).

Table 3.6 Assay-specific estimated lower reference limits for post-Synacthen cortisol in healthy adult volunteers.

Assay	Males	Females	Adults
GC-MS	417	422	420
Centaur	448	446	446
Abbott	430	416	430
Roche	574	524	NA
Immulite	469	478	474
Beckman	459	455	459

For the GC-MS, Centaur and Immulite assays, male and female means were combined to allow calculation of a single adult lower limit. For the Abbott and Beckman assays, both with significantly different gender means, the male lower limit, which was the higher of the two, was used as applicable to both.

3.5.5 The effect of assay reformulation on cortisol measurement

The results reported above are those from the reformulated Centaur assay. In contrast to the new assay, baseline and post-Synacthen cortisol concentrations measured by the old assay were normally distributed about the mean in both female and male subjects. The old assay detected a statistically significant gender difference in baseline cortisol concentrations which was also apparent with the new assay (figure 3.4). Statistically significant differences in mean cortisol concentration at baseline and post-Synacthen, in assay bias and in the lower reference limit for the cortisol response to Synacthen were also detected between the two assays (table 3.7).

Table 3.7 Changes in serum total cortisol measurement following reformulation of the Siemens Centaur assay.

	Mean [cortisol] (nmol/L)				Overall assay bias ratio	Lower reference limit		
	Baseline		Post-Synacthen					
	Males	Females	Males	Females		Males	Females	Adults
Old Centaur	336	289	656	638	1.18	496	502	499
New Centaur	299	257	599	578	1.06	448	446	446

Mean serum total cortisol concentration, assay bias ratio and lower reference limits differed significantly (P<0.0) between the two assays. The gender-difference in mean baseline cortisol concentration seen with the old assay persisted with the new (p <0.05) but was not present post-Synacthen stimulation. A single adult lower reference limit was therefore calculated from male and female samples combined.
For each assay, overall mean bias ratio was calculated from male and female, baseline and post-Synacthen cortisol concentrations.

3.5.6 Comparison between healthy volunteers and patients with potential adrenal disease

3.5.6.1 Mean cortisol concentration

The patients with potential adrenal disease were a heterogeneous group consisting of those with longstanding Addison's disease, patients at risk of hypoadrenalism but little in the way of clinical symptoms and others with symptoms consistent with hypoadrenalism but no clear clinical risk (table 2.1). Despite this, mean cortisol concentration in this group was significantly lower than in healthy volunteers at baseline (218 (112 – 424) nmol/L vs 261 (132 – 515) nmol/L; P=0.012) and post-Synacthen stimulation (461 (279 – 763) nmol/L vs 557 (418 – 742 nmol/L); P=0.001), even after exclusion of the three patients with undetectable cortisol concentrations (figure 3.7).

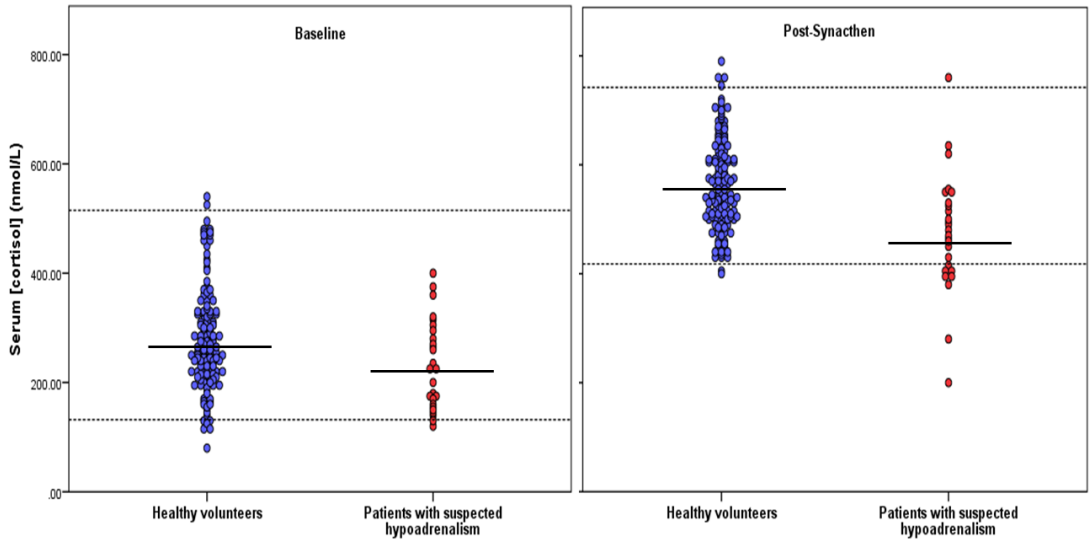


Figure 3.7 Baseline and post-Synacthen serum total cortisol in healthy volunteers and patients with suspected hypoadrenalism. Each dot represents an individual subject. The upper and lower dotted lines in each graph represent the 2.5th and 97.5th percentiles in healthy volunteers, while the dark continuous lines represent mean concentration.

3.5.6.2 Diagnosing hypoadrenalism

Using the cut-offs determined by this study, 40% (12/30) of the patients within the "hypoadrenal" group failed the short Synacthen test, compared to 1.4% of healthy volunteers; however not all patients within this group had definite disease. The outcome of the Synacthen test in each patient was, therefore, compared to the pre-test probability of disease, calculated using the criteria set out in table 2.2.

This score was not a definitive measure of the presence or absence of adrenal insufficiency nor has it been clinically validated; but as a definitive diagnosis was not available at the time of testing, it provided some indication of how well the new cut-off correlated with clinical suspicion of disease. Table 3.8 shows the predicted outcome of the Synacthen test based on the pre-test probability score and the actual test outcome, whilst the agreement between the pre-test score and the Synacthen test is summarised in table 3.9.

Of the 10 patients with a high pre-test probability of adrenal insufficiency, 9 (90%) failed the Synacthen test, using the lower reference limit for the Abbot assay of 430 nmol/L; while 13 (87%) of the 15 patients with a low pre-test probability passed the Synacthen test.

The pre-test score was indeterminate (4) in five patients (table 3.8). Three of these (patients 13, 16 and 26) were referred for testing as a result of a pituitary adenoma and partial anterior hypopituitarism (growth hormone deficiency and/or hypogonadotrophic hypogonadism) and all 3 passed the Synacthen test. One of the other two patients (patient 15) had type 1 diabetes with recurrent hypoglycaemia, weight loss and fatigue and the other (patient 18) had primary hypothyroidism and vitamin B12 deficiency. Patient 15 failed the Synacthen test whilst patient 18 passed, and whilst both clinical scenarios could be consistent with an increased risk of autoimmune hypoadrenalism, the former is more suggestive.

Table 3.8 Pre-test probability of hypoadrenalism and Synacthen test outcome in patients with suspected hypoadrenalism.

Patient number	Gender	Age	Likelihood score	Post-Synacthen [cortisol]	Synacthen test - outcome
1	F	67	10	<28	Fail
2	M	63	10	<28	Fail
3	F	65	2	637	Pass
4	F	82	2	530	Pass
5	F	57	6	515	Pass
6	M	61	2	431	Pass
7	M	74	2	459	Pass
8	F	46	2	490	Pass
9	M	54	2	502	Pass
10	M	62	7	279	Fail
11	M	64	8	414	Fail
12	F	40	10	<28	Fail
13	M	35	4	451	Pass
14	F	39	2	406	Fail
15	M	43	4	379	Fail
16	M	50	4	468	Pass
17	M	64	2	396	Fail
18	F	47	4	478	Pass
19	F	47	2	524	Pass
20	M	46	2	550	Pass
21	F	81	8	373	Fail
22	F	70	6	404	Fail
23	M	36	8	201	Fail
24	M	55	2	554	Pass
25	F	23	2	622	Pass
26	F	43	4	551	Pass
27	F	46	2	465	Pass
28	F	54	2	762	Pass
29	F	28	10	396	Fail
30	F	29	2	495	Pass

Pre-test probability was calculated from criteria set out in table 2.2. A low score was defined as 2 or less, a high score 5 or more and a score of 3 or 4 as indeterminate. A post-Synacthen serum cortisol concentration ≥ 430 nmol/L was defined as a “pass”.

Table 3.9 Agreement between pre-test probability and Synacthen test outcome.

	Synacthen test			
Pre-test probability	**Pass**	**Fail**	**Totals**	**Percentage agreement**
High	1	9	**10**	**90%**
Low	13	2	**15**	**87%**
Indeterminate	4	1	**5**	**-**
Totals	**18**	**12**	**30**	**88%**

Summary of the results from table 3.8 showing the agreement between pre-test probability and the result of the Synacthen test. Agreement is defined as a failed Synacthen test when pre-test probability is high and a Synacthen test pass when pre-test probability is low. A serum cortisol concentration ≥ 430 nmol/L was defined as a "pass".

3.6 Discussion

3.6.1 Cortisol characteristics

Previous studies have shown conflicting results with respect to both the distributional form of baseline and post-Synacthen cortisol concentration and the effect of gender. Clark *et al* found the cortisol response to Synacthen to be non-Gaussian, and both method- and gender-dependent (Clark et al. 1998), whilst Klose and colleagues confirmed clinically significant inter-assay differences but did not find a gender effect or evidence for non-normal distribution (Klose et al. 2007). By using GC-MS to measure total cortisol it was possible to establish which of these features were genuine and which attributable to the imperfections of immunoassay.

Thus, there was no significant gender difference in baseline cortisol concentration or in the cortisol response to Synacthen, although cortisol concentrations were normally distributed about the mean in males only. This difference in the distributional form between genders might be explained by variations in endogenous oestradiol concentration, and hence CBG, in the female population. This is supported by the finding that the free cortisol index (total cortisol divided by CBG), which has been shown to correlate well with measured serum free cortisol (le Roux et al. 2002) was normally distributed in both our male and female groups (chapter 5). However,

against this is the work by Stewart *et al* and Kirschbaum *et al*, which failed to show an effect of menstrual cycle phase (follicular versus luteal) on mean cortisol concentrations throughout the day or on the cortisol response to Synacthen stimulation respectively (Stewart et al. 1993; Kirschbaum et al. 1999), despite evidence of significantly different oestradiol concentrations between the two groups. Menopausal status also appears to play some part in the skewed distribution in women, as evidenced by the higher proportion of menopausal women at the high end of the concentration range.

Immunoassay cortisol showed gender differences and differences in distribution about the mean when compared to GC-MS cortisol measurement. Most notably, for all assays apart from the Centaur, post-Synacthen cortisol concentrations were normally distributed in females as well as in males. This may be explained by the poor specificity of immunoassay antibodies for cortisol, which results in steroid hormones other than cortisol being measured. It is, therefore, not surprising that "immunoassay cortisol" behaves differently to the actual cortisol measured by GC-MS.

Another significant deviation from GC-MS was seen with the Abbott, Roche and Beckman assays. These showed a significant gender difference in baseline and post-Synacthen cortisol concentrations, with significantly higher concentrations seen in male subjects. These differences were statistically significant for all three assays, and for the Roche assay, with a 78 nmol/L difference at baseline and 60 nmol/L difference post-Synacthen, an equally relevant clinical difference. However, for the Abbot and Beckman assays, whose concentrations differed by 40-45 nmol/L and 30-35 nmol/L at baseline and post-Synacthen respectively, the clinical significance is less clear, particularly if allowance is made for assay imprecision at these concentrations.

3.6.2 Immunoassay and GC-MS correlation

All immunoassays studied correlated well with GC-MS, but showed slight positive bias for both male and female subjects. Overall bias ranged from 2 to 30% and was greatest for the Roche assay followed by the Immulite, Beckman, Centaur and Abbott assays, respectively. This positive bias was not unexpected given what is known about antibody specificity but was in contrast to the findings of a UK-wide audit of the short Synacthen test which showed immunoassay cortisol was negatively biased relative to isotope dilution – GC-MS assigned values (Chatha et al. 2010). Negative bias was also reported for the Abbott and Beckman assays by Dodd *et al* who went on to demonstrate that mean assay bias had, in fact, gone from being positive in 2010 to negative in 2013 (table 1.2) (Dodd et al. 2014). As there had been no reports of major assay reformulations over this period, it is likely that minor reformulations had been carried out by manufacturers but had, possibly, been considered too insignificant to warrant disclosure.

Siemens' reformulation of their Centaur assay allowed the effects of assay changes to be studied. Mean cortisol concentrations fell significantly for all subjects with the new assay (table 3.7) and mean assay bias fell from 18 to 6%. Users were informed of the assay change but there was no accompanying recommendation for them to alter their reference ranges. These ongoing changes to cortisol immunoassays raise questions about how long these assays can remain useful in clinical practice. The Synacthen test requires valid cortisol cut-offs to allow it to discriminate between a normal and abnormal HPA axis, but with ever changing assays it is increasingly difficult to define the true cortisol response to Synacthen.

3.6.3 Assay specific cut-offs

This study is the largest to examine the cortisol response to the standard dose Synacthen test and the first to compare results from five widely used immunoassays

to a reference mass spectrometry method. The results illustrate the influence of assay performance on cortisol measurement, which is such that mean cortisol concentrations post-Synacthen vary by as much as 230 nmol/L (542 nmol/l in females with the Abbott assay to 772 nmol/l in males with the Roche). This finding may be explained by differences in assay calibration or in the specificity of assay antibodies, but irrespective of its origin, it has the potential to cause patient misclassification; hence the need for assay-specific reference ranges and lower reference limits.

In contrast to the mean cortisol response to Synacthen, the lower limit of the normative range, defined as the 2.5th percentile value of the log transformed data, showed much less of a gender difference, ranging from 2 to 14 nmol/L for the Centaur, Abbott, Immulite and Beckman assays. As these differences were too small to be of clinical significance and fell within the precision limits of their assays, there was no need for gender related reference limits for these assays. However, for the Roche assay, the difference between male and female lower limits was 50 nmol/L and, consequently, gender-specific lower limits are required.

Prior to this study, a cortisol cut-off of 550 nmol/L was used locally to define an adequate response to Synacthen stimulation. This limit was based on the method-specific cortisol response to Synacthen defined by Clark *et al* in 1998 and the assay that was in use in the laboratory at that time. However, the laboratory has changed its cortisol assay twice since 1998, with no associated change in cortisol cut-off. This study shows that continuing to use this cut-off would potentially result in a misdiagnosis of adrenal insufficiency in a significant number of healthy individuals (27%, 42%, 16% & 21% with the Centaur, Abbott, Immulite and Beckman assays respectively, falling to 12%, 19%, 4% & 9% respectively at a cut-off of 500 nmol/l). Conversely, by applying assay-specific lower reference limits, the percentage of false positive Synacthen tests in healthy volunteers reduces to 2% or less for all the above assays. The newly defined reference ranges thus have the potential to reduce

inappropriate treatment and costly follow-up investigations for patients with a low pre-test probability of hypoadrenalism.

3.6.4 Patients with suspected hypoadrenalism

The validity of our assay-specific lower reference limits in a population with potential adrenal insufficiency was put to the test in a group of patients referred for a Synacthen test from specialist Endocrine clinics. Twelve of the 30 patients tested (40%) failed the Synacthen test, which, at first glance, might suggest the lower reference limit is not valid in a diseased population; particularly as 80% of patients (24/30) would have failed the test were the 550 nmol/L cut-off used instead.

However, this patient group was markedly heterogeneous and included patients with longstanding Addison's disease on daily glucocorticoid replacement and others with non-specific symptoms such as fatigue, but little else (table 2.1). Each patient was, therefore, stratified into a high or low pre-test probability of disease on the basis of criteria set out in table 2.2 before the outcome of the Synacthen test was known.

Subsequent analysis, using the Abbott-specific lower reference limit, showed good agreement between the Synacthen test result and the pre-test probability of disease, with 90% (9/10) of patients with a high pre-test probability failing the Synacthen test and 87% (13/15) of those with a low probability, passing the test. In contrast, the 550 nmol/L cut-off would have resulted in all 10 patients with a high pre-test probability failing the Synacthen test, along with 11 of the 15 patients (73%) with a low pre-test probability as well.

Only 3 patients showed disagreement between the pre-test probability score and the outcome of the Synacthen test when the Abbott-specific cut-off was used. Patient 5 had a high pre-test probability of disease but passed the Synacthen test, whilst patients 14 and 17 had a low pre-test score but failed the test; all are considered in more detail below.

Patient 5 was a 57-year-old woman who gave a history of asthma for which she took a daily Seretide inhaler (2)[10], intermittent courses of oral glucocorticoids (3) and her only clinical complaint was of fatigue (1), giving her a total score of 6. She passed her Synacthen test convincingly, with a post-Synacthen cortisol of 515 nmol/L, although her baseline cortisol was relatively high amongst the patient group, at 360 nmol/L, raising the possibility of interference from exogenous glucocorticoids as the cause.

Another possible explanation for the discordance between pre-test probability and Synacthen test outcome in this patient is incorrect risk scoring, as ascribing a score of 3 to intermittent oral glucocorticoid use may be over-cautious, particularly as exact details of dosage, frequency and compliance were lacking. Thus, a more accurate total score might have been 3 or 4, reflecting the fact that her risk is not entirely clear-cut. This also highlights the limitations of the proposed clinical risk score, which clearly requires further validation before it can be considered for clinical practice. Subsequent follow-up confirmed she did not have hypoadrenalism despite the high pre-test probability.

Patient 14 was a 39-year-old woman who complained of dizziness and postural hypotension, and was assigned a score of 2. Nevertheless, she failed the Synacthen test convincingly, with a cortisol of 406 nmol/L and, perhaps, with hindsight her score was inappropriately low, given that postural hypotension is an unusual symptom in a young patient not treated with anti-hypertensive medication. Subsequent review of her notes, however, confirmed that she did not have hypoadrenalism at the time of testing, nor had she developed it since.

Patient 17, a 64-year-old man with Crohn's disease and recurrent oral steroid use, was also a potential candidate for iatrogenic secondary hypoadrenalism, despite the

[10] The score for each risk factor is included in brackets, as described in table 2.2.

low pre-test score. He also failed the Synacthen test convincingly, with a cortisol of 396 nmol/L, but was subsequently lost to follow-up and does not appear to have been commenced on glucocorticoid replacement.

Subsequent long-term follow up of the remaining patients with suspected hypoadrenalism confirmed adrenal insufficiency in patients 1, 2, 10, 12 and 23, all of whom had a pre-test probability of 7 or higher and failed the Synacthen test. Patients 11 and 29 scored 8 and 10, respectively, and had iatrogenic secondary adrenal insufficiency at the time of the study, confirmed on Synacthen testing, but they subsequently recovered and stopped their glucocorticoid treatment.

Patient 22 also had a high pre-test probability score and failed her salivary Synacthen test; however several subsequent Synacthen tests excluded hypoadrenalism until she underwent further pituitary surgery in 2016. Nevertheless, she was clearly at high risk of pituitary disease and her failed Synacthen test may have been a true reflection of mild secondary adrenal insufficiency.

Comparing the outcome of the Synacthen test and pre-test probability to actual outcome 5 years later provides a more robust tool for determining their diagnostic accuracy. Thus, both are shown to perform well, with sensitivities and negative predictive values of 100%, assuming that patients who were lost to follow-up had not received treatment for hypoadrenalism elsewhere or died from the disease, and a specificity of 78% for the Synacthen test, compared to 65% for the pre-test probability score, whose specificity suffers as a result of the inclusion of an “indeterminate” risk categorisation.

The validity of using a single cut-off to diagnose adrenal insufficiency also needs to be addressed. Disease of the HPA axis is a continuum and cortisol values lying just above the lower reference limit may well represent impaired HPA axis function in patients where the clinical features are suggestive. In such cases, a single cut-off is

no replacement for good clinical judgement and it is recommended that the lower limits defined in this study are used as a guide to aid in the diagnosis of adrenal insufficiency rather than as an absolute.

3.6.4 Conclusions

Measuring cortisol by immunoassay affects not only its concentration but also its distributional form and can create gender differences where, in fact, there are none. Comparing immunoassay cortisol to GC-MS measurements identified assay effects that could account for some of the contradictory findings reported in previous studies of the short Synacthen test. Recognising the extent to which cortisol concentrations are affected by assay is essential for the correct interpretation of the Synacthen test.

This study developed lower reference limits for the cortisol response to Synacthen for five widely used automated immunoassays and demonstrated that the 550 nmol/L cortisol cut-off is too high for most assays. As a result of this work, local laboratories adjusted their cortisol cut-off in line with newly defined reference ranges and it is recommended that other laboratories do the same. However, adopting more appropriate reference ranges does not alter the fact that inter-individual variation and the accuracy and precision of cortisol assays will limit the value of a single, fixed cut-off concentration for identifying adrenal insufficiency. It is, therefore, advisable that a more flexible approach to diagnosis is adopted; one which combines the cortisol response to Synacthen with the pre-test probability of disease.

Endocrinologists investigating patients with suspected adrenal insufficiency need to be aware of the limitations of immunoassay cortisol measurement and should clarify which method is in use in their laboratories before interpreting post-Synacthen cortisol concentration. In light of these findings, it is likely that the cortisol response to the insulin stress test will also need to be redefined.

CHAPTER 4

DETERMINING THE SALIVARY CORTISOL RESPONSE TO SYNACTHEN STIMULATION

4.1 Introduction

Salivary cortisol offers a non-invasive, relatively stress-free alternative to measurement of cortisol in serum and has the potential advantage over the latter of directly evaluating free, bioavailable cortisol. Salivary cortisol assays have been available for over 35 years; however early radioimmunoassays lacked specificity, current immunoassays lack sensitivity at the lower end of the range and LC-MS/MS assays have, until relatively recently, lacked the necessary turnaround time for urgent assessment of a patient with suspected adrenal failure (Owen et al. 2013a).

Developments in LC-MS/MS technology and their increasing availability in clinical laboratories have led to renewed interest in using salivary cortisol to evaluate the HPA axis. In particular, late night salivary cortisol is now well established as a screen for Cushing's syndrome; and despite several different diagnostic cut-offs being reported in the literature (Baid et al. 2007; Zerikly et al. 2010; Erickson et al. 2012), for most assays the cut-off falls somewhere below 3 nmol/L. Nevertheless, until LC-MS/MS salivary cortisol assays are better standardised, laboratories will need to define their own cut-offs.

Salivary cortisol as an alternative to serum total cortisol in evaluating glucocorticoid sufficiency in patients with low protein secondary to liver disease or an acute phase response has also received significant interest (Cohen et al. 2004; Arafah 2006; Galbois et al. 2010; Elbuken et al. 2016) and has been recommended as the method of choice in these scenarios. However, the salivary cortisol response to Synacthen stimulation in a healthy population is not yet defined and until this has been done it will not be possible to thoroughly evaluate its performance in disease.

4.2 Aims

The work described in this chapter was undertaken to:

- Define the salivary cortisol response to Synacthen in healthy volunteers and establish the lower reference limit for this response using an in-house LC-MS/MS assay.
- Evaluate salivary cortisol as an alternative to serum total cortisol in the interpretation of the short Synacthen test in patients with potential hypoadrenalism.
- Explore and compare the relationship between serum total and salivary free cortisol responses to Synacthen stimulation in health and disease.

4.3 Subjects and sample analysis

4.3.1 Healthy volunteers

This group consisted of 60 male and 79 female volunteers, as described in chapter 2. Salivary cortisol was measured by the, previously described, in-house LC-MS/MS assay within 1 year of sample collection.

4.3.2 Patients with suspected hypoadrenalism

This group consisted of 13 male and 17 female recruits, whose clinical details are summarised in table 2.1 and their salivary cortisol was also measured by LC-MS/MS within 1 year of sample collection. One patient (patient 21; table 2.1) was excluded from this arm of the study due to a lack of sufficient post-Synacthen saliva for analysis. The salivary cortisol results from the three patients with suppressed baseline serum cortisol were not included in calculations of the mean or comparisons with healthy volunteers, to avoid negatively skewing the results.

4.4 Statistics

4.4.1 Healthy volunteers

Mean salivary cortisol and reference ranges, defined as "mean ± 1.96*SD", were calculated from log-transformed data at baseline and 30 minutes post-Synacthen. The 2.5th centile (mean – 1.96*SD), or lower reference limit, of this response was used as a cut-off concentration to differentiate between glucocorticoid sufficiency and deficiency. Despite log-transformation, salivary cortisol was not normally distributed in males or females at either time point so differences between genders and time-points were evaluated using the Mann-Whitney U test.

Baseline and post-Synacthen results were combined before the correlation between serum and salivary cortisol was examined using scatter plots.

4.4.2 Patients with established or potential adrenal disease

Mean salivary cortisol concentrations in this group were calculated from log-transformed data at baseline and 30 minutes post-Synacthen and were compared to healthy volunteers using the Mann-Whitney U test. The pre-test probability of HPA axis disease in each patient was compared to the salivary cortisol response and a percentage agreement calculated. Agreement between serum and salivary cortisol responses to Synacthen was similarly evaluated.

Scatter plots were used to explore the correlation between combined baseline and post-Synacthen serum and salivary cortisol concentrations in this population. In all comparisons, a p-value <0.05 was considered significant.

4.5 Results

4.5.1 Salivary cortisol in healthy volunteers

4.5.1.1 At baseline

Salivary cortisol was not normally distributed in males or females (figure 4.1) and was skewed to the right, reflecting a similar pattern to that seen in total serum cortisol in women but not in men (figure 3.1). There was no age effect and no significant concentration difference between genders (table 4.1, figure 4.2), with concentrations ranging between 0.6 and 12.0 nmol/L in men and 0.8 to 9.2 nmol/L in women.

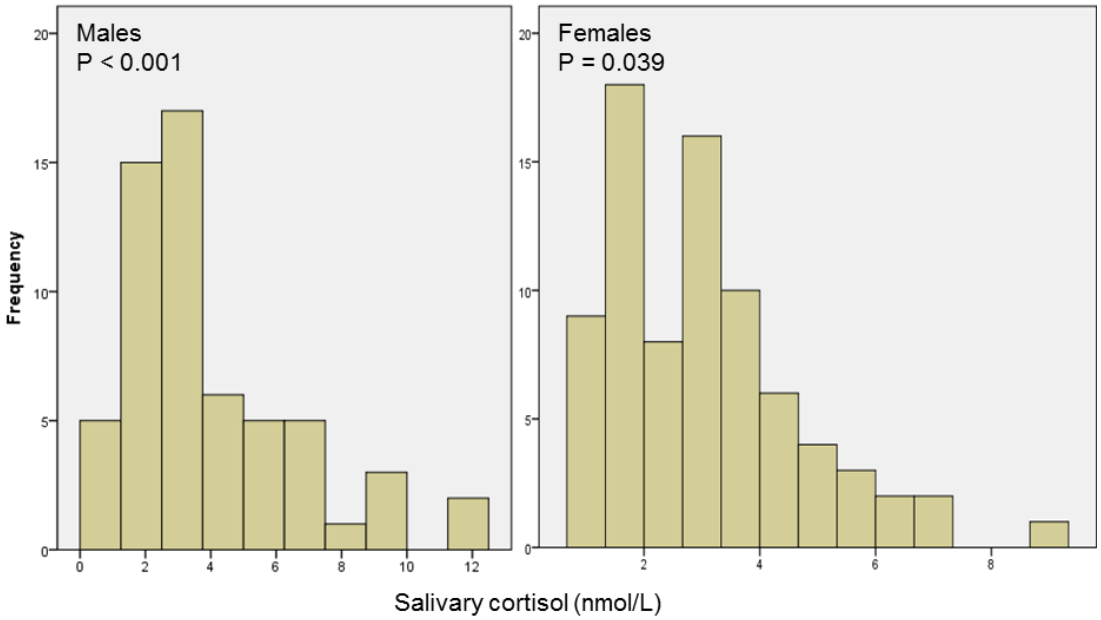


Figure 4.1 Distribution of baseline salivary cortisol concentration in males and females. The distribution was examined using the Kolmogorov-Smirnov test, with a P-value >0.05 indicating a normal distribution.

Table 4.1 Salivary cortisol concentration in healthy volunteers.

	Mean salivary [cortisol] (2.5th – 97.5th percentile) – nmol/L			
	Male	**Female**	**P value***	**Combined Adult**
0 Minute	3.2 (0.8 – 12.0)	2.7 (1.0 – 7.5)	0.125	2.9 (0.9 – 9.2)
30 Minute	19.1 (9.8 – 37.3)	19.6 (10.9 – 36.2)	0.443	19.3 (10.3 – 36.2)

* P-value <0.05 indicates a significant difference between gender means.

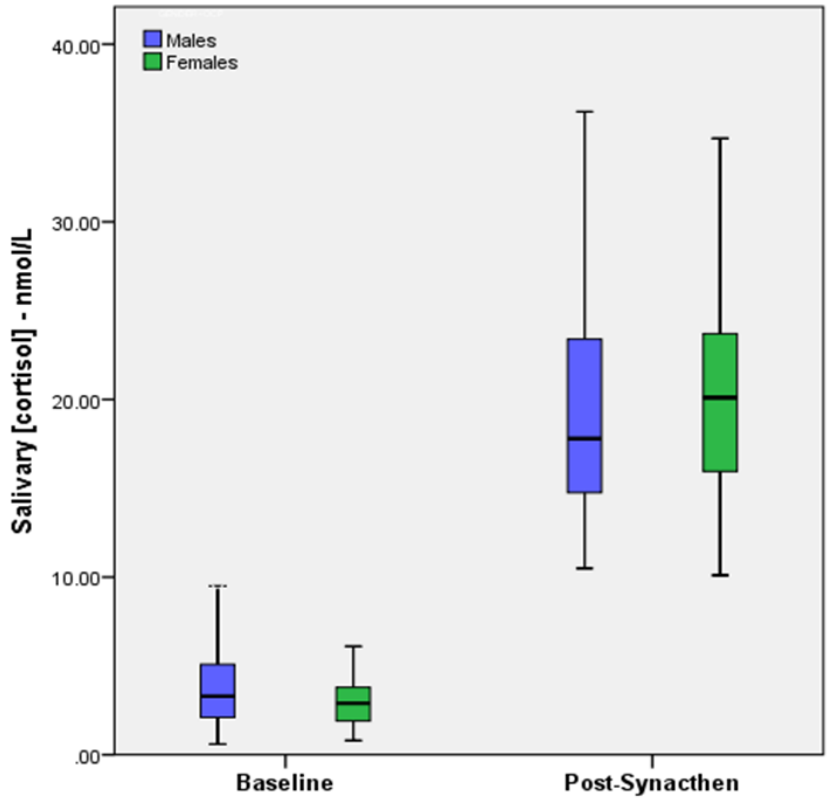


Figure 4.2 Mean salivary cortisol concentration at 0 and 30 minutes by gender. There was no significant gender difference at either timepoint (p = 0.115 at baseline; p = 0.591 post-Synacthen), but mean salivary cortisol concentration was significantly lower at baseline than post-Synacthen concentration (P <0.001).

4.5.1.2 Post-Synacthen stimulation

Salivary cortisol continued to be skewed to the right following Synacthen stimulation (figure 4.3), ranging from 10.5 to 39.7 nmol/L in men and 10.1 to 34.8 nmol/L in women. There was no significant gender difference in concentration (table 4.1, figure 4.2). The post-Synacthen salivary cortisol concentration was, on average, 29 times

lower than the serum cortisol response (range 14 to 60 times). It increased an average of 7.9 times (range 1.4 – 27.6 times) the baseline salivary concentration and the mean incremental rise was 17.4 nmol/L (range 4.5 – 37.5 nmol/L) across the entire group.

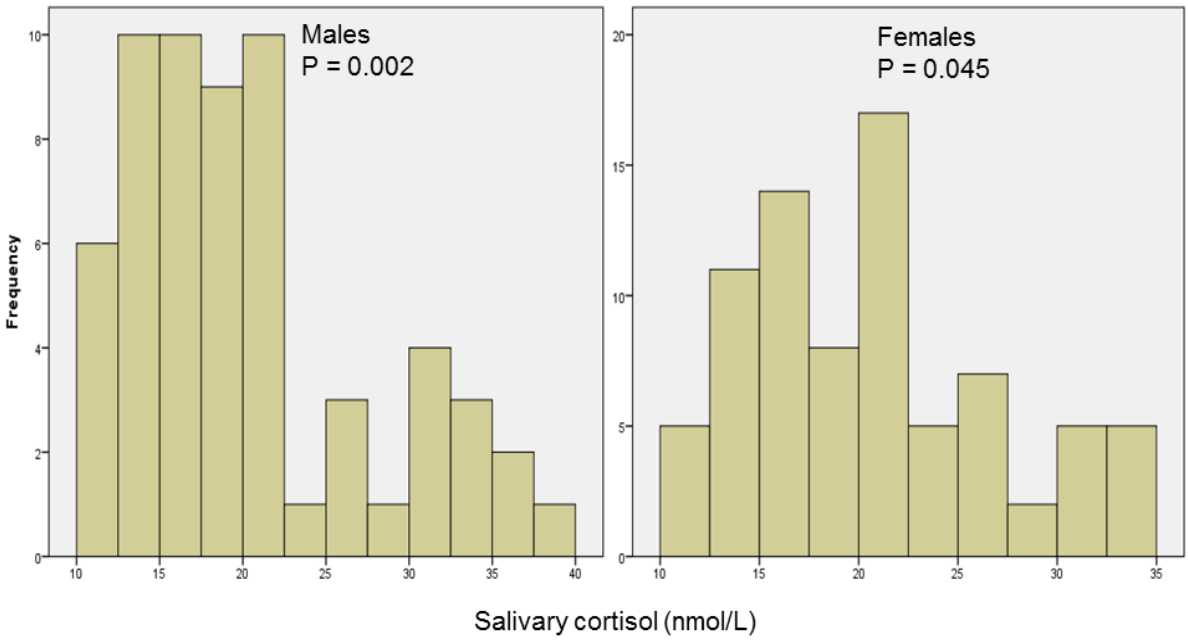


Figure 4.3 Distribution of post-Synacthen salivary cortisol concentration in males and females. The distribution was examined using the Kolmogorov-Smirnov test, with a P-value >0.05 indicating a normal distribution.

4.5.1.3 Reference range and lower limit of normal

The mean salivary cortisol response to Synacthen in healthy adults was 19.3 nmol/L, with a 2.5th to 97.5th centile range of 10.3 to 36.2 nmol/L. The lower limit of this reference range was arbitrarily used to define the cut-off concentration above which adrenal insufficiency can be excluded, in line with the definition used for the serum cut-off (section 3.4.1).

4.5.2 Salivary cortisol in patients with potential adrenal insufficiency

4.5.2.1 Mean salivary cortisol concentration at baseline and post-Synacthen stimulation

Mean baseline salivary cortisol concentration in patients with suspected adrenal insufficiency was 2.6 nmol/L and did not differ significantly from the mean baseline concentration of 2.9 nmol/L seen in healthy volunteers (p = 0.418) (figure 4.4). In contrast, the mean concentration post-Synacthen stimulation in this cohort was 11.3 nmol/L which was significantly lower than the corresponding concentration of 19.3 nmol/L in healthy volunteers (p<0.001), despite excluding patients with undetectable serum total cortisol concentrations (<28 nmol/L) (figure 4.4).

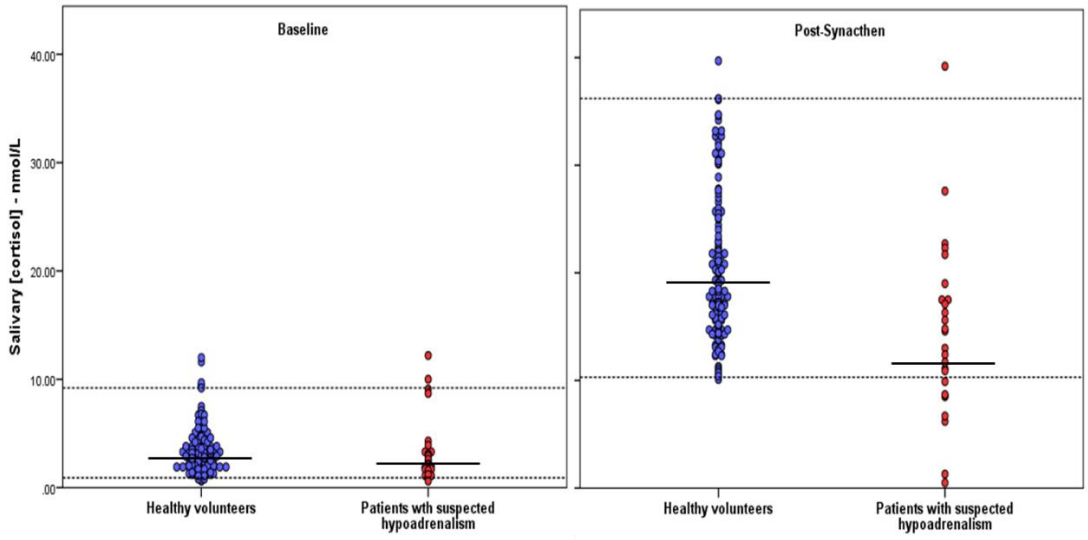


Figure 4.4 Baseline and post-Synacthen salivary cortisol in healthy volunteers and patients with suspected hypoadrenalism. Each dot represents an individual subject. The upper and lower dotted lines in each graph represent the 2.5th and 97.5th percentiles in healthy volunteers, while the dark continuous lines represent mean concentration.

4.5.2.2 Salivary cortisol to diagnose adrenal disease

Eleven patients (38%) in the group with potential hypoadrenalism had salivary cortisol concentrations below the 10.3 nmol/L cut-off and failed the Synacthen test (table 4.2).

In contrast, only 3 (1.8%) of the healthy volunteers failed to achieve a concentration above this cut-off.

Ten of the 30 patients studied had a pre-test probability score of 5 or more and were considered at high risk for adrenal insufficiency. One of these was the patient in whom the post-Synacthen saliva sample was of insufficient volume to allow analysis. Of the remaining nine, eight (88.9%) were identified by the salivary cortisol response, while one patient passed the Synacthen test convincingly, with a salivary cortisol concentration of 17.5 nmol/L.

Fifteen patients were considered low risk for hypoadrenalism, with pre-test probability scores of 2 or less. Fourteen of these (93.3%) passed the Synacthen test and one failed by a small margin, with a post-Synacthen cortisol concentration of 9.9 nmol/L. Five patients scored 4, placing them in the indeterminate risk group; three patients passed the Synacthen test (patients 15, 18 and 26), with concentrations of 10.9 nmol/L, 11.7 nmol/L and 27.6 nmol/L, whilst 2 patients failed (patients 13 and 16) with concentrations of 8.7 nmol/L and 8.6 nmol/L, respectively. Table 4.3 shows the agreement between pre-test probability and the result of the salivary Synacthen test.

Table 4.2 Pre-test probability of hypoadrenalism and outcome of salivary Synacthen test in patients with suspected hypoadrenalism.

Patient number	Gender	Age	Likelihood score	Post-Synacthen [saliva]	Saliva outcome
1	F	67	10	1.0	Fail
2	M	63	10	0.2	Fail
3	F	65	2	19.0	Pass
4	F	82	2	39.3	Pass
5	F	57	6	17.5	Pass
6	M	61	2	17.1	Pass
7	M	74	2	14.8	Pass
8	F	46	2	17.5	Pass
9	M	54	2	9.9	Fail
10	M	62	7	1.3	Fail
11	M	64	8	6.2	Fail
12	F	40	10	0.3	Fail
13	M	35	4	8.7	Fail
14	F	39	2	16.3	Pass
15	M	43	4	10.9	Pass
16	M	50	4	8.6	Fail
17	M	64	2	15.6	Pass
18	F	47	4	11.7	Pass
19	F	47	2	22.3	Pass
20	M	46	2	12.4	Pass
21	F	81	8	-	-
22	F	70	6	6.7	Fail
23	M	36	8	0.5	Fail
24	M	55	2	13.0	Pass
25	F	23	2	22.7	Pass
26	F	43	4	27.6	Pass
27	F	46	2	14.6	Pass
28	F	54	2	21.7	Pass
29	F	28	10	8.5	Fail
30	F	29	2	11.1	Pass

Pre-test probability was calculated from criteria set out in table 2.2. A low score was defined as 2 or less, a high score 5 or more and a score of 3 or 4 as indeterminate. A salivary cortisol concentration ≥ 10.3 nmol/L was defined as a “pass”.

Table 4.3 Agreement between pre-test probability and Synacthen test salivary result.

	Synacthen test			
Pre-test probability	**Pass**	**Fail**	**Total**	**Percentage agreement**
High	1	8	**9**	**89**
Low	14	1	**15**	**93**
Indeterminate	3	2	**5**	**-**
Total	**18**	**11**	**29**	**92**

Summary of the results from table 4.2 showing the agreement between pre-test probability and the result of the salivary Synacthen test. Agreement is defined as a failed Synacthen test when pre-test probability is high and a Synacthen test pass when pre-test probability is low. A salivary cortisol concentration ≥ 10.3 nmol/L was defined as a "pass".

4.5.2.3 Total serum cortisol versus salivary cortisol in diagnosing adrenal insufficiency

Ten of the patients studied had a high pre-test probability score of hypoadrenalism compared to twelve who failed the Synacthen test using serum cortisol measurements and eleven who failed using salivary cortisol. Tables 3.9 and 4.3 show the agreement between the pre-test probability of disease and serum and salivary test outcomes, respectively.

Comparison between the performance of the two tests shows similar agreement with the pre-test probability score when the likelihood of disease is high (89% for the salivary test vs 90% for serum), but that the salivary cortisol test shows better agreement with the score in patients considered to be at lower risk (93 vs 87% concordance). Overall agreement between the salivary cortisol test and pre-test likelihood of disease was also slightly higher at 92%, versus 88% for serum total cortisol.

Table 4.4 shows the agreement between the two tests, irrespective of their agreement with the pre-test prediction. The result of the serum test was considered the gold standard for the purposes of this comparison and the salivary test shows better agreement with this result in patients who pass the Synacthen test (15/18) than it

does in those who fail the test (8/12). Overall agreement between the two tests is 79%.

Table 4.4 Agreement between serum and saliva Synacthen test results.

	Synacthen result (Saliva)			
Synacthen result (Serum)	**Pass**	**Fail**	**Total**	**Percentage agreement**
Pass	15	3	**18**	**83%**
Fail	3	8	**11**	**67%**
Total	**18**	**11**	**29**	**79%**

Summary of the results from table 4.2 showing agreement between serum and salivary Synacthen tests, irrespective of their agreement with pre-test probability. A serum cortisol concentration ≥ 430 nmol/L and a salivary cortisol concentration ≥ 10.3 nmol/L were defined as a "pass".

There were 6 patients with discrepant serum and salivary results (table 4.5). Three had an indeterminate pre-test likelihood score of 4, but three patients were considered low risk for adrenal insufficiency and subsequently failed either the serum or salivary Synacthen test. The patient who was considered high risk but passed both Synacthen tests was considered in some detail in chapter 3.

Table 4.5 Discrepancies between predicted and/or Synacthen test outcome and salivary cortisol outcome.

Patient number	**Gender**	**Age**	**Likelihood score**	**Predicted outcome**	**Synacthen test outcome**	**Saliva outcome**
9	***M***	***54***	***2***	***Pass***	***Pass***	***Fail***
14	*F*	*39*	*2*	*Pass*	*Fail*	*Pass*
17	*M*	*64*	*2*	*Pass*	*Fail*	*Pass*
13	M	35	4	Indeterminate	Pass	Fail
15	M	43	4	Indeterminate	Fail	Pass
16	M	50	4	Indeterminate	Pass	Fail
5	**F**	**57**	**6**	**Fail**	**Pass**	**Pass**

Predicted outcome was determined from the pre-test probability, with a low score defined as 2 or less, a high score 5 or more and an indeterminate score 3 or 4.
A serum cortisol concentration ≥ 430 nmol/L and a salivary cortisol concentration ≥ 10.3 nmol/L were defined as a "pass".

Figure 4.5 shows post-Synacthen salivary and serum cortisol concentrations in healthy volunteers and patients with possible hypoadrenalism. The range of salivary cortisol responses is wider than that of serum cortisol, but both exhibit considerable overlap between healthy volunteers and patients with potential adrenal disease, despite the lower mean concentration in the latter group. In light of the wider range of "normal" responses, it is somewhat surprising that salivary cortisol shows better agreement with the pre-test probability of disease than serum cortisol. Nevertheless, the results presented here suggest a single salivary cortisol cut-off can be as reliable as a serum cut-off in the interpretation of the Synacthen test.

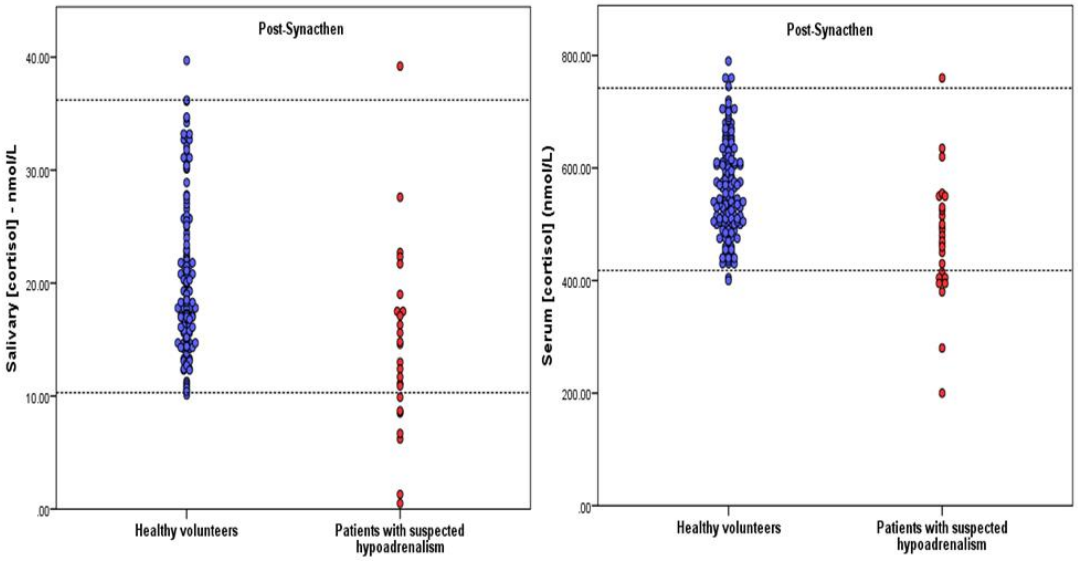


Figure 4.5 Post-Synacthen salivary and serum cortisol concentrations in healthy volunteers and patients with suspected hypoadrenalism. Each dot represents an individual subject. The upper and lower dotted lines in each graph represent the 2.5th and 97.5th percentiles in healthy volunteers.

4.5.3 Correlation between salivary free and serum total cortisol in health and disease

Baseline and post-Synacthen results were combined before the correlation between salivary and serum total cortisol was studied. Salivary cortisol correlated well with serum total cortisol measured by both GC-MS (figure 4.6A) and the Abbott immunoassay (figure 4.6A), with no real difference in the strength of the correlation

between the two assays. The correlation was not affected by gender, but there was less scatter at higher cortisol concentrations, suggesting better correlation.

The relationship between serum and salivary cortisol concentration was maintained in patients with suspected adrenal disease and was not visibly different to the relationship seen in healthy volunteers (figure 4.6C and D).

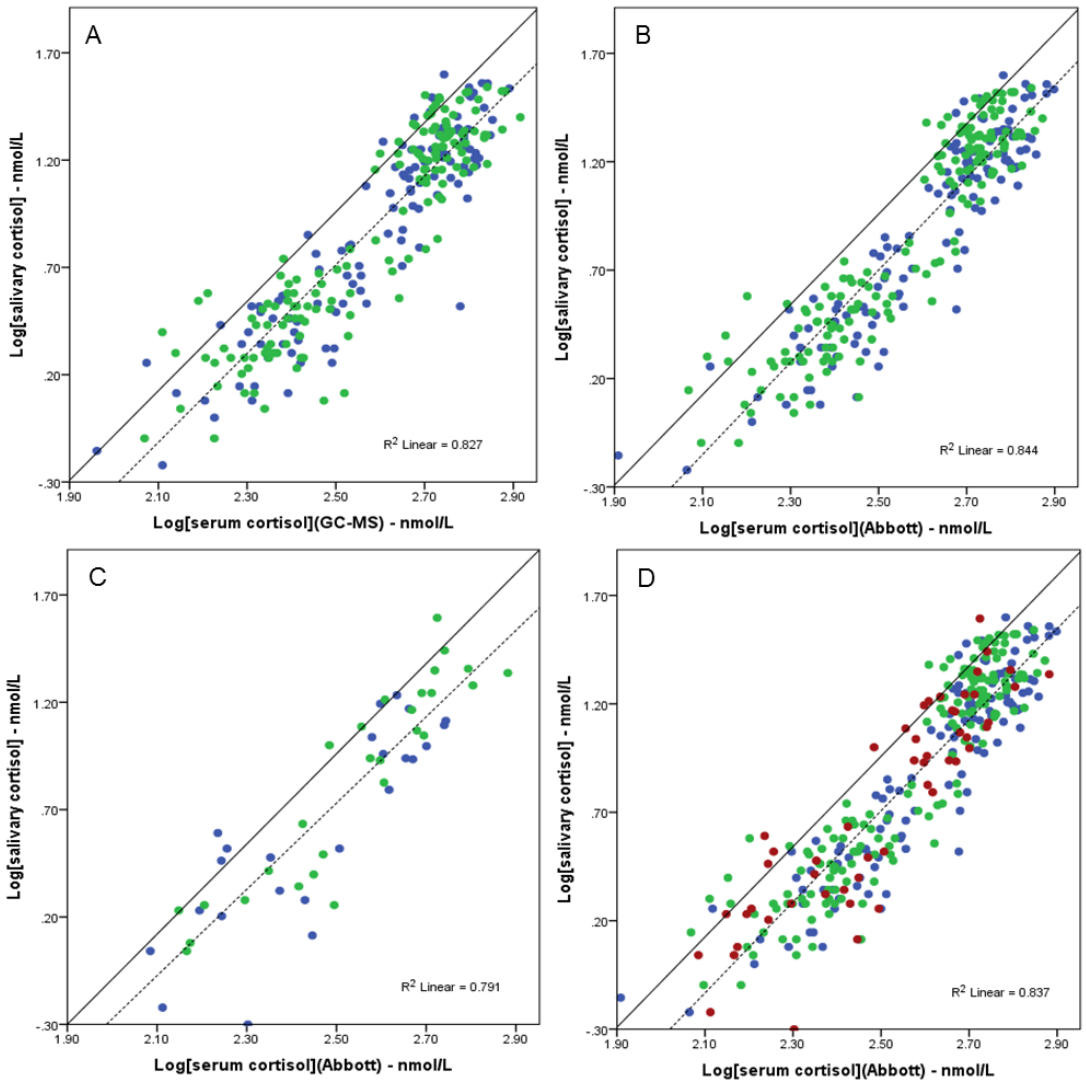


Figure 4.6 Correlation between log-transformed salivary cortisol and serum total cortisol measured by GC-MS and the Abbott immunoassay in healthy volunteers and patients with suspected hypoadrenalism. Panel (A) shows the relationship with GC-MS cortisol in healthy volunteers, (B) the relationship with Abbott cortisol in healthy volunteers, (C) the relationship with Abbott cortisol in patients with suspected hypoadrenalism and (D) the relationship with Abbott cortisol in the two groups combined. The solid line represents equivalence between the two; the dotted line is the line of best fit. ● males; ● females; ● patients with suspected hypoadrenalism

4.6 Discussion

4.6.1 Defining salivary cortisol in healthy volunteers

Measurement of serum free cortisol, the bioactive fraction of total cortisol, remains beyond the reach of routine clinical laboratories, and serum total cortisol remains the analyte of choice for evaluating suspected hypoadrenalism, despite well-recognised assay and biological limitations. Salivary cortisol, which is free and protein-unbound, offers a possible surrogate measure of serum free cortisol for the evaluation of the hypothalamic-pituitary-adrenal axis but also provides an opportunity to gain a better understanding of how bioactive cortisol behaves in a healthy population.

The skewed distribution of baseline salivary cortisol in male and female volunteers contrasts with the normal distribution of serum total cortisol reported in males, but not females, in this study and the normal distribution noted in other studies of the serum cortisol response to Synacthen stimulation (Klose et al. 2007). This distribution is not surprising given the influence of external factors, such as BMI, ethnicity, stress and disrupted sleep on cortisol concentration (Veldhuis et al. 2009; Amirian et al. 2015; DeSantis et al. 2015), none of which were used as exclusion criteria for recruitment into this study.

Salivary cortisol has been shown to predict mortality in elderly patients with chronic disease (Schoorlemmer et al. 2009) and, in contrast to serum total cortisol, early morning salivary cortisol is significantly higher in patients followed up one year after suffering a subarachnoid haemorrhage than in healthy volunteers (Poll et al. 2013). This has led to the suggestion that salivary cortisol gives different information to serum total cortisol. Thus, a better understanding of baseline salivary cortisol in health is needed to help with its interpretation in patients with non-adrenal disease which is, nonetheless, influenced by cortisol. This study contributes to existing

knowledge by defining the distribution and concentration of baseline salivary cortisol in a large cohort of healthy volunteers.

Baseline salivary cortisol concentrations ranged between 0.6 and 12.0 nmol/L and showed significant overlap with concentrations in patients with suspected hypoadrenalism, which ranged between 0.3 and 12.2 nmol/L. The mean concentration of 2.9 nmol/L in healthy volunteers was lower than the 6.7 nmol/L reported by Perogamvros *et al* in their study of the salivary cortisol response to Synacthen stimulation in a group of 14 control subjects (gender undefined) (Perogamvros et al. 2010a), but not too dissimilar to the 3.9 nmol/L reported by the same author in a group of 68 patients who presented to a specialised Endocrine unit for investigation of adrenal function and were subsequently found to have a normal response to Synacthen stimulation (Perogamvros et al. 2010b). In both studies, salivary cortisol was measured by an in-house LC-MS/MS assay.

Sample timing, collection technique (Poll et al. 2007) and method differences (Patel et al. 2004a; Kosak et al. 2014) may explain some of the concentration difference observed between the results of this study and those in the literature. Although as this is the largest study to date to use an LC-MS/MS cortisol assay, the results presented here are likely to be the most representative of concentrations in the general population. However, given the lower concentrations reported in this study, further evaluation of the salivary cortisol assay, to ensure it meets the criteria required for validation of an LC-MS/MS assay (Honour 2011), is needed, as is comparison with other LC-MS/MS assays through participation in an EQA scheme.

In fact, such a scheme exists already. It includes 72 assays, 15 of which are LC-MS/MS; with ELISA and luminescent immunoassay making up the majority of the remainder. The scheme runs twice a year, distributing 3 salivary samples on each occasion. Variability in cortisol concentration across all assays is huge, with results

between 1.8 nmol/L and 15.5 nmol/L reported for a sample with a mean LC-MS/MS concentration of 3.5 nmol/L; results between 15.7 nmol/L and 56.3 nmol/L for a sample with a mean LC-MS/MS concentration of 16.9 nmol/L and between 8.3 nmol/L and 34.3 nmol/L for a sample with a mean LC-MS/MS concentration of 9.9 nmol/L. LC-MS/MS assays show much less variability, with concentrations ranging between 1.8 nmol/L and 4.2 nmol/L, 15.7 nmol/L and 19.3 nmol/L and 8.3 nmol/L and 11.3 nmol/L, respectively, for the same three samples. The results from the assay used in this study, however, are typically at the lower end of this range, with concentrations of 3.5 nmol/L, 15.7 nmol/L and 8.3 nmol/L, respectively.

This study also explored the possibility of using baseline salivary cortisol concentration as a predictor of the Synacthen test result. Patel *et al* showed in a study of salivary cortisol measurement in patients taking steroid nose drops that a subnormal morning salivary cortisol (defined as a result <2.8 nmol/L) identified an impaired Synacthen test result with a positive predictive value of 100% (Patel et al. 2001). In contrast, in this study, baseline salivary cortisol concentrations in patients with confirmed hypoadrenalism at the time of testing (patients 1, 2, 10, 11, 12, 23 and 29) fell between 0.3 and 3.3 nmol/L, compared to concentrations between 1.1 and 12.2 nmol/L in patients who passed the test and concentrations between 0.6 and 12.0 nmol/L in healthy volunteers. Any cut-off designed to identify all hypoadrenal patients would, therefore, overlap with results from healthy volunteers and patients with normal adrenal function, meaning a positive predictive value of 100% is impossible.

Thus, this study suggests that baseline salivary cortisol is of limited value as a predictor of hypoadrenalism. This might be explained by differences in sampling time, as participants in this study underwent Synacthen testing at any time between 8.0 and 11.0 am, whereas Patel *et al* used samples collected between 6.0 and 8.0 am to define their early morning cortisol concentrations (Patel et al. 2001; Patel et al. 2004b) and Deutschbein *et al* used samples collected between 8.0 and 9.0 am (Deutschbein

et al. 2009). Furthermore, substantial falls in salivary cortisol concentration throughout the course of the morning have been reported (Dorn et al. 2007) and hence the heterogeneity of sample collection times might explain the wide range of baseline salivary cortisol concentrations and the poor correlation between baseline concentration and response to Synacthen test. However, further study in a larger group of patients with confirmed hypoadrenalism is needed to confirm that baseline salivary cortisol concentrations in patients and healthy volunteers are not significantly different.

Post-Synacthen:

The post-Synacthen salivary cortisol response was similarly broad – ranging from 10.1 to 39.7 nmol/L, with little correlation between baseline and post-Synacthen concentration (r^2=0.06). Nevertheless, it was possible to define a "cut-off" for diagnosing hypoadrenalism using the lower limit of the normal range (the 2.5th percentile of the log transformed data) which when applied to the results in healthy volunteers excluded hypoadrenalism in all but one of them. This "cut-off" performed slightly better than both the Abbott and GC-MS serum cortisol "cut-offs" which diagnosed hypoadrenalism in 2 and 4 healthy volunteers respectively. In each of these cases cortisol concentration fell below the cut-off with only one test i.e. either salivary, serum Abbott or serum GC-MS cortisol.

This confirms the results of other studies that have explored salivary cortisol as an alternative to serum for evaluating adrenal disease; although the reference ranges and cut-off derived in this study are considerably different to those quoted in the literature. This may be explained by assay differences; a study by Deutschbein *et al* used a modified serum cortisol RIA to measure salivary cortisol in 21 healthy controls and reported a mean peak salivary cortisol of 60.6 nmol/L at 30 minutes (Deutschbein et al. 2009). Different sampling times may also be to blame; Perogamvros *et al*

reported a mean peak salivary cortisol at 60 minutes of 36.7 nmol/L in 68 patients with suspected adrenal insufficiency in whom disease was subsequently excluded (Perogamvros et al. 2010b), although the 30 minute response was 24.6 nmol/L which compares well to the mean of 19.3 nmol/L reported in this study. Cohort size is also likely to play a part in the differences observed between mean concentrations due to the wide range of responses and the greater impact this will have on mean concentration with fewer subjects.

Differences between salivary cortisol LC-MS/MS assays also make it difficult to establish a single, universal reference range; however, the range established by this study for the post-Synacthen response is narrower than many of those reported in the literature and includes the largest cohort of healthy volunteers yet studied. This provides a good starting point for further work to confirm the validity of the "normal" response, as defined in healthy volunteers, in a group of patients with unequivocal adrenal disease.

Further evaluation of salivary cortisol, using a well-defined healthy population with no co-morbidities such as hypertension, depression or high BMI, which can confound baseline cortisol concentration, would also be of use as it may provide a narrower range and distribution of baseline concentrations with less overlap with patients with possible hypoadrenalism, enabling the definition of a baseline salivary cortisol concentration which distinguishes adrenal insufficiency from adequate adrenal function without the need for a Synacthen test. This would be invaluable as at present salivary cortisol offers little advantage over serum total cortisol in the evaluation of adrenal disease as both require a Synacthen stimulation test for meaningful interpretation, and serum cortisol immunoassays are currently quicker and less labour intensive than salivary LC-MS/MS assays.

4.6.2 Salivary cortisol as an alternative to serum total cortisol

Patients with suspected hypoadrenalism

The salivary Synacthen test compared well to the pre-test probability score in evaluating patients with suspected hypoadrenalism, identifying 8 of the 9 patients considered at high risk, and 14 of the 15 patients at low risk, correctly. A further 5 patients had an indeterminate score, highlighting the limitations of the pre-test probability scoring criteria; but these will not be considered in more detail here.

Only 2 patients had a Synacthen test result that was at odds with the pre-test probability score: patient 5, who has been considered in some detail in chapter 3 but whose salivary cortisol result warrants some further discussion here and patient 9.

Patient 5 had a high pre-test probability score of 6 but passed the salivary Synacthen test. Her baseline salivary cortisol concentration was 12.2 nmol/L, rising by 5.3 nmol/L, to 17.5 nmol/L post-Synacthen. This contrasts sharply with the corresponding mean baseline cortisol and post-Synacthen increment of 2.7 nmol/L and 17.4 nmol/L, respectively, in healthy female volunteers, raising the possibility of sample contamination with oral or inhaled glucocorticoid as the cause of the discrepant outcome. However, as LC-MS/MS assays are specific for cortisol, the patient would need to have taken hydrocortisone specifically, just before producing the baseline sample for interference to affect the 30-minute sample as well (Perogamvros et al. 2010a), and she did not report this being the case. Furthermore, subsequent review of her records 5 years later confirmed that she had not required treatment for adrenal insufficiency nor had she undergone further testing, thus confirming the outcome of the test.

Patient 9 was defined as low risk for adrenal insufficiency, scoring 2 on the pre-test likelihood scale. He had hypogonadotrophic hypogonadism and gynaecomastia, with

no biochemical or radiological evidence of pituitary or autoimmune disease. His baseline salivary cortisol was 2.9 nmol/L, rising to 9.9 nmol/L post-Synacthen; a borderline fail. Again this raises questions about the validity of the pre-test probability scoring system as his hypogonadotrophic hypogonadism could be an indication of pituitary disease, although the lack of other radiological or biochemical evidence of disease is somewhat against this.

A review of his medical notes 5 years after the Synacthen test revealed he missed a follow-up appointment 1 year after the test and had no further contact with the hospital. This makes it impossible to either confirm or exclude subsequent adrenal insufficiency; although it is unlikely given that he has not re-presented to hospital in the intervening years. Nevertheless, this highlights the need for some flexibility when using a single cut-off concentration to interpret the Synacthen test, particularly at concentrations close to the cut-off. It all demonstrates the value of using a pre-test probability of disease when interpreting the result, particularly one that has been properly validated.

As discussed in chapter 3, subsequent follow up of the remaining patients with suspected hypoadrenalism confirmed adrenal insufficiency in 7 patients (1, 2, 10, 11, 12, 23 and 29), all of whom had a pre-test probability of 7 or higher and were correctly identified by the salivary Synacthen test.

Comparison between serum and salivary Synacthen tests:

This study showed good correlation between salivary and serum total cortisol concentrations in both healthy volunteers and patients with suspected hypoadrenalism, despite the former reflecting free cortisol concentration and the latter free and protein-bound cortisol combined.

Direct comparison between the two tests showed a different outcome for the salivary Synacthen test in 3/139 (2%) healthy volunteers when compared to serum cortisol

measured by the Abbott immunoassay and in 5/139 (4%) volunteers when compared to GC-MS cortisol. This difference rose to 21% in patients with suspected hypoadrenalism, with disagreement between salivary cortisol and serum cortisol measured by the Abbott immunoassay in 6 of the 29 patients studied.

Nevertheless, based on clinical outcome 5 years later and an assumption that patients who were lost to follow-up had not succumbed to hypoadrenalism, and using the cut-offs derived in this thesis, both tests had 100% sensitivity for detecting genuine hypoadrenalism and a negative predictive value of 100% for excluding disease. Specificity was 82% for salivary cortisol and 78% for serum cortisol, whilst the positive predictive value was 64% for salivary cortisol and 54% for serum cortisol. These results suggest, therefore, that the salivary Synacthen test may be marginally superior to serum for evaluating hypoadrenalism.

Limitations:

This work aimed to define and validate a salivary cortisol cut-off for use with the Synacthen test and whilst it has achieved this to a certain extent, inclusion of a larger group of patients with confirmed adrenal insufficiency would have allowed more robust testing of the cut-off. Further work comparing the salivary Synacthen test to the insulin tolerance test, which is still considered the gold-standard for diagnosing hypoadrenalism, would also help confirm that the cut-off performs as well as this study suggests. However, such a study would prove challenging to undertake, especially in healthy volunteers, given the unpleasant effects of hypoglycaemia.

The scoring system used in this study to assign a pre-test likelihood of hypoadrenalism was designed to create a reference point for each patient to which the serum and salivary cortisol results could be compared, but it has not been properly validated as a tool for this purpose. Comparison of the pre-test score to the confirmed diagnosis 5 years after the Synacthen test showed it had 100% sensitivity for

detecting hypoadrenalism but only 65% specificity and positive and negative predictive values of 47% and 100% respectively. Thus, although it matches both serum and salivary Synacthen tests' sensitivity and negative predictive value, it has worse specificity because of the inclusion of an indeterminate risk group. A better alternative might, therefore, be to seek consensus from a cohort of Endocrinologists as to the criteria that constitute high risk and use these to develop a more robust, universal scoring system which can then be tested in a patient population.

Finally, as both salivary cortisol and cortisone have been shown to correlate well with serum free cortisol (Perogamvros et al. 2010a), including the latter would have allowed direct comparison between the two to determine which performed best. However, as salivary cortisol has been shown to correlate better with serum free cortisol when patients treated with hydrocortisone are excluded, it was considered to be the better option. Salivary cortisone concentration is also influenced by 11β-HSD2 activity, which adds further potential for misinterpretation, particularly in patients with impaired enzyme activity.

4.6.3. Conclusions

Saliva is a valid alternative to serum cortisol for interpreting the Synacthen test in suspected adrenal insufficiency and may offer marginally better specificity than serum. It also offers the advantage of less invasive sample collection, although as Synacthen is administered by either intravenous or intramuscular injection, this remains a relatively invasive test.

Salivary cortisol concentrations correlate well with serum cortisol, with good agreement between the serum and salivary cortisol response to Synacthen stimulation in healthy volunteers. However, despite equally good correlation between serum and salivary concentrations in patients with suspected hypoadrenalism, there is disagreement between the results of the two tests in 1 in 5 patients studied. Further

comparison between the two tests and with the insulin tolerance test in a larger cohort of hypoadrenal patients would help to establish if either test clearly out-performed the other.

LC-MS/MS is the method of choice for measuring salivary cortisol but until there is better assay standardisation, a single salivary cortisol reference range is unlikely, which may limit the use of salivary cortisol assays in routine clinical practice. A single salivary cortisol cut-off for interpreting the Synacthen test is also an attractive prospect, but its validity needs to be confirmed in patients with hypoadrenalism before it can be used widely.

CHAPTER 5

SERUM FREE CORTISOL IN EVALUATING THE ADRENAL AXIS

5.1 Introduction

The initial aims of this study were to define assay-specific serum total cortisol reference ranges and a salivary cortisol reference range for the response to Synacthen in healthy volunteers and to explore the effect of high CBG on this response. However, the limitations of serum total cortisol in evaluating adrenal function are well known and the availability of serum total cortisol and CBG measurements in such a large cohort presented an opportunity to explore the calculated serum free cortisol response to Synacthen stimulation in healthy volunteers as well.

Two calculated surrogates of serum free cortisol were selected: Coolens' calculated serum free cortisol (SFC) and the free cortisol index (FCI) (Coolens et al. 1987; le Roux et al. 2002). SFC was the first such calculation to be derived and has been shown to correlate well with free cortisol measured by ultrafiltration and by LC-MS/MS (Coolens et al. 1987; Pretorius et al. 2011). FCI was subsequently derived as a simpler alternative to SFC and has been shown to correlate well with both SFC and serum free cortisol measured by gel-filtration in healthy volunteers (le Roux et al. 2002).

Calculated serum free cortisol is of particular value in patients with altered serum protein concentration (le Roux et al. 2002; Vincent et al. 2009; Fede et al. 2014), although its performance in critically ill patients is somewhat less reliable (Bendel et al. 2008; Cohen et al. 2013; Molenaar et al. 2015). There are also concerns about the validity of adopting a single reference range, due to poor comparison between methods and variable correlation between calculated and measured serum free cortisol (Pretorius et al. 2011). It is, therefore, essential that a suitable "reference range" for calculated free cortisol is defined in a healthy population before using it to evaluate adrenal function in other patient populations.

5.2 Aims

The work described in this chapter aimed to:

- Evaluate the role of calculated serum free cortisol in the interpretation of the short Synacthen test in healthy volunteers and to define a lower reference limit for this response, if possible.
- Explore the relationship between calculated serum free cortisol, serum total cortisol and salivary cortisol in healthy volunteers.
- Explore serum free cortisol as an alternative to serum total cortisol and salivary cortisol in the interpretation of the short Synacthen test in patients with presumed disease of the HPA axis.

5.3 Subjects and sample analysis

5.3.1 Subjects

This arm of the study included samples from 139 healthy volunteers (60 male and 79 female) and 30 patients with possible hypoadrenalism.

5.3.2 Laboratory investigations

Serum total cortisol was measured by GC-MS and the five automated immunoassays described in chapter 2 in the healthy volunteers and by the Abbott assay only in the suspected hypoadrenal patient group. CBG was measured by RIA (Biosource, Nivelles, Belgium) in both groups, although analysis was undertaken 4 years later in the patient group than in the healthy volunteers.

5.3.3 Calculated serum free cortisol

Serum free cortisol (SFC), was calculated from the Coolens' equation as follows:

$$U = \sqrt{(Z^2 + 0.0122T)} - Z$$

where $Z = 0.0167 + 0.182(G - T)$ and U, G and T are SFC, CBG and total cortisol in μmol/L, respectively.

The free cortisol index (FCI) was calculated by dividing measured total cortisol by CBG and is reported in nmol/mg.

5.3.4 Statistical analysis

SFC and FCI were log-transformed for calculation of mean concentrations in the two groups and reference ranges in the healthy volunteers. The three patients with established Addison's disease were excluded from calculations of the mean to avoid skewing the results. The unpaired t-test was used to compare means between the two patient groups and between genders within the healthy volunteer group, while the paired t-test was used to compare mean SFC and FCI at each time point.

Assay-specific reference ranges were determined in healthy volunteers and were compared using the paired t-test. The lower limit of the reference range for the Abbott assay was used as a cut-off to differentiate between normal adrenal function and deficiency, and comparison was made between the free cortisol Synacthen test outcome and the pre-test probability of HPA axis disease, as defined in chapter 2. A similar comparison was made between the outcome of the free cortisol Synacthen test and the serum and salivary cortisol results. The relationships between SFC and FCI, calculated free cortisol and serum total cortisol and calculated free cortisol and salivary cortisol were explored using scatter plots. In all cases, a p-value <0.05 was considered significant.

5.4 Results

5.4.1 CBG concentration

CBG concentration was significantly lower in males than in females, with no significant age effect and no difference between concentrations at baseline and post-Synacthen. Mean concentrations (2.5^{th} – 97.5^{th} centile) were 58 (42 – 81) mg/L in men and 64 (43 – 95) mg/L in women.

5.4.2 Serum free cortisol and free cortisol index

5.4.2.1 Baseline

Coolens' serum free cortisol (SFC) calculated from GC-MS cortisol correlated well with the free cortisol index (FCI), although the relationship was non-linear (figure 5.1). Baseline SFC more closely resembled serum total cortisol in that it was normally distributed in males, but not females, whereas FCI was normally distributed in both genders. In contrast to serum total cortisol, differences between genders were significant ($p<0.005$) for both SFC and FCI at baseline (tables 5.1 and 5.2, figure 5.2); and for each gender, baseline SFC was significantly higher than FCI ($p<0.005$) (figures 5.1 & 5.2); with mean concentrations of 9.5 and 4.6 nmol/mg respectively in males and 7.7 and 3.9 nmol/mg in females (tables 5.1 and 5.2).

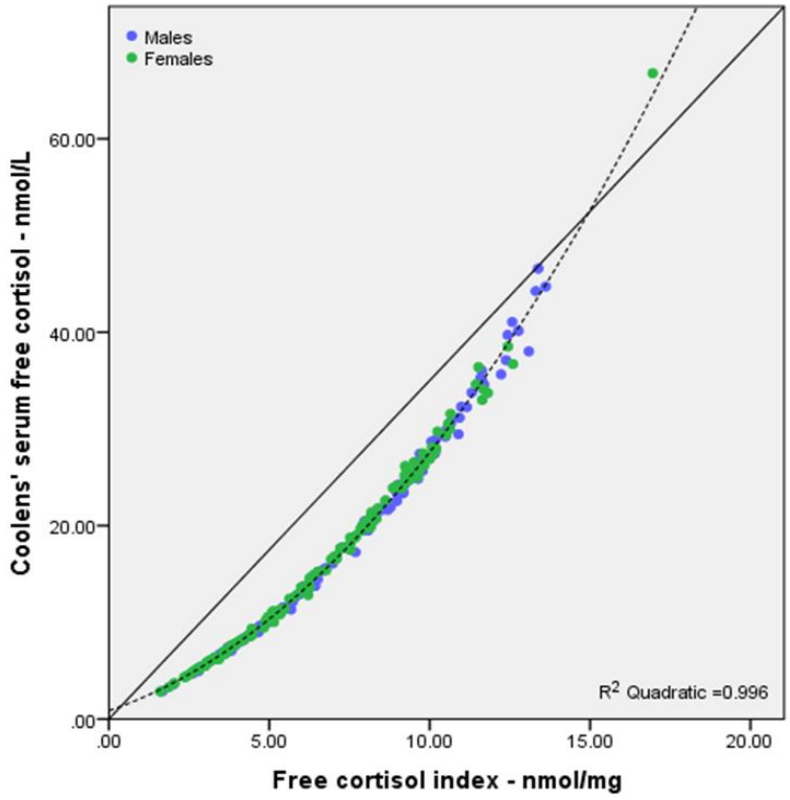


Figure 5.1 Correlation between GC-MS SFC and FCI. The solid line shows perfect correlation between methods whilst the dotted line is the line of best fit which is non-linear.

Table 5.1 Assay-specific baseline serum free cortisol in healthy volunteers.

	Mean Serum Free Cortisol (2.5th – 97.5th centile) (nmol/L)		
Assay	**Males n = 60**	**Females n = 79**	**P-value***
GC-MS	9.5 (4.0 – 22.6)	7.7 (3.5 – 16.7)	P<0.005
Centaur	10.4 (4.5 – 23.9)	8.1 (3.6 – 18.4)	P<0.001
Abbott	10.0 (4.2 – 24.0)	7.7 (3.5 – 17.0)	P<0.001
Roche	13.9 (5.1 – 37.5)	9.6 (4.0 – 23.1)	P<0.001
Immulite	11.8 (4.6 – 30.1)	8.5 (3.8 – 19.0)	P<0.001
Beckman	10.2 (4.6 – 22.5)	7.9 (3.5 – 17.5)	P<0.001

*P-value <0.05 indicates a significant difference between means.

Table 5.2 Assay-specific baseline FCI in healthy volunteers.

	Mean Free Cortisol Index (2.5th – 97.5th centile) (nmol/mg)		
Assay	**Males n = 60**	**Females n = 79**	**P-value***
GC-MS	4.6 (2.3 – 9.3)	3.9 (2.0 – 7.4)	P<0.005
Centaur	5.0 (2.6 – 9.7)	4.0 (2.0 – 8.0)	P<0.001
Abbott	4.8 (2.4 – 9.8)	3.9 (2.0 – 7.6)	P<0.001
Roche	6.2 (2.9 – 13.0)	4.6 (2.3 – 9.4)	P<0.001
Immulite	5.5 (2.7 – 11.0)	4.2 (2.2 – 8.2)	P<0.001
Beckman	4.9 (2.6 – 9.2)	4.0 (2.0 – 7.7)	P<0.001

*P-value <0.05 indicates a significant difference between means.

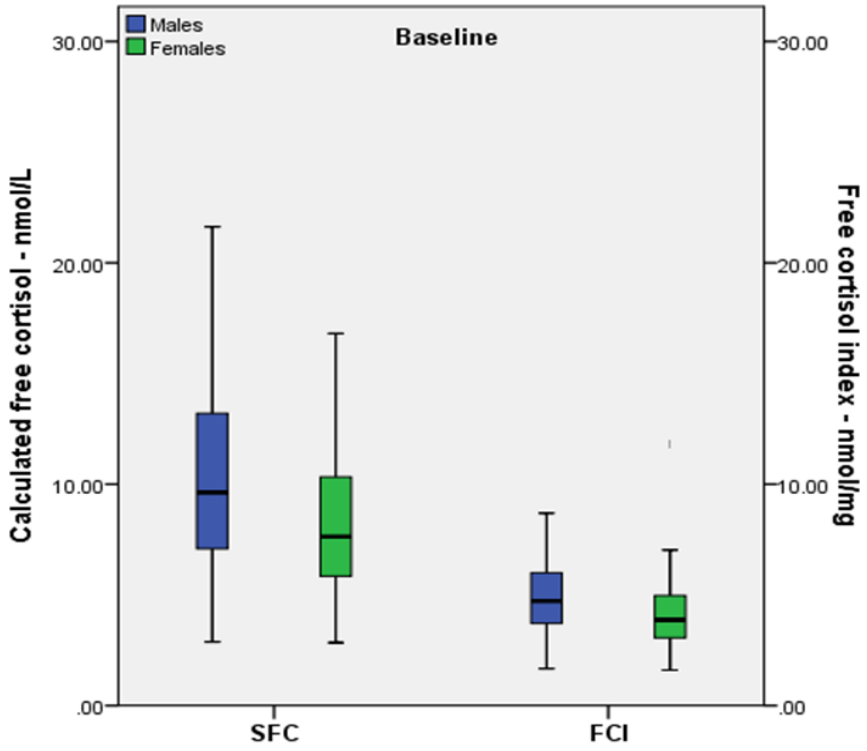


Figure 5.2 Baseline GC-MS SFC and FCI in healthy volunteers. The difference between genders was significant (p<0.005) for both SFC and FCI and SFC was significantly higher than FCI (p<0.005).

5.4.2.2 Post-Synacthen

Following Synacthen stimulation, SFC was not normally distributed in males or females, while FCI remained normally distributed in both. Gender differences remained significant (p<0.005) (tables 5.3 and 5.4, figure 5.3) and SFC remained significantly higher than FCI (p<0.005) (figure 5.3); with mean concentrations of 26.5

and 9.7 nmol/mg in males and 23.0 and 8.8 nmol/mg in females, respectively (tables 5.3 and 5.4).

Table 5.3 Assay-specific post-Synacthen SFC in healthy volunteers.

	Mean Serum Free Cortisol (2.5th – 97.5th centile) (nmol/L)		
Assay	**Males n = 60**	**Females n = 79**	**P-value***
GC-MS	26.5 (15.3 – 45.7)	23.0 (13.4 – 39.5)	P<0.005
Centaur	29.7 (16.6 – 53.2)	24.4 (13.4 – 44.3)	P<0.001
Abbott	27.9 (15.8 – 49.2)	22.0 (12.5 – 38.8)	P<0.001
Roche	46.9 (24.2 – 91.1)	34.7 (18.0 – 66.9)	P<0.001
Immulite	33.4 (17.6 – 63.3)	27.9 (15.1 – 51.5)	P=0.001
Beckman	32.0 (16.9 – 60.6)	25.5 (13.4 – 48.6)	P<0.001

*P-value <0.05 indicates a significant difference between means.

Table 5.4 Assay-specific post-Synacthen FCI in healthy volunteers

	Mean Free Cortisol Index (2.5th – 97.5th centile) (nmol/mg)		
Assay	**Males n = 60**	**Females n = 79**	**P-value**
GC-MS	9.7 (6.7 – 14.0)	8.8 (6.0 – 12.8)	P<0.005
Centaur	10.4 (7.2 – 15.1)	9.1 (6.1 – 13.6)	P<0.001
Abbott	10.0 (6.9 – 14.5)	8.5 (5.7 – 12.7)	P<0.001
Roche	13.4 (9.1 – 19.6)	11.2 (7.5 – 16.8)	P<0.001
Immulite	11.1 (7.5 – 16.5)	9.9 (6.6 – 14.8)	P=0.001
Beckman	10.9 (7.3 – 16.2)	9.4 (6.1 – 14.4)	P<0.001

*P-value <0.05 indicates a significant difference between means.

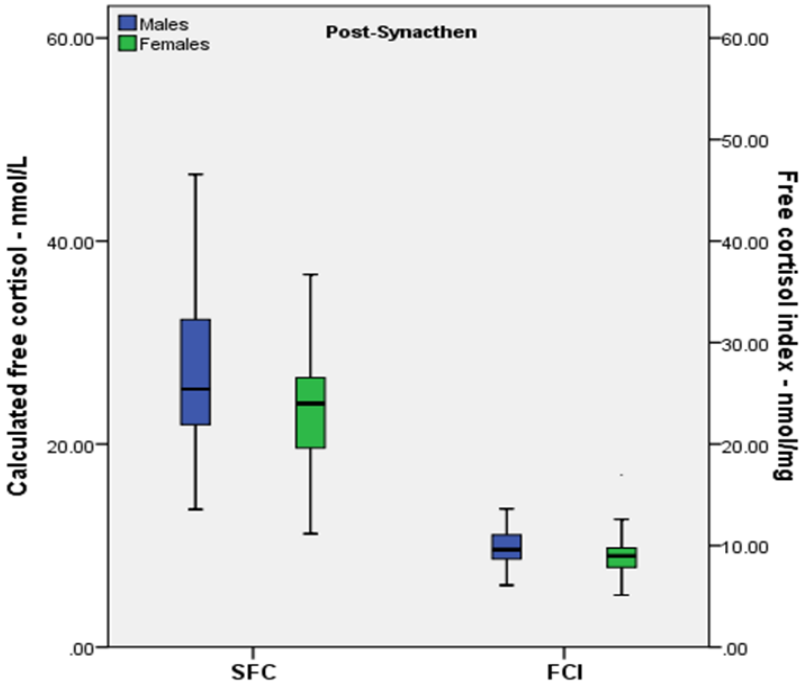


Figure 5.3 Post-Synacthen GC-MS SFC and FCI in healthy volunteers. The difference between genders was significant ($p<0.005$) for both SFC and FCI; and SFC was significantly higher than FCI ($p<0.005$).

5.4.2.3 Assay-specific ranges

Separate male and female SFC and FCI were calculated for each immunoassay at each time-point (tables 5.1 – 5.4) and, for males, all were shown to differ significantly from those calculated from GC-MS cortisol ($p<0.05$) (figures 5.4 and 5.5). Comparison between immunoassays in males showed no difference between either parameter for the Beckman assay when compared to the Centaur and Abbott assays at baseline (figure 5.4), although all assays were significantly different from each other post-Synacthen (figure 5.5).

For females, and in line with serum total cortisol, baseline Abbott and Beckman SFC and FCI did not differ significantly from each other or from those calculated from GC-MS cortisol (figure 5.4). Post-Synacthen, all immunoassays differed from each other and from GC-MS (figure 5.5).

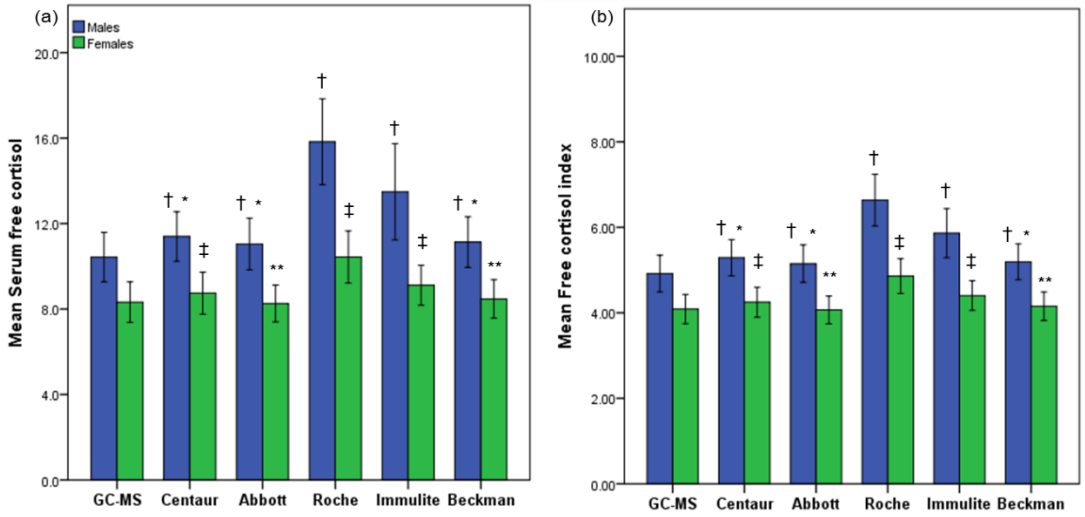


Figure 5.4 Mean assay-specific baseline SFC (a) and FCI (b) in males and females**.** Comparisons between assays were made within gender groups. † denotes statistically significant difference (P<0.05) from GC-MS cortisol in males; ‡ denotes statistically significant difference (P<0.05) from GC-MS cortisol in females; * and ** denote immunoassays which do not differ significantly from each other in males and females, respectively.

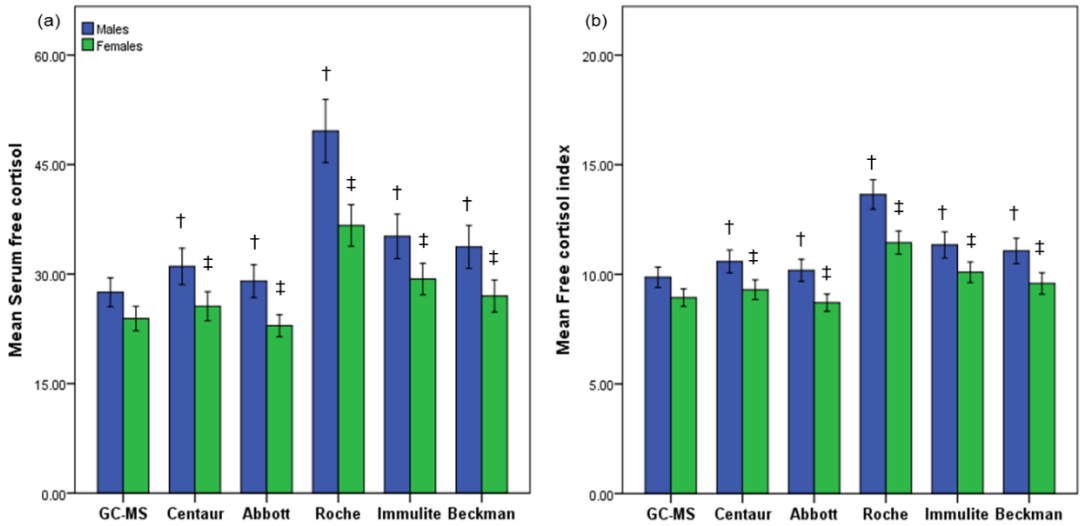


Figure 5.5 Mean assay-specific post-Synacthen SFC (a) and FCI (b) in males and females. † denotes statistically significant difference (P<0.05) from GC-MS cortisol in males; ‡ denotes statistically significant difference (P<0.05) from GC-MS cortisol in females

5.4.2.4 Lower reference limits

The 2.5th percentile, or lower reference limit, of the GC-MS-calculated free cortisol response to Synacthen stimulation was 15.3 in males and 13.4 in females for SFC and 6.7 nmol/mg in males and 6.0 nmol/mg in females for FCI. The corresponding lower reference limits for immunoassay SFC ranged between 15.8 and 24.2 in males and 12.5 to 18.0 in females (table 5.3) and for FCI, between 6.9 and 9.1 nmol/mg in males and 5.7 and 7.5 nmol/mg in females (table 5.4).

5.4.3 Correlation with serum total cortisol and salivary free cortisol

5.4.3.1 Serum

SFC and FCI correlated equally well with GC-MS cortisol across the range of concentrations studied, with no obvious gender difference (figures 5.6A and 5.7A), although the relationship between SFC and GC-MS cortisol was non-linear. The relationship between log-transformed calculated free cortisol and serum cortisol concentration was also explored, and showed better correlation (figures 5.6B and 5.7B) than the raw data, with less scatter and a clear linear relationship in both cases. Subsequent comparisons between free cortisol calculated from immunoassay cortisol were therefore made using log-transformed data.

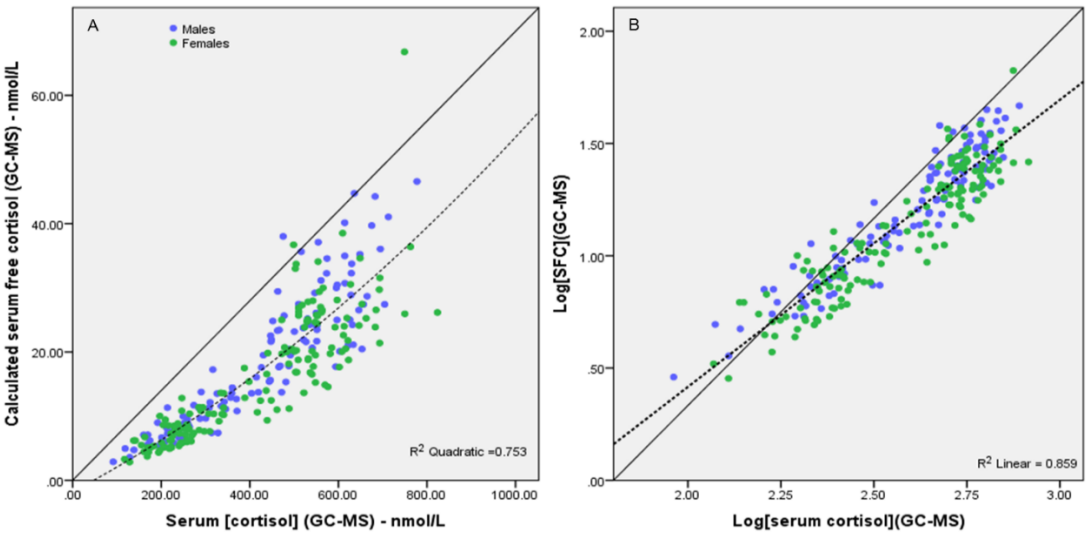


Figure 5.6 Correlation between GC-MS SFC and serum total cortisol measured by GC-MS (A) and between their log-transformed derivatives (B). The solid line represents perfect correlation; dotted line is the line of best fit.

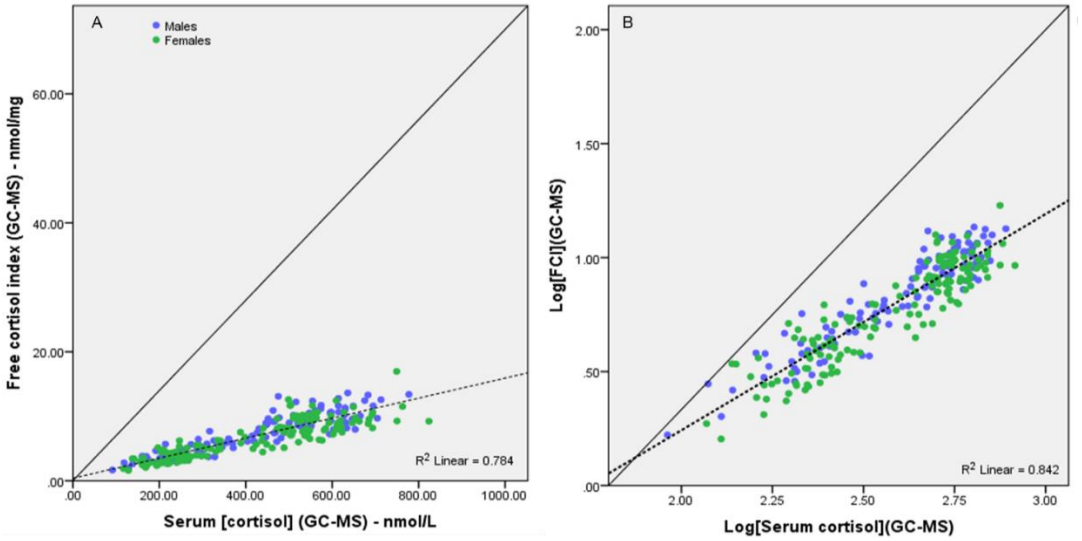


Figure 5.7 Correlation between GC-MS FCI and serum total cortisol measured by GC-MS (A) and between their log-transformed derivatives (B). The solid line represents perfect correlation; dotted line is the line of best fit.

The correlation between immunoassay SFC and FCI and serum total cortisol measured by GC-MS (figures 5.8 & 5.9) was broadly similar to that between GC-MS SFC and FCI and serum total cortisol, with minor differences in the overall relationships due to assay differences, particularly for SFC.

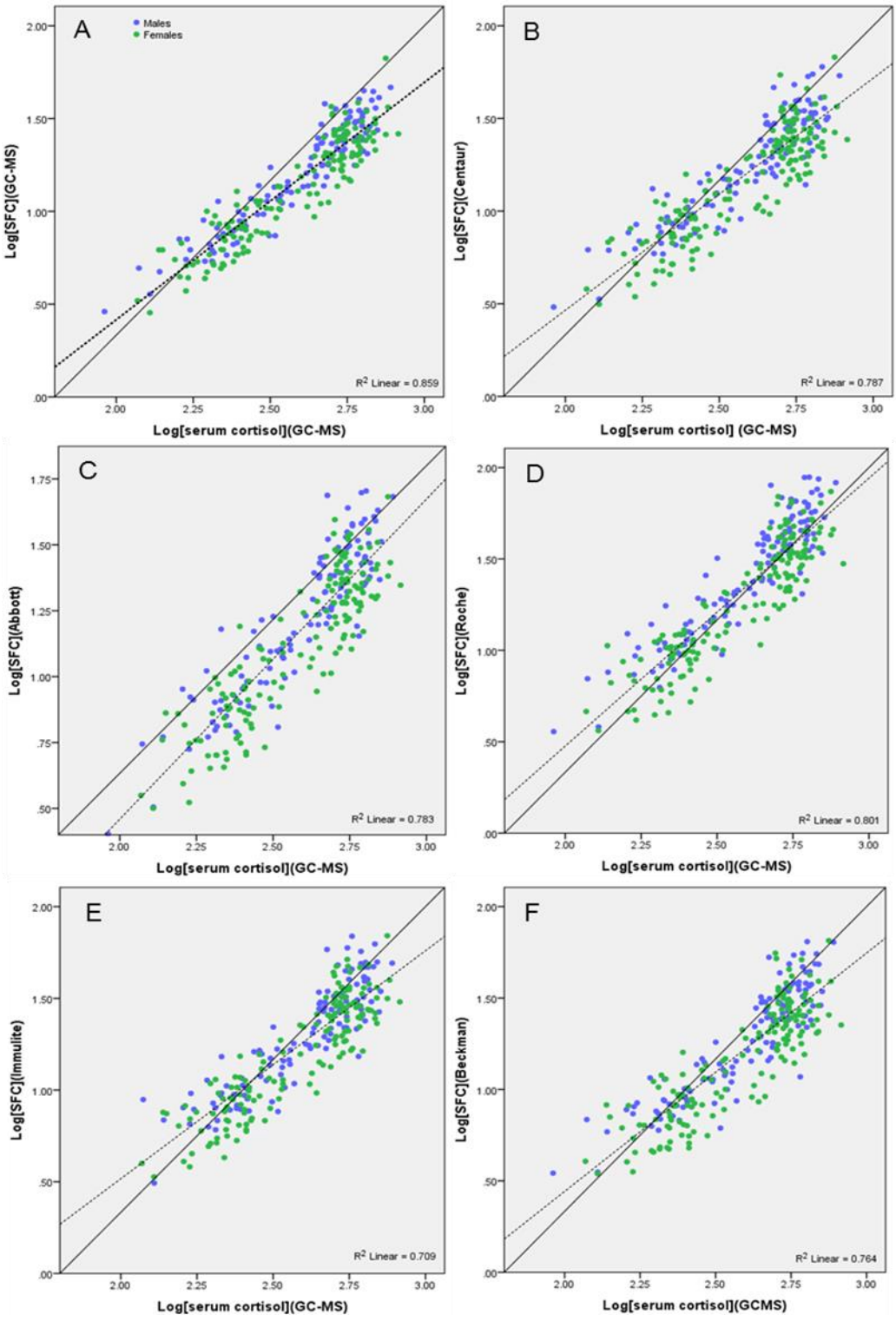


Figure 5.8 Correlation between log-transformed SFC and serum total cortisol measured by GC-MS. The graphs show the relationship between SFC calculated from cortisol measured by (A) GC-MS, (B) Centaur, (C) Abbott, (D) Roche, (E) Immulite and (F) Beckman assays, respectively. The solid line represents perfect correlation; dotted line is the line of best fit.

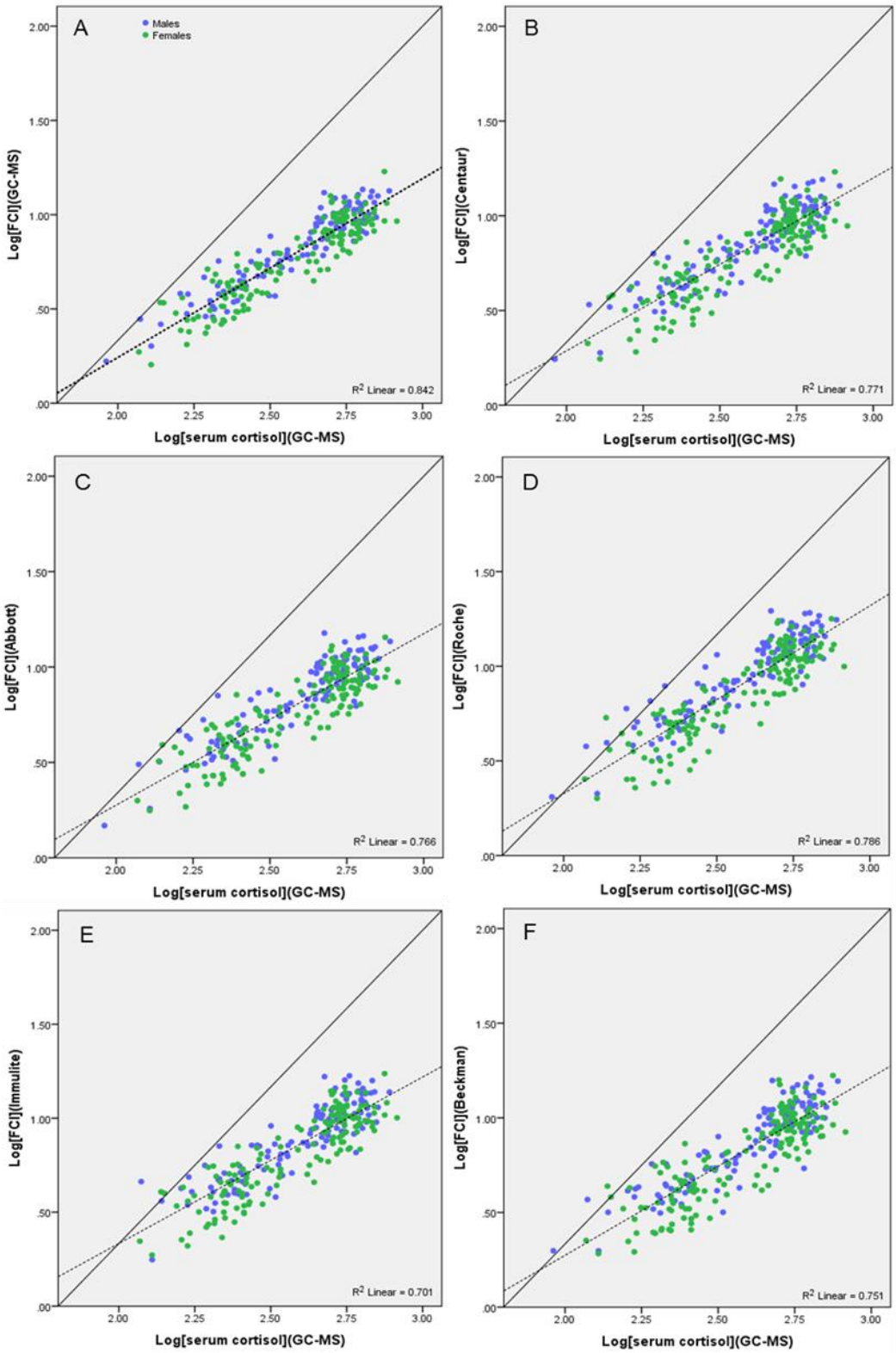


Figure 5.9 Correlation between log-transformed FCI and serum total cortisol measured by GC-MS. The graphs show the relationship between FCI calculated from cortisol measured by (A) GC-MS, (B) Centaur, (C) Abbott, (D) Roche, (E) Immulite and (F) Beckman assays, respectively. The solid line represents perfect correlation; dotted line is the line of best fit.

5.4.3.2 Saliva

GC-MS SFC and FCI correlated reasonably well with salivary cortisol (figures 5.10 & 5.11), although with wider scatter than the correlation with serum total cortisol, despite similar correlation coefficients. Correlation between the log-transformed data was, again, slightly improved over the raw data and was used for subsequent comparisons.

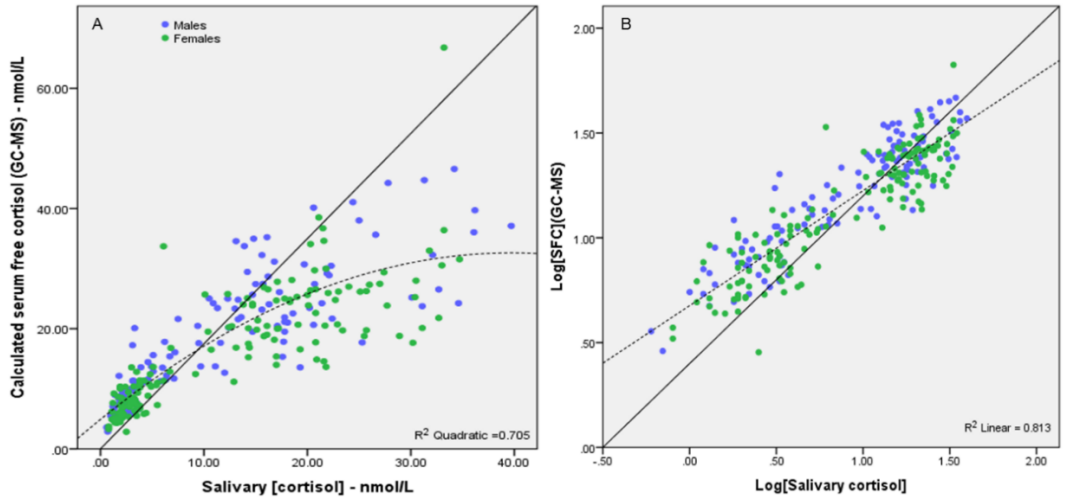


Figure 5.10 Correlation between GC-MS SFC and salivary cortisol (A) and between their log-transformed derivatives (B). The solid line represents equivalence between methods; dotted line is the line of best fit, which is non-linear for the direct correlation.

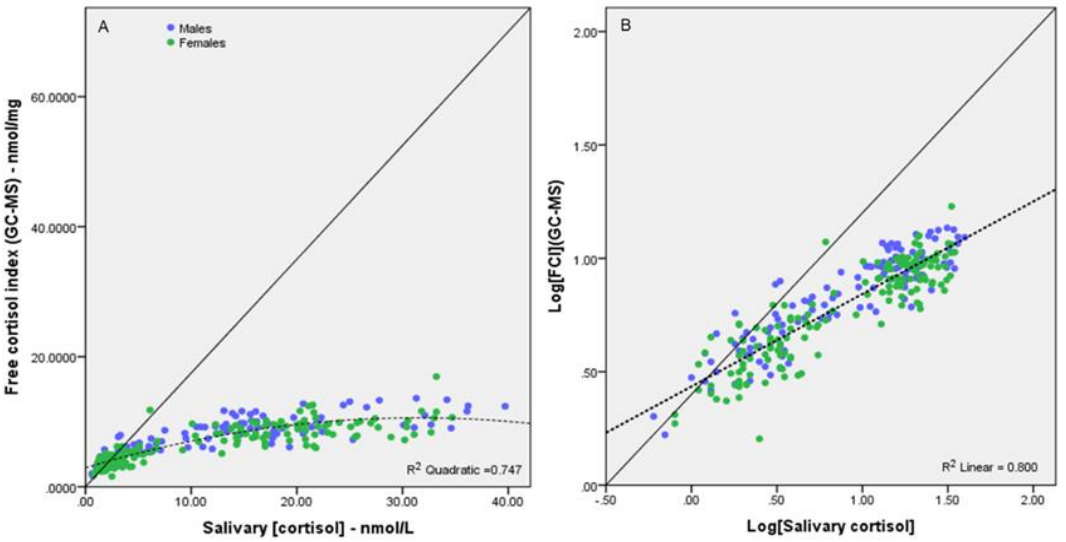


Figure 5.11 Correlation between GC-MS FCI and salivary cortisol (A) and between their log-transformed derivatives (B). The solid line represents equivalence between methods; dotted line is the line of best fit, which is non-linear for the direct correlation.

Immunoassay SFC and FCI also correlated well with salivary cortisol, with little difference between assays, apart from the Roche SFC which showed positive bias relative to GC-MS and the other immunoassays (figure 5.12). Correlation coefficients were similar (R^2 = 0.7 – 0.8) for most assays for both parameters; nevertheless, FCI appeared to show less in the way of assay-specific differences in its correlation with salivary cortisol than SFC (figure 5.12 & 5.13).

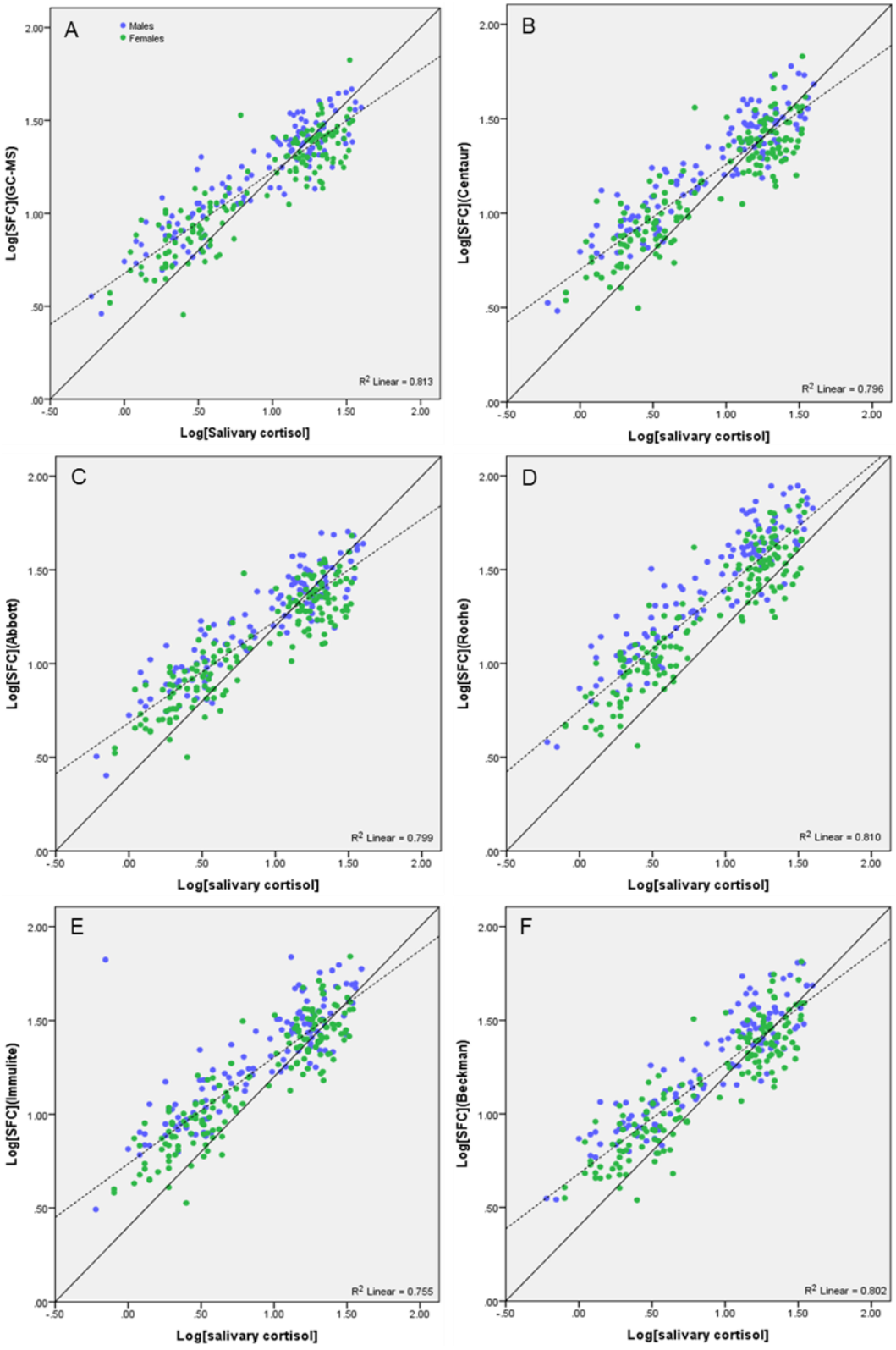


Figure 5.12 Correlation between log-transformed SFC and salivary cortisol. The graphs show the relationship between SFC calculated from cortisol measured by (A) GC-MS, (B) Centaur, (C) Abbott, (D) Roche, (E) Immulite and (F) Beckman assays, respectively. The solid line represents equivalence between methods; dotted line is the line of best fit.

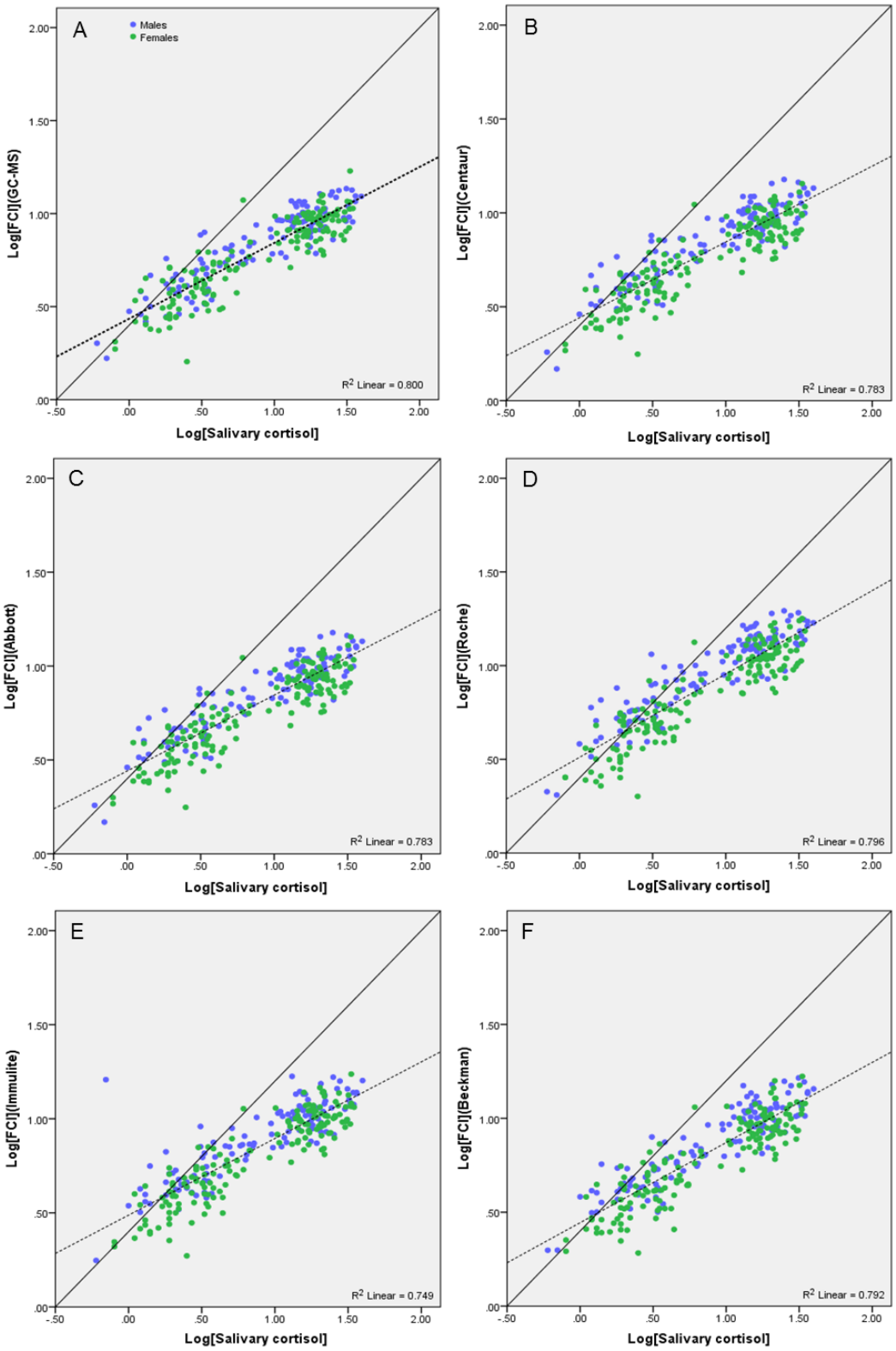


Figure 5.13 Correlation between log-transformed FCI and salivary cortisol. The graphs show the relationship between FCI calculated from cortisol measured by (A) GC-MS, (B) Centaur, (C) Abbott, (D) Roche, (E) Immulite and (F) Beckman assays, respectively. The solid line represents equivalence between methods; dotted line is the line of best fit.

5.4.4 Comparison between healthy volunteers and patients with potential hypoadrenalism

5.4.4.1 CBG and calculated free cortisol

CBG was significantly lower in patients with potential hypoadrenalism than in healthy volunteers; mean concentration 46 mg/L (28 – 73 mg/L) compared to 58 mg/L in healthy males and 64 mg/L in females ($p<0.001$). However, mean SFC and FCI in male patients were no different to those in healthy volunteers at baseline (10.7 vs 10.0; $p=0.648$ and 4.8 nmol/mg vs 4.8 nmol/mg; $p=0.960$) or post-Synacthen (31.3 vs 27.9, $p=0.247$ and 10.4 nmol/mg vs 10.0 nmol/mg; $p=0.584$).

Mean SFC and FCI in female patients were not significantly different to those in male patients at baseline ($p = 0.983$ and 0.880, respectively) or post-Synacthen ($p = 0.751$ and 0.949) and were, consequently, significantly higher than those in healthy female volunteers (10.8 vs 7.7; $p<0.005$ and 4.7 nmol/mg vs 3.9 nmol/mg; $p<0.001$ at baseline and 32.7 vs 22.0; $p<0.001$ and 10.3 nmol/mg vs 8.5 nmol/mg; $p<0.005$ post-Synacthen).

When male and female samples were combined, mean baseline SFC, but not FCI, was significantly higher in patients with suspected hypoadrenalism than in healthy volunteers (10.8 nmol/L vs 8.6 nmol/L, $p<0.05$; and for FCI, 4.8 nmol/mg vs 4.3 nmol/mg, $p = 0.138$). Post-Synacthen, both SFC and FCI were significantly higher in the patient group (32.0 nmol/L vs 24.4, $p<0.001$ and 10.3 nmol/mg vs 9.1 nmol/mg, $p<0.01$, respectively) (figures 5.14 & 5.15).

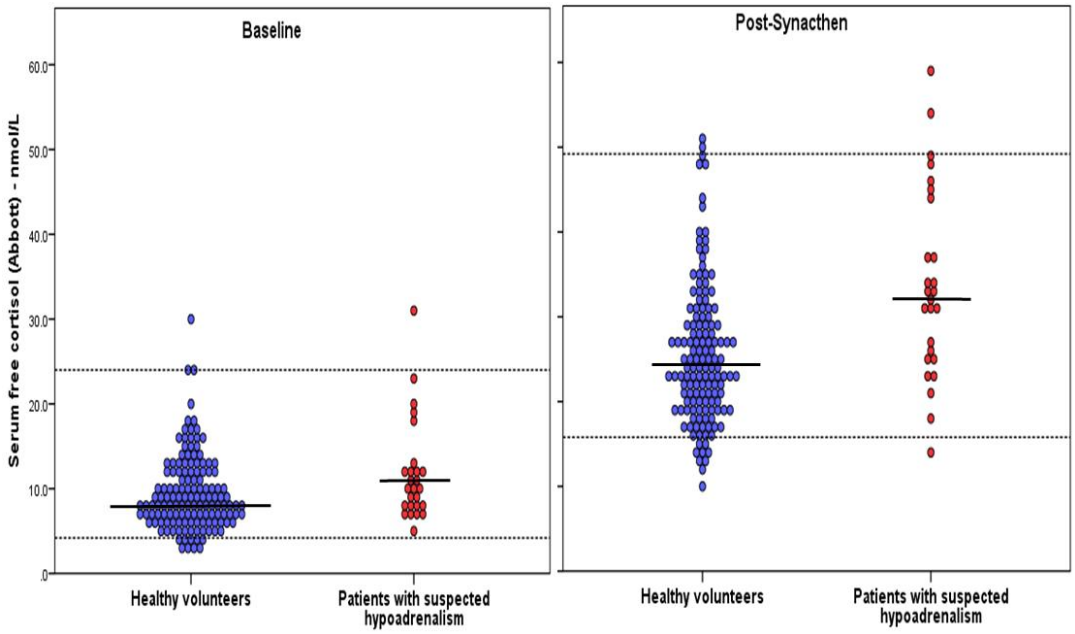


Figure 5.14 Baseline and post-Synacthen SFC in healthy volunteers and patients with suspected hypoadrenalism. Each dot represents an individual subject. The upper and lower dotted lines in each graph represent the 2.5th and 97.5th percentiles in healthy volunteers, while the dark continuous lines represent mean concentration.

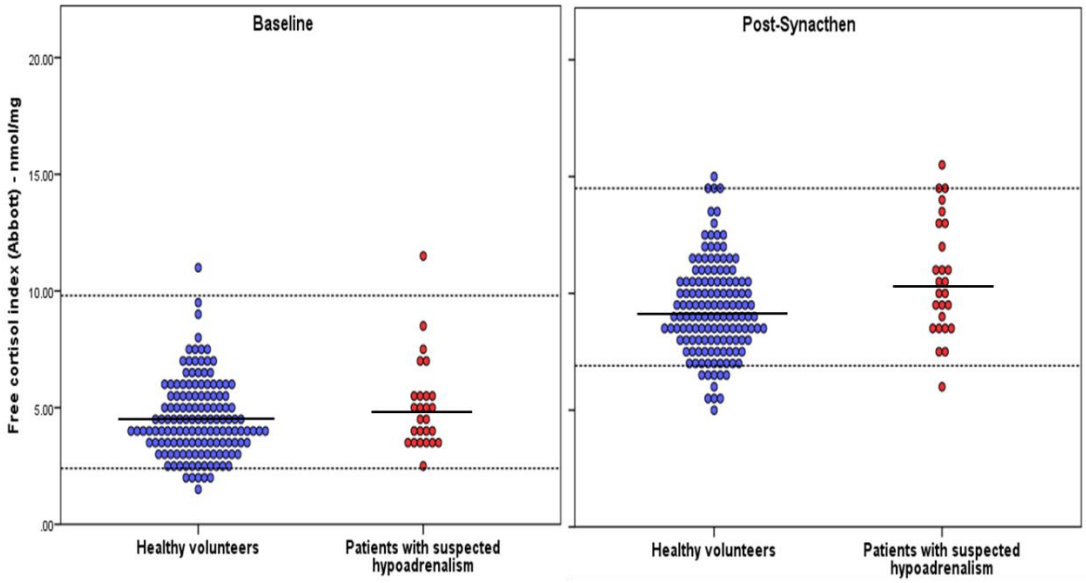


Figure 5.15 Baseline and post-Synacthen FCI in healthy volunteers and patients with suspected hypoadrenalism. Each dot represents an individual subject. The upper and lower dotted lines in each graph represent the 2.5th and 97.5th percentiles in healthy volunteers, while the dark continuous lines represent mean concentration.

5.4.4.2 Correlation with serum and salivary cortisol

SFC and FCI correlated well with serum total cortisol in patients with possible hypoadrenalism, with a similar pattern to that seen in healthy volunteers (figure 5.16), but with slight positive bias, particularly with SFC.

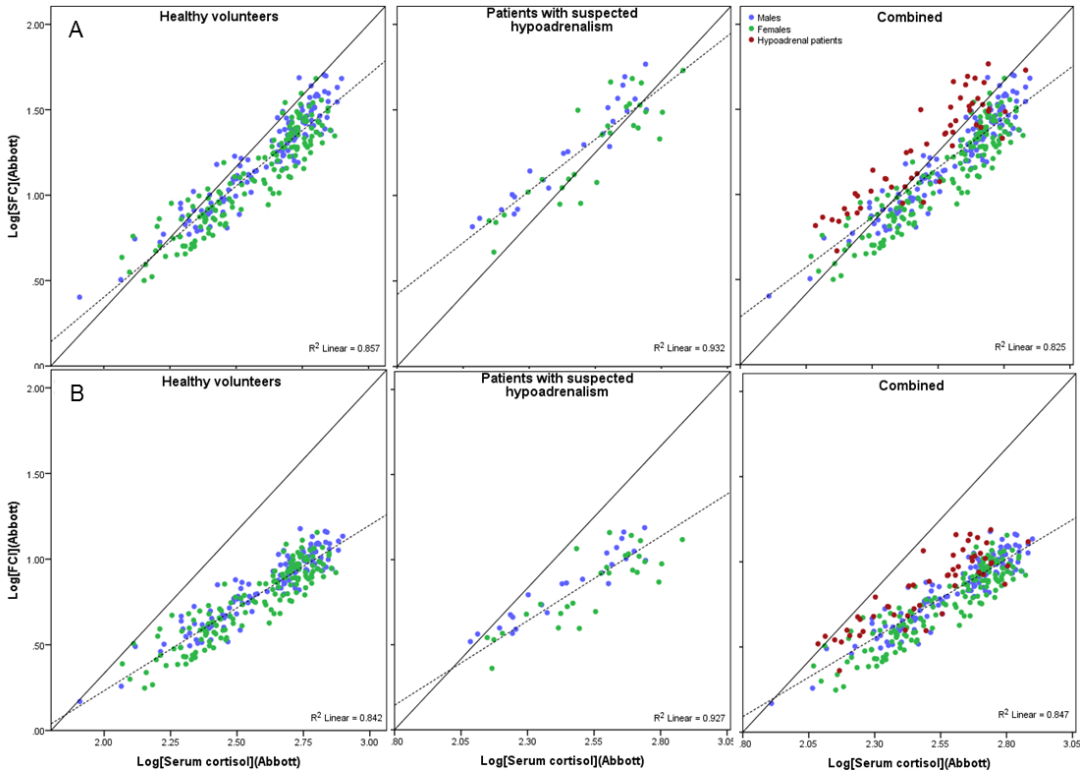


Figure 5.16 Correlation between log-transformed SFC and FCI and serum total cortisol. The graphs show the relationship between Abbott (SFC) and serum total cortisol measured by the Abbott assay (A) and Abbott FCI and serum total cortisol measured by the Abbott assay (B) in healthy volunteers and patients with suspected hypoadrenalism, separately and combined. The solid line represents equivalence between methods; the dotted line is the line of best fit.

The correlation between SFC and FCI and salivary cortisol in potentially hypoadrenal patients was also similar to that seen in healthy volunteers (figure 5.17) and the slight positive bias noted in the relationship between the calculated free cortisol parameters and serum total cortisol persisted.

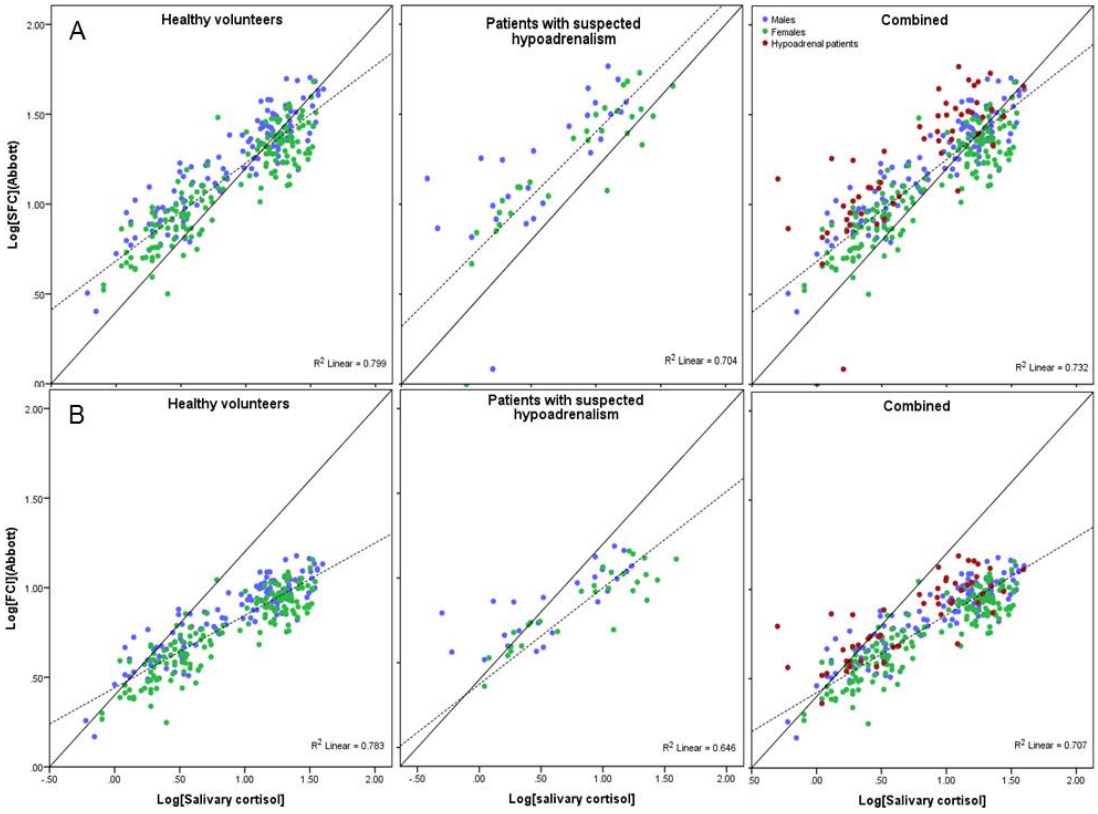


Figure 5.17 Correlation between log-transformed SFC and FCI and salivary cortisol. The graphs show the relationship between Abbott (SFC) and salivary cortisol (A) and Abbott FCI and salivary cortisol (B) in healthy volunteers and patients with suspected hypoadrenalism, separately and combined. The solid line represents equivalence between methods; the dotted line is the line of best fit.

5.4.4.3 Diagnosing hypoadrenalism

The outcome of the Synacthen test was the same in patients with possible adrenal insufficiency irrespective of whether SFC or FCI was used. Thus, using the lower reference limits derived above, four patients (14%) would have failed the Synacthen test with either parameter (table 5.5), compared to 2 male volunteers (3%) and 2 or 3 female volunteers (3-4%), depending on whether SFC or FCI was used. This contrasts with the outcomes seen with both serum and salivary cortisol whereby 12 and 10 patients, respectively, failed the Synacthen test, compared with only 3 healthy volunteers.

As discussed in chapters 3 and 4, a pre-test likelihood of adrenal insufficiency was determined for each patient, based on their presenting symptoms and medical history, and was used to explore agreement between high clinical suspicion of adrenal disease and test outcome. Calculated free cortisol performed poorly in this comparison, detecting only 4 of the 10 patients with a high pre-test probability (table 5.6). However, the pre-test probability over-estimated the presence of disease, as only 7 of the high risk patients were subsequently shown to have hypoadrenalism, either known at the time of testing or confirmed later. SFC and FCI correctly identified only 4 of these 7 patients (patients 1, 2, 12 and 23), compared to serum and salivary cortisol which identified them all. Calculated free cortisol thus appears to have greater specificity than either serum or salivary cortisol but at the expense of sensitivity.

Table 5.5 Pre-test probability of hypoadrenalism and outcome of post-Synacthen SFC and FCI in patients with suspected adrenal insufficiency.

Patient number	Gender	Age	Likelihood score	Post-Synacthen (SFC)	Outcome	Post-Synacthen (FCI)	Outcome
1	F	67	10	1.0	Fail	0.5	Fail
2	M	63	10	1.3	Fail	0.7	Fail
3	F	65	2	30.6	Pass	9.5	Pass
4	F	82	2	45.5	Pass	12.9	Pass
5	F	57	6	24.8	Pass	8.4	Pass
6	M	61	2	36.9	Pass	11.8	Pass
7	M	74	2	49.4	Pass	14.5	Pass
8	F	46	2	48.2	Pass	13.9	Pass
9	M	54	2	36.6	Pass	11.2	Pass
10	M	62	7	18.0	Pass	7.3	Pass
11	M	64	8	27.1	Pass	9.4	Pass
12	F	40	10	0.9	Fail	0.5	Fail
13	M	35	4	44.1	Pass	13.3	Pass
14	F	39	2	46.0	Pass	14.4	Pass
15	M	43	4	23.0	Pass	8.4	Pass
16	M	50	4	31.2	Pass	10.1	Pass
17	M	64	2	32.6	Pass	10.9	Pass
18	F	47	4	34.3	Pass	10.8	Pass
19	F	47	2	33.8	Pass	10.5	Pass
20	M	46	2	58.5	Pass	15.4	Pass
21	F	81	8	33.2	Pass	11.3	Pass
22	F	70	6	23.2	Pass	8.3	Pass
23	M	36	8	13.9	Fail	6.2	Fail
24	M	55	2	31.6	Pass	9.9	Pass
25	F	23	2	21.4	Pass	7.4	Pass
26	F	43	4	30.8	Pass	9.7	Pass
27	F	46	2	33.0	Pass	10.6	Pass
28	F	54	2	53.8	Pass	13.1	Pass
29	F	28	10	25.5	Pass	9.0	Pass
30	F	29	2	25.6	Pass	8.7	Pass

Pre-test probability was calculated from criteria set out in table 2.2. A low score was defined as 2 or less, a high score 5 or more and a score of 3 or 4 as indeterminate. SFC and FCI were calculated from serum cortisol measured by the Abbott assay with a “pass” defined as an SFC ≥ 15.8 in males and ≥ 12.5 in females or an FCI ≥ 6.9 nmol/mg in males and ≥ 5.7 nmol/mg in females.

Table 5.6 Agreement between pre-test probability of hypoadrenalism and the outcome of the Synacthen test.

	Synacthen test			
Pre-test probability	**Pass**	**Fail**	**Totals**	**Percentage agreement**
High	6	4	10	**40%**
Low	15	0	15	**100%**
Indeterminate	5	0	5	**-**
Totals	**26**	**4**	**30**	**76%**

This table depicts the outcome of the Synacthen test with both SFC and FCI, as their performance did not differ. Agreement is defined as a failed Synacthen test when pre-test probability is high and a Synacthen test pass when pre-test probability is low.

5.5 Discussion

5.5.1 Cortisol binding globulin

5.5.1.1 Healthy volunteers

CBG concentrations in healthy volunteers ranged between 38 and 88 mg/L in men and 38 and 99 mg/L in women, with a significantly lower mean concentration in men (58 mg/L) than in women (64 mg/L). The clinical relevance of this difference is debatable, given the wide range of concentrations observed and it contrasts with the results of several other studies which show a single reference range for healthy volunteers (Coolens et al. 1987; Davidson et al. 2006; Poomthavorn et al. 2009; Vincent et al. 2009; Barlow et al. 2010; Perogamvros et al. 2011). These ranges are typically wide, but the mean concentration tends to be lower than that reported in this study, often falling between 45 and 50 mg/L (Qureshi et al. 2007; Poomthavorn et al. 2009; Tan et al. 2010) or even lower (Coolens et al. 1987).

However, patient numbers in these studies are small and CBG assays are not standardised and have no external quality assurance scheme. Thus, it is not possible to confirm that different assays produce comparable results or, indeed, that different batches of the same assay are comparable. Manufacturer-derived reference ranges

can also be unreliable, due to small patient numbers and inadequate selection criteria e.g. inclusion of women taking exogenous oestrogens.

It is, therefore, not surprising that few studies show agreement with the results presented here. Although, in their study of serum free cortisol in a group of patients with bronchiectasis, Barlow *et al* reported equally wide CBG concentrations, ranging between 28 and 85 mg/L, noting, as with this study, that this was wider and somewhat higher than the manufacturer's quoted range of 31 – 53 mg/L (Barlow et al. 2010).

5.5.1.2 Patients with suspected hypoadrenalism

CBG concentrations in the potentially hypoadrenal patient group ranged between 26 and 84 mg/L, with a mean concentration of 46 mg/L which was significantly lower than that in healthy volunteers. This was an unexpected finding for which there are several possible explanations. Samples from this patient group were collected 1 – 2 years later than the samples from healthy volunteers and were stored frozen at -80°C for up to 4 years prior to analysis. Thus, it is possible that the lower concentrations are due to sample degradation, although this explanation is thought to be unlikely as CBG has been shown to be stable for at least six freeze-thaw cycles when stored at -25°C for 2 years (Barlow et al. 2010). A more likely explanation and one that cannot be excluded is that this difference simply represents the variable performance of different batches of the CBG assay.

It is also possible that this represents a genuine difference, arising as an adaptive response to lower serum cortisol concentrations, in an attempt to preserve bioavailable cortisol in the face of diminishing synthesis. Tan *et al* reported significantly lower CBG concentrations in patients with severe liver disease who had evidence of adrenal insufficiency compared to those who did not (Tan et al. 2010), but this has not been reported elsewhere and other studies have shown no difference in CBG concentration between patients with adrenal insufficiency and those with

normal adrenal function (Coolens et al. 1987; Barlow et al. 2010). Furthermore, if this were a genuine effect it would be expected that CBG concentrations would be lowest in the five patients with confirmed hypoadrenalism (patients 1, 2, 10, 12 and 23); but this was not the case.

The positive bias in the correlation between SFC and FCI and salivary cortisol in patients with suspected hypoadrenalism relative to the relationship in healthy volunteers also points towards artefactually low CBG concentrations in this group, as there is no clear explanation why this should be the case.

5.5.2 Calculated free cortisol

Free cortisol calculated from the Coolens' equation (SFC) and free cortisol index (FCI) behaved similarly in healthy volunteers, with no particular advantage of one over the other, except that FCI is easier to calculate and thus less subject to error. Using the lower limit of the reference range as a cut-off for identifying adrenal insufficiency, SFC incorrectly diagnosed disease in 2 (Immulite) to 6 (Roche) healthy volunteers, while FCI diagnosed disease in 4 (Centaur, Immulite, Beckman) to 6 (Roche) subjects. This was similar to serum total cortisol which would have diagnosed adrenal insufficiency in 2 (Abbott) to 5 (Roche) healthy volunteers, but contrasted with salivary cortisol which diagnosed disease in only 1 healthy volunteer.

A further disadvantage of calculated free cortisol over salivary cortisol was its assay-dependence, meaning that no single "reference range" is applicable. It also requires measurement of CBG, the limitations of which have already been discussed. Thus, of the three, calculated free cortisol is the least robust. Nevertheless, both SFC and FCI exhibited better specificity in patients with suspected hypoadrenalism than either serum total cortisol or salivary cortisol, although this came at the expense of sensitivity, which, given the life-threatening potential of undiagnosed adrenal insufficiency, is unacceptable.

Overall agreement between post-Synacthen calculated free cortisol and serum total or salivary cortisol in patients with possible hypoadrenalism was also a cause for concern. Nine patients (30%) who passed the Synacthen test with SFC and FCI failed with serum cortisol and 7 patients (24%) failed with salivary cortisol. This suggests that despite the good correlation between free and total serum cortisol and between free cortisol in serum and saliva, agreement is poor and the relationship is more complex than simple correlation plots show.

A similar suggestion was made by Mallat in response to a study by Estrada-Y-Martin *et al* which reported good correlation between serum free and salivary cortisol and recommended using the latter as an alternative to serum free cortisol in patients with septic shock (Estrada and Orlander 2011; Mallat 2012). Mallat made the point that the strength of the relationship between two variables does not in itself make the two interchangeable. This is further highlighted in this study by the finding that mean SFC and FCI concentrations in patients with suspected hypoadrenalism were significantly higher than in healthy volunteers, while the opposite was true of serum total cortisol and salivary cortisol. Thus, calculated free cortisol is not a simple surrogate for either serum total or salivary free cortisol, and interpretation of the differences between them is limited because of the possible inaccuracies in CBG measurement.

Mean post-Synacthen SFC concentrations in healthy volunteers ranged between 22.0 nmol/L (female, Abbott) and 46.9 nmol/L (male, Roche) in this study, with an equivalent mean FCI range of 8.5 nmol/mg to 13.4 nmol/mg, depending on assay and gender. These results were lower than other studies suggest, with LeRoux *et al* reporting a mean post-Synacthen FCI of 15.2 nmol/mg in healthy volunteers, using total cortisol measured by a Roche RIA and Barlow *et al*, defining a mean post-Synacthen SFC of 69 in a group of patients with lung disease, using the original Centaur assay (le Roux et al. 2002; Barlow et al. 2010). Other ranges of calculated and measured serum free cortisol have been reported (Bonte et al. 1999; Pretorius

et al. 2011), and calculated free cortisol has been shown at times to be both higher (Barlow et al. 2010) and lower (Tan et al. 2010) than measured free cortisol, making it impossible to extract a single, meaningful reference range from the existing literature. A particular strength of this study, however, is the number of healthy volunteers included and the wide range of assays studied.

5.5.3 Limitations

Given the differences between the results of this study and others and the different behaviour exhibited by calculated free cortisol compared to serum total or salivary free cortisol in the "hypoadrenal" patients, it would have been useful to have included measured serum free cortisol in addition to calculated SFC and FCI. This would have eliminated the uncertainty surrounding CBG measurement and would have shown conclusively whether serum free cortisol differs between healthy volunteers and this group of patients.

The delay in analysing CBG in the patient group also introduced unnecessary questions about the reliability of the results, although the delay itself is less likely to be the real problem, rather the use of different batches of the CBG assay. This could have been easily overcome by including independent quality control material with each batch of samples as an internal standard to confirm comparable assay performance.

Another weakness of this study is the mixed nature of the patient group and its small size. Including patients with established adrenal disease is unhelpful in establishing cut-offs to differentiate between normal adrenal function and the early stages of a failing gland, but is essential because it demonstrates the disparate concentrations between this group and healthy volunteers and provides reassurance that severe disease will not be missed, irrespective of diagnostic cut-off. The other groups of patients included – those with non-specific symptoms but no clear risk factors for

hypoadrenalism, those with risk factors but no clear symptoms and those with a combination of both – are precisely the ones who need to be studied, but their numbers were not large enough to allow any definite conclusions to be made about their cortisol response to Synacthen stimulation. Nevertheless, the lower serum total cortisol and salivary free cortisol in these patients hint at a possible role for cortisol in their symptoms.

Furthermore, the differences between serum total cortisol, salivary cortisol and possibly calculated free cortisol in this group suggest all three need to be studied to better understand early changes in the HPA axis and to explore a possible role for minor changes in bioavailable cortisol in patients with the non-specific symptoms of adrenal insufficiency but without evidence of actual disease.

5.5.4 Further work

A useful extension to this study would be to compare the serum total cortisol, salivary free cortisol and measured and calculated serum free cortisol concentrations and responses to Synacthen between cohorts of patients as outlined above and a small group of healthy volunteers. However, although this may help further our understanding of cortisol and the HPA axis, serum free cortisol is not a viable alternative to serum total cortisol or salivary cortisol at present and cannot replace either until better CBG and free cortisol assays are available.

CHAPTER 6

THE EFFECT OF ALTERED PROTEIN STATES ON THE CORTISOL RESPONSE TO SYNACTHEN

6.1 Introduction

Serum cortisol is predominantly protein-bound and consequently measurement of total cortisol is affected by changes in serum protein concentration. Patients with the nephrotic syndrome and liver cirrhosis have low serum albumin and CBG concentrations and show a lower serum total cortisol response to ACTH stimulation than do healthy volunteers (Klose et al. 2007; Fede et al. 2014). Acutely unwell patients also have low serum protein, but interpretation of serum total cortisol measured by immunoassay is complicated by increased concentrations of endogenous steroids, high prevalence of heterophilic antibodies (Briegel et al. 2009) and accentuation of the differences between immunoassays (Cohen et al. 2006), particularly following Synacthen stimulation.

This latter group of patients was therefore excluded from this study to avoid confounding its primary aim of exploring the effect of altered protein concentrations on cortisol measurement. Nevertheless, the results presented here are likely to prove useful in future studies attempting to further the understanding of cortisol measurement in these patients. The inclusion of serum cortisol measured by GC-MS and of salivary cortisol measurement in this study was designed to identify whether either of these could provide an alternative to immunoassay cortisol for evaluating the HPA axis in patients with low serum protein concentrations. However, it is also likely to provide a useful baseline for those considering the value of these measures in assessing the HPA axis in acutely unwell patients.

In contrast to patients with low serum protein, those with high concentrations, as seen in states of oestrogen excess e.g. pregnancy and combined oral contraceptive use have higher mean post-ACTH cortisol concentrations than the general population (Meulenberg et al. 1987; Jung et al. 2011). Despite this overall increase, immunoassay cortisol in this group shows an unexpected negative bias relative to

cortisol measured by LC-MS/MS (Jung et al. 2011). This difference can be eliminated by heat-treatment and is thought to be due to reduced cortisol detection by immunoassay antibodies due to its increased binding to CBG.

6.2 Aims

The work described in this chapter was undertaken:

- To investigate the effects of increased CBG concentration in women taking an oestrogen-containing oral contraceptive pill on serum total cortisol measurement, salivary cortisol measurement and calculated serum free cortisol.
- To examine the relationship between immunoassay and GC-MS cortisol measurement in women taking exogenous oestrogens to identify any interference and determine how this can be overcome.
- To investigate the effects of reduced CBG concentration in patients with the nephrotic syndrome or liver cirrhosis on serum total cortisol measurement, salivary cortisol measurement and calculated serum free cortisol.
- To explore the effects of altered protein states on the serum total cortisol, salivary cortisol and calculated free cortisol responses to Synacthen stimulation.

6.3 Subjects and sample analysis

6.3.1 Healthy volunteers

This group consisted of 60 male and 79 female volunteers, as described in chapter 2. Serum total cortisol was measured by GC-MS and the Centaur, Abbot, Roche, Immulite and Beckman immunoassays; salivary cortisol was measured by LC-MS/MS and CBG was measured by RIA.

6.3.2 Patients with low serum albumin

There were 7 men and 3 women in the low protein group. Mean albumin concentration at the time of the Synacthen test was 30.3 g/L (range 29 – 34). Serum total cortisol was measured by GC-MS and the Abbot immunoassay; salivary cortisol was measured by LC-MS/MS and CBG was measured by RIA; analyses were carried out 4 - 5 years later than those in healthy volunteers.

6.3.3 Female volunteers taking exogenous oestrogen

This group comprised 24 young, premenopausal women, ranging in age between 21 and 40 years. Serum total cortisol was measured by GC-MS and the Centaur, Abbot, Roche, Immulite and Beckman immunoassays; salivary cortisol was measured by LC-MS/MS and CBG was measured by RIA.

6.3.4 Subjects for interference studies

The laboratory information management system (LIMS) was used to identify 9 women in the third trimester of pregnancy (≥ 28 weeks) who attended a hospital antenatal clinic or ward and 9 male (M), 9 female (F) and 9 young female (YF) patients whose general practitioner had requested serum cortisol measurement as part of their routine care. Female patients were defined as women over the age of 45 who were considered unlikely to be taking an oestrogen containing oral contraceptive pill, whilst the young females were under the age of 35 and hence considered a better match for the pregnant women. As no clinical history was available, it was impossible to exclude those taking exogenous steroids or oestrogens.

Samples were retrieved from storage 5 days after the initial request (i.e. just prior to being discarded) and anonymised. The effects of heat and Danazol treatment on serum total cortisol measurement using the Abbott Architect assay were then studied.

Once the initial investigations had been completed, a random selection of 20 samples each from the male and female volunteers and 19 samples from the OCP-female volunteers were chosen and a further 20 samples from women in the third trimester of pregnancy were retrieved as described above.

6.4 Interference studies

6.4.1 Initial investigations

6.4.1.1 Heat treatment

Serum from 3 each of the pregnant, male, female and young female subjects was split into four 300 μL aliquots. Baseline serum total cortisol was measured in 1 aliquot and the three remaining aliquots were incubated in a 60°C water bath for 30, 60 and 120 minutes respectively. Samples were then left to cool for 30 minutes at room temperature before being analysed.

6.4.1.2 Danazol treatment

Baseline cortisol was measured in serum from 3 each of the pregnant, male, female and young female subjects. Separate aliquots of serum were created from each sample and a 200 mg/L stock solution of Danazol dissolved in 100% Methanol was used to prepare samples containing concentrations of 0.2 mg/L, 0.5 mg/L, 1.0 mg/L and 10 mg/L of Danazol, corresponding to 592.6 nmol/L, 1481.5 nmol/L, 2963 nmol/L and 29,630 nmol/L respectively. A blank sample was prepared by spiking a further aliquot with an equivalent volume of Methanol. Total cortisol was re-measured after incubation at room temperature for 1 hour and again after 24 hours.

The experiment was repeated in the remaining 12 samples but with Danazol concentrations of 10 mg/L, 20 mg/L and 40 mg/L, corresponding to 29,630 nmol/L, 59,259 nmol/L and 118,519 nmol/L, and an incubation period of 1 hour only.

6.4.2 Study samples

Samples were run in batches of 20 (5 each of male, female, OCP female and pregnant subjects) over 4 separate days. Baseline cortisol was measured in each sample, following which one aliquot was incubated at 60°C for 30 minutes and allowed to cool to room temperature for 30 minutes before cortisol was re-measured. Two further aliquots were created from each sample – one was spiked with Danazol to give a final concentration of 10 mg/L and the other with an equivalent volume of 100% Methanol. These aliquots were incubated at room temperature for 1 hour before analysis.

6.5 Statistics

6.5.1 Main study subjects

Serum, salivary and calculated free cortisol results were log-transformed for calculation of mean concentrations and reference ranges in each group at each time point, although the OCP-female and low protein groups were too small for valid reference ranges to be derived. Differences between cortisol concentrations in the three groups were compared using the unpaired t-test, or the Mann-Whitney U test for non-parametric data. In all cases, a p-value <0.05 was considered significant.

Differences in the relationship between immunoassay and GC-MS serum cortisol, salivary cortisol and GC-MS serum cortisol and calculated free cortisol and GC-MS serum cortisol between the 3 groups were studied using scatter plots, and, in the case of serum cortisol, Bland-Altman plots.

6.5.2 Interference studies

The percentage increase in serum cortisol from baseline was calculated for heat treated samples by dividing the concentration post-treatment by the pre-treatment result. For the Danazol-treated samples the percentage increase was calculated by dividing the Danazol-spiked concentration by the result of the Methanol-spiked blank.

The unpaired t-test was used to compare the increases in cortisol concentration following heat or Danazol treatment between the different patient groups with a p-value of < 0.05 indicating a significant difference.

6.6 Results

6.6.1 Serum total cortisol

6.6.1.1 Mean concentration and lower reference limits

Serum cortisol measured by GC-MS was normally distributed at baseline and following Synacthen stimulation in women taking exogenous oestrogen (OCP-females) and in patients with low serum protein. Mean baseline concentrations were 537 nmol/L in OCP-females and 305 nmol/L in low protein patients, with mean post-Synacthen concentrations of 869 nmol/L and 552 nmol/L, respectively (table 6.1).

For women taking an OCP, cortisol was measured by all 5 of the immunoassays studied, with mean concentrations ranging between 429 nmol/L and 646 nmol/L at baseline and 747 nmol/L and 1026 nmol/L post-Synacthen, depending on the assay used. Mean concentrations measured by the Abbott immunoassay in the low protein group were 282 nmol/L at baseline and 514 nmol/L post-Synacthen (table 6.1).

Both groups were too small for a robust lower reference limit for the response to Synacthen to be determined; however, an assay-specific 95^{th} percentile range was calculated for each group (table 6.1) and was used to compare responses in these patient types with those in healthy volunteers, with the aim of highlighting any obvious differences.

Table 6.1 Assay-specific serum total cortisol concentrations in OCP females and patients with low serum protein at baseline and post-Synacthen.

	Mean [Cortisol] (2.5th – 97.5th centile) nmol/L			
	OCP-females (n = 24)		**Low-protein patients (n = 10)**	
Assay	**Baseline**	**Post-Synacthen**	**Baseline**	**Post-Synacthen**
GC-MS	537 (315 - 914)	869 (649 - 1162)	305 (173 – 537)	552 (393 – 776)
Centaur	488 (323 - 738)	763 (619 – 940)	-	-
Abbott	465 (301 - 718)	747 (577 - 967)	282 (167 – 476)	514 (384 – 688)
Roche	646 (383 - 1090)	1026 (791 - 1330)	-	-
Immulite	510 (330 - 788)	850 (688 - 1051)	-	-
Beckman	429 (286 - 643)	757 (604 - 948)	-	-

6.6.1.2 Comparison with healthy volunteers

Mean cortisol concentration in women taking exogenous oestrogen was significantly higher at baseline (table 6.2, figure 6.1) and post-Synacthen (table 6.3, figure 6.1) than in healthy female volunteers for all assays studied. The Beckman assay showed the smallest differences – 177 nmol/L at baseline and 163 nmol/L post-Synacthen, whilst the equivalent differences for the Roche assay were 354 nmol/L and 314 nmol/L, respectively. This translated into a post-Synacthen 2.5th centile concentration up to 267 nmol/L higher (Roche assay) in women taking exogenous oestrogen than in those who were not.

Table 6.2 Assay-specific serum total cortisol concentrations in female volunteers and OCP-females at baseline.

	Baseline Mean [Cortisol] (2.5th – 97.5th centile) nmol/L		
Assay	**Non –OCP females n = 79**	**OCP females n = 24**	**P-value***
GC-MS	248 (134 - 459)	537 (315 - 914)	P < 0.001
Centaur	257 (138 - 477)	488 (323 - 738)	P < 0.001
Abbott	247 (134 - 455)	465 (301 - 718)	P < 0.001
Roche	292 (147 - 581)	646 (383 - 1090)	P < 0.001
Immulite	267 (144 - 495)	510 (330 - 788)	P < 0.001
Beckman	252 (143 - 444)	429 (286 - 643)	P < 0.001

*P-value <0.05 indicates a significant difference between means.

Table 6.3 Assay-specific serum total cortisol concentrations in female volunteers and OCP-females post-Synacthen.

	Post-Synacthen Mean [Cortisol] (2.5th – 97.5th centile) nmol/L		
Assay	**Non-OCP females n = 79**	**OCP females n = 24**	**P-value***
GC-MS	558 (422 - 737)	869 (649 - 1162)	P < 0.001
Centaur	578 (446 - 750)	763 (619 – 940)	P < 0.001
Abbott	542 (416 - 707)	747 (577 - 967)	P < 0.001
Roche	712 (524 - 967)	1026 (791 - 1330)	P < 0.001
Immulite	628 (478 - 826)	850 (688 - 1051)	P < 0.001
Beckman	594 (455 - 777)	757 (604 - 948)	P < 0.001

*P-value <0.05 indicates a significant difference between means.

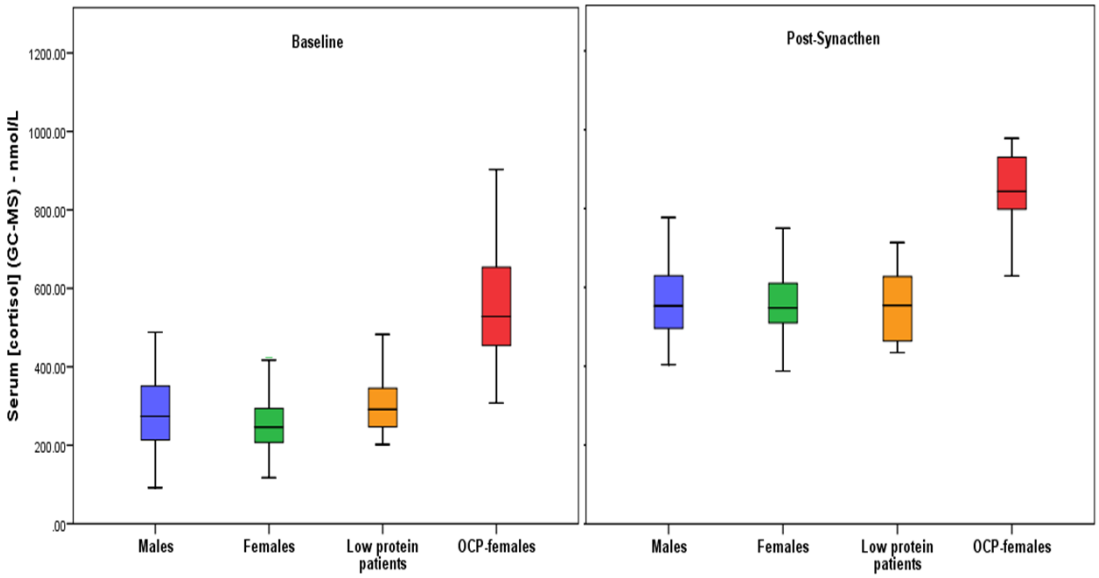


Figure 6.1 Mean serum cortisol in healthy volunteers, patients with low serum protein concentration and women taking an OCP at baseline and post-Synacthen.

Mean baseline cortisol concentration measured by GC-MS was not significantly different in low protein patients than in healthy male or female volunteers (figure 6.1, table 6.4). This remained the case following Synacthen stimulation. When serum cortisol was measured using the Abbott assay, however, post-Synacthen cortisol concentration in low protein patients was significantly lower than in healthy male volunteers (table 6.4).

Table 6.4 Serum total cortisol concentration in patients with low serum protein and healthy volunteers by GC-MS and the Abbott immunoassay.

	Mean serum [cortisol] (2.5th – 97.5th percentile) nmol/L				
	Low protein patients	**Males**	**P-value**	**Females**	**P-value**
Baseline					
GC-MS	305 (173 – 537)	274 (131 - 575)	0.363	254 (139 - 463)	0.052
Abbott	282 (167 – 476)	289 (151 - 556)	0.967	247 (134 - 455)	0.201
Post-Synacthen					
GC-MS	552 (393 – 776)	563 (418 - 757)	0.804	555 (421 - 731)	0.838
Abbott	514 (384 – 688)	577 (430 - 773)	0.027*	542 (416 - 707)	0.251

P-value <0.05 indicates a significant difference between means.

6.6.1.3 Correlation between immunoassay and GC-MS

Cortisol measured by immunoassay was plotted against GC-MS cortisol for healthy volunteers and OCP-females and showed good correlation, but with some evidence of negative bias at higher concentrations for all assays apart from the Roche (figure 6.2). Further examination of the plots suggested that this negative bias arose predominantly in samples from women taking exogenous oestrogens, which appeared to behave differently to those from healthy volunteers.

This was confirmed on the corresponding Bland-Altman plots for each immunoassay (figure 6.3). Immunoassay cortisol from OCP-females showed negative bias relative to GC-MS except with the Roche assay where bias remained positive, albeit to a lesser extent than in non-OCP subjects. Given the lack of specificity of cortisol immunoassays, particularly in comparison to GC-MS, this was an unexpected finding which suggested under-recovery of cortisol in this group.

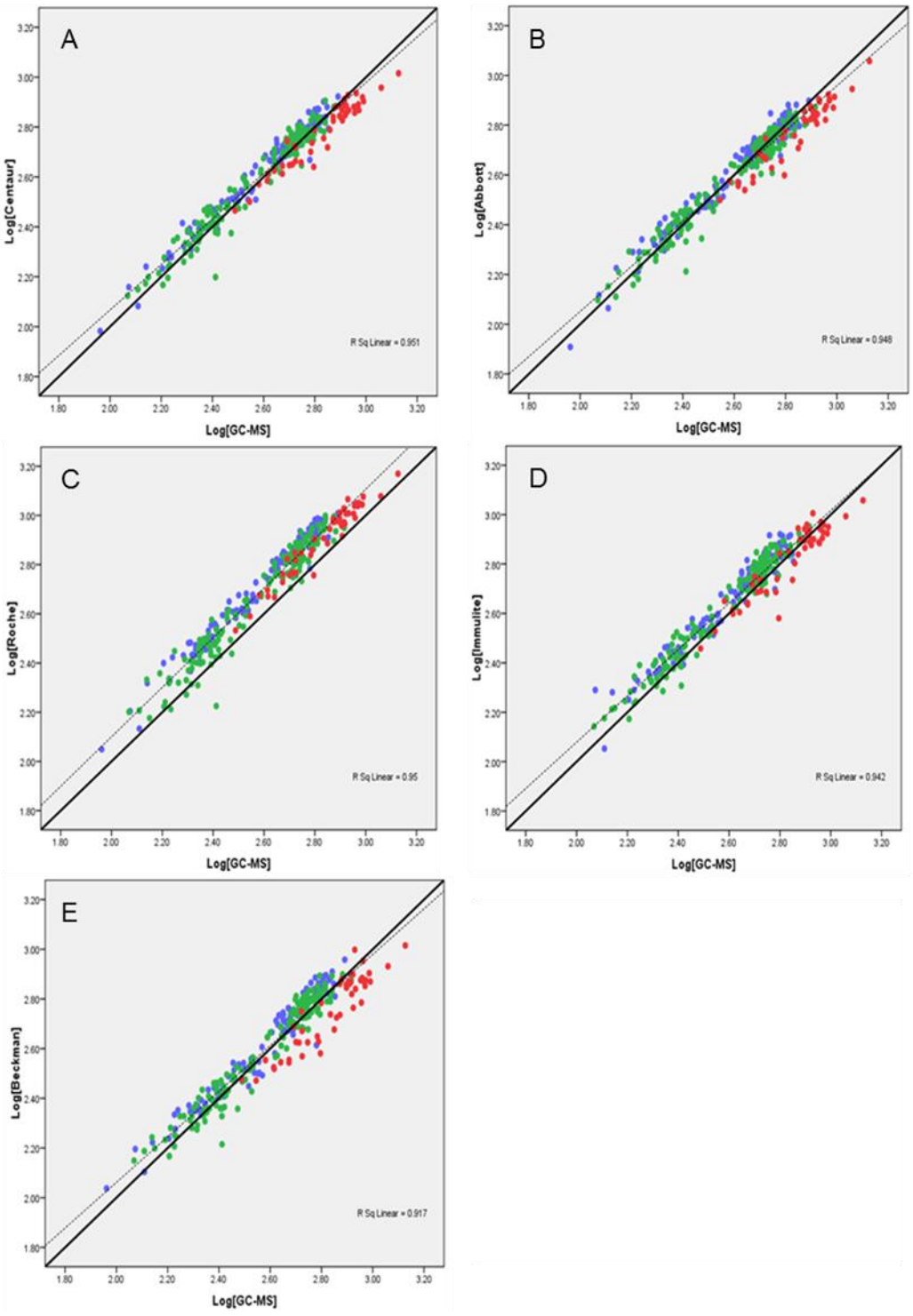


Figure 6.2 Correlation between log-transformed immunoassay and GC-MS cortisol in healthy volunteers and OCP-females. A) Centaur, B) Abbott, C) Roche, D) Immulite and (E) Beckman assays. The solid line represents equivalence between methods; dotted line equals line of best fit.
● = males; ● = non-OCP females; ● = OCP-females.

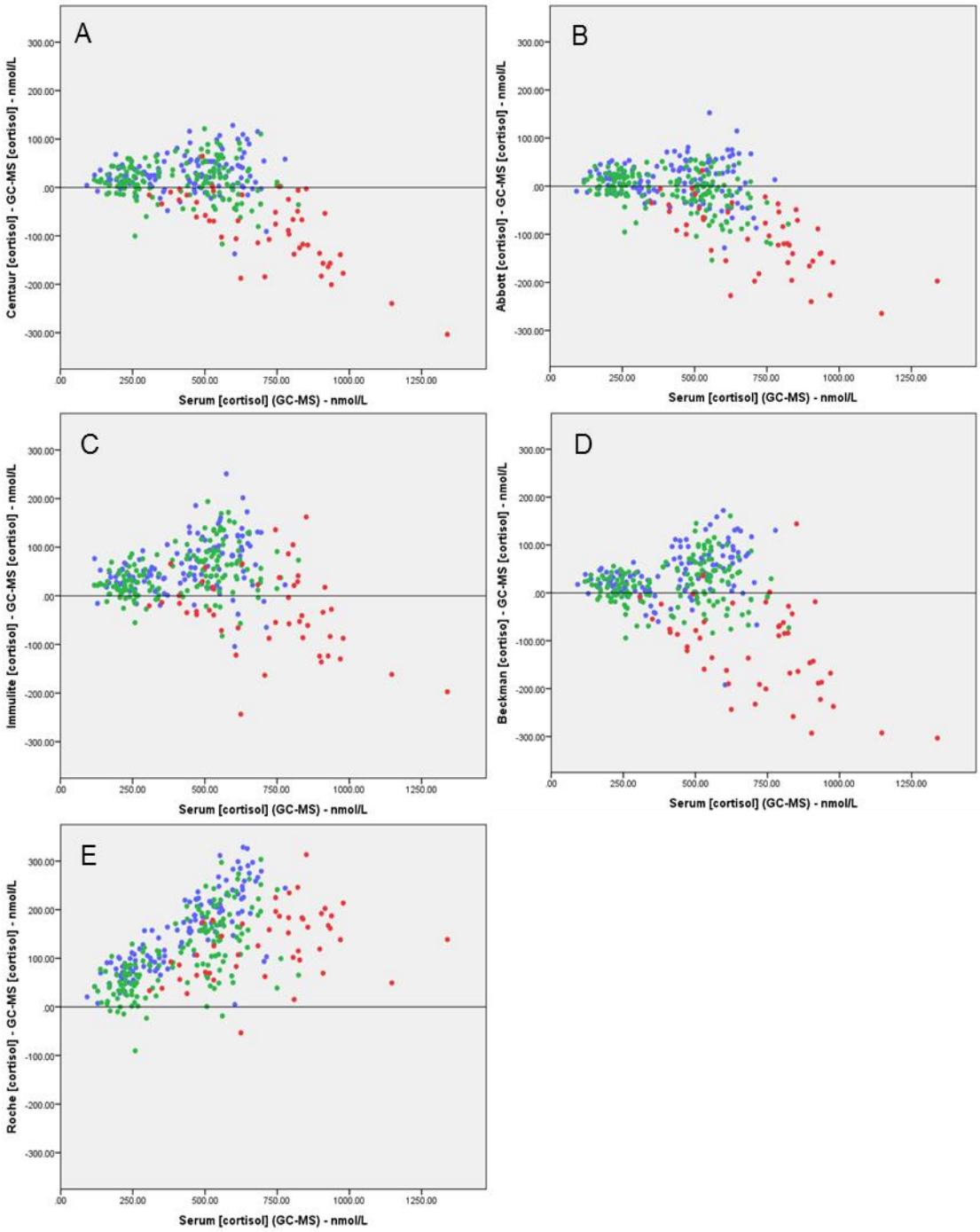


Figure 6.3 Bland-Altman plots comparing immunoassay and GC-MS cortisol in healthy volunteers and OCP-females. A) Centaur, B) Abbott, C) Immulite, D) Beckman and E) Roche. x-axis = difference between immunoassay and GC-MS cortisol concentrations, y-axis = GC-MS cortisol concentration

The solid line represents equivalence between methods. Scatter of results above and below the line represents positive and negative bias of the immunoassay respectively. ● = males; ● = non-OCP females; ● = OCP-females.

To further evaluate the observed bias, mean bias ratios were calculated for the OCP female group for each immunoassay relative to GC-MS, as described in section 3.4. In line with the findings from the correlation and Bland-Altman plots, the mean bias ratio for this group was negative for all immunoassays apart from the Roche, ranging from 0.96 (a negative bias of 4%) for the Immulite assay to 0.83 for the Beckman. This contrasted with the bias in non-OCP subjects which was positive for all assays (table 6.5).

The calculated mean bias ratio was also helpful in demonstrating that the Roche assay was not excluded from the negative bias affecting samples from OCP females in the other assays, although the effect was diminished. Thus, although remaining positive, the bias exhibited by the Roche assay was almost halved; falling from 30% in non-OCP subjects to 18% in OCP females (table 6.5).

Table 6.5 Mean assay-specific bias ratios for healthy volunteers and OCP females.

	Overall Mean Bias Ratio		
Assay	**Healthy volunteers**	**OCP Females**	**P-value***
Centaur	1.06	0.89	P<0.001
Abbott	1.02	0.86	P<0.001
Roche	1.30	1.18	P<0.001
Immulite	1.13	0.96	P<0.001
Beckman	1.07	0.83	P<0.001

Overall mean bias ratio was calculated using combined baseline and post-Synacthen cortisol concentrations. * P-value <0.05 indicates a significant difference.

Cortisol in patients with low serum protein measured by the Abbott immunoassay also correlated well with GC-MS cortisol, albeit with a slight negative bias across the range of concentrations studied (figure 6.4A). On initial inspection, there was no suggestion that the relationship was any different to that in healthy volunteers (figure 6.4B), and this was largely supported by the Bland-Altman plot (figure 6.5). The mean bias ratio, however, was significantly lower in this group than in male volunteers ($p<0.05$) although not in comparison to female volunteers ($p=0.642$) (figure 6.6). This makes

a genuine difference between immunoassay cortisol measurement in low protein patients and healthy volunteers unlikely.

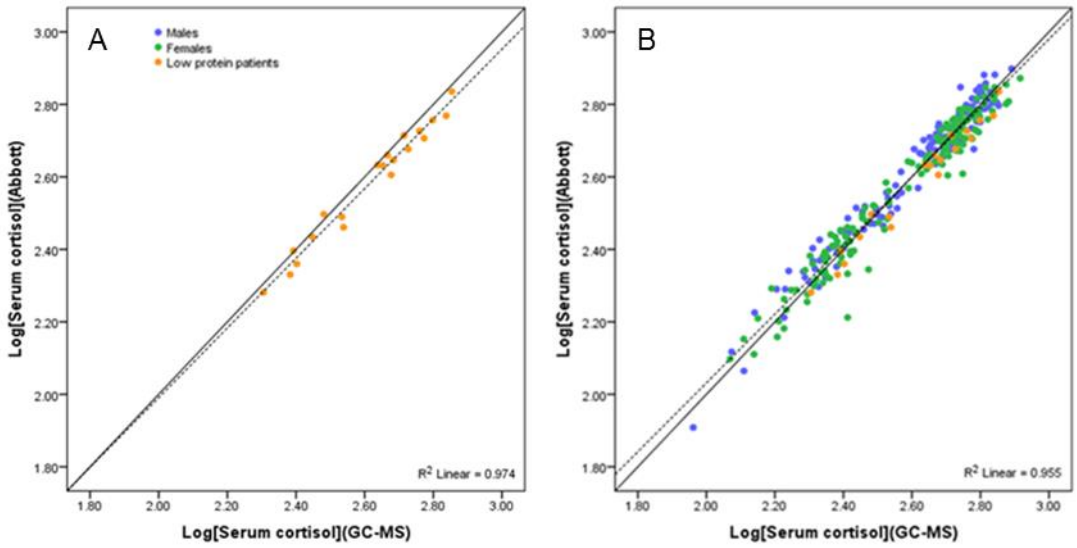


Figure 6.4 Correlation between log-transformed Abbott and GC-MS cortisol in healthy volunteers and low protein patients. Panel A shows the correlation in low protein patients and B shows the two groups combined. The solid line represents equivalence between methods; dotted line equals line of best fit.

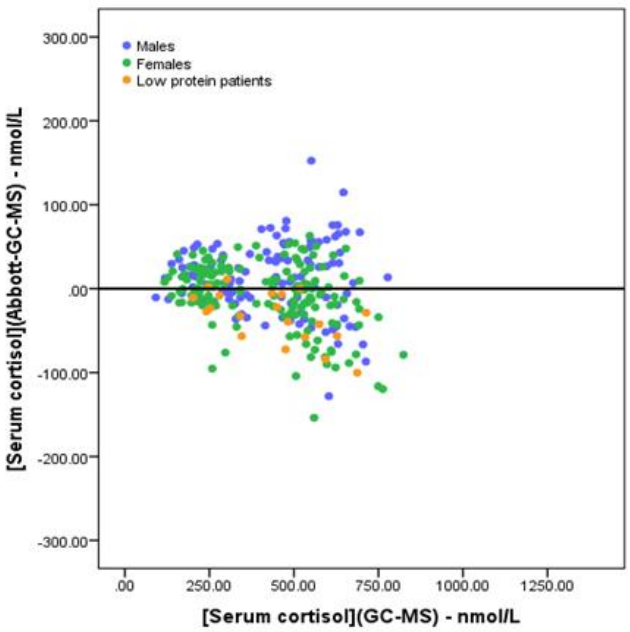


Figure 6.5 Bland-Altman plot comparing Abbott and GC-MS cortisol in healthy volunteers and low protein patients. The solid line represents equivalence between methods.

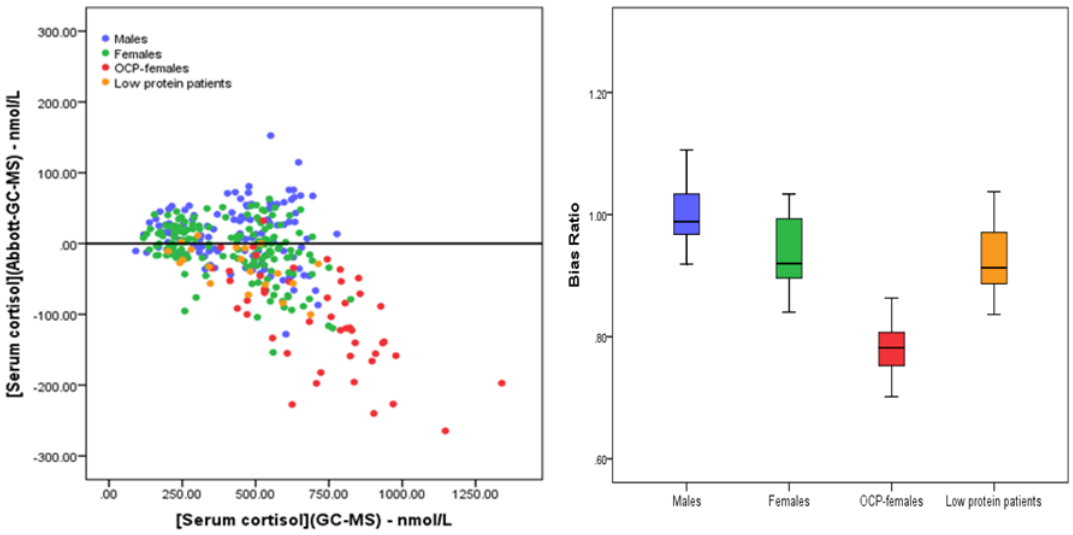


Figure 6.6 Bland-Altman plot comparing Abbott and GC-MS cortisol and mean bias ratios for all patient groups. The solid line represents equivalence between methods.

6.6.1.4 Interference studies

It was assumed that the negative bias observed with samples from OCP-females was due to excess CBG binding to cortisol and blocking the immunoassay antibodies' ability to find their binding sites on the cortisol molecule. Further investigations, aimed at displacing cortisol from CBG were therefore undertaken.

Initial investigations

Studies using heat treatment to denature CBG (and other proteins) and Danazol to competitively displace cortisol from CBG were designed and carried out in serum from pregnant and non-pregnant patients to determine the optimum conditions to use with study samples. Serum cortisol increased by up to 124% in pregnant women, 26% in men, 33% in women and 52% in young women following heat treatment at 60°C for 30, 60 or 120 minutes (table 6.6). This effect was observed at all time points and was not appreciably greater with longer incubation times.

A further 3 samples from each patient group were spiked with Danazol at concentrations ranging from 0.2 mg/L to 10 mg/L, which corresponded to concentrations 1 to 3 times and 55 to 157 times higher than the baseline cortisol concentrations in these samples, respectively. Serum cortisol did not increase significantly from baseline with Danazol concentrations below 10 mg/L, but increases of up to 43%, 24%, and 23% were seen in pregnant subjects, males and females, respectively, at this concentration (table 6.7).

Increasing the incubation time from 1 to 24 hours resulted in a further increase in cortisol concentrations from baseline in pregnant samples but not in the other groups (table 6.8). Similarly, increasing the concentration of Danazol to 20 mg/L or 40 mg/L had little effect on the post-treatment cortisol concentration (table 6.9).

Table 6.10 shows a summary of the effects of the different treatments and identifies the two which were selected for use with patient samples – namely heat treatment at 60°C for 30 minutes and spiking with 10 mg/L of Danazol for 1 hour at room temperature.

Table 6.6 Effect of heat treatment on immunoassay cortisol.

	P1	**P2**	**P3**	**M1**	**M2**	**M3**	**F1**	**F2**	**F3**	**YF1**	**YF2**	**YF3**
Baseline [cortisol] nmol/L	334.8	544.3	319.9	359.7	363.8	228.8	212.8	283.3	340.0	329.7	329.5	480.6
[Cortisol] nmol/L @: **30 minutes**	519.4	1125.1	462.5	431.6	453.9	277.8	272.5	359.3	446.3	392.4	500.3	594.6
60 minutes	514.5	1219.7	482.4	433.2	437.1	289.4	260.1	357.8	453.5	400.8	486	612.1
120 minutes	520.5	1151.7	458.8	420.4	438.5	280.9	265.9	354.4	445.3	398.4	490.6	557.1
Mean [cortisol] nmol/L	518.1	1165.5	467.9	428.4	443.2	282.7	266.2	357.2	448.4	397.2	492.3	587.9
Percentage increase @: **30 minutes**	55	107	45	20	25	21	28	27	31	19	52	24
60 minutes	54	124	51	20	20	26	22	26	33	22	47	27
120 minutes	55	112	43	17	21	23	25	25	31	21	49	16
Mean percentage increase	55	114	46	19	22	24	25	26	32	20	49	22

The table shows serum total cortisol concentrations measured before and after samples were incubated at 60°C for 30, 60 and 120 minutes respectively and the corresponding percentage increase in concentration from baseline. The increase in sample YF2 was more in keeping with that seen in pregnant women, suggesting the patient may have been pregnant or taking an oestrogen-containing OCP, so these results were excluded from final calculations of the overall mean percentage increase. The increase in sample P2 also appeared out of step with the results from the other pregnant samples and was similarly excluded from further analysis. P – pregnant women, M – male, F – female (aged ≥ 40 years) and YF – young female (aged ≤ 35 years).

.

Table 6.7 Effect of Danazol treatment on immunoassay cortisol (1-hour incubation).

	P1	**P2**	**P3**	**M1**	**M2**	**M3**	**F1**	**F2**	**F3**	**YF1**	**YF2**	**YF3**
Spiked baseline [cortisol] nmol/L	439.5	398.5	437.9	296.4	217.4	253.5	379.5	295.7	199.3	171.3	510.8	244.1
[Cortisol] nmol/L with: **0.2 mg/L Danazol**	443.3	396.4	432.4	301.9	229.0	-	377.0	292.1	202.7	171.4	527.4	245.3
0.5 mg/L Danazol	444.9	396.5	444.2	294.2	213.8	259.3	389.9	288.0	196.8	176.5	528.9	245.5
1.0 mg/L Danazol	479.5	414.4	457.5	307.1	230.6	260.3	397.2	307.9	207.6	176.1	526.8	246.7
10 mg/L Danazol	629.9	529.8	574.6	368.3	260.0	287.5	465.2	336.8	230.5	210.0	603.9	293.0
Percentage increase with: **0.2 mg/L Danazol**	1	-1	-1	2	5	-	-1	-1	2	0	3	0
0.5 mg/L Danazol	1	-1	1	-1	-2	2	3	-3	-1	3	4	1
1.0 mg/L Danazol	9	4	4	4	6	3	5	4	4	3	3	1
10 mg/L Danazol	43	33	31	24	20	13	23	14	16	23	18	20

The table shows serum total cortisol concentrations measured before and after samples were incubated with varying concentrations of Danazol (0.2 mg/L, 0.5 mg/L, 1.0 mg/L and 10 mg/L, respectively) for 1 hour at room temperature and the corresponding percentage increase in concentration from baseline. The 12 samples treated with Danazol were not the same ones described in the heat treatment studies.
P – pregnant women, M – male, F – female (aged ≥ 40 years) and YF – young female (aged ≤ 35 years).

Table 6.8 Effect of Danazol treatment on immunoassay cortisol (24-hour incubation).

	P1	**P2**	**P3**	**M1**	**M2**	**M3**	**F1**	**F2**	**F3**	**YF1**	**YF2**	**YF3**
Spiked baseline [cortisol] nmol/L	441.2	402.8	432.3	303.0	227.3	272.7	399.1	302.0	204.6	180.4	538.3	252.1
[Cortisol] nmol/L with: **0.2 mg/L Danazol**	458.0	399.6	449.8	299.5	235.2	-	399.7	307.4	210.5	186.3	548.6	262.7
0.5 mg/L Danazol	456.2	412.4	474.8	310.6	242.8	268.5	417.6	303.1	213.5	192.8	562.4	262.1
1.0 mg/L Danazol	487.0	427.3	475.2	335.8	244.6	287.7	420.9	317.1	215.5	194.9	555.2	273.5
10 mg/L Danazol	656.4	548.2	591.0	366.8	265.0	321.4	474.4	351.6	238.6	211.2	625.7	306.0
Percentage increase with: **0.2 mg/L Danazol**	4	-1	4	-1	3	-	0	2	3	3	2	4
0.5 mg/L Danazol	3	2	10	3	7	-2	5	0	4	7	4	4
1.0 mg/L Danazol	10	6	10	11	8	6	5	5	5	8	3	8
10 mg/L Danazol	49	36	37	21	17	18	19	16	17	17	16	21

The table shows serum total cortisol concentrations measured before and after samples were incubated with varying concentrations of Danazol (0.2 mg/L, 0.5 mg/L, 1.0 mg/L and 10 mg/L, respectively) for 24 hours at room temperature and the corresponding percentage increase in concentration from baseline. These were the same 12 samples from the 1-hour incubation studies. P – pregnant women, M – male, F – female (aged ≥ 40 years) and YF – young female (aged ≤ 35 years).

Table 6.9 Effect of high concentration Danazol treatment on immunoassay cortisol (1-hour incubation).

	P1	**P2**	**P3**	**M1**	**M2**	**M3**	**F1**	**F2**	**F3**	**YF1**	**YF2**	**YF3**
Spiked baseline [cortisol] nmol/L	256	1288	490.8	171.8	177.7	246.7	200.3	273.3	-	111.3	244.6	190.6
[Cortisol] nmol/L with: **20 mg/L Danazol**	347.1	1740.6	624.3	207.7	222.2	298.9	250.7	337.6	213.9	133.2	298.4	221.6
40 mg/L Danazol*	-	-	577.1	-	-	245.6	-	-	185.3	-	-	183.7
Percentage increase with: **20 mg/L Danazol**	36	35	27	21	25	21	25	24	-	20	22	16
40 mg/L Danazol	-	-	31	-	-	16	-	-	23	-	-	13

The table shows serum total cortisol concentrations measured before and after samples were incubated for 1-hour at room temperature with 20 mg/L or 40 mg/L concentrations of Danazol, respectively, and in samples P3, M3 and YF3, with both. These 12 samples were different to those used in the other studies. P – pregnant woman, M – male, F – female (aged ≥ 40 years) and YF – young female (aged ≤ 35 years).

Table 6.10 Mean percentage increase in serum cortisol concentration measured by the Abbott assay following heat and Danazol treatment.

	Pregnant	Non-pregnant
Heat Treatment		
30 minutes	1.50	1.24
60 minutes	1.52	1.25
120 minutes	1.49	1.22
Danazol concentration – 1-hr incubation		
0.2 mg/L	1.00	1.01
0.5 mg/L	1.01	1.01
1.0 mg/L	1.06	1.04
10 mg/L	1.36	1.19
20 mg/L	1.33	1.22
40 mg/L	1.31	1.17
Danazol concentration – 24-hr incubation		
0.2 mg/L	1.02	1.02
0.5 mg/L	1.05	1.04
1.0 mg/L	1.09	1.07
10 mg/L	1.41	1.18

The non-pregnant results were calculated from male, female and young female patients combined. The percentage increase in cortisol concentration following heat treatment was calculated from unspiked baseline samples, while the increase following Danazol treatment was calculated from a baseline that had been spiked with an equivalent volume of 100% methanol. The red boxes highlight the pre-treatment conditions that were used to investigate study subjects.

Study samples:

Baseline cortisol concentrations did not differ significantly between pregnant and OCP-females ($p=0.844$) nor between male and female volunteers ($p=0.430$) (table 6.11). The increase in cortisol concentration from baseline was significantly higher following heat treatment than with Danazol for all patient groups ($p<0.001$) and was higher for both treatments in pregnant patients than in any other group ($p<0.01$). OCP-females also showed significantly greater increases in cortisol following treatment than both male and female volunteers ($p<0.001$). The differences between the effects of heat and Danazol treatment on male and female samples were small but statistically significant (heat treatment, $p=0.045$; Danazol $p=0.048$).

Table 6.11 Effect of heat and Danazol treatment on cortisol measurement by the Abbott assay in pregnant patients and healthy male, female and OCP-female volunteers.

	Mean baseline [cortisol] in nmol/L		Mean [cortisol] in nmol/L following:		Percentage increase following:	
	Unspiked	**Spiked**	**Heat**	**Danazol**	**Heat**	**Danazol**
Pregnant women	434.3	412.6	717.5	560.5	66 (15)	35 (5)
OCP-females	427.7	403.6	656.4	526.0	53 (10)	30 (5)
Females	237.4	220.3	303.1	262.8	27 (8)	19 (5)
Males	260.6	241.2	316.3	279.9	22 (8)	17 (4)

Serum total cortisol was measured in 20 samples each from anonymised pregnant women, male and female healthy volunteers and 19 samples from healthy OCP females before and after heat and Danazol treatment, and mean concentrations were calculated. The percentage increase in cortisol concentration following heat treatment was calculated from the unspiked baseline samples, while the increase following Danazol treatment was calculated from a baseline that had been spiked with an equivalent volume of 100% methanol.

Figure 6.7 shows a series of Bland-Altman plots examining the effect of Danazol and heat treatment on the performance of the Abbott assay relative to GC-MS. Figure 6.7a shows the negative bias of the Abbott assay relative to GC-MS with samples from OCP-females. Following treatment with Danazol, cortisol concentration is significantly higher than at baseline for all sample types ($p<0.001$), although some OCP females continue to show negative bias relative to GC-MS (figure 6.7b). Heat treatment overcomes the negative bias almost completely, revealing a proportional positive bias across all concentrations (figure 6.7c).

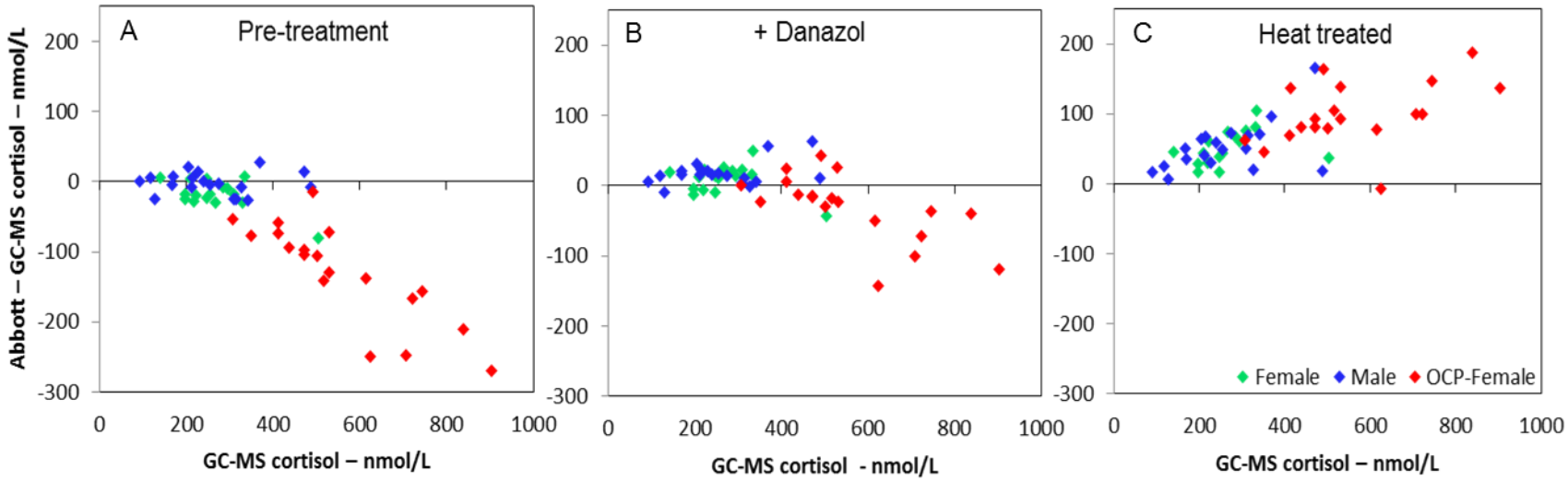


Figure 6.7 Bland-Altman plots comparing baseline cortisol measured by the Abbott immunoassay with GC-MS cortisol in healthy male, female and OCP-female volunteers. Panel (A) shows the relationship between samples before either treatment, (B) shows the effect of incubating the samples with Danazol at a concentration of 10 mg/L for 1 hour and (C) shows the effect of incubating the samples at 60°C for 30 minutes before analysis.

6.6.2. Salivary cortisol

6.6.2.1 Mean concentration and lower reference limits

Salivary cortisol was not normally distributed in women taking exogenous oestrogen at baseline or post-Synacthen but was normally distributed at both time points in patients with low serum protein. Mean baseline concentration was 5.1 nmol/L in OCP-females and 5.3 nmol/L in low protein patients, rising to 19.7 nmol/L and 19.0 nmol/L, respectively, following Synacthen stimulation (table 6.12).

Table 6.12 Salivary cortisol concentration in all patient groups at baseline and post-Synacthen.

	Mean salivary [cortisol] (2.5^{th} – 97.5^{th} percentile) – nmol/L		
	OCP-Females	**Healthy volunteers**	**Low protein patients**
0 Minute	5.1 (1.9 – 14.0)	2.9 (0.9 – 9.2)	5.3 (1.1 – 26.2)
30 Minute	19.7 (9.5 – 41.2)	19.3 (10.3 – 36.2)	19.0 (7.7 – 46.9)

6.6.2.2 Comparison with healthy volunteers

Baseline salivary cortisol was significantly higher in women taking exogenous oestrogen and in patients with low serum protein than in male or female volunteers ($p<0.001$) but there was no significant difference between the two altered protein groups ($p=0.897$) (figure 6.8). There was also no significant difference between any of the groups (males, females, OCP-females or low protein patients) following Synacthen stimulation (p-values 0.591 to 0.971) (figure 6.8).

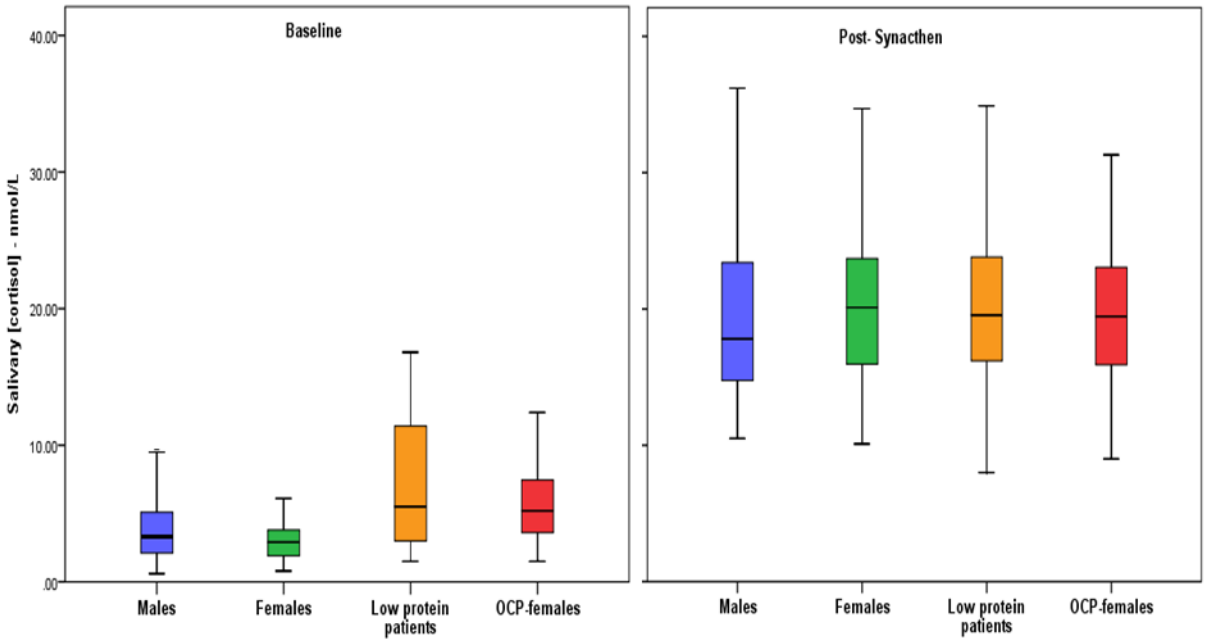


Figure 6.8 Mean salivary cortisol concentration in healthy volunteers, patients with low serum protein concentration and women taking an OCP at baseline and post-Synacthen.

6.6.2.3 Correlation with serum cortisol

Salivary cortisol in women taking exogenous oestrogen and in those with low serum protein concentration correlated well with serum cortisol measured by GC-MS and the Abbott immunoassay, although the correlation was weaker than that in healthy volunteers (Figure 6.9 A & B). Salivary cortisol was negatively biased relative to serum cortisol in all groups at lower concentrations but this was most marked in the exogenous oestrogen group and least in patients with low serum protein. Positive bias was observed at higher concentrations, particularly in healthy volunteers and low protein patients, less in women taking exogenous oestrogens (Figure 6.9).

Comparing the relationship between salivary and serum cortisol in women taking exogenous oestrogens with that in patients with low serum protein showed overall negative bias in the oestrogen group compared to overall positive bias in the low-protein group (figure 6.10), despite a similar range of salivary cortisol concentrations.

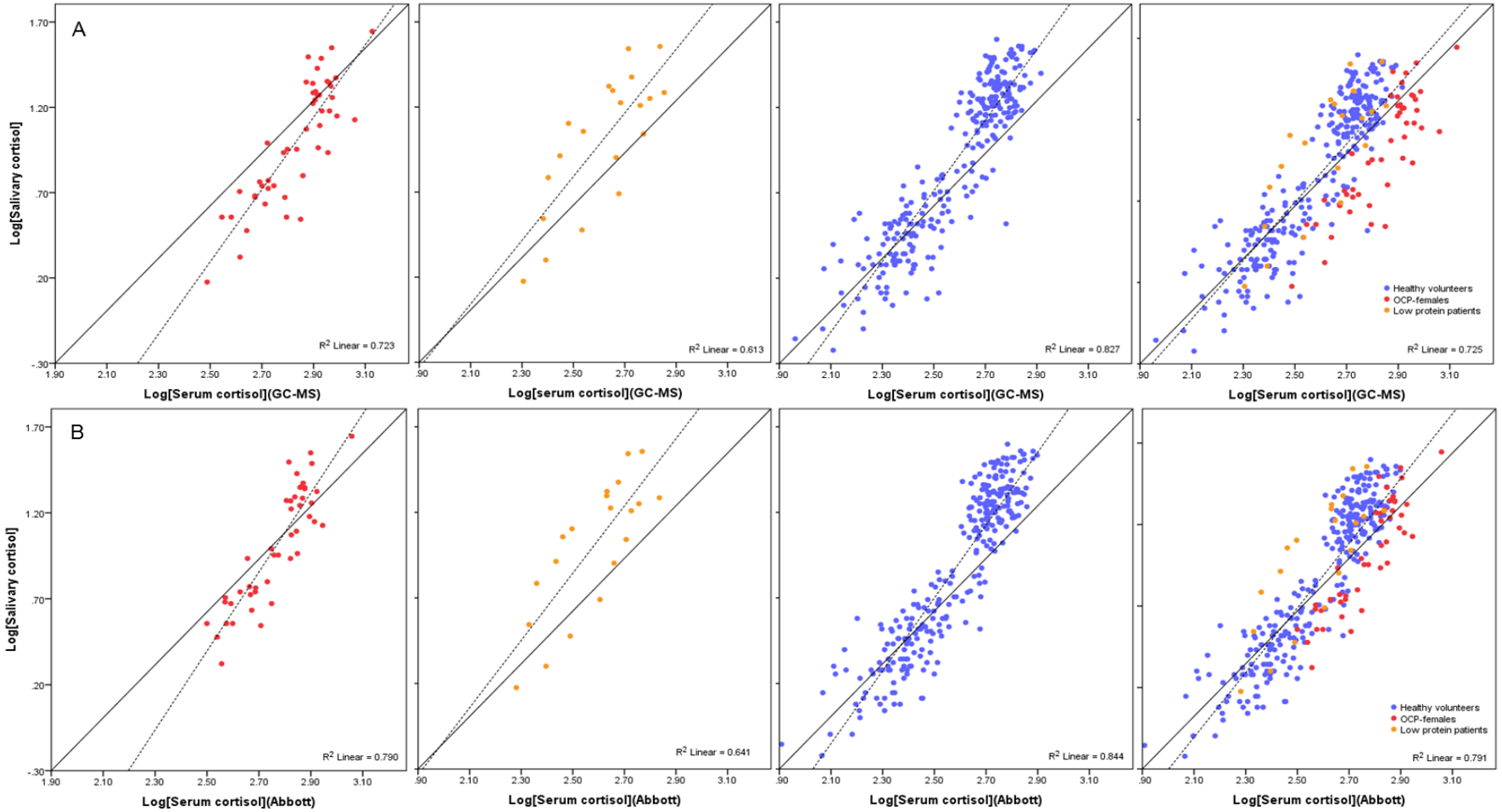


Figure 6.9 Correlation between log-transformed salivary cortisol and serum cortisol measured by GCMS and immunoassay (Abbott) cortisol in all groups studied. Panel (A) shows the relationship between salivary cortisol and GC-MS cortisol in OCP-females, low-protein patients, healthy volunteers and the three combined and (B) shows the respective relationships between salivary cortisol and Abbott cortisol. The solid line represents equivalence between methods; dotted line equals line of best fit.

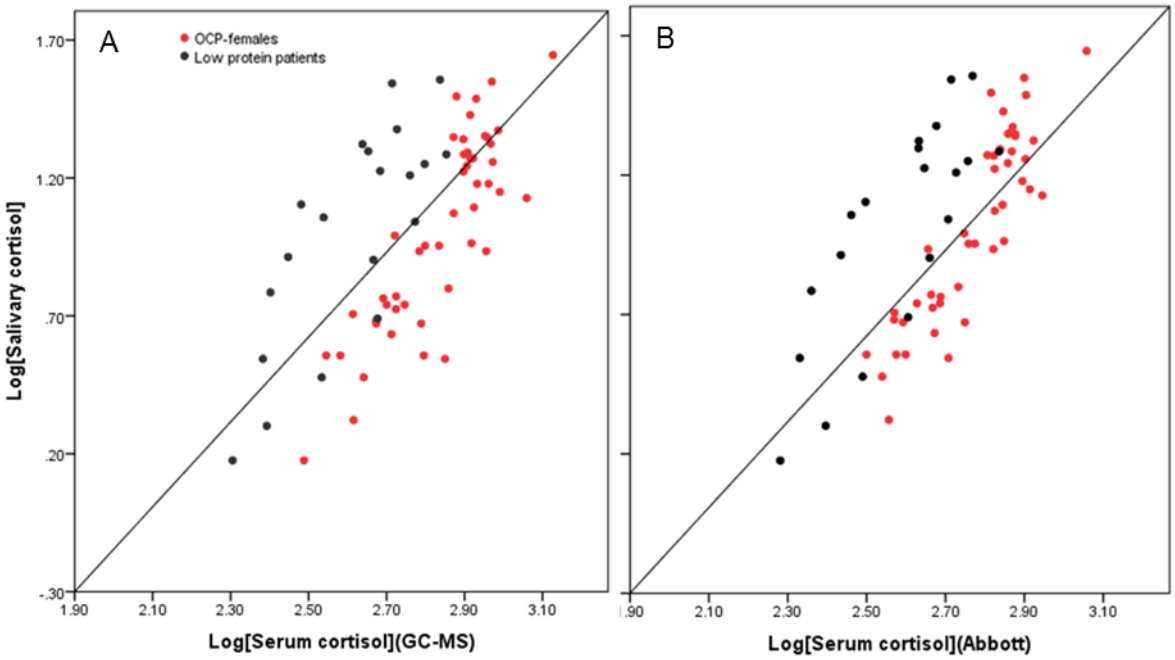


Figure 6.10 Correlation between log-transformed salivary cortisol and serum total cortisol in OCP-females and low-protein patients. Panel (A) shows the relationship with GC-MS cortisol and (B) the relationship with Abbott cortisol.

6.6.3 Calculated free cortisol

6.6.3.1 Mean concentration and lower reference limits

Serum free cortisol (SFC) and free cortisol index (FCI) were calculated from serum total cortisol measured by GC-MS (GC-MS SFC; GC-MS FCI) and the five immunoassays studied in women taking exogenous oestrogen and from GC-MS and Abbott cortisol in patients with low serum protein. Both GC-MS parameters were normally distributed at baseline but not post-Synacthen stimulation in OCP-females and were normally distributed at both time points in patients with low serum protein.

In both patient groups SFC was significantly higher than FCI at each time-point ($p<0.001$) (tables 6.13 to 6.16) and both parameters were assay-dependent, although in women taking exogenous oestrogen there was no significant difference between either SFC or FCI with the GC-MS and Centaur assays at baseline or between GC-

MS, Centaur and Immulite assays or Abbott and Beckman assays post-Synacthen stimulation (tables 6.13 to 6.16).

Table 6.13 Baseline SFC in male and female healthy volunteers, low-protein patients and OCP-females.

	Mean Baseline Serum Free Cortisol (2.5th – 97.5th centile)			
Assay	**Low protein patients n = 10**	**Males n = 60**	**Females n = 79**	**OCP females n = 24**
GC-MS	18.7 (9.2 – 38.0)	9.5 (4.0 – 22.6)	7.7 (3.5 – 16.7)	10.1 (6.4 – 16.1)
Centaur	-	10.4 (4.5 – 23.9)	8.1 (3.6 – 18.4)	9.6 (6.3 – 14.7)
Abbott	16.7(8.4 – 33.1)	10.0 (4.2 – 24.0)	7.7 (3.5 – 17.0)	8.2 (5.0 – 13.3)
Roche	-	13.9 (5.1 – 37.5)	9.6 (4.0 – 23.1)	12.5 (7.4 – 21.4)
Immulite	-	11.8 (4.6 – 30.1)	8.5 (3.8 – 19.0)	9.2 (5.6 – 15.0)
Beckman	-	10.2 (4.6 – 22.5)	7.9 (3.5 – 17.5)	7.4 (4.7 – 11.7)

Table 6.14 Baseline FCI in male and female healthy volunteers, low-protein patients and OCP-females**.**

	Mean Baseline Free Cortisol Index (2.5th – 97.5th centile)			
Assay	**Low protein patients n = 10**	**Males n = 60**	**Females n = 79**	**OCP females n = 24**
GC-MS	7.3 (4.3 – 12.4)	4.6 (2.3 – 9.3)	3.9 (2.0 – 7.4)	4.7 (3.3 – 6.7)
Centaur	-	5.0 (2.6 – 9.7)	4.0 (2.0 – 8.0)	4.5 (3.2 – 6.3)
Abbott	6.8 (4.0 – 11.5)	4.8 (2.4 – 9.8)	3.9 (2.0 – 7.6)	4.0 (2.7 – 5.9)
Roche	-	6.2 (2.9 – 13.0)	4.6 (2.3 – 9.4)	5.5 (3.7 – 8.2)
Immulite	-	5.5 (2.7 – 11.0)	4.2 (2.2 – 8.2)	4.3 (2.9 – 6.4)
Beckman	-	4.9 (2.6 – 9.2)	4.0 (2.0 – 7.7)	3.7 (2.5 – 5.3)

Table 6.15 SFC in male and female healthy volunteers, low-protein patients and OCP-females post-Synacthen.

	Mean Post-Synacthen Serum Free Cortisol (2.5th – 97.5th centile)			
Assay	**Low protein patients n = 10**	**Males n = 60**	**Females n = 79**	**OCP females n = 24**
GC-MS	50.0 (34.8 – 71.9)	26.5 (15.3 – 45.7)	23.0 (13.4 – 39.5)	20.1 (12.1 – 33.2)
Centaur		29.7 (16.6 – 53.2)	24.4 (13.4 – 44.3)	19.0 (11.2 – 32.1)
Abbott	44.0 (31.4 – 61.7)	27.9 (15.8 – 49.2)	22.0 (12.5 – 38.8)	16.1 (9.91– 26.1)
Roche	-	46.9 (24.2 – 91.1)	34.7 (18.0 – 66.9)	26.4 (15.2 – 45.8)
Immulite	-	33.4 (17.6 – 63.3)	27.9 (15.1 – 51.5)	19.5 (11.2 – 33.9)
Beckman	-	32.0 (16.9 – 60.6)	25.5 (13.4 – 48.6)	16.4 (10.0 – 27.0)

Table 6.16 FCI in male and female healthy volunteers, low-protein patients and OCP-females post-Synacthen.

	Mean Post-Synacthen Free Cortisol Index (2.5th – 97.5th centile)			
Assay	**Low protein patients n = 10**	**Males n = 60**	**Females n = 79**	**OCP females n = 24**
GC-MS	13.7 (10.9 – 17.4)	9.7 (6.7 – 14.0)	8.8 (6.0 – 12.8)	7.7 (5.5 – 10.6)
Centaur	-	10.4 (7.2 – 15.1)	9.1 (6.1 – 13.6)	7.4 (5.2 – 10.5)
Abbott	12.8 (10.0 – 16.4)	10.0 (6.9 – 14.5)	8.5 (5.7 – 12.7)	6.6 (4.7 – 9.3)
Roche	-	13.4 (9.1 – 19.6)	11.2 (7.5 – 16.8)	9.1 (6.5 – 12.6)
Immulite	-	11.1 (7.5 – 16.5)	9.9 (6.6 – 14.8)	7.5 (5.2 – 10.9)
Beckman	-	10.9 (7.3 – 16.2)	9.4 (6.1 – 14.4)	6.7 (4.7 – 9.5)

6.6.3.2 Comparison with healthy volunteers

Mean CBG concentration was 116 mg/L in women taking exogenous oestrogen which was significantly higher ($p<0.001$) than the corresponding concentrations of 58 mg/L and 64 mg/L in healthy male and female volunteers, respectively. Mean concentration in patients with low serum protein was 41 mg/L which was significantly lower ($p<0.001$) than in healthy volunteers.

The higher CBG concentration in OCP-females, however, was not enough to off-set the increased serum total cortisol concentration than in female volunteers and thus mean baseline SFC and FCI were also significantly higher when total cortisol was measured by GC-MS or the Roche assay ($p<0.05$). For all other assays, there was no significant difference (tables 6.13 & 6.14). Following Synacthen stimulation, mean SFC and FCI were significantly lower in OCP-females than in healthy female volunteers ($p<0.05$) (tables 6.15 & 6.16), although this effect was less marked with GC-MS than with any of the immunoassays studied.

In the low protein patients, mean SFC and FCI were significantly higher than in healthy volunteers and OCP-females at baseline and post-Synacthen (tables 6.13 – 6.16)

6.6.3.3 Correlation with serum and salivary cortisol

SFC and FCI correlated well with GC-MS cortisol in women taking exogenous oestrogen and in patients with low serum protein (figure 6.11), although the relationships differed from those in healthy volunteers. In OCP-females both relationships showed negative bias relative to those in healthy volunteers, whilst in low protein patients the relationships were positively biased – particularly in the case of SFC (figure 6.11)

There was also good correlation with salivary cortisol (figure 6.12), with better agreement with healthy volunteers in each relationship, for both OCP-females and low protein patients (figure 6.12).

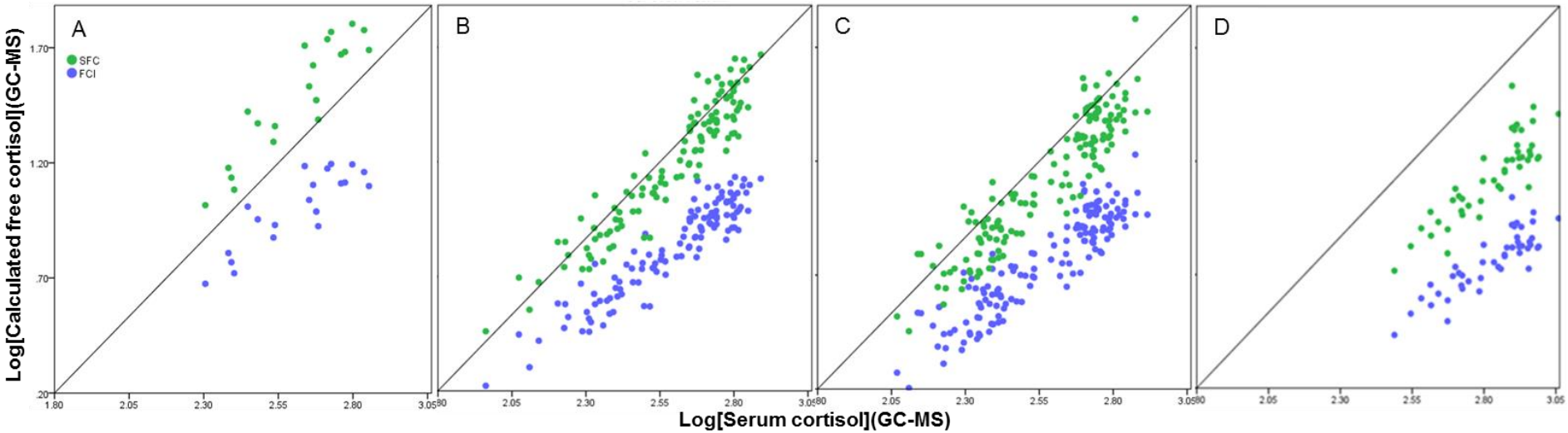


Figure 6.11 Correlation between log-transformed calculated free cortisol (SFC and FCI) and serum cortisol measured by GCMS in all groups studied. Panel (A) shows the relationship in low-protein patients, (B) male volunteers, (C) female volunteers and (D) OCP-females. The solid line represents equivalence between methods.

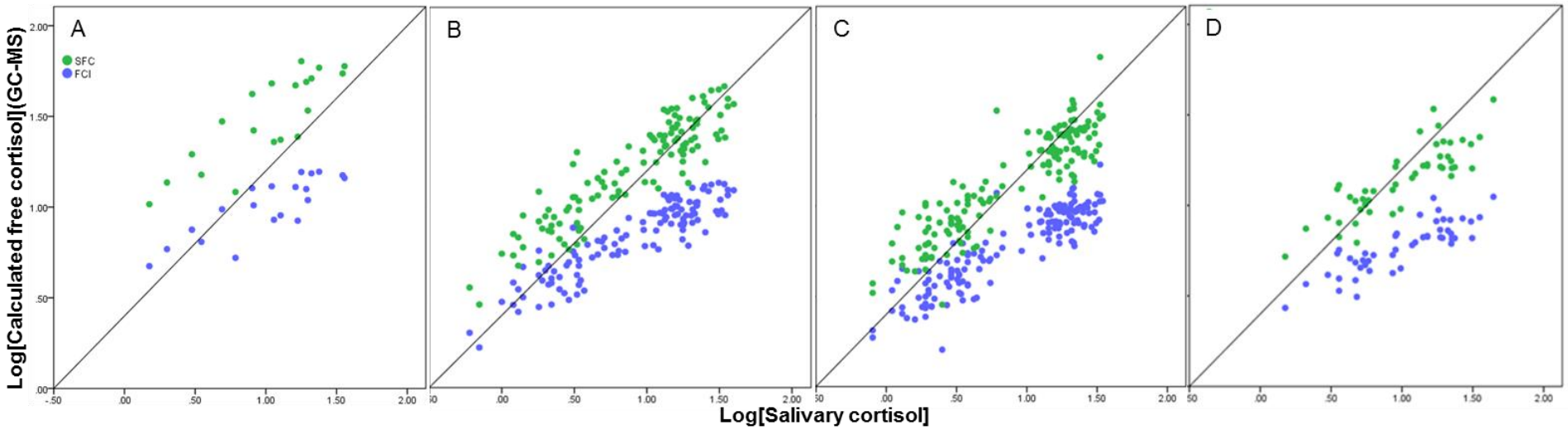


Figure 6.12 Correlation between log-transformed calculated free cortisol (SFC and FCI) and salivary cortisol in all groups studied. Panel (A) shows the relationship in low-protein patients, (B) male volunteers, (C) female volunteers and (D) OCP-females. The solid line represents equivalence between methods.

6.7 Discussion

6.7.1 Serum total cortisol

Serum total cortisol is affected by changes in serum protein concentration (Meulenberg et al. 1987; Klose et al. 2007; Galbois et al. 2010), leading to the potential for adrenal insufficiency to be misdiagnosed or missed in certain patients. Women taking exogenous oestrogens have significantly higher CBG concentrations than women not taking exogenous oestrogens and mean post-Synacthen cortisol concentrations up to 314 nmol/L higher; so using the same cortisol cut-off to diagnose adrenal insufficiency in this group would lead to some patients being missed. Some authors have suggested this effect is best overcome by stopping the OCP for 6 weeks before testing (Simunkova et al. 2008); however, this is neither convenient for patients nor clinically feasible when expedient testing is required and a preferable approach would be to define a more accurate cut-off in this group.

This study has demonstrated that the magnitude of this difference is assay-dependent and, consequently, assay-specific cut-offs are needed for this patient group; however patient numbers were too small to allow a lower reference limit or cut-off to be defined. This study also identified a previously unreported negative bias, relative to GC-MS, in immunoassay cortisol measurement in patients taking exogenous oestrogen. A similar effect was described in pregnant women in a study by Jung *et al* and was shown to be overcome by heating samples to 60°C prior to analysis (Jung et al. 2011).

This bias was presumed to be secondary to excessive cortisol binding to CBG which is denatured by heat, following which cortisol is released, becoming available for detection by anti-cortisol antibodies. However, the effects of heat treatment are non-specific, affecting many, if not all, serum proteins, so eliminating bias in this manner does not confirm excess cortisol-CBG binding as the cause. In contrast, by using Danazol, a synthetic steroid which binds to CBG, thus displacing cortisol, and which

does not cross-react with cortisol antibodies, to reduce the negative bias, it was possible to confirm CBG-cortisol binding as the cause. This is further supported by the evidence of reduced negative bias in the Roche assay, which uses Danazol to dissociate cortisol from CBG prior to analysis, compared to the Abbott assay which relies on a pH change and its antibody's affinity for cortisol, which may be sufficient when CBG concentrations are within normal limits, but not when CBG is present in excess.

The negative bias in samples from women taking exogenous oestrogens further complicates measurement of cortisol in this patient group, and, in particular, raises questions about the utility of immunoassays. It is not feasible to subject all samples from oestrogen-treated or pregnant women to a heat-treatment step prior to analysis, particularly as clinical information is frequently lacking, meaning labs could not reliably determine which samples required treatment. Neither is it feasible for all clinical laboratories to replace immunoassays with LC-MS/MS. Therefore, selecting immunoassays that include a steroid displacement step prior to analysis or putting pressure on manufacturers to routinely implement such a step would seem to be the best available options.

The evidence surrounding the serum cortisol response to Synacthen stimulation in patients with low serum albumin concentrations is somewhat confusing. Many of the studies have been undertaken in critically ill patients in whom the effect of hypoproteinaemia on cortisol measurement and interpretation of the Synacthen test is confounded by an activated HPA axis and disturbance of the free:bound cortisol equilibrium. Others have selected patients on the basis of their Child-Pugh score, rather than serum albumin concentration, resulting in the recruitment of patients with serum albumin concentrations that are no different to those in the healthy volunteer group (Rauschecker et al. 2016). Klose *et al* included a small group of patients with the nephrotic syndrome in their study of the factors affecting interpretation of the

Synacthen test; however, they do not make clear what the serum albumin or CBG concentrations were in their patient group and in any case Galbois *et al* showed little correlation between serum albumin and CBG concentrations in their study of adrenal function in hospitalised cirrhotic patients (Klose et al. 2007; Galbois et al. 2010).

This study selected a group of outpatients with a diagnosis of either cirrhosis or the nephrotic syndrome who had a serum albumin concentration between 24 and 34 g/L. Serum CBG was shown to be significantly lower in this group than in healthy volunteers but there was no clinically significant difference in serum total cortisol at baseline or post-Synacthen between the two groups. This suggests that serum total cortisol can be reliably used to evaluate the adrenal axis in well patients with moderately low serum albumin, although as only one of the ten patients studied had a serum albumin concentration below 29 g/L, further confirmation in patients with lower serum albumin is still required.

6.7.2 Salivary cortisol

Salivary cortisol has been suggested as an alternative to serum total cortisol for evaluating the adrenal axis in healthy patients (Deutschbein et al. 2009), acutely unwell hospitalised patients (Raff et al. 2008), patients with cirrhosis (Galbois et al. 2010) and women taking an oral contraceptive oestrogen-containing pill (Simunkova et al. 2008). It shows good agreement with serum total cortisol in healthy patients and correlates better with HPA axis disease than serum total cortisol in patients with altered protein states.

This study confirmed that there is no significant difference between the mean salivary cortisol response to Synacthen in healthy volunteers, women taking an oestrogen-containing OCP and patients with mild hypoproteinaemia secondary to cirrhosis or the nephrotic syndrome and suggests that this could be used as a suitable alternative to serum total cortisol for identifying patients with adrenal insufficiency without the

need for specific cut-offs in different populations. However, salivary cortisol was not 100% specific in patients with altered protein concentrations, as two of the OCP-females and one patient with low serum protein failed the salivary Synacthen test when the 10.3 nmol/L lower reference limit derived in chapter 4 was used. None of these patients had symptoms of adrenal insufficiency at the time of testing, so the risk of over-diagnosis in some patients remains.

Comparison with serum cortisol revealed that none of these patients would have failed the Synacthen test based on their serum cortisol concentration if the lower reference limit derived in healthy volunteers was used as a cut-off. However, using the lower reference limit of the serum response in the OCP group; one of the OCP-females would also have failed with serum cortisol measured by GC-MS and the Centaur assay, although not with any of the other 4 immunoassays, but the other would not.

Mean baseline salivary cortisol was significantly higher in women taking exogenous oestrogen and in patients with low serum protein than in healthy volunteers. This was not entirely unexpected in either patient group as both findings have been reported elsewhere. Boisseau *et al* showed that women taking an oestrogen-containing OCP had higher basal salivary cortisol concentrations than their untreated counterparts and reduced HPA responsiveness to exercise (Boisseau et al. 2013). The mechanism behind these findings is not known nor is its significance; however, it indicates that the higher serum total cortisol concentrations seen in women taking an OCP are not simply due to increased serum protein concentration but represent a genuine physiological, or pathological, effect. Further study to try to understand the significance of this effect is now needed.

Several studies have explored free cortisol concentration in patients with low serum protein due to cirrhosis or secondary to acute disease, although most used serum

free cortisol rather than salivary cortisol in their investigations (Tan et al. 2010; Degand et al. 2015). Thevenot *et al* measured both serum free cortisol and salivary cortisol and showed a non-significant increase in both in patients with liver disease of increasing severity (progression from Child-Pugh score A through to C, corresponding to serum albumin concentrations of 37.3 ± 4.8 g/L, 26.9 ± 5.3 g/L and 24.3 ± 5.5 g/L), despite progressively lower serum total cortisol concentrations (Thevenot et al. 2011).

Serum free cortisol has been shown to be significantly higher in patients with cirrhosis than in healthy volunteers, despite equivalent baseline total cortisol concentrations (Tan et al. 2010; Degand et al. 2015), and although salivary free cortisol was not measured in these studies a similar finding would be expected. The results of this study suggest that this is, indeed, the case – salivary free cortisol is higher in patients with low serum protein than in healthy volunteers. It has been suggested that this reflects disease severity and associated underlying inflammation; although the presence of this finding in a relatively healthy outpatient population with mild hypoalbuminaemia suggests that this may not be the case. Further study, designed to explore differences in free cortisol rather than total cortisol in this population is, therefore, needed.

The correlation between salivary free cortisol and serum total cortisol was slightly weaker in women taking an OCP and patients with low serum protein than in healthy volunteers, with the relationship in OCP-females showing overall negative bias relative to that in healthy volunteers and low protein patients showing positive bias. This finding can be explained by the differing CBG concentrations between the three groups which result in differences in the fraction of unbound cortisol available for diffusion into saliva (Vining and McGinley 1987). Thus, a larger proportion of serum total cortisol is free in patients with low serum protein compared to healthy volunteers,

who in turn have a larger proportion of free cortisol than women taking exogenous oestrogen. A similar pattern would not be expected in the relationship between serum free cortisol and salivary free cortisol.

6.7.3 Calculated free cortisol

Serum free cortisol was higher than free cortisol index for all assays, at all time points and for all patient groups but they paralleled each other in behaviour so will be considered together as calculated free cortisol (CFC). Baseline calculated free cortisol was shown to be higher in women taking exogenous oestrogens and patients with low serum protein than in healthy volunteers. However, the difference between OCP-females and healthy volunteers was only significant for CFC derived from cortisol measured by GC-MS and the Roche assay; while the difference between low protein patients and healthy volunteers which, based on salivary cortisol concentration, was expected to parallel that seen in OCP-females was significantly higher, at roughly twice the value in healthy volunteers.

The inconsistency in these differences is likely to be explained by measurement inaccuracies rather than being a genuine effect. Serum total cortisol in women taking exogenous oestrogen is subject to negative bias with all immunoassays apart from the Roche, hence the lack of a significant difference with the other assays. The reliability of CBG measurement in the low protein group has already been called into question, and the high CFC lends support to the suggestion that differences between the two CBG assays might account for some of the difference in CBG concentration between healthy volunteers and the low protein group.

6.7.4 Limitations

Recruitment criteria for the patients with low serum protein included a diagnosis of cirrhosis or the nephrotic syndrome and serum albumin concentration below the lower

limit of the reference range (35 mg/L), with the expectation that this group would have lower serum CBG and total cortisol concentrations than healthy volunteers. When it became apparent that there was no significant difference between baseline serum total cortisol in the two groups there was concern that the wrong patient group had been selected. However, the mean albumin concentration of 30.3 ± 2.9 g/L was only slightly higher than that used by Tan *et al* in their study of adrenal function in stable severe liver disease (28.8 ± 1.0 g/L) and was lower than the 36 ± 4 g/L used in the study by Rauschecker *et al* to define the free cortisol response to Synacthen in mildly cirrhotic patients, suggesting this group was reasonably representative of moderate hypoalbuminaemia (Tan et al. 2010; Rauschecker et al. 2016). Nevertheless, comparing mean serum albumin concentration in this group to that in the healthy volunteers and OCP-females would have been a useful addition to this work.

As expected, serum CBG was significantly lower in patients with low serum protein than in healthy volunteers. However, in common with the samples from patients with suspected hypoadrenalism, samples from this group were stored at -80°C for 4 years prior to analysis, raising similar questions to those addressed in chapter 5 about sample stability and assay variation contributing to the lower concentrations seen. Consequently, measurement of serum free cortisol would have been preferable to calculated surrogates as it would provide better understanding of the relationship between serum total and free cortisol and between salivary cortisol and serum free cortisol in patients with altered protein states. It would also allow better evaluation of how well the relationship between measured and calculated free cortisol is maintained in these states.

6.7.5 Further work

This study has shown that serum total cortisol can be used to evaluate the adrenal axis in well patients with mild to moderate hypoproteinaemia but not in women taking

an OCP, in whom a separate reference range is needed if oestrogen cannot be discontinued for 6 weeks in advance of the test. Salivary cortisol is valid in both groups, whilst calculated free cortisol performs less well in either group than in healthy volunteers. Further work is now needed to explore the serum total cortisol response to Synacthen in well patients with severe hypoproteinaemia and the salivary cortisol response in unwell patients – both those with cirrhosis and those in a critical care setting.

CHAPTER 7

GENERAL DISCUSSION AND CONCLUSIONS

The initial premise of this work was that adrenal insufficiency was being over-diagnosed in patients with an intact HPA axis undergoing a Synacthen test because the wrong cortisol cut-off was being used. The results presented here confirm that this was the case, with 10% of healthy volunteers failing the Synacthen test when the 550 nmol/L cut-off was used. This would have risen to a 27% false positive rate, had the cut-off remained the same when the Centaur cortisol assay was reformulated in 2010, and 42% false positive rate when the laboratory switched to the Abbott assay later that same year. An assay-specific cortisol lower reference limit is thus essential for the Synacthen test to retain any value as a diagnostic tool in the assessment of adrenal insufficiency or HPA axis dysfunction.

The stated aims of this thesis were:

- To establish valid serum total cortisol cut-offs for use with the short Synacthen test in patients with normal and altered CBG concentrations.
- To investigate, using current assays, the effect of assay differences on these cut-offs.
- To explore salivary free cortisol and calculated serum free cortisol as alternatives to serum total cortisol in interpreting the short Synacthen test in patients with altered CBG concentrations.

The following discussion sets out how these aims were to be addressed and explores the extent to which they were achieved.

7.1 Serum total cortisol response to Synacthen stimulation

The limitations of cortisol immunoassays, including their poor specificity (Curtis 2009; Owen et al. 2013a) and lack of standardisation, make it difficult to define the true nature of the cortisol response to Synacthen stimulation. Mass spectrometry, on the other hand, suffers none of these limitations and provides definitive identification of

cortisol, and is, therefore, ideally placed to reliably define the response. Using GC-MS, the cortisol response to Synacthen stimulation was shown to be unaffected by age or gender. A single reference range and lower reference limit (defined as the 2.5th percentile) were calculated in healthy volunteers and, at 420 nmol/L, the latter was noted to be significantly lower than the 550 nmol/L cut-off historically used to diagnose adrenal insufficiency.

Reference ranges and lower reference limits were subsequently calculated for the five immunoassays included in the study. The Centaur and Immulite assays most closely resembled GC-MS, in that they showed no gender effect, and for each, a single lower reference limit was calculated. This was determined as 446 nmol/L for the Centaur and 474 nmol/L for the Immulite assay. The Abbott and Beckman assays showed significantly different male and female ranges, but their lower reference limits were close enough for the male limit, which was higher than the female limit, to be adopted as the assay-specific cut-off. Thus, a lower reference limit of 430 nmol/L was recommended for the Abbott assay and 459 nmol/L for the Beckman. For the Roche assay, the differences between genders were too great, necessitating gender-specific lower reference limits of 574 nmol/L in men and 524 nmol/L in women.

The newly-defined Abbott lower reference limit was subsequently applied to a group of patients with suspected adrenal insufficiency undergoing Synacthen testing, to establish its ability to correctly identify disease. It performed well, with a sensitivity and negative predictive value of 100% and a specificity of 78%. In comparison, the 550 nmol/L cut-off would have had an equivalent sensitivity and negative predictive value, but with a specificity of only 26%, and, given that long-term exogenous glucocorticoids are not without side-effects, such a high failure rate is unacceptable. These findings demonstrate the superiority of an assay-specific lower reference limit over the 550 nmol/L cut-off, particularly in terms of avoiding over-diagnosis.

This study also included a pre-test probability score derived from patients' symptoms and clinical history. This was a subjective score based entirely on the author's clinical judgement and has not been validated in any way. Nevertheless, it had the same sensitivity and negative predictive value as the Synacthen test, with a specificity of 65%, which did not overlap entirely with the false positive results from the Synacthen test. Thus it may be possible to improve the specificity of the Synacthen test by combining its result with a pre-test probability of disease score, albeit one which has been derived by consensus from a group of Consultant Endocrinologists, either locally or at a national level, and extensively validated in an appropriate population.

One of the objectives of this study was to determine the validity of the newly defined cut-offs in patients with suspected adrenal insufficiency, and to a certain extent this has been achieved. However, the patient group was not large or homogeneous enough to unequivocally state that the cut-offs established in healthy volunteers are equally valid in patients with adrenal insufficiency. Furthermore, testing in the patient group was only performed with the Abbott assay, so even with a larger, better defined group of patients it would not be possible to confirm that all the lower reference limits determined in this study could accurately differentiate between adrenal insufficiency and normal adrenal function. Nevertheless, the results from the Abbott assay are encouraging, the lower reference limits for 4 of the immunoassays studied fall within 50 nmol/L of each other and their performance in healthy volunteers is broadly similar. There is, thus, no reason to suggest their performance in patients with adrenal disease would be significantly worse than that of the Abbott.

Another of the objectives was to determine whether a cortisol cut-off determined in patients with normal serum CBG concentrations was equally valid in patients with high and low serum CBG concentration. To this end a group of women taking an oestrogen-containing oral contraceptive pill and a group of patients with low serum

protein as a result of the nephrotic syndrome or cirrhosis underwent Synacthen testing. The lower reference limit in women taking exogenous oestrogen was between 145 nmol/L and 267 nmol/L higher than in healthy volunteers with normal CBG concentrations, depending on the assay used. None of the 24 women studied failed the Synacthen test, which was not unexpected as they were healthy volunteers; however, this contrasted with a 1.4% false positive rate in healthy volunteers not taking oestrogens, and, given how much higher the cortisol response in the oestrogen group was, there is a risk that the cut-offs derived in this study would miss adrenal insufficiency, particularly in its early stages.

In patients with low serum protein, the opposite was true. Mean serum cortisol concentration post Synacthen was 514 nmol/L for the Abbott assay and 552 nmol/L with GC-MS, with corresponding lower limits of 384 nmol/L and 393 nmol/L, respectively. Two of the ten patients studied (20%) would have failed the Synacthen test using the Abbott-specific cut-off of 430 nmol/L, compared to 2 out of 139 healthy volunteers (1.4%) and none of the 24 women taking oestrogens. This false positive rate would have risen to 70% if the 550 nmol/L cut-off were used instead.

Similar findings have been reported elsewhere, with Klose et al showing a 50% failure rate with the Synacthen test in a group of patients with the nephrotic syndrome when a 500 nmol/L cut-off was used (Klose et al. 2007). A similar proportion (46%) of cirrhotic patients, with mean serum albumin concentration of 21 g/L, failed to achieve a total cortisol of 550 nmol/L post Synacthen stimulation (Vincent et al. 2009).

Thus, this study confirms the need for assay-specific serum total cortisol reference ranges and lower reference limits and succeeds in defining these in healthy volunteers and, to some extent, in validating them in patients with suspected adrenal insufficiency. It demonstrates that these cut-offs, if used in patients with altered CBG concentrations, would result in either under- or over-diagnosis of adrenal

insufficiency, depending on whether CBG was high or low, indicating a need for separate validated reference ranges in these groups if serum total cortisol is to be used. Alternatively, consideration should be given to using either salivary or serum free cortisol in the assessment of the HPA axis in these patients.

7.2 Salivary free cortisol

A further objective of this study was to establish a reference range and lower reference limit for the salivary cortisol response to Synacthen in healthy volunteers and determine their validity in patients with suspected adrenal insufficiency and high and low serum CBG concentration. Salivary free cortisol has the advantage over serum total cortisol in that it should be unaffected by changes in serum protein concentration and samples are easier to collect. A mean salivary cortisol concentration of 19.3 nmol/L was shown to be valid in both male and female healthy volunteers, with a lower reference limit of 10.3 nmol/L. Applying this cut-off in patients with suspected adrenal insufficiency showed it performed better than serum total cortisol, matching the latter's sensitivity and negative predictive value of 100%, but with a higher specificity of 82%. It also out-performed serum total cortisol in healthy volunteers, with only 1 of the 139 recruits (0.7%) failing to reach the cut-off.

As expected, in patients with low serum protein and in women taking an OCP, the post-Synacthen response was independent of serum CBG concentration. Despite slight differences between the range of responses, there was no significant difference between mean salivary cortisol in healthy volunteers and those with either high or low CBG, although false positive results were more prevalent in both the latter groups. In women taking oestrogens, 2 out of 24 (8%) failed the Synacthen test using the cut-off defined in healthy volunteers, whilst 1 of the 10 patients with low protein (10%) failed. Salivary cortisol, therefore, offers a very real alternative to serum total cortisol; but does not eliminate the potential for misdiagnosis entirely.

However, salivary cortisol has some way to go before it can be used routinely to evaluate patients with suspected adrenal insufficiency. It is measured by non-standardised LC-MS/MS assays, and its external quality assurance scheme, which includes 15 assays, shows these vary substantially. The results from the assay used in this study typically fall at, or near, the lower end of the range, which may help explain why the concentrations presented here are lower than many of those in the literature, which are, themselves, equally diverse (Perogamvros et al. 2009; Montsko et al. 2014; Cornes et al. 2015). These results are, consequently, not applicable to other assays and cannot be put forward as a single, valid cut-off for use with the Synacthen test. Laboratories will, therefore, need to determine their own reference ranges and cut-offs, but for these to be accurate they will need to include large numbers of patients and controls, and ideally should also include a comparison to the insulin tolerance test.

Such large-scale studies are not feasible for most clinical laboratories and it is difficult to envisage saliva ever replacing serum total cortisol for routine evaluation of patients with suspected adrenal insufficiency. Its main role is, therefore, likely to be found in evaluating the HPA axis in critically unwell patients, given the inaccuracy in serum total cortisol in these patients (Hamrahian et al. 2004; Cohen et al. 2006; Briegel et al. 2009). It may also have a role to play in evaluating adrenal function in patients with altered serum CBG concentrations, particularly those with low serum protein.

Thus, this study was able to establish a reference range and lower reference limit for the salivary cortisol response to Synacthen and was able to demonstrate that this cut-off was valid in patients with suspected adrenal insufficiency and in those with altered CBG concentration. However, differences between salivary cortisol assays limit the widespread utility of this cut-off and further work exploring its response to the insulin

stress test and its performance in larger groups of patients with altered serum protein concentration and hypoadrenalism is needed.

7.3 Calculated serum free cortisol

The final objective of this study was to establish a reference range and lower reference limit for the calculated serum free cortisol response to Synacthen in healthy volunteers and determine its validity in the groups of patients mentioned above. Calculated, rather than measured, serum free cortisol was chosen because of the lack of availability of routine serum free cortisol assays, and both Coolens' serum free cortisol (SFC) and the free cortisol index (FCI) were calculated to establish whether either showed superiority over the other.

The results presented here show that both calculated parameters were assay and gender- dependent, thus offering a slight disadvantage over serum total cortisol, which exhibits assay, but not gender, dependence. Twelve different reference ranges and lower reference limits were, therefore, required for each of the two calculated parameters and the outcome of the Synacthen test depended on which assay was being used. Thus, the Immulite SFC cut-off misdiagnosed adrenal insufficiency in 2 healthy volunteers compared to 6 with the Roche cut-off.

Calculated serum free cortisol also performed worse in patients with suspected adrenal insufficiency. This was because serum CBG concentration in this group was found to be significantly lower than in healthy volunteers, resulting in higher mean calculated serum free cortisol. This finding may have been genuine, but it did not fit with the results of serum total and salivary free cortisol in this group, which were both lower than in healthy volunteers, and could not be verified because serum free cortisol had not been measured. This effect lowered the sensitivity of the Synacthen test to 57% when calculated serum free cortisol was used, although the corresponding specificity was 100%.

The utility of calculated serum free cortisol in interpreting the Synacthen test in patients with altered CBG concentrations was also difficult to establish. The mean response in women taking exogenous oestrogens was significantly lower than in healthy volunteers, whilst in patients with low serum protein it was significantly higher. Again, the lack of measured serum free cortisol for comparison made it impossible to ascertain whether this was a genuine effect or an artefact of measurement, although its disagreement with the salivary free cortisol response in these groups, supported it being artefactual. It has also been suggested that calculated serum free cortisol parameters, which were determined in patients with normal CBG concentrations, may not be valid in patients with altered CBG concentration (Cohen et al. 2013; Molenaar et al. 2015), or, indeed, in patients undergoing a Synacthen test (Barlow et al. 2010).

Thus, although it was possible to calculate reference ranges and lower reference limits for SFC and FCI in healthy volunteers, these could not be validated in patients with suspected adrenal insufficiency and the utility of calculated serum free cortisol in patients with altered CBG concentration was not established.

7.4 General conclusions

7.4.1 Summary of findings

The main findings of this thesis can, therefore, be summarised as follows:

- Assay-specific reference ranges for the serum total cortisol response to Synacthen are needed to avoid over-diagnosing adrenal insufficiency. However, even with assay-specific ranges, the Synacthen test will potentially misdiagnose adrenal insufficiency in over 1 in 5 unaffected patients when serum total cortisol is used.
- Salivary cortisol is a suitable alternative to serum total cortisol for evaluating suspected adrenal insufficiency and offers better specificity. Unlike serum

total cortisol, it does not appear to be affected by changes in CBG concentration and, hence, offers the potential for a single cut-off to be used in the interpretation of the Synacthen test.

- Calculated serum free cortisol does not present a reliable alternative to either serum total or salivary cortisol in the interpretation of the Synacthen test.

In addition to its main findings, this study also suggested a potential role for a pre-test likelihood of disease score in improving the diagnostic specificity of the Synacthen test, it identified ongoing changes to cortisol immunoassays, difficulties in measuring serum CBG and potential differences between serum total and serum free cortisol in healthy volunteers. These are briefly discussed below.

7.4.2 Pre-test likelihood of disease

The increase in failed Synacthen tests in patients with few clinical features of adrenal insufficiency that first prompted this study, suggests that the pre-test probability of disease is not always considered before testing is undertaken. This study derived a pre-test score which showed 88% concordance with the outcome of the Synacthen test. Combining this score with the result of the Synacthen test was subsequently shown to improve diagnostic specificity from 78% with the Synacthen test alone to 87%. There is, therefore, likely to be some diagnostic gain from deriving and applying such a score to patients undergoing Synacthen testing, particularly when testing is being undertaken by non-specialists.

7.4.3 Cortisol immunoassays

Despite 50 years of cortisol assays, measurement of serum total cortisol continues to present significant challenges. Differences between immunoassays necessitate assay-specific reference ranges and cut-offs, but ongoing changes in assay performance mean these do not remain valid indefinitely. This was highlighted in this

study by the reformulation of the Siemens Centaur assay. This lowered the assay-specific cut-off from 499 nmol/L to 446 nmol/L, but despite alerting users to a change in the assay, the manufacturer did not indicate that a reference range change would be necessary. Similar changes have been noted with the Abbott assay over the past 3 years. EQA returns have identified a gradual reduction in cortisol concentration, although there has been no formal notification of any change by the manufacturer. The Roche assay has also seen significant change recently. This has eliminated the gender difference but has left Roche users without a valid cortisol reference range or cut-off.

LC-MS/MS assays offer a potential solution to this problem, particularly with the improved throughput and turnaround times of current assays (Owen et al. 2013a). They also provide a more accurate measure of serum total cortisol than immunoassay in patients with high serum protein concentrations, those taking exogenous glucocorticoids (Owen et al. 2013a) and the critically ill (Briegel et al. 2009). Nevertheless, LC-MS/MS is not yet prevalent enough in clinical laboratories to recommend abandoning cortisol immunoassays in its favour.

7.4.4 CBG measurement

This study highlights the pitfalls of using non-standardised assays without external quality assurance checks. Serum CBG assays fall into this category and, as a result, it was impossible to determine whether differences between calculated serum free cortisol responses in different patient groups were genuine. In particular, low CBG concentrations in patients with suspected adrenal disease, which resulted in equivalent serum free cortisol concentrations in patients with potential adrenal disease and healthy volunteers, could not be adequately explained. These results did not match those found with serum total or salivary cortisol and contradicted the findings of other studies which have shown lower measured serum free cortisol in

patients with adrenal insufficiency than in healthy volunteers (Rauschecker et al. 2016).

7.4.5 Serum free cortisol

Serum total cortisol differentiates between gross adrenal insufficiency and "normal" adrenal function well, but does not permit an understanding of the nuances of fluctuations in bioavailable cortisol. This study demonstrated significantly lower calculated serum free cortisol concentrations in men than in women, despite there being no significant gender difference in baseline serum total cortisol or salivary cortisol concentration. A similar gender difference in measured serum free cortisol has been reported elsewhere (Sofer et al. 2016), suggesting it is likely to be a genuine finding that is undetectable with serum total cortisol.

Thus, serum total cortisol may not reflect serum free cortisol concentration accurately enough to allow a clear understanding of cortisol's full physiological role. To date, it has been suggested that cortisol excess plays a part in depression (Scott and Dinan 1998), the metabolic syndrome (Reinehr et al. 2014), cardiovascular disease and hypertension (Morelli et al. 2010). Low urinary free cortisol and hypofunction of the HPA axis have been linked with chronic fatigue (Scott and Dinan 1998; Scott et al. 1998). However, to establish whether cortisol has a direct causative role in these conditions requires a thorough understanding of free cortisol in health, which can only be gained by measuring serum free cortisol directly.

7.5 Further work

This study has shown that calculated free cortisol is not reliable enough to replace measured serum free cortisol in the evaluation of the HPA axis. Further work using measured serum free cortisol in this setting is therefore needed. Defining a reference range for serum free cortisol in a well-defined "healthy" population is also needed, in

order to understand cortisol's contribution to disease. However, this will require improved assays.

Salivary cortisol has the potential to replace serum total cortisol for evaluating the HPA axis in patients with altered serum protein concentrations. Further studies exploring the salivary cortisol response to the insulin stress test and larger studies in patients with altered serum protein concentration and hypoadrenalism are necessary before it can be safely used for this purpose.

Ongoing changes to cortisol immunassays threaten the reliability of their reference ranges and cut-offs. By comparing immunoassay results to those measured by GC-MS, this study has provided a reference point against which future assay changes can be quantified. This should allow valid reference ranges to be derived in the future, without a need to repeat the work undertaken here. External quality assurance laboratories have access to high quality mass methods and are ideally placed to undertake this work, although the cooperation of assay manufacturers will be required.

PUBLICATIONS

PAPERS:

El-Farhan N, Pickett A, Ducroq D, Bailey C, Mitchem K, Morgan N, Armston A, Jones L, Evans C, Rees DA. Method-specific serum cortisol responses to the adrenocorticotrophin test: comparison of gas-chromatography-mass spectrometry and five automated immunoassays. *Clin Endocrinol (Oxf)* 2013;**78**:673-80.

Owen LJ, Adaway JE, Davies S, Neale S, **El-Farhan N**, Ducroq D, Evans C, Rees DA, MacKenzie F, Keevil BG. Development of a rapid assay for the analysis of serum cortisol and its implementation into a routine service laboratory. *Ann Clin Biochem* 2013;**50**:345-52.

Evans C, **El-Farhan N**, Ducroq D, Rees A. Response to 'Determining the utility of the 60-min cortisol measurement in the short Synacthen test'. *Clin Endocrinol (Oxf)* 2014;**80**:766-7.

Dodd AJ, Ducroq DH, Neale SM, Wise MP, Mitchem KL, Armston A, Barth JH, **El-Farhan N**, Rees DA, Evans C. The effect of serum matrix and gender on cortisol measurement by commonly used immunoassays. *Ann Clin Biochem* 2014;**51**:379-85.

El-Farhan N, Rees DA, Evans C. Measuring cortisol in serum, urine and saliva – are our assays good enough? *Ann Clin Biochem* 2017;**54**:308-22

REFERENCES

UKNEQAS Annual review 2010.

UKNEQAS Annual review 2012.

UKNEQAS Annual review 2013.

Abdu, T. A. et al. 1999. Comparison of the low dose short synacthen test (1 microg), the conventional dose short synacthen test (250 microg), and the insulin tolerance test for assessment of the hypothalamo-pituitary-adrenal axis in patients with pituitary disease. *J Clin Endocrinol Metab* 84(3), pp. 838-843.

Allende, F. et al. 2014. LC-MS/MS Method for the Simultaneous Determination of Free Urinary Steroids. *Chromatographia* 77, pp. 637-642.

Amirian, I. et al. 2015. Working night shifts affects surgeons' biological rhythm. *Am J Surg* 210(2), pp. 389-395.

Annesley, T. M. 2003. Ion suppression in mass spectrometry. *Clin Chem* 49(7), pp. 1041-1044.

Arafah, B. M. 2006. Hypothalamic pituitary adrenal function during critical illness: limitations of current assessment methods. *J Clin Endocrinol Metab* 91(10), pp. 3725-3745.

Arlt, W. and Allolio, B. 2003. Adrenal insufficiency. *Lancet* 361(9372), pp. 1881-1893.

Bacarese-Hamilton, T. et al. 1992. A fully automated enzyme immunoassay for the measurement of cortisol in biological fluids. *Eur J Clin Chem Clin Biochem* 30(9), pp. 531-535.

Baid, S. K. et al. 2007. Radioimmunoassay and tandem mass spectrometry measurement of bedtime salivary cortisol levels: a comparison of assays to establish hypercortisolism. *J Clin Endocrinol Metab* 92(8), pp. 3102-3107.

Barlow, N. L. et al. 2010. An evaluation of measured and calculated serum free cortisol in a group of patients with known adrenal suppression. *Ann Clin Biochem* 47(Pt 3), pp. 200-204.

Bendel, S. et al. 2008. Free cortisol in sepsis and septic shock. *Anesth Analg* 106(6), pp. 1813-1819.

Boisseau, N. et al. 2013. Oral contraception but not menstrual cycle phase is associated with increased free cortisol levels and low hypothalamo-pituitary-adrenal axis reactivity. *J Endocrinol Invest* 36(11), pp. 955-964.

Bonte, H. A. et al. 1999. The use of free cortisol index for laboratory assessment of pituitary-adrenal function. *Clin Chem Lab Med* 37(2), pp. 127-132.

Booth, M. et al. 1961. Protein binding of cortisol in health and in pregnancy. *J Endocrinol* 23, pp. 25-35.

Bornstein, S. R. et al. 2016. Diagnosis and Treatment of Primary Adrenal Insufficiency: An Endocrine Society Clinical Practice Guideline. *J Clin Endocrinol Metab* 101(2), pp. 364-389.

Borst, G. C. et al. 1982. Discordant cortisol response to exogenous ACTH and insulin-induced hypoglycemia in patients with pituitary disease. *N Engl J Med* 306(24), pp. 1462-1464.

Breuer, H. and Siekmann, L. 1975. Mass fragmentography as reference method in clinical steroid assay. *J Steroid Biochem* 6(5), pp. 685-688.

Briegel, J. et al. 2009. Multicenter comparison of cortisol as measured by different methods in samples of patients with septic shock. *Intensive Care Med* 35(12), pp. 2151-2156.

Brock, P. et al. 1978. Direct solid-phase 125I radioimmunoassay of serum cortisol. *Clin Chem* 24(9), pp. 1595-1598.

Brossaud, J. et al. 2009. Cross reactions elicited by serum 17-OH progesterone and 11-desoxycortisol in cortisol assays. *Clin Chim Acta* 407(1-2), pp. 72-74.

Burke, C. W. 1969a. Accurate measurement of steroid-protein binding by steady-state gel filtration. *Biochim Biophys Acta* 176(2), pp. 403-413.

Burke, C. W. 1969b. Biologically active cortisol in plasma of oestrogen-treated and normal subjects. *Br Med J* 2(5660), pp. 798-800.

Campuzano, H. C. et al. 1973. A radioimmunoassay for cortisol in human plasma. *Biochem Med* 7(3), pp. 350-362.

Chatha, K. K. et al. 2010. National UK audit of the short synacthen test. *Ann Clin Biochem* 47(Pt 2), pp. 158-164.

Chitale, A. et al. 2013. Determining the utility of the 60 min cortisol measurement in the short synacthen test. *Clin Endocrinol (Oxf)* 79(1), pp. 14-19.

Chrousos, G. P. 1995. The hypothalamic-pituitary-adrenal axis and immune-mediated inflammation. *N Engl J Med* 332(20), pp. 1351-1362.

Clark, P. M. et al. 1998. Defining the normal cortisol response to the short Synacthen test: implications for the investigation of hypothalamic-pituitary disorders. *Clin Endocrinol (Oxf)* 49(3), pp. 287-292.

Clayton, R. N. 1989. Diagnosis of adrenal insufficiency. *BMJ* 298(6669), pp. 271-272.

Cohen, J. et al. 2004. Salivary cortisol concentration in the intensive care population: correlation with plasma cortisol values. *Anaesth Intensive Care* 32(6), pp. 843-845.

Cohen, J. et al. 2013. Comparison of the diagnostic accuracy of measured and calculated free cortisol in acutely ill patients using the Coolens equation. *Crit Care Resusc* 15(1), pp. 39-41.

Cohen, J. et al. 2006. Variability of cortisol assays can confound the diagnosis of adrenal insufficiency in the critically ill population. *Intensive Care Med* 32(11), pp. 1901-1905.

Coolens, J. L. et al. 1987. Clinical use of unbound plasma cortisol as calculated from total cortisol and corticosteroid-binding globulin. *J Steroid Biochem* 26(2), pp. 197-202.

Cooper, M. S. and Stewart, P. M. 2003. Corticosteroid insufficiency in acutely ill patients. *N Engl J Med* 348(8), pp. 727-734.

Cornes, M. P. et al. 2015. Salivary cortisol and cortisone responses to tetracosactrin (synacthen). *Ann Clin Biochem* 52(Pt 5), pp. 606-610.

Courtney, C. H. et al. 2000. Comparison of one week 0900 h serum cortisol, low and standard dose synacthen tests with a 4 to 6 week insulin hypoglycaemia test after pituitary surgery in assessing HPA axis. *Clin Endocrinol (Oxf)* 53(4), pp. 431-436.

Crowley, S. et al. 1991. The use of low doses of ACTH in the investigation of adrenal function in man. *J Endocrinol* 130(3), pp. 475-479.

Cunningham, S. K. et al. 1983. Normal cortisol response to corticotropin in patients with secondary adrenal failure. *Arch Intern Med* 143(12), pp. 2276-2279.

Curtis, G. E., Stutchfield, P.R. 2009. A case to illustrate diagnostic pitfalls in a district general hospital: analytical and interpretative. *Ann Clin Biochem* 46(Suppl 1), p. 130.

Davidson, J. S. et al. 2006. A case of low cortisol-binding globulin: use of plasma free cortisol in interpretation of hypothalamic-pituitary-adrenal axis tests. *Ann Clin Biochem* 43(Pt 3), pp. 237-239.

Degand, T. et al. 2015. Assessment of adrenal function in patients with acute hepatitis using serum free and total cortisol. *Dig Liver Dis* 47(9), pp. 783-789.

Demers, L. 2005. The adrenal cortex. In: CA, B. et al. eds. *Tietz textbook of clinical chemistry*. Elsevier Saunders, pp. 2003-2052.

DeSantis, A. S. et al. 2015. Racial and ethnic differences in diurnal cortisol rhythms: are they consistent over time? *Psychosom Med* 77(1), pp. 6-15.

Deutschbein, T. et al. 2012. Salivary cortisol as a diagnostic tool for Cushing's syndrome and adrenal insufficiency: improved screening by an automatic immunoassay. *Eur J Endocrinol* 166(4), pp. 613-618.

Deutschbein, T. et al. 2009. Diagnosis of secondary adrenal insufficiency in patients with hypothalamic-pituitary disease: comparison between serum and salivary cortisol during the high-dose short synacthen test. *Eur J Endocrinol* 160(1), pp. 9-16.

Dhillo, W. S. et al. 2002. Cortisol-binding globulin is important in the interpretation of dynamic tests of the hypothalamic--pituitary--adrenal axis. *Eur J Endocrinol* 146(2), pp. 231-235.

Dickstein, G. 2001. Hypothalamo-pituitary-adrenal axis testing: nothing is sacred and caution in interpretation is needed. *Clin Endocrinol (Oxf)* 54(1), pp. 15-16.

Dickstein, G. et al. 1991. Adrenocorticotropin stimulation test: effects of basal cortisol level, time of day, and suggested new sensitive low dose test. *J Clin Endocrinol Metab* 72(4), pp. 773-778.

Dodd, A. J. et al. 2014. The effect of serum matrix and gender on cortisol measurement by commonly used immunoassays. *Ann Clin Biochem* 51(Pt 3), pp. 379-385.

Dorin, R. I. et al. 2003. Diagnosis of adrenal insufficiency. *Ann Intern Med* 139(3), pp. 194-204.

Dorn, L. D. et al. 2007. Salivary cortisol reflects serum cortisol: analysis of circadian profiles. *Ann Clin Biochem* 44(Pt 3), pp. 281-284.

Elbuken, G. et al. 2016. Comparison of total, salivary and calculated free cortisol levels in patients with severe sepsis. *J Intensive Care* 4, p. 3.

Erickson, D. et al. 2012. Late-night salivary cortisol for diagnosis of Cushing's syndrome by liquid chromatography/tandem mass spectrometry assay. *Clin Endocrinol (Oxf)* 76(4), pp. 467-472.

Estrada, Y. M. R. M. and Orlander, P. R. 2011. Salivary cortisol can replace free serum cortisol measurements in patients with septic shock. *Chest* 140(5), pp. 1216-1222.

Fede, G. et al. 2014. Comparison of total cortisol, free cortisol, and surrogate markers of free cortisol in diagnosis of adrenal insufficiency in patients with stable cirrhosis. *Clin Gastroenterol Hepatol* 12(3), pp. 504-512.e508; quiz e523-504.

Fiad, T. M. et al. 1994. The overnight single-dose metyrapone test is a simple and reliable index of the hypothalamic-pituitary-adrenal axis. *Clin Endocrinol (Oxf)* 40(5), pp. 603-609.

Galbois, A. et al. 2010. Assessment of adrenal function in cirrhotic patients: salivary cortisol should be preferred. *J Hepatol* 52(6), pp. 839-845.

Grinspoon, S. K. and Biller, B. M. 1994. Clinical review 62: Laboratory assessment of adrenal insufficiency. *J Clin Endocrinol Metab* 79(4), pp. 923-931.

Hamrahian, A. H. et al. 2004. Measurements of serum free cortisol in critically ill patients. *N Engl J Med* 350(16), pp. 1629-1638.

Haning, R. V., Jr. et al. 1982. Danazol and its principal metabolites interfere with binding of testosterone, cortisol, and thyroxin by plasma proteins. *Clin Chem* 28(4 Pt 1), pp. 696-698.

Hawley, J. M. et al. 2016. Candidate Reference Measurement Procedure for the Quantification of Total Serum Cortisol with LC-MS/MS. *Clin Chem* 62(1), pp. 262-269.

Ho, J. T. et al. 2006. Septic shock and sepsis: a comparison of total and free plasma cortisol levels. *J Clin Endocrinol Metab* 91(1), pp. 105-114.

Holder, G. 2006. Measurement of glucocorticoids in biological fluids. *Methods Mol Biol* 324, pp. 141-157.

Honour, J. W. 2011. Development and validation of a quantitative assay based on tandem mass spectrometry. *Ann Clin Biochem* 48(Pt 2), pp. 97-111.

Hurel, S. J. et al. 1996. The short Synacthen and insulin stress tests in the assessment of the hypothalamic-pituitary-adrenal axis. *Clin Endocrinol (Oxf)* 44(2), pp. 141-146.

Inder, W. J. et al. 2012. Measurement of salivary cortisol in 2012 - laboratory techniques and clinical indications. *Clin Endocrinol (Oxf)* 77(5), pp. 645-651.

Jones, S. L. et al. 1994. An audit of the insulin tolerance test in adult subjects in an acute investigation unit over one year. *Clin Endocrinol (Oxf)* 41(1), pp. 123-128.

Jonsson, B. A. et al. 2003. Determination of cortisol in human saliva using liquid chromatography-electrospray tandem mass spectrometry. *J Chromatogr B Analyt Technol Biomed Life Sci* 784(1), pp. 63-68.

Jung, C. et al. 2011. A longitudinal study of plasma and urinary cortisol in pregnancy and postpartum. *J Clin Endocrinol Metab* 96(5), pp. 1533-1540.

Kamel, R. M. 2010. The onset of human parturition. *Arch Gynecol Obstet* 281(6), pp. 975-982.

Kao, P. C. et al. 2001. Diagnosis of adrenal cortical dysfunction by liquid chromatography-tandem mass spectrometry. *Ann Clin Lab Sci* 31(2), pp. 199-204.

Kazlauskaite, R. et al. 2008. Corticotropin tests for hypothalamic-pituitary- adrenal insufficiency: a metaanalysis. *J Clin Endocrinol Metab* 93(11), pp. 4245-4253.

Keefe, C. C. et al. 2014. Simultaneous measurement of thirteen steroid hormones in women with polycystic ovary syndrome and control women using liquid chromatography-tandem mass spectrometry. *PLoS One* 9(4), p. e93805.

Keevil, B. G. 2013. Novel liquid chromatography tandem mass spectrometry (LC-MS/MS) methods for measuring steroids. *Best Pract Res Clin Endocrinol Metab* 27(5), pp. 663-674.

Kehlet, H. and Binder, C. 1973. Value of an ACTH test in assessing hypothalamic-pituitary-adrenocortical function in glucocorticoid-treated patients. *Br Med J* 2(5859), pp. 147-149.

Kehlet, H. et al. 1976. Short ACTH test in assessing hypothalamic-pituitary-adrenocortical function. *Br Med J* 1(6004), pp. 249-251.

Kirchhoff, F. et al. 2011. Quantification of free serum cortisol based on equilibrium dialysis and isotope dilution-liquid chromatography-tandem mass spectrometry. *Clin Biochem* 44(10-11), pp. 894-899.

Kirschbaum, C. et al. 1999. Impact of gender, menstrual cycle phase, and oral contraceptives on the activity of the hypothalamus-pituitary-adrenal axis. *Psychosom Med* 61(2), pp. 154-162.

Klose, M. et al. 2005. Adrenocortical insufficiency after pituitary surgery: an audit of the reliability of the conventional short synacthen test. *Clin Endocrinol (Oxf)* 63(5), pp. 499-505.

Klose, M. et al. 2007. Factors influencing the adrenocorticotropin test: role of contemporary cortisol assays, body composition, and oral contraceptive agents. *J Clin Endocrinol Metab* 92(4), pp. 1326-1333.

Koal, T. et al. 2012. Standardized LC-MS/MS based steroid hormone profile-analysis. *J Steroid Biochem Mol Biol* 129(3-5), pp. 129-138.

Kominami, G. et al. 1980. A sensitive enzyme immunoassay for plasma cortisol. *Clin Chim Acta* 103(3), pp. 381-391.

Kosak, M. et al. 2014. Serum cortisol seems to be a more appropriate marker for adrenocortical reserve evaluation in ACTH test in comparison to salivary cortisol. *Physiol Res* 63(2), pp. 229-236.

Kyriakopoulou, L. et al. 2013. A sensitive and rapid mass spectrometric method for the simultaneous measurement of eight steroid hormones and CALIPER pediatric reference intervals. *Clin Biochem* 46(7-8), pp. 642-651.

Landon, J. et al. 1963. THE ADRENOCORTICAL RESPONSE TO INSULIN-INDUCED HYPOGLYCAEMIA. *J Endocrinol* 27, pp. 183-192.

le Roux, C. W. et al. 2002. Free cortisol index as a surrogate marker for serum free cortisol. *Ann Clin Biochem* 39(Pt 4), pp. 406-408.

Lebrethon, M. C. et al. 1994. Regulation of corticotropin receptor number and messenger RNA in cultured human adrenocortical cells by corticotropin and angiotensin II. *J Clin Invest* 93(4), pp. 1828-1833.

Lee, S. et al. 2014. Simultaneous determination of cortisol and cortisone from human serum by liquid chromatography-tandem mass spectrometry. *J Anal Methods Chem* 2014, p. 787483.

Lindholm, J. et al. 1978. Reliability of the 30-minute ACTH test in assessing hypothalamic-pituitary-adrenal function. *J Clin Endocrinol Metab* 47(2), pp. 272-274.

Lindstrom, L. et al. 1982. The measurement of serum cortisol by a solid-phase chemiluminescence immunoassay. *J Steroid Biochem* 16(4), pp. 577-580.

MacMahon, W. et al. 1983. A simplified ultrafiltration method for determination of serum free cortisol. *Clin Chim Acta* 131(3), pp. 171-184.

Mallat, J. 2012. High correlation between salivary cortisol and free serum cortisol measurements does not mean good agreement. *Chest* 141(1), pp. 273; author reply 273-274.

Mansoor, S. et al. 2007. Sixty-minute post-Synacthen serum cortisol level: a reliable and cost-effective screening test for excluding adrenal insufficiency compared to the conventional short Synacthen test. *Singapore Med J* 48(6), pp. 519-523.

Mattingly, D. 1962. A simple fluorimetric method for the estimation of free 11-hydroxycorticoids in human plasma. *J Clin Pathol* 15, pp. 374-379.

Mayenknecht, J. et al. 1998. Comparison of low and high dose corticotropin stimulation tests in patients with pituitary disease. *J Clin Endocrinol Metab* 83(5), pp. 1558-1562.

McConway, M. G. and Chapman, R. S. 1986. Development and evaluation of a simple, direct, solid-phase radioimmunoassay of serum cortisol from readily available reagents. *Clin Chim Acta* 158(1), pp. 59-70.

Meulenberg, P. M. et al. 1987. The effect of oral contraceptives on plasma-free and salivary cortisol and cortisone. *Clin Chim Acta* 165(2-3), pp. 379-385.

Mills, I. H. 1962. Transport and metabolism of steroids. *Br Med Bull* 18, pp. 127-133.

Moisey, R. et al. 2006. Interpretation of the short Synacthen test in the presence of low cortisol-binding globulin: two case reports. *Ann Clin Biochem* 43(Pt 5), pp. 416-419.

Molenaar, N. et al. 2015. Three calculations of free cortisol versus measured values in the critically ill. *Clin Biochem* 48(16-17), pp. 1053-1058.

Montsko, G. et al. 2014. A validated method for measurement of serum total, serum free, and salivary cortisol, using high-performance liquid chromatography coupled with high-resolution ESI-TOF mass spectrometry. *Anal Bioanal Chem* 406(9-10), pp. 2333-2341.

Morelli, V. et al. 2010. Subclinical hypercortisolism: correlation between biochemical diagnostic criteria and clinical aspects. *Clin Endocrinol (Oxf)* 73(2), pp. 161-166.

Nieman, L. K. et al. 2008. The diagnosis of Cushing's syndrome: an Endocrine Society Clinical Practice Guideline. *J Clin Endocrinol Metab* 93(5), pp. 1526-1540.

Oelkers, W. 1996. Adrenal insufficiency. *N Engl J Med* 335(16), pp. 1206-1212.

Owen, L. J. et al. 2013a. Development of a rapid assay for the analysis of serum cortisol and its implementation into a routine service laboratory. *Ann Clin Biochem* 50(Pt 4), pp. 345-352.

Owen, L. J. et al. 2013b. Is calibration the cause of variation in liquid chromatography tandem mass spectrometry testosterone measurement? *Ann Clin Biochem* 50(Pt 4), pp. 368-370.

Park, Y. J. et al. 1999. Reproducibility of the cortisol response to stimulation with the low dose (1 microg) of ACTH. *Clin Endocrinol (Oxf)* 51(2), pp. 153-158.

Patel, R. S. et al. 2004a. Production of gender-specific morning salivary cortisol reference intervals using internationally accepted procedures. *Clin Chem Lab Med* 42(12), pp. 1424-1429.

Patel, R. S. et al. 2004b. Morning salivary cortisol versus short Synacthen test as a test of adrenal suppression. *Ann Clin Biochem* 41(Pt 5), pp. 408-410.

Patel, R. S. et al. 2001. Preliminary results of a pilot study investigating the potential of salivary cortisol measurements to detect occult adrenal suppression secondary to steroid nose drops. *Clin Otolaryngol Allied Sci* 26(3), pp. 231-234.

Perogamvros, I. et al. 2011. Serum regulates cortisol bioactivity by corticosteroid-binding globulin-dependent and independent mechanisms, as revealed by combined bioassay and physicochemical assay approaches. *Clin Endocrinol (Oxf)* 75(1), pp. 31-38.

Perogamvros, I. et al. 2010a. Salivary cortisone is a potential biomarker for serum free cortisol. *J Clin Endocrinol Metab* 95(11), pp. 4951-4958.

Perogamvros, I. et al. 2010b. Measurement of salivary cortisol with liquid chromatography-tandem mass spectrometry in patients undergoing dynamic endocrine testing. *Clin Endocrinol (Oxf)* 72(1), pp. 17-21.

Perogamvros, I. et al. 2009. Simultaneous measurement of cortisol and cortisone in human saliva using liquid chromatography-tandem mass spectrometry: application in basal and stimulated conditions. *J Chromatogr B Analyt Technol Biomed Life Sci* 877(29), pp. 3771-3775.

Playfair, J. H. et al. 1974. Radioimmunoassay and saturation analysis. Production of antibodies and binding reagents. *Br Med Bull* 30(1), pp. 24-26.

Plumpton, F. S. and Besser, G. M. 1969. The adrenocortical response to surgery and insulin-induced hypoglycaemia in corticosteroid-treated and normal subjects. *Br J Surg* 56(3), pp. 216-219.

Poll, E. M. et al. 2013. Blunted serum and enhanced salivary free cortisol concentrations in the chronic phase after aneurysmal subarachnoid haemorrhage--is stress the culprit? *Stress* 16(2), pp. 153-162.

Poll, E. M. et al. 2007. Saliva collection method affects predictability of serum cortisol. *Clin Chim Acta* 382(1-2), pp. 15-19.

Poomthavorn, P. et al. 2009. Serum free cortisol index, free cortisol, and total cortisol in critically ill children. *Intensive Care Med* 35(7), pp. 1281-1285.

Pretorius, C. J. et al. 2011. Free cortisol method comparison: ultrafiltation, equilibrium dialysis, tracer dilution, tandem mass spectrometry and calculated free cortisol. *Clin Chim Acta* 412(11-12), pp. 1043-1047.

Qureshi, A. C. et al. 2007. The influence of the route of oestrogen administration on serum levels of cortisol-binding globulin and total cortisol. *Clin Endocrinol (Oxf)* 66(5), pp. 632-635.

Raff, H. et al. 2008. Cosyntropin-stimulated salivary cortisol in hospitalized patients with hypoproteinemia. *Endocrine* 34(1-3), pp. 68-74.

Raff, H. and Singh, R. J. 2012. Measurement of late-night salivary cortisol and cortisone by LC-MS/MS to assess preanalytical sample contamination with topical hydrocortisone. *Clin Chem* 58(5), pp. 947-948.

Rauschecker, M. et al. 2016. Cosyntropin-Stimulated Serum Free Cortisol in Healthy, Adrenally Insufficient, and Mildly Cirrhotic Populations. *J Clin Endocrinol Metab* 101(3), pp. 1075-1081.

Reinehr, T. et al. 2014. Relationships between 24-hour urinary free cortisol concentrations and metabolic syndrome in obese children. *J Clin Endocrinol Metab* 99(7), pp. 2391-2399.

Reynolds, R. M. et al. 2006. Assessing the HPA axis in patients with pituitary disease: a UK survey. *Clin Endocrinol (Oxf)* 64(1), pp. 82-85.

Riad-Fahmy, D. et al. 1982. Steroids in saliva for assessing endocrine function. *Endocr Rev* 3(4), pp. 367-395.

Ruiter, A. F. et al. 2012. Determination of unbound prednisolone, prednisone and cortisol in human serum and saliva by on-line solid-phase extraction liquid chromatography tandem mass spectrometry and potential implications for drug monitoring of prednisolone and prednisone in saliva. *Biomed Chromatogr* 26(7), pp. 789-796.

Schoorlemmer, R. M. et al. 2009. Relationships between cortisol level, mortality and chronic diseases in older persons. *Clin Endocrinol (Oxf)* 71(6), pp. 779-786.

Scott, L. V. and Dinan, T. G. 1998. Urinary free cortisol excretion in chronic fatigue syndrome, major depression and in healthy volunteers. *J Affect Disord* 47(1-3), pp. 49-54.

Scott, L. V. et al. 1998. Blunted adrenocorticotropin and cortisol responses to corticotropin-releasing hormone stimulation in chronic fatigue syndrome. *Acta Psychiatr Scand* 97(6), pp. 450-457.

Sesay, A. M. et al. 2013. Development of a competitive immunoassay for the determination of cortisol in human saliva. *Anal Biochem* 434(2), pp. 308-314.

Shah, A. et al. 1992. Hazards of pharmacological tests of growth hormone secretion in childhood. *Bmj* 304(6820), pp. 173-174.

Simunkova, K. et al. 2008. Comparison of total and salivary cortisol in a low-dose ACTH (Synacthen) test: influence of three-month oral contraceptives administration to healthy women. *Physiol Res* 57 Suppl 1, pp. S193-199.

Smith, R. E. et al. 1996. Localization of 11 beta-hydroxysteroid dehydrogenase type II in human epithelial tissues. *J Clin Endocrinol Metab* 81(9), pp. 3244-3248.

Sofer, Y. et al. 2016. GENDER DETERMINES SERUM FREE CORTISOL: HIGHER LEVELS IN MEN. *Endocr Pract* 22(12), pp. 1415-1421.

Stewart, P. M. et al. 1988. A rational approach for assessing the hypothalamo-pituitary-adrenal axis. *Lancet* 1(8596), pp. 1208-1210.

Stewart, P. M., Krone, N.P. 2011. The adrenal cortex. In: Melmed S, P.K.S., Larsen P.R, Kronenberg H.M ed. *Williams Textbook of Endocrinology*. Saunders, Elsevier Inc, pp. 479-544.

Stewart, P. M. et al. 1993. The hypothalamo-pituitary-adrenal axis across the normal menstrual cycle and in polycystic ovary syndrome. *Clin Endocrinol (Oxf)* 38(4), pp. 387-391.

Streeten, D. H. et al. 1996. The potential for serious consequences from misinterpreting normal responses to the rapid adrenocorticotropin test. *J Clin Endocrinol Metab* 81(1), pp. 285-290.

Tai, S. S. and Welch, M. J. 2004. Development and evaluation of a candidate reference method for the determination of total cortisol in human serum using isotope dilution liquid chromatography/mass spectrometry and liquid chromatography/tandem mass spectrometry. *Anal Chem* 76(4), pp. 1008-1014.

Tan, T. et al. 2010. Characterising adrenal function using directly measured plasma free cortisol in stable severe liver disease. *J Hepatol* 53(5), pp. 841-848.

Taylor, P. J. 2005. Matrix effects: the Achilles heel of quantitative high-performance liquid chromatography-electrospray-tandem mass spectrometry. *Clin Biochem* 38(4), pp. 328-334.

Thaler, L. M. and Blevins, L. S., Jr. 1998. The low dose (1-microg) adrenocorticotropin stimulation test in the evaluation of patients with suspected central adrenal insufficiency. *J Clin Endocrinol Metab* 83(8), pp. 2726-2729.

Thevenot, T. et al. 2011. Assessment of adrenal function in cirrhotic patients using concentration of serum-free and salivary cortisol. *Liver Int* 31(3), pp. 425-433.

Veldhuis, J. D. et al. 2009. Basal, pulsatile, entropic (patterned), and spiky (staccato-like) properties of ACTH secretion: impact of age, gender, and body mass index. *J Clin Endocrinol Metab* 94(10), pp. 4045-4052.

Vincent, R. P. et al. 2009. Serum total cortisol and free cortisol index give different information regarding the hypothalamus-pituitary-adrenal axis reserve in patients with liver impairment. *Ann Clin Biochem* 46(Pt 6), pp. 505-507.

Vining, R. F. and McGinley, R. A. 1987. The measurement of hormones in saliva: possibilities and pitfalls. *J Steroid Biochem* 27(1-3), pp. 81-94.

Vining, R. F. et al. 1983. Hormones in saliva: mode of entry and consequent implications for clinical interpretation. *Clin Chem* 29(10), pp. 1752-1756.

Vogeser, M. et al. 2002. Dialyzable free cortisol after stimulation with Synacthen. *Clin Biochem* 35(7), pp. 539-543.

Walker, R. F. et al. 1978. Adrenal status assessed by direct radioimmunoassay of cortisol in whole saliva or parotid saliva. *Clin Chem* 24(9), pp. 1460-1463.

Weintrob, N. et al. 1998. Standard and low-dose short adrenocorticotropin test compared with insulin-induced hypoglycemia for assessment of the hypothalamic-pituitary-adrenal axis in children with idiopathic multiple pituitary hormone deficiencies. *J Clin Endocrinol Metab* 83(1), pp. 88-92.

Wong, V. et al. 2004. Saliva and bloodspot cortisol: novel sampling methods to assess hydrocortisone replacement therapy in hypoadrenal patients. *Clin Endocrinol (Oxf)* 61(1), pp. 131-137.

Wood, J. B. et al. 1965. A RAPID TEST OF ADRENOCORTICAL FUNCTION. *Lancet* 1(7379), pp. 243-245.

Yates, A. M. et al. 2008. Interlaboratory variation in 25-hydroxyvitamin D2 and 25-hydroxyvitamin D3 is significantly improved if common calibration material is used. *Clin Chem* 54(12), pp. 2082-2084.

Zerikly, R. K. et al. 2010. Diagnostic characteristics of late-night salivary cortisol using liquid chromatography-tandem mass spectrometry. *J Clin Endocrinol Metab* 95(10), pp. 4555-4559.

APPENDICES

Appendix 1 – Recruitment Poster:

Aged 18 -80 and in good health?

We are looking for volunteers to undergo a 30 minute test.

This will involve blood sampling before and after being given a synthetic hormone.

The information we obtain will be used to derive clinical laboratory reference ranges which will help us to diagnose adrenal disease.

If you are interested or would like more information please telephone:
Dr Nadia EI-Farhan (UHW Ext. 8346),
Sister Nikki Davies (UHW Ext. 3495)
or Dr Aled Rees (UHW Ext. 2309)

Thank you

Appendix 2 – Volunteer information sheet

VOLUNTEER INFORMATION SHEET
(Version 4 October 2008)

1. The Study

Determination of method-specific normal cortisol and adrenal hormone responses to the short Synacthen test.

2. Introduction

You are being invited to take part in a clinical research study. Before you decide whether you wish to become involved it is important that you understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with others if you wish. Please do not hesitate to ask us if there is anything that is not clear or if you would like more information. Take time to consider whether or not you would wish to take part.

Thank you for reading this.

3. What is the purpose of this study?

Your body contains two glands, known as the adrenal glands, which produce a hormone called cortisol. Cortisol has many roles in the body including regulating blood pressure and blood sugar levels, keeping your skin and bones healthy and strengthening your immune system. The Synacthen test is used to assess the adrenal gland. In this test, patients are given a synthetic hormone, which stimulates the gland to produce cortisol. A blood sample is taken before and after Synacthen is given and the amount of cortisol produced is measured on both samples. The level of cortisol in the second sample is used to diagnose an under-active adrenal gland. However, we do not yet know for certain what absolute level of cortisol we should use to rule out abnormal adrenal glands. One of the purposes of this study is to define the normal cortisol response to Synacthen in healthy volunteers.

Another purpose of this study is to establish whether the cortisol level used to diagnose under-active adrenal glands is the same no matter where and how it is measured. Currently in the United Kingdom there are several different ways of measuring cortisol. These all rely on the same principle but use different machines to make the measurement. One study,

carried out 10 years ago, showed that the different ways of measuring cortisol did give significantly different cortisol levels, and suggested that the level of cortisol needed to rule out under-active adrenal glands should depend on the method used to measure cortisol. Our laboratory now uses a new method and as a result we do not know for sure whether the cortisol level we rely on to make the diagnosis of adrenal under-activity is valid. We hope to establish the appropriate levels of cortisol for use in our laboratory as well as for some of the other common cortisol methods in use in the United Kingdom today.

We will also be looking at the normal salivary cortisol response to Synacthen as recent work has suggested that measuring cortisol in saliva may be better than measuring it in blood. Patients might prefer this way too as it would mean they would not have to give blood.

Finally, in women under the age of 40 we will also measure some of the hormones that precede the formation of cortisol (17-hydroxyprogesterone,11-deoxycortisone, androstenedione and testosterone). In a rare condition known as congenital adrenal hyperplasia one of the steps in cortisol synthesis is blocked resulting in low cortisol and raised levels of these other hormones. We hope to establish a reference range for these cortisol precursors in normal individuals.

4. Why have I been chosen?

You have been chosen for this study because you are fit and healthy with no medical conditions that prevent you from undergoing a short Synacthen test.

5. Do I have to take part?

It is up to you to decide whether or not to take part. If you decide to take part you will be given this information sheet to keep and be asked to sign a consent form. If you decide to take part you are still free to withdraw at any time and without giving a reason. A decision to withdraw at any time, or a decision not to take part, will not affect the standard of medical care you receive or your legal rights in any way.

6. What will happen to me if I do take part?

You will be asked to attend the clinical research facility/endocrine unit for a total of 90 minutes for this test to be carried out.
On arrival you will have your pulse and blood pressure measured and a thorough inspection of the inside of your mouth will be made. You will be given a small pot and asked

to provide a sample of saliva (spit). An intravenous cannula (drip) will be inserted into your arm. You will have the option of a small local anaesthetic injection before this if you prefer. We will take 20 ml of blood (approximately 4 teaspoonfuls) from the cannula and then inject 1 ml of Synacthen. This will be washed through with 10 ml of salty water. 30 minutes later a further 20 ml of blood (approximately 4 teaspoonfuls) will be removed from the cannula. This will not require a further injection. The cannula will then be removed. You will also be asked to provide a second sample of saliva at this time.
30 minutes later a further 10ml of blood (approximately 2 teaspoonfuls) will be collected. This will require a further needle-prick.
You will be asked to remain on the unit for a further 30 minutes. At the end of this time, your pulse and blood pressure will be measured again and a repeat inspection of the inside of your mouth will be made.
One week after your short synacthen test you will be telephoned at home by the person who carried out the test, or a colleague also involved in the study, to ensure you are well and have not experienced any side effects.

7. What if my response to Synacthen is abnormal?

It is very rare to have under-active adrenal glands. However, if your cortisol response to Synacthen is lower than expected, we will contact you by telephone to let you know. We will then write to your GP to inform him/her of your result and we will arrange for you to be seen in the Endocrine clinic for further investigation and treatment.

8. What do I have to do?

You will not need to be fasting and there is no specific preparation for this test. However, you will need to refrain from smoking for the 30 minutes before the test and during the test itself.

9. What is Synacthen?

Your brain produces a hormone known as adrenocorticotrophin (ACTH). This hormone stimulates the adrenal glands to produce cortisol. Synacthen is a man-made equivalent of ACTH and acts in a similar way to your body's own hormone, i.e. stimulates the adrenal glands to produce cortisol.

10. What are the side effects of Synacthen?

Synacthen is a very safe drug with very few side-effects. However, as it is a man-made hormone there is a small risk

that you may be allergic to it. If this occurs, you may experience marked redness, swelling and itching around the injection site. In more severe cases you may feel light-headed and dizzy and experience nausea and/or vomiting. You may also notice a generalised itching with flushing and swelling. Very rarely, a full allergic reaction may develop with swelling of the lips and airways and shortness of breath. Because of this risk, all necessary treatments for an allergic reaction will be available on the unit for the duration of your visit.
Although allergy to Synacthen has been reported, it is very rare in people who are not allergic to other things and who do not have asthma. If you do have asthma, suffer from allergies or have previously had an allergic reaction to Synacthen, you must tell the doctor or nurse looking after you, as this will prevent you from taking part in this study.

11. What are the possible disadvantages and risks of taking part?

There is a very small risk of an allergic reaction to Synacthen. This is described above. If you do develop dizziness or become light-headed, you will be advised not to drive or operate machinery until this has passed. You may experience possible discomfort (temporary pain, swelling, bruising and rarely infection) from the insertion of the venous cannula and withdrawal of blood. No other side effects are anticipated from the study procedures.

The risks to an unborn human foetus or a breastfed child from Synacthen are not fully known. Because of this, women who are pregnant or breastfeeding a child may not participate in this trial. All women of child-bearing age will be asked to have a pregnancy test before taking part to exclude the possibility of pregnancy.

12. What are the potential benefits of taking part?

You are unlikely to benefit directly by taking part in this study. However, the information we get will ensure that the short synacthen test used in diagnosing under-active adrenal glands is reliable and accurate.

13. What if new information becomes available?

Sometimes during the course of a research project, new information becomes available about the drug that is being studied. If this happens, your research doctor will tell you about it and discuss with you whether you want to continue in the study. If you decide to withdraw, your research doctor will make arrangements for your care to continue. If you decide to

continue in the study you will be asked to sign an updated consent form.

Also, on receiving new information your research doctor might consider it to be in your best interests to withdraw you from the study. He/she will explain the reasons and arrange for your care to continue.

14. What if something goes wrong?

The study is indemnified by Cardiff University and if you are harmed by taking part in this research project, compensation will be available.
If you are harmed due to someone's negligence, then you may have grounds for a legal action but you may have to pay for it. Regardless of this, if you wish to complain, or have any concerns about any aspect of the way you have been approached or treated during the course of this study, the normal National Health Service complaints mechanisms should be available to you.

15. Will my taking part in this study be kept confidential?

All information which is collected about you during the course of the research will be kept strictly confidential. Any information about you which leaves the hospital will have your name and address removed so that you cannot be recognised from it. With your permission your GP will be informed of your participation in this study. With your permission we may also look at sections of your medical notes which are relevant to the research study.

16. What will happen to the samples I give?

The blood and saliva samples taken from you for this study will be analysed and then stored for up to 2 months in the department of medical biochemistry. They will then be destroyed by incineration according to standard clinical practice.

17. What will happen to the results of the research study?

The results of the research study will be prepared for publication in appropriate medical journals together with presentation at medical conferences. They will also be submitted to the Royal College of Pathologists as part of Dr Nadia El-Farhan's final dissertation. Patients participating in the study will be able to obtain a copy of the results after they have been published in the relevant journal(s). Patients will not be identified in any report/publication.

18. Who is organising and funding the research?

The study is being organised by Dr Aled Rees in the Section of Endocrinology (the Principal Investigator) Dr Carol Evans (Principal Biochemist) and Dr Nadia El-Farhan (Specialist Registrar) from the department of Medical Biochemistry. Funding for the study is provided from the Department of Medical Biochemistry and the Section of Endocrinology in the Heath hospital. The doctors conducting the research are not being paid for recruiting and looking after patients in the study.

19. Who has reviewed the study?

The study has been reviewed by the Cardiff and Vale NHS Trust/Cardiff University peer and risk review committee and by the South East Wales Local Research Ethics Committee.

20. Contact for further information

Should you have any further questions or queries regarding this research study, then please do not hesitate to contact me on 029 20748346 or Dr Aled Rees on 029 20745002 during working hours. If anything arises outside of normal working hours then I can be contacted directly on (07889) 159792.

Thank you for taking part in this study.

Dr Nadia El-Farhan
Specialist Registrar in Chemical Pathology (Metabolic Medicine)

You will be given a copy of this Patient Information Sheet and a signed consent form to keep.

Appendix 3 – Study protocol

Determination of method-specific normal cortisol and adrenal hormone reponses to the short Synacthen® test

Short Title: Short Synacthen® test Study

Trust reference 06/DTD/3791E

EudraCT number 2007-000056-14

Healthcare Professionals involved in the study

Name	Signature	Date
Dr Nadia El-Farhan		
Dr Carol Evans		
Dr Aled Rees		
Dr Steve Davies		
Prof Maurice Scanlon		
Sister Niki Davies		
Sister Janet Lewis		
Dr Atilla Turkes		
Mr David Ducroq		
Mr Alan Pickett		
Mr Robert Henley		
Mr Gareth Jones		

Contents Page

List of Abbreviations

ACTH	adrenocorticotropic hormone
CBG	cortisol binding globulin
17OHP	17 Hydroxyprogesterone
UHW	University Hospital of Wales

1. BRIEF SUMMARY

The Synacthen® test is used to assess adrenal sufficiency. The adrenal steroid hormone, cortisol, is measured before and after injection of Synacthen®. In normal individuals serum cortisol rises to an arbitrary value (or cut off). Method dependent cut-offs for the assays currently in use in clinical laboratories in the UK are not available. We will perform Synacthen® tests in 120 healthy subjects (60 female and 60 male) aged 18-80. Since oestrogens increase cortisol levels we will also assess 60 females taking oral contraceptive pills containing oestrogen. We will derive method dependent cut-offs for the serum and saliva cortisol response and compare serum cortisol methods.

Deficiency of enzymes needed for cortisol biosynthesis cause congenital adrenal hyperplasia. This is characterised by increased serum concentrations of 17 Hydroxyprogesterone (17OHP), an intermediate in the pathway. In mild cases, measurement of 17OHP after Synacthen® stimulation can be used to assist in diagnosis. As a secondary objective, we will derive a normal 17OHP response using serum collected as part of the Synacthen® test study.

2. BACKGROUND

Short Synacthen® test cortisol response

Synacthen® is a synthetic analogue of ACTH which has been used since the 1960s to assess adrenal sufficiency. It is now well established as a first line test to investigate diseases of the hypothalamo-pituitary-adrenal axis and to assess adrenal function in patients on long-term corticosteroid therapy. Briefly, cortisol is measured before and after injection of 250 micrograms of Synacthen®. In a normal individual serum cortisol will rise to concentrations greater than an arbitrary value (typically 550 nmol/l) 30 minutes after administration of Synacthen®.

In 2004 the All Wales Clinical Biochemistry Audit group surveyed protocols for performing and interpreting short Synacthen® tests (1). This identified wide differences in practice within Wales. As a result standards were drawn up for performance of the test. It was noted that there was considerable variability or bias between cortisol immunoassays and that the cortisol cut-off chosen for interpretation of the short Synacthen® test should be method dependent.

Clark et al., in 1998 reported cortisol cut-offs following Synacthen® using 4 well established commercially available cortisol immunoassays (2). This study demonstrated considerable differences between the cortisol immunoassays used in clinical laboratories at the time. It was also apparent that there were differences in gender-related responses to Synacthen® although there was no dependence on age. In the 8 years since publication of this study there have been advances in formulation of cortisol immunoassays as well as the instrumentation used to perform analyses. At UHW cortisol is currently assayed using the Bayer Centaur automated immunoassay analyser. This assay was not available at the time of the study by Clark et al.,. Our current short Synacthen® test cut-offs therefore rely on historical reference ranges

which have become outdated. A re-evaluation of the cortisol cut-off is required to ensure that patients are not incorrectly classified.

It has been long been recognised that oestrogens (including ethinyloestradiol prescribed in combined oral contraceptive pills) increase total (but not free) serum cortisol levels. The degree of increase is related to the dose used and is thought to be due to an elevation in cortisol binding globulin (CBG) (3). Although the rise in free serum cortisol to supra-physiological doses of ACTH (i.e. the Synacthen® test) is unlikely to be affected by concomitant oestrogen therapy, no comparisons of total serum cortisol in response to Synacthen® have been performed between women taking oestrogens and those who are not. Knowledge of the salivary cortisol response may also be useful in patients with decreased serum CBG concentrations e.g. severe nephrotic syndrome in whom the serum cortisol response may be misleading (4). We therefore plan to measure salivary cortisol as part of our study protocol to assess the response of free cortisol.

Short Synacthen® test 17 hydroxyprogesterone response

17 Hydroxyprogesterone (17OHP) is an intermediate in the biosynthesis of cortisol. Deficiency of 21-hydroxylase enzyme activity leads to an increased concentration of 17OHP in the peripheral circulation. The short Synacthen® test can be used to assist in diagnosis of mild cases of congenital adrenal hyperplasia. Current reference ranges are taken from the literature.

3. SUMMARY OF SYNACTHEN® CHARACTERISTICS

Name of the Medicinal Product	®Ampoules 250mcg
Qualitative and Quantitative Composition	Tetracosactide acetate PhEur 250micrograms per ampoule.
Pharmaceutical form	A clear colourless sterile solution in a clear glass ampoule.
Therapeutic indications	Diagnostic test for the investigation of adrenocortical insufficiency.
Posology and method of administration	*Adults*: This preparation of Synacthen® is intended for administration for diagnostic purposes only as a single intramuscular or intravenous dose; it is not to be used for repeated therapeutic administration. The 30-minute Synacthen® diagnostic test: This test is based on measurement of the plasma cortisol concentration immediately before and exactly 30 minutes after an intramuscular or intravenous injection of 250 micrograms (1ml) Synacthen®. *Elderly*: There is no evidence to suggest that dosage should be different in the elderly.
Contraindications	History of hypersensitivity to ACTH, Synacthen® or Synacthen® Depot. Synacthen® is contra-indicated in patients with allergic disorders (e.g. asthma).
Special warnings and precautions for use	Before using Synacthen®, the doctor should make every effort to find out whether the patient is suffering from, or has a history of, allergic disorders (see Section 4.3 “Contra-indications”). In particular, he should enquire whether the patient has previously experienced adverse reactions to ACTH, Synacthen® or other drugs. Synacthen® should only be administered under the supervision of appropriate senior hospital medical staff (e.g. consultants). If local or systemic hypersensitivity reactions occur after the injection (for example, marked redness and pain at the injection site, urticaria, pruritus, flushing, faintness or dyspnoea), Synacthen® or other ACTH preparations should be avoided in the future. Hypersensitivity reactions tend to occur within 30 minutes of an injection. The patient should therefore be kept under observation during this time. Preparation should be made in advance to combat any anaphylactic reaction that may occur after an injection of Synacthen®. In the event of a serious anaphylactic reaction occurring, the following measures must be taken immediately: administer adrenaline (0.4 to 1 ml of a 0.1% solution intramuscularly or 0.1 to 0.2 ml of a 0.1% solution in 10 ml physiological saline slowly intravenously) as well as a large intravenous dose of a corticosteroid (for example 100 mg to 500 mg hydrocortisone, three or four times in 24 hours), repeating the dose if necessary. The hydrocortisone product information prepared by the manufacturer should also be consulted.

Interaction with other medicinal products and other forms of interaction	None known.
Pregnancy and lactation	The Synacthen® test should not be utilised during pregnancy and lactation unless there are compelling reasons for doing so.
Effects on ability to drive and use machines	Patients should be warned of the potential hazards of driving or operating machinery if they experience side effects such as dizziness.
Undesirable effects	*Hypersensitivity reactions:* Synacthen® may provoke hypersensitivity reactions. In patients suffering from, or susceptible to, allergic disorders (especially asthma) this may take the form of anaphylactic shock (see Section "Contra-indications"). Hypersensitivity may be manifested as skin reactions at the injection site, dizziness, nausea, vomiting, urticaria, pruritus, flushing, malaise, dyspnoea, angioneurotic oedema and Quinke's oedema. Other side effects are unlikely to be observed with short-term use of Synacthen® as a diagnostic tool. Should information be required on the side effects reported with therapeutic use of tetracosactide acetate, see Synacthen® Depot Summary of Product Characteristics.
Overdose	Overdosage is unlikely to be a problem when the product is used as a single dose for diagnostic purposes.
Pharmacodynamic properties	Tetracosactide acetate consists of the first 24 amino acids occurring in the natural corticotropic hormone (ACTH) sequence and displays the same physiological properties as ACTH. In the adrenal cortex, it stimulates the biosynthesis of glucocorticoids, mineralocorticoids, and, to a lesser extent androgens. The site of action of ACTH is the plasma membrane of the adrenocortical cells, where it binds to a specific receptor. The hormone-receptor complex activates adenylate cyclase, stimulating the production of cyclic AMP (adenosine monophosphate) and so promoting the synthesis of pregnenolone from cholesterol. From pregnenolone the various corticosteroids are produced via different enzymatic pathways.
Pharmacokinetic properties	Following an intravenous injection, elimination of tetracosactide acetate from the plasma consists of 3 phases. The half-lives of these phases are approximately 7 minutes (0 to 1 hour), 37 minutes (1 to 2 hours) and 3 hours thereafter. Tetracosactide acetate has an apparent volume of distribution of approximately 0.4L/kg. In the serum, tetracosactide acetate is broken down by serum endopeptidases into inactive

	oligopeptides and then by aminopeptidases into free amino acids. The rapid elimination from plasma is probably not attributable to this relatively slow cleavage process, but rather to the rapid concentration of the active substance in the adrenal glands and kidneys. Following an iv dose of ^{131}I-labelled tetracosactide acetate, 95 to 100% of the radioactivity is excreted in the urine within 24 hours.
Preclinical safety data	There are no pre-clinical data of relevance to the prescriber, which are additional to those already included in other sections of the Summary of Product Characteristics.
List of excipients	Acetic acid, sodium acetate, sodium chloride and water.
Incompatibilities	None known.
Shelf life	5 years.
Special precautions for storage	Synacthen® should be protected from light and stored in a refrigerator (2 - 8°C).
Nature and contents of container	The ampoules are colourless glass PhEur type I. Five ampoules are packed in a cardboard box.
Instructions for use, handling and disposal	Shake well before use.
Marketing authorisation holder	Alliance Pharmaceuticals Ltd Avonbridge House Bath Road Chippenham Wiltshire SN15 2BB
Marketing Authorisation number(s)	PL 16853/0017
Date of first authorisation/renewal of the authorisation	25 June 1998
Date of revision of the text	February 2005
Legal Status	Alliance, Alliance Pharmaceuticals and associated devices are registered Trademarks of Alliance Pharmaceuticals Ltd.

4. TRIAL OBJECTIVES AND PURPOSE

1. To define the cortisol response to Synacthen® in normal volunteers using commercially available cortisol immunoassays in use in clinical laboratories in the UK.
2. To compare cortisol results against the gold standard method gas chromatography - mass spectrometry.
3. To investigate the effect of oral contraceptives containing ethinyloestradiol on the cortisol cut-off in response to Synacthen®.
4. To define the salivary cortisol response to Synacthen® in normal volunteers
5. To define the 17OHP response to Synacthen® in normal female volunteers.

5. TRIAL DESIGN

End-points

1. The primary end-point of the study will be to establish method dependent cortisol cut offs for the normal response to Synacthen® using the 5th percentile.
2. Cortisol measurements by immunoassay will be compared with the GC-MS gold standard method for normal volunteers and patients with hypopituitarism and hypoadrenalism.
3. Synacthen® responses in women taking ethinyloestradiol-containing contraceptive pills will be compared with those who are not.
4. We will establish cut offs for the salivary cortisol response to Synacthen® in normal volunteers using the 5th percentile.
5. We will establish a 17OHP cut off in response to Synacthen® in normal female volunteers using the 5th percentile.

Subjects

We aim to recruit
120 normal volunteers (60 male and 60 female).
60 females taking an ethinyloestradiol-containing oral contraceptive (details of the preparation to be noted).
60 patients with hypoadrenalism or hypopituitarism having a short Synacthen® test as part of their clinical care.

Short Synacthen® Test Procedure

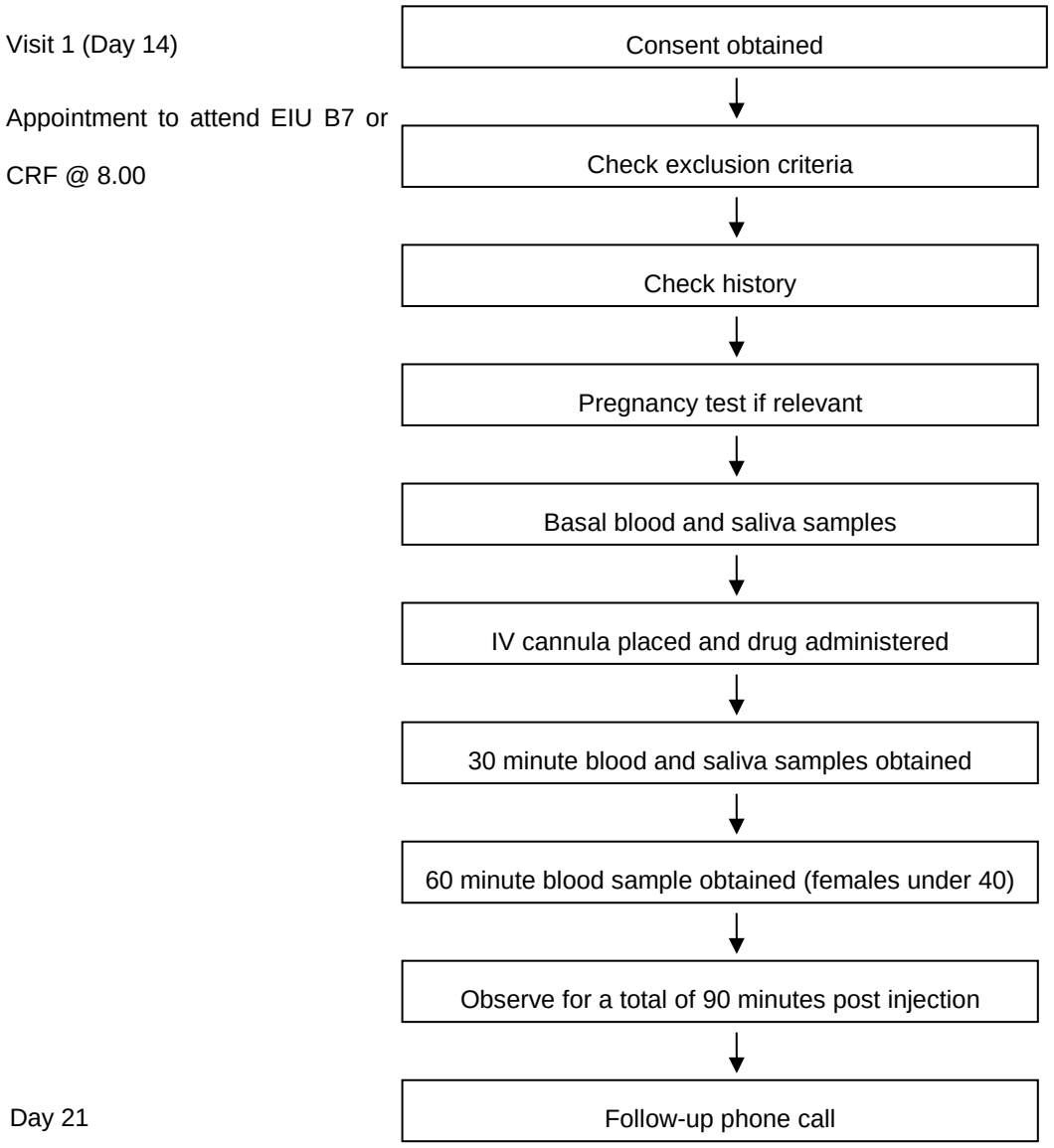
1. The short Synacthen® test will be performed by experienced nursing staff (Sisters Nikki Davies and Janet Lewis) in the Endocrine Investigation Unit on B7. The test will be performed between 9 and 10 am. No smoking is allowed in the 30 minutes prior to or during the test. The patient should remain at rest. Subjects are not required to be fasting and can drink freely during the test. The subject should be asked to rinse their mouth with water 15 minutes before commencing the test.
2. Pulse, blood pressure and respiratory rate as indicators for any possible anaphylaxis in addition to clinical inspection of the mouth to look for signs of oedema will be recorded before the test.
3. Synacthen® testing has been very rarely associated with an anaphylactic response (none recorded at UHW in over 20 years). Full resuscitation facilities are available on the ward. Any adverse drug reactions would be reported in the standard manner.
4. Collection of the basal saliva sample should be done prior to collection of blood or injection by asking the subject to passively drool into a universal container (5mL or 1 teaspoon of saliva should be collected).
5. Subjects will be offered a subcutaneous local anaesthetic injection (lidocaine 0.5%) before insertion of the IV cannula.
6. Collect 20 ml blood into plain (red top) vacutainers (no anticoagulant and no gel). Label the tube with the collection time.
7. Inject Synacthen® (250 µg in 1 ml) intravenously.
8. Collect 20 ml blood into plain vacutainers (no anticoagulant and no gel) at 30 minutes after the injection.
9. Saliva should also be collected 30 minutes after administration of Synacthen® as previously described.
10. For female volunteers under the age of 40, collect a further 10 ml of blood into plain vacutainers (no anticoagulant and no gel) 60 minutes after the injection.

11. Ensure each tube is carefully labelled, including the collection time. Send the samples to the laboratory with 1 request form, stating the analyses required.
12. Observe subjects for a total of 90 minutes after the Synacthen injection. Monitor pulse, blood pressure and respiratory rate as indicators for any possible anaphylaxis in addition to clinical inspection of the mouth to look for signs of oedema before allowing patient to leave.

Synacthen® Supply, Storage, Packaging and Labelling

On receipt of a requisition form the pharmacy clinical trials department will supply the investigators with sufficient stock of Tetracosactide 250 mcg/ml injection (Synacthen® manufactured by Alliance – 5 ampoules per box). Each box will be labelled 'For Clinical Trial Use Only' and with 'For use in Synacthen® project ID 2007-000056-14'. These boxes of Synacthen® should be stored separately from usual B7 Endocrine unit stock of Synacthen®. The drug will be stored under recommended conditions (between 2ºC and 8ºC). The pharmacy will also prepare an accountability form for the investigators to complete whenever they administer a dose of Synacthen® to study volunteers. This accountability form will enable recording of date of administration, subject name and/or number, batch number and expiry date of ampoule administered.

Flow chart of procedure



Biochemical Analyses

Serum cortisol will be measured on the basal and 30 minute serum samples using the Bayer Centaur automated immunoassay analyser on the day of sampling in the Department of Medical Biochemistry UHW. Further aliquots of serum will be stored at -50ºC pending batched analysis of cortisol by various methods. Cortisol will be assayed using the Roche E170 automated immunoassay analyser, the Abbott automated immunoassay analyser, Immulite automated immunoassay analyser, Immulite 2000 automated immunoassay analyser, Beckmann Access automated immunoassay analyser, an in-house tandem mass spectrometry assay and the gold standard gas chromatography - mass spectrometry. A small number of samples (n=25) will be re-measured using the Bayer Centaur cortisol assay to ensure that there has been no degradation of serum cortisol during storage.

Salivary cortisol will be measured using a commercially available immunoassay and by tandem MS.

17OHP will be measured in basal, 30 and 60 minute samples taken from 60 women under the age of 40 using an in-house radioimmunoassay. In addition cortisol, 17 OHP, 11 deoxycortisol, 21 deoxycortsiol, androstenedione and testosterone will be measured by tandem MS.

6. STUDY POPULATION

Subject recruitment

Volunteers will be recruited from Staff in the Directorates of Medicine and Pathology in the University Hospital of Wales and Medical Students. Posters will be displayed to invite participation. The Welsh Blood Service will be asked to display posters at their donation sessions and an investigator will attend these sessions to provide further information to any interested individuals.

Patients with hypoadrenalism or hypopituitarism will be recruited from patients under the care of Professor M.F. Scanlon, Dr J.S. Davies or Dr D.A. Rees having a short Synacthen® test as part of their clinical care. Between 5 and 10 such patients undergo short synacthen tests per week in the Endocrine Investigation Unit.
Potential participants (those with suspected primary or secondary adrenal insufficiency) will initially be identified from the Endocrinology analyser database at the University Hospital of Wales. Eligible subjects will be invited to participate through a letter from the clinician responsible for their care (Dr J S Davies, Professor M F Scanlon, Dr D A Rees). Potential participants will also be identified as they pass through the Endocrine clinic systems at UHW (Professor M F Scanlon, Dr J S Davies, Dr D A Rees) and will be invited to participate by the clinician responsible for their care. All patients with adrenal insufficiency identified by means of the synacthen test will be treated and followed up in the Endocrinology clinic according to standard clinical practice.

Screening log

Investigators will maintain a record of all volunteers who were considered for the study but who were not enrolled. The reason for exclusion will be recorded.
Investigators will also maintain a record of all volunteers who were enrolled onto the study (*i.e.* who signed the consent form). In the event that the test was not performed the reason will be recorded.

It is planned to recruit 180 healthy volunteers and 60 patients with hypoadrenalism or pituitary insufficiency.

Subject inclusion criteria
Volunteers will be in self-proclaimed good health. Subjects will be free of illness on the day of testing and will not be taking drug therapy.
Female subjects under the age of 40 will undergo the test in the follicular phase of the menstrual cycle. Ideally in the first five days following menstruation.

Females taking an ethinyloestradiol-containing oral contraceptive will be recruited. Details of the preparation will be noted.

Subject exclusion criteria
Subjects will not be included in the study if he/she:

1) Is pregnant or lactating. Females of childbearing potential must have a negative pregnancy test before enrollment onto the study. Non-child bearing potential is defined as post-menopausal for at least 1 year, surgical sterilisation or hysterectomy at least three months before the start of the study,
2) Is using corticosteroids,
3) has any significant intercurrent disease,
4) has a history of thyroid or other autoimmune disease,
5) has a previous history of hypersensitivity to Synacthen®,
6) has a previous history of asthma
7) has a history of allergic disorder
8) has any mental condition rendering the patient unable to understand the nature or possible consequences of the study, and/or evidence of an uncooperative attitude.

Patient inclusion criteria
Patients will be free of intercurrent illness on the day of testing.

Patient exclusion criteria
Subjects being investigated for possible adrenal insufficiency will not be included in the study if he/she:
1) Is pregnant or lactating. Females of childbearing potential must have a negative pregnancy test before enrolment onto the study. Non-child bearing potential is defined as post-menopausal for at least 1 year, surgical sterilisation or hysterectomy at least three months before the start of the study.
2) Is using corticosteroids,
3) has a previous history of hypersensitivity to Synacthen®,
4) has a history of asthma or allergic disorder
5) has any mental condition rendering the patient unable to understand the nature or possible consequences of the study, and/or evidence of an uncooperative attitude.

Withdrawal of subjects
Subjects will be free to withdraw from the study at any time without giving reasons. Where reasons are given these will be recorded on a case report form (CRF).
Recruitment will continue until the agreed number of participants has been achieved.

7. TREATMENT OF SUBJECTS

Consent

Potential participants will initially be identified by poster advertisement at the University Hospital of Wales. The volunteer information sheet will explain the nature of the study, its purpose, the procedures involved, the expected duration, the potential risks and benefits involved and any discomfort it may entail. Eligible subjects will be invited to participate by Dr Aled Rees, Dr Carol Evans, Dr Nadia El-Farhan, Sister Nikki Davies or Sister Janet Lewis. A subject screening log will be used to record all contacts (Appendix 1). Each subject will be informed that participation in the study is voluntary and that they may withdraw from the study at any time. Participants will be given adequate time to review the volunteer information sheet and will have sufficient opportunity to understand the objectives, risks and inconveniences of the study and the conditions under which it is to be conducted. The following individuals will be permitted to take consent:

Dr Aled Rees
Dr Nadia El-Farhan
Sister Nikki Davies
Sister Janet Lewis

This informed consent will be given by means of a standard written statement, written in non-technical language. The subject should read and consider the statement before signing and dating it, and will be given a copy of the signed document. No subject can enter the study before his/her informed consent has been obtained. Participation will be recorded in a subject enrolment log (Appendix 2). Participants will be assigned a subject number to ensure anonymity.

Patient care
With their consent we will write to the patients GPs to inform them that they are participating in this trial.
Although subjects by definition will be healthy it is possible that some subjects will have clinically unsuspected and significant adrenal suppression. If this is identified then they will be offered review by Dr Rees.

Follow up of subjects and patients
A follow up phone call will be made by Sister Nikki Davies, Sister Janet Lewis, Dr Aled Rees or Dr Nadia El-Farhan.

8. ASSESSMENT OF SAFETY

Expected adverse reactions:
Hypersensitivity reactions: Synacthen® may provoke hypersensitivity reactions. In patients suffering from, or susceptible to, allergic disorders (especially asthma) this may take the form of anaphylactic shock (see Section 3 “Contra-indications”).
Hypersensitivity may be manifested as skin reactions at the injection site, dizziness, nausea, vomiting, urticaria, pruritus, flushing, malaise, dyspnoea, angioneurotic oedema and Quinke's oedema. Any increased incidence of any adverse reaction (expected or unexpected) will be discussed by the trial management group and a decision taken to terminate the study where appropriate.

Instructions on the reporting procedures for serious adverse events
Information about all non-serious and serious adverse events (SAEs), irrespective of causality, whether volunteered by the subject, discovered by investigator questioning, or detected through physical examination, laboratory test or other means, will be collected and recorded in the case report form. Medical conditions/diseases present before starting the study are only considered adverse events if they worsen after the

study procedures. Abnormal laboratory values or test results constitute adverse events only if they induce clinical signs or symptoms or require therapy, or are considered clinically significant for any reason, in which case they will be recorded on a case report form (appendix 3 for subjects, appendix 4 for patients).

Where possible, each adverse event will also be described by:
1. Its duration
2. The severity grade (mild, moderate, severe)
3. The action(s) taken

Serious adverse events
Information about all SAEs will be collected and recorded. All SAEs will be reported in accordance with Cardiff University's Policy and Procedure for Reporting Research Related Adverse Events.
A SAE is an undesirable sign, symptom or medical condition which:
1. is fatal or life-threatening
2. required or prolonged hospitalisation
3. was significantly or permanently disabling or incapacitating
4. constitutes a congenital anomaly or a birth defect
5. are medically significant, may jeopardise the subject and may require medical or surgical intervention to prevent one of the outcomes listed above.

Any SAE, irrespective of causality, occurring in a patient after providing informed consent and until one week after ending study participation will be recorded.

The Principal Investigator, Dr Rees, will report all SAEs/SUSARs to the Sponsor within 24 hours of being made aware of the event. The immediate report may be made orally or in writing and shall be promptly followed by a detailed written report. This report will be sent to the Sponsor via email or facsimile. Where email is used, this must be followed up with a faxed signed version. The Principal Investigator must report all SAEs/SUSARs using the SAE Report Form agreed with the sponsor and must be accompanied by a completed CIOMS form.

Complete information may not be available within the required timeframes for reporting. In this case, the Initial Report will include as much information as is available at the time. The Principal Investigator will be required to submit a follow-up report as soon as complete information becomes available and in any case no later than 24hrs after submission of the Initial Report.

On receipt of each SAE Report Form, the Sponsor will provide a separate assessment of the Causality and Expectedness of that SAE. Should any part of the Sponsor's assessment be in conflict with that of the Principal Investigator, then the Sponsor can not overrule the PI's assessment. Both opinions must be recorded and both opinions must be submitted to the relevant authorities.

All adverse events (AEs) related to the IMP that are both serious and unexpected (i.e. a SUSAR) are subject to expedited reporting:
1. Fatal or life threatening. The Sponsor must ensure that SUSARs are reported to the main REC and the MHRA as soon as possible, but no later than **7 days** after the Sponsor had information that the case fulfilled the minimum criteria for initial expedited reporting. Any additional follow up information should be reported within a further **8 days** of sending the initial report. All reports should normally be submitted by the PI (on behalf of the Sponsor).

2. Non fatal or non life-threatening. The Sponsor must ensure that SUSARs are reported to the main REC and the MHRA as soon as possible, but no later than **15 days** after the Sponsor had information that the case fulfilled the minimum criteria for initial expedited reporting. Any additional follow up information should be reported as soon as it becomes available. All reports should normally be submitted by the PI (on behalf of the Sponsor).

One year following the granting of a CTA, and thereafter annually, the PI must compile an annual safety report, consisting of a list of all the suspected serious adverse reactions which have occurred during that year in relation to the trial and submit this to the:

- MHRA
- Sponsor
- REC that granted approval

9. STATISTICS

Sample size

There is a large literature on establishing reference ranges for analyte values. The complexity and minimum sample size required to determine reliable reference ranges depends on whether or not the analyte values are age dependent. For the cortisol response to synacthen®, there is evidence (Clark et al., 1998) that the distributional properties are independent of age and so univariate reference ranges are appropriate. In such a situation, the IFCC and the FDA recommend that a minimum sample size of 60 be used for a Gaussian distribution of values, or for data that can be transformed to Gaussian form. In all other situations non-parametric techniques should be used with a minimum sample size of 120.

The distribution of Synacthen® stimulated cortisol results is positively skewed (Clark et al., 1998). If the data can not be transformed to Gaussian form, then non-parametric methods will require the sample size to be increased to 120 males and 120 females to estimate the specified centiles.

Statistical Analyses

The effect of gender differences between the females taking ethinyl oestradiol and age and gender matched control subjects on cortisol responses will be assessed using either the Mann-Whitney U test, or unpaired t-test for non-Gaussian and Gaussian distributions respectively. A p value of less than 0.05 will be considered statistically significant.

Cortisol immunoassays will be compared to the GC-MS method by Altman Bland plots and linear regression. The relative bias will be assessed by comparing results from each immunoassay with the GC-MS result for that sample expressed as a bias ratio. Comparison of the bias ratios will be performed using the students t-test.

10. PROJECT MANAGEMENT

Study Management

Dr Aled Rees, Dr Carol Evans and Dr Nadia El-Farhan will be responsible for the day-to-day conduct of the study.

Dr Aled Rees (Senior Lecturer in Endocrinology) will co-ordinate recruitment and obtain subject consent, along with Dr Nadia El-Farhan, Sister Nikki Davies and Sister Janet Lewis who will also perform the short Synacthen® tests.
Mr Alan Pickett (BMS2) will be responsible for receipt and storage of samples in the Department of Medical Biochemistry.

Dr Nadia El-Farhan will be responsible for coordinating biochemical analyses, data entry and will at a later stage conduct statistical analysis on the data obtained, in conjunction with the Department of Medical Statistics. The database will be held on the H: drive of a Trust computer in the department of Medical Biochemistry. Cardiff and Vale IT department operate a daily incremental back up of the H: drive. Data will be archived securely for 15 years after the trial has ended.

A formal meeting will take place on a fortnightly basis to discuss all aspects of the study and to plan for the forthcoming fortnight. Informal discussions regarding any difficulties that arise with the project will also take place as required. A further bi-monthly meeting involving Professor Scanlon, Dr Rees, Dr Evans and Dr El-Farhan will occur to discuss progress and to focus on manuscript / abstract preparation as the results allow.

Procedures for data collection and recording
The investigator must record all data relating to protocol procedures, study drug administration and laboratory data in the Case report form (CRF). The investigator may designate authority to complete CRFs to appropriately qualified staff by completing the signature log.
The investigator must sign CRFs to attest to their accuracy and completeness.
All corrections must be inserted in such a way as to not obscure the original entry. The correct data must be inserted, initialled and authorised by site personnel. If it is not obvious why a change has been made, a reason must be provided.

Source Data Verification
All original records and reports will be retained. Quality control data will be recorded for biochemical assays.

Data Quality
Reasons should be given on the relevant CRF for any missing data or other protocol deviations. Any data management queries will be returned to the principal investigator who should ensure that these are dealt with promptly.

Changes to the protocol
Any change or addition to this protocol will require a written protocol amendment. Amendments significantly affecting the scope of the investigation or the scientific quality of the study will be submitted for additional approval by the Local Research Ethics Committee, Cardiff University Research and Development department and Cardiff and Vale NHS Trust Research and Development department.

Auditing procedures
As part of Good Clinical Practice, Dr Rees and Dr Evans will ensure that the study protocol and documentation are closely monitored. All study documentation will be available for inspection at any time by appropriate regulatory authorities including internal audits by the Cardiff University Research and Development audit officer.
The trial will be conducted in compliance with the protocol, GCP and the applicable regulatory requirements.

The study will be monitored by Cardiff University R&D office as per the monitoring schedule outlined in the sponsorship agreement between the principal investigator and the University.

Publication of results

Any formal presentation or publication of data from this study will be considered as a joint publication by the investigators and authorship will be determined by mutual agreement. The research findings will be disseminated through peer reviewed publications and presentations at regional/national/international meetings. Following scientific peer review and publication, the results of the study will be communicated directly to all participants and, additionally, to local patient support groups where appropriate.

11. REFERENCES

1. All Wales Clinical Biochemistry Audit Group. Standards for the Performance of the Short Synacthen test in Investigating Suspected Adrenocorticol Insufficiency. http://www.acbwales.org.uk
2. Clark PM *et al.,*1998. Defining the normal cortisol response to the short Synacthen test; implications for the investigation of hypothalamic-pituitary disorders. Clinical Endocrinology. 49: 287-292.
3. Wiegratz, I., *et al.,* 2003. Effect of four oral contraceptives on thyroid hormones, adrenal and blood pressure parameters. Contraception. 67: 361-6.
4. Brennan A *et al.,* 2004. Nephrotic syndrome: cause of an abnormal response to the rapid ACTH stimulation test. Nephrol Dial Transplant; 19: 477-8.

Appendices

Appendix 4 – Sample analysis protocol

A total of 340 patient and 30 iQC samples will be sent for batched analysis over a 10 week period. The prolonged analysis time is intentional and it is important that batches are not analysed more frequently than once a week. It is also useful for the same person (who is familiar with this protocol) to run the samples every week. So if you are unable to run samples any particular week then it is preferable for you to leave them until you are next available.

1. Sample Preparation:

- Samples will be transported frozen and should remain frozen at -20º until analysis
- On the day of analysis remove one batch from freezer and allow samples to thaw and reach room temperature before proceeding
- Mix samples thoroughly and transfer to tubes suitable for automated analysis

2. Sample Analysis:

- Run one batch of samples per week
- Each batch will contain up to 17 paired patient samples and 3 iQC samples (High, Med, Low)
- Samples should be enrolled on your lab system as for normal patient samples and run in the order they are sent
- Use a single analyser for all study samples i.e. if your lab has more than one analyser used for cortisol analysis ensure that it is always the same analyser that is used for these samples
- Record results in the SST Trial Excel spreadsheet supplied
- On the day of the run please record the following analyser and assay details:
    1) Results and targets of your laboratory's iQC levels
    2) Reagent and Calibration Lot Numbers
    3) Date of most recent calibration.
    4) Analyser serial number

3. Post-analysis:

- Refreeze samples post analysis until results have been reviewed
- Send results of analysis to Mr Alan Pickett (Email: Alan.Pickett@cardiffandvale.wales.nhs.uk, Tel: 029 20748368) or Dr Carol Evans (Email: Carol.Evans@CardiffandVale.wales.nhs.uk, Tel: 029 2074 8367) within the week of analysis and before analysis of any further batches, but continue weekly analysis unless advised otherwise
- Once all samples have been analysed please provide details of all cortisol EQA returns and UKNEQAS ABC scores received by your lab over the entire period of analysis

In case of discordant results (e.g. result of 0 min sample greater than 30 min sample), analyser fault or accidental loss of sample:

- Re-assay discordant samples and record both the original result and the result of the re-assay.

- If an analyser fault occurs then re-assay samples post repair and record both pre- and post-repair results
- If a further sample is required then please contact Mr Alan Pickett (Tel. 029 20748368) for further advice.

Appendix 5 – Publications

Clinical Endocrinology (2013) **78**, 673–680 doi: 10.1111/cen.12039

ORIGINAL ARTICLE

Method-specific serum cortisol responses to the adrenocorticotrophin test: comparison of gas chromatography-mass spectrometry and five automated immunoassays

Nadia El-Farhan*,†, Alan Pickett*, David Ducroq‡, Catherine Bailey§, Kelly Mitchem¶, Nicola Morgan**, Annie Armston††, Laila Jones†, Carol Evans* and D. Aled Rees†

*Department of Medical Biochemistry and Immunology, University Hospital of Wales, †Centre for Endocrine and Diabetes Sciences, Institute for Molecular and Experimental Medicine, School of Medicine, Cardiff University, ‡Quality Laboratory, WEQAS, Cardiff, §Biochemistry Department, Royal Gwent Hospital, Newport, ¶Biochemistry Department, Prince Charles Hospital, Merthyr Tydfil, **Biochemistry Department, Bristol Royal Infirmary, Bristol and ††Biochemistry Department, Southampton General Hospital, Southampton, UK

Summary

Objective The serum cortisol response to the adrenocorticotrophin (ACTH) test is known to vary significantly by assay, but lower reference limits (LRL) for this response have not been established by the reference gas chromatography-mass spectrometry (GC-MS) method or modern immunoassays. We aimed to compare the normal cortisol response to ACTH stimulation using GC-MS with five widely used immunoassays.

Design, Patients and Measurements An ACTH test (250 μg iv $ACTH_{1-24}$) was undertaken in 165 healthy volunteers (age, 20–66 years; 105 women, 24 of whom were taking an oestrogen-containing oral contraceptive pill [OCP]). Serum cortisol was measured using GC-MS, Advia Centaur (Siemens), Architect (Abbott), Modular Analytics E170 (Roche), Immulite 2000 (Siemens) and Access (Beckman) automated immunoassays. The estimated LRL for the 30 min cortisol response to ACTH was derived from the 2·5th percentile of log-transformed concentrations.

Results The GC-MS-measured cortisol response was normally distributed in males but not females, with no significant gender difference in baseline or post-ACTH cortisol concentration. Immunoassays were positively biased relative to GC-MS, except in samples from women on the OCP, who showed a consistent negative bias. The LRL for cortisol was method-specific [GC-MS: 420 nM; Architect: 430 nM; Centaur: 446 nM; Access 459 nM; Immulite (2000) 474 nM] and, for the E170, also gender-specific (female: 524 nM; male 574 nM). A separate LRL is necessary for women on the OCP.

Conclusions Normal cortisol responses to the ACTH test are influenced significantly by assay and oestrogen treatment. We recommend the use of separate reference limits in premenopausal women on the OCP and warn users that cortisol measurements in this subgroup are subject to assay interference.

(Received 24 July 2012; returned for revision 2 August 2012; finally revised 4 September 2012; accepted 5 September 2012)

Introduction

The adrenocorticotrophin (ACTH) stimulation test is widely used in the evaluation of the hypothalamic-pituitary-adrenal (HPA) axis.[1,2] However, there are well-recognised limitations, most notably its failure to detect acute pituitary failure.[3–5] There is also debate as to whether the standard or low-dose test performs better in evaluating the adrenal axis.[6] Despite these limitations, the ACTH stimulation test is often chosen in preference to the insulin stress test, due to the risks associated with the latter and its contraindication in several patient groups[7]; and much work has been undertaken to establish appropriate cortisol cut-offs to distinguish normality from disease of the HPA axis. More recently, it has become clear that differences in the assays used to measure serum cortisol impact significantly on the interpretation of post-ACTH cortisol values.[8,9] Indeed, when the ACTH stimulation test was first described, serum cortisol was universally measured using a nonspecific, fluorimetric assay that measured both cortisol and corticosterone.[10] Since then automated serum cortisol immunoassays have become standard in most clinical laboratories, each with its own performance characteristics and specificity for cortisol. Studies evaluating the ACTH stimulation test have demonstrated the need for assay-specific serum cortisol cut-offs.[8,9] However, they have differed in

Correspondence: Aled Rees, Centre for Endocrine and Diabetes Sciences, Institute for Molecular and Experimental Medicine, School of Medicine, Cardiff University, Heath Park, Cardiff CF14 4XN, UK.
Tel.: +44 (0) 2920 742341; Fax: +44 (0) 2920 744671;
E-mail: reesda@cf.ac.uk

CE and DAR contributed equally to this study.

their findings, variably describing normal and non-normal distributional forms of stimulated cortisol concentration in healthy volunteers. These differences have been attributed to population differences between studies, most notably the inclusion or not of women taking exogenous oestrogens, in addition to assay effect. Furthermore, although the significance of assay differences is now well recognised, there are no published data defining the lower limits of the cortisol response in healthy volunteers for the most popular cortisol immunoassays in current use. Thus, users of the ACTH stimulation test are aware of the need for assay-specific lower limits, but are unable to implement them.

By measuring total cortisol using gas chromatography-mass spectrometry (GC-MS), a reference method for cortisol measurement and the method on which the increasingly popular, less labour-intensive LC-MS/MS, is based,[11] we sought to define a lower reference limit for the cortisol response to ACTH stimulation in healthy volunteers, which is unaffected by immunoassay variability. We also set out to establish a method-specific lower reference limit for five widely used automated serum cortisol immunoassays and to identify assay-specific characteristics that might explain some of the previously observed differences in results.

Subjects and methods

Subjects

One hundred and sixty-five healthy volunteers aged between 20 and 66 years were recruited from Hospital and University staff. Pregnant and breastfeeding women were excluded as was anyone reporting significant intercurrent disease, a history of thyroid or other autoimmune disease, previous sensitivity to ACTH testing, asthma or an allergic disorder, and treatment with corticosteroids. Of those taking part, 60 were men and 105 women (90 premenopausal and 15 postmenopausal). Twenty-four of the premenopausal female participants were taking an oestrogen-containing oral contraceptive pill (OCP), containing between 20 and 35 μg of ethinyloestradiol, at the time of testing. Mean age (and range) was 37·1 (22–62), 40·7 (20–66) and 28·7 (21–40) years in male, non-OCP female and OCP female groups, respectively. The study protocol was approved by the South East Wales Research Ethics Committee, Cardiff University (study sponsors) and the UK Competent Authority (the Medicines and Healthcare Products Regulatory Authority). All subjects gave written informed consent before study commencement.

Sample collection and handling

The ACTH stimulation tests were carried out in the morning between 08·30 and 11·30 h. Subjects were not required to fast overnight, but were restricted from eating, drinking or smoking for the 30 min before the test. There were no restrictions on physical exercise prior to the test, but participants rested in a sitting position for 15 min before the test began and then for the duration of the test. Women under the age of 40 were tested in days 1–7 of the follicular phase of their menstrual cycle. An indwelling catheter was inserted into a superficial antecubital vein, and the ACTH stimulation test was performed by intravenous administration of a 250-μg bolus of synthetic $ACTH_{1-24}$ (Synacthen; Alliance Pharmaceuticals Ltd, Wiltshire, UK). Blood was collected at baseline and 30 min into plain tubes (free of separator gel). Samples were allowed to clot, and following separation, one aliquot of separated serum was analysed immediately, using the Siemens Centaur cortisol assay, with the remainder stored frozen at −20 °C. Samples were later removed for analysis in batches consisting of 34 patient and three internal quality control samples (target concentrations: 76, 528 and 696 nM, respectively), over a 10-week period. This included repeat analysis on the Centaur which excluded sample degradation (R^2 = 0·96 prestorage *vs* poststorage cortisol).

Analytical methods

Total cortisol was measured by the Welsh External Quality Assurance Scheme (WEQAS) Reference Laboratory using a modified version of their GC-MS reference method[12] which differed in that the bracketed standard curve used in the reference method was replaced with a conventional 6-point standard curve. Interassay CVs for this assay were 5·4%, 6·2% and 6·8% at concentrations of 120, 363 and 657 nM, respectively. Total cortisol was also analysed by five automated immunoassays: the Advia Centaur (Siemens AG, Erlangen, Germany) with interassay CVs of 5·2%, 4·5% and 2·9% at concentrations of 105, 571 and 784 nM, respectively, Modular Analytics E170 (Roche, Mannheim, Germany) with interassay CVs of 9·8%, 6·1% and 6·6% at concentrations of 116, 613 and 818 nM, respectively, Immulite 2000 (Siemens AG) with interassay CVs of 10·2%, 7·7% and 7·2% at concentrations of 117, 695 and 978 nM, respectively, Access (Beckman Coulter, Brea, CA, USA) with interassay CVs of 9·9%, 5·0% and 2·5% at concentrations of 124, 620 and 828 nM, respectively, and Architect (Abbott Laboratories, Chicago, IL, USA) with interassay CVs of 10·4%, 5·4% and 6·8% at concentrations of 97, 549 and 840 nM, respectively. All assays were solid-phase competitive binding immunoassays using chemiluminescent detection, except the E170 which uses electrochemiluminescent detection.

All participating laboratories were enrolled in the UK National External Quality Assurance Scheme at the time of analysis, and their cortisol assays were performing well within their method groups.

Statistics

All statistical analyses were performed using SPSS version 16.0 (SPSS Inc., Chicago, IL, USA). The nature of the distribution of total cortisol in healthy volunteers was evaluated using the Kolmogorov–Smirnov test, with a P-value > 0·05 confirming a Normal distribution. The distributional form was found to vary by assay, time point and gender; hence, all data were log-transformed to give a Normal distribution prior to analysis. A mean cortisol concentration was determined at each time point for each assay. A lower reference limit for cortisol was calculated for each assay

from the mean cortisol concentration at 30 min as the 2·5th percentile, that is, mean – 1·96*SD. These values were then back-transformed to give the geometric mean, 2·5th and 97·5th centile values and lower reference limits presented here.

An assay-specific bias ratio was calculated for each sample (at each time point) by dividing the untransformed immunoassay result by the matched GC-MS result. Differences between method means and bias ratios were compared using the paired *t*-test with Bland-Altman plots also used to demonstrate assay differences. Gender differences and differences between non-OCP females and OCP females were evaluated using the unpaired *t*-test. In all cases, differences were considered to be significant when $P < 0{\cdot}05$. Bland-Altman and assay correlations were plotted using untransformed immunoassay and GC-MS results.

Results

Baseline cortisol

When measured by GC-MS, baseline cortisol was normally distributed in male but not in non-OCP female volunteers, with no significant gender difference in mean values (Table 1) or any effect of age ($R^2 = 0{\cdot}003$, *P*-value = 0·26). Baseline cortisol was also normally distributed in males for all immunoassays apart from the Immulite (2000), and in non-OCP females, the non-normal distribution persisted for all immunoassays (data not shown). In contrast to GC-MS, all immunoassays showed a statistically significant gender difference in mean baseline cortisol concentration (Table 1). Mean cortisol concentration as measured by GC-MS was significantly lower than immunoassay cortisol in males for all assays and in non-OCP females for all but the Architect and Access assays. In females taking an OCP, baseline cortisol was normally distributed with a significantly higher mean cortisol concentration than in either male or non-OCP females (Table 1). This finding was consistent across GC-MS and all immunoassays studied.

Post-ACTH cortisol

After stimulation with ACTH, GC-MS cortisol remained normally distributed in male but not in non-OCP female volunteers with no significant concentration difference between genders (Table 2). Post-ACTH cortisol was also normally distributed in males for all immunoassays and, in contrast to GC-MS, in non-OCP females for all immunoassays apart from the Centaur (data not shown). The gender difference identified in baseline cortisol persisted in poststimulation cortisol concentrations with the Architect, E170 and Access assays, with a significantly higher cortisol response in males (Table 2). For all assays, the mean cortisol concentration differed significantly from GC-MS cortisol in both male and non-OCP female subjects. Stimulated cortisol in OCP females retained a normal distribution with markedly higher mean cortisol concentrations than in males or non-OCP females (Table 2).

Assay correlation with GC-MS

As anticipated, all assays showed good correlation with GC-MS over the full range of baseline and stimulated cortisol concentrations (92–1339 nM), although this relationship varied based on gender and exogenous oestrogen use. Figure 1 demonstrates the differences observed, with samples from male subjects showing a slightly higher positive bias than samples from non-OCP female subjects for all assays apart from the Architect for which there was virtually no assay bias for male samples, but a small negative bias for samples from non-OCP females. This relationship is quantified as the mean bias ratio (Table 3), which confirms that for all assays, overall assay bias for both male and non-OCP female samples is positive. This is in contrast to the mean bias ratio for samples from OCP females that is negative for all assays apart from the E170 – although this assay also shows the same overall bias pattern (male > non-OCP female > OCP female) as that of the other assays (Fig. 1). This OCP-dependent difference in assay behaviour was further explored in Bland-Altman plots (Fig. 2), which clearly demonstrated the overall negative bias of cortisol assays relative to GC-MS when samples collected from OCP females were used. Once again, the exception was the E170 assay in which overall bias remained positive for all subjects, albeit somewhat lower in OCP females than non-OCP subjects.

Table 1. Geometric mean of baseline cortisol concentrations in male, non- oral contraceptive pill (OCP) female and OCP-female subjects

Assay	Males $n = 60$	Non-OCP females $n = 79$	*P*-value*	OCP females $n = 24$	*P*-value†
GC-MS	274 (131–575)	254 (139–463)	0·193	542 (318–922)	<0·001
Centaur	298 (158–565)‡	257 (138–477)‡	0·023	488 (323–738)	<0·001
Architect	289 (151–556)‡	247 (134–455)§	0·018	465 (301–718)	<0·001
E170	370 (182–750)‡	292 (147–581)‡	0·001	646 (383–1090)	<0·001
Immulite (2000)	316 (165–604)‡	267 (144–495)‡	0·003	510 (330–788)	<0·001
Access	293 (160–538)‡	252 (143–444)§	0·011	429 (286–643)	<0·001

Results are expressed as geometric mean (2·5th–97·5th percentile) in nM.
**P*-value for difference between genders.
†*P*-value for difference between women taking an oral contraceptive pill and those who were not.
‡*P*-value for immunoassay *vs* GC-MS < 0·005.
§*P*-value 0·95 and 0·21 for Architect and Access assays *vs* GC-MS respectively.

Table 2. Geometric mean of post-adrenocorticotrophin stimulation cortisol concentrations in male, non-oral contraceptive pill (OCP) female and OCP-female subjects

Assay	Males $n = 60$	Non-OCP females $n = 79$	P-value*	OCP females $n = 24$	P-value†
GC-MS	563 (418–757)	555 (421–731)	0·594	870 (643–1177)	<0·001
Centaur	599 (448–802)‡	578 (446–750)‡	0·138	763 (619–940)	<0·001
Architect	577 (430–773)‡	542 (416–707)‡	0·012	747 (577–967)	<0·001
E170	772 (574–1039)‡	712 (524–967)‡	0·003	1026 (791–1330)	<0·001
Immulite (2000)	641 (469–874)‡	628 (478–826)‡	0·449	850 (688–1051)	<0·001
Access	625 (459–852)‡	594 (455–777)‡	0·045	757 (604–948)	<0·001

Results are expressed as geometric mean (2·5th–97·5th percentile) in nM.
*P-value for difference between genders.
†P-value for difference between women taking an oral contraceptive pill and those who were not.
‡P-value for immunoassay *vs* gas chromatography-mass spectrometry (GC-MS) < 0·02.

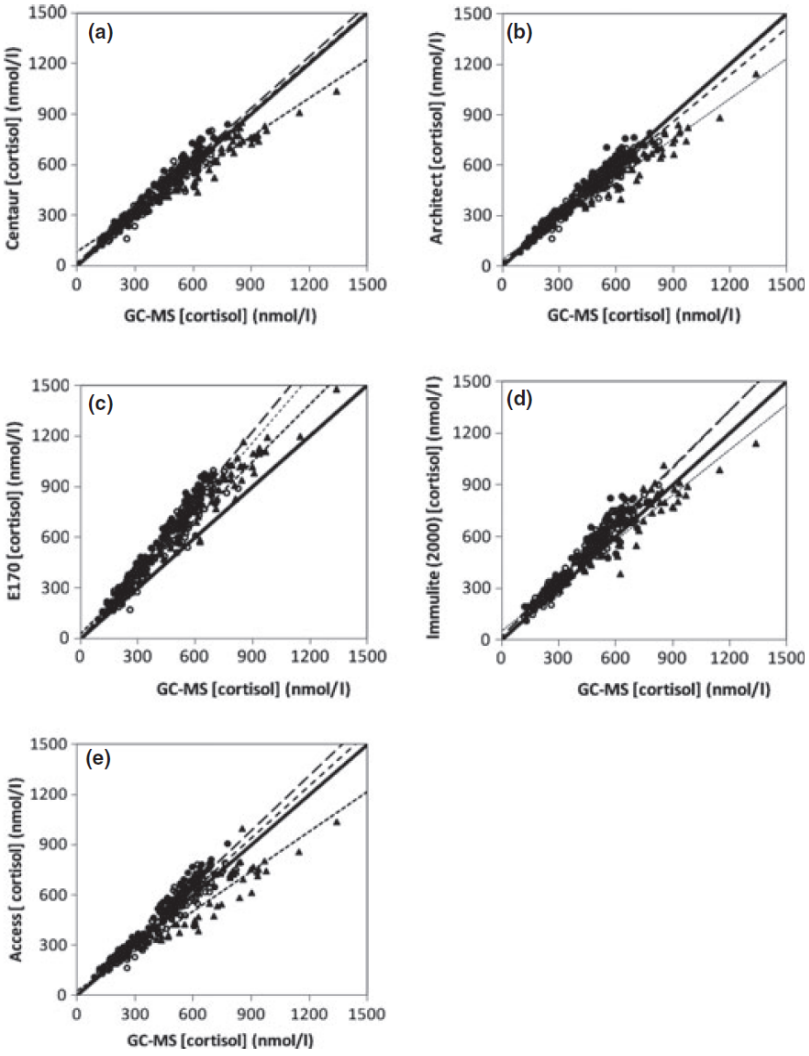


Fig. 1 Correlation plots demonstrating the relationship between immunoassay and gas chromatography-mass spectrometry (GC-MS) cortisol for all baseline and post-adrenocorticotrophinstimulation cortisol measurements combined, for male (●), non-oral contraceptive pill (OCP) female (○) and OCP female (▲) subjects. (a) Centaur; (b) Architect; (c) E170; (d) Immulite (2000) and (e) Access assays. The solid black line indicates equivalence between methods; _ _ _ line of best fit, males; - - - - - line of best fit, non-OCP females; —— line of best fit, OCP females.

Assay-specific lower reference limits of cortisol post-ACTH

The lower reference limit of cortisol concentration 30 min post-ACTH for GC-MS was 417, 422 and 649 nM in males, non-OCP females and OCP females, respectively. A gender-specific lower reference limit for cortisol was also determined for each immunoassay (Table 4). As mean cortisol concentration post-ACTH stimulation was not gender dependent for the GC-MS, Centaur

Table 3. Bias ratios for cortisol immunoassay compared to gas chromatography-mass spectrometry (GC-MS)

Assay	Males	Non-OCP females	*P*-value*	OCP females	*P*-value†
Centaur	1·08	1·05	0·012	0·88	<0·001
Architect	1·04	1·00	<0·001	0·85	<0·001
E170	1·36	1·25	<0·001	1·18	<0·001
Immulite (2000)	1·15	1·11	0·009	0·96	<0·001
Access	1·09	1·05	0·002	0·83	<0·001

Overall mean bias ratio was calculated using untransformed baseline (0 min) and post-ACTH (30 min) cortisol concentrations.
**P*-value for difference between genders.
†*P*-value for difference between women taking an oral contraceptive pill and those who were not.

and Immulite (2000) assays (Table 2), we propose a single lower limit calculated by combining all (male and female) non-OCP subjects. This was not possible for the E170, Architect and Access assays because mean cortisol concentration post-ACTH showed evidence of gender dependence. However, inspection of the lower limit for the Architect and Access assays showed the difference between genders (14 and 4 nM, respectively) to be within expected assay variability (assay precision was 5·4% and 5·0%, respectively, at 549 and 620 nM), so gender-related reference ranges may not be necessary in clinical practice. However, for the E170, the difference between the male and female lower limits was significant (50 nM) making gender-related reference ranges necessary.

Discussion

This is the largest study to examine normative responses to the standard dose ACTH test and the first to compare results

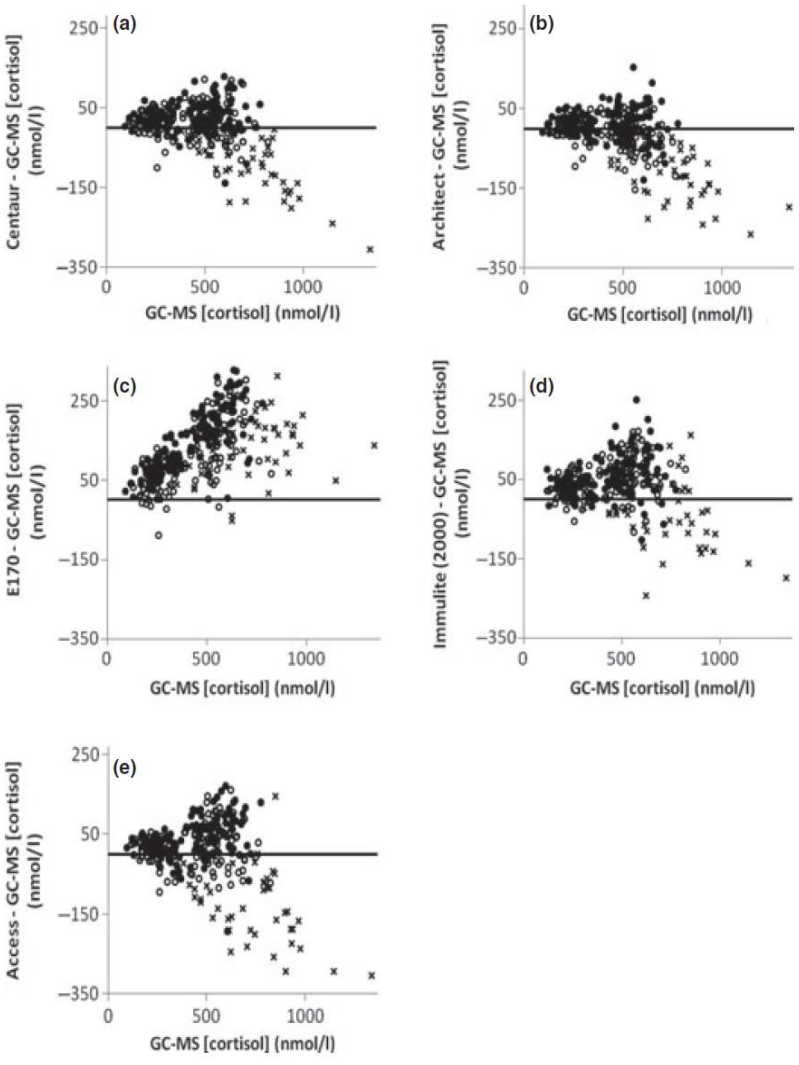


Fig. 2 Bland-Altman plots showing the difference between immunoassay and gas chromatography-mass spectrometry (GC-MS) cortisol plotted against GC-MS cortisol concentration for all baseline and post-adrenocorticotrophin stimulation cortisol measurements. (a) Centaur; (b) Architect; (c) E170; (d) Immulite (2000); (e) Access. The solid black line indicates no difference between assays. ●, Male subjects; ○, non-oral contraceptive pill (OCP) female subjects; ×, OCP-female subjects.

Table 4. Assay-specific estimated lower reference limits for post-adrenocorticotrophin cortisol according to gender and oral contraceptive pill (OCP)-status

Assay	Males	Non-OCP females	Combined male and Non-OCP female subjects*	OCP females
GC-MS	418	421	420	643
Centaur	448	446	446	619
Architect	430	416	NA	577
E170	574	524	NA	791
Immulite (2000)	469	478	474	688
Access	459	455	NA	604

The estimated lower reference limit was determined by back transformation of the 2·5th percentile value (mean − 1·97*SD) of the log-transformed data. Results are expressed in nM.
NA, not applicable.
*A single assay-specific combined LRL for males and non-OCP females was calculated by pooling all post-ACTH cortisol results in these subgroups.

using a reference mass spectrometry method with those from five modern immunoassays in widespread use. We show that responses to ACTH stimulation in healthy volunteers may be influenced significantly by assay and exogenous oestrogens, and less consistently by gender in an assay-dependent manner. Previous studies have shown conflicting results with respect to both the distributional form of pre- and post-ACTH cortisol and the influence of gender. Clark *et al.*[8] found that the response of serum cortisol to ACTH was non-Gaussian, and both method- and gender-dependent, with some variation on gender influence between methods. In contrast, Klose *et al.*,[9] who confirmed clinically significant interassay differences in stimulated cortisols, did not find an effect of gender nor evidence for non-normal distribution. By using GC-MS to measure total cortisol, we were able to establish which of these features were genuine and which attributable to the imperfections of immunoassay. When measured by GC-MS, we found that the cortisol response to ACTH is normally distributed in men, but not in women, with the exception of those taking an oestrogen-containing OCP. Whilst others have shown that ACTH-induced increases in total cortisol are not influenced by phase of the menstrual cycle,[13,14] we speculate that this gender difference in distributional form may relate to variations in endogenous oestrogen and consequent effects on cortisol binding globulin (CBG) concentration. Mean cortisol concentrations at baseline and post-ACTH did not differ between men and women when measured by GC-MS. However, for all immunoassays, a gender difference was found for baseline concentrations and for the Architect, E170 and Access assays this gender difference persisted post-ACTH stimulation. This assay variability suggests that gender differences in the cortisol response to ACTH may be due to analytical factors rather than gender differences in sensitivity of the adrenal cortex to ACTH stimulation.[13,15,16]

In contrast to the marginal influences of gender, we confirmed a marked effect of oestrogen treatment on the cortisol response to ACTH stimulation whether measured by GC-MS or immunoassay. Compared with women not taking oestrogen, women treated with the OCP displayed a 1·7- to 2·2-fold and 1·3- to 1·6-fold elevation in total cortisol at baseline and 30 min, respectively. This accords with our understanding of a marked stimulatory influence of oestrogen on CBG production[17] and is in agreement with the findings of Klose *et al.*[9] who recommended that this test should only be undertaken after the OCP has been discontinued. The comparison with GC-MS was of further value as it confirmed the negative assay bias reported by Jung *et al.*[18] in women in the third trimester of pregnancy and on the OCP when cortisol measurement by the Advia Centaur was compared with LC-MS/MS. Our findings in women on the OCP demonstrate that this bias is not universal across all immunoassays, varies in magnitude by assay and is also seen when immunoassay cortisol is compared with GC-MS. This negative bias is in stark contrast to the overall positive bias that was identified for samples from subjects (male and female) not on the OCP, and given the increase in total cortisol concentration in these subjects, this further complicates the interpretation of the test and raises additional concerns about its validity in this group. We believe this effect may relate to increased cortisol-CBG binding in the presence of increased serum concentrations of the latter, which results in reduced availability of cortisol epitopes for binding to assay antibody. However, further work is needed to ascertain the mechanism of this effect, as it has been shown that heat treatment is effective at eliminating it,[18] but it is unclear whether this is due to a direct effect on CBG. It would also be valuable to establish the duration of this effect and its reversibility after stopping the OCP. In the meantime, our findings reinforce the potential risks of misclassifying hypoadrenal women taking oestrogen as eucortisolaemic if oestrogen status is not taken into account. We thus share Klose *et al.*' views that consideration should be given to stopping oestrogen replacement before ACTH testing, but if this is not feasible, then the lower reference limits provided here may assist clinicians in their decision-making.

Our results illustrate the influence that assay performance can have on cortisol measurements, such that mean cortisol values post-ACTH ranged from 542 nM in non-OCP females with the Architect assay to 772 nM in males with the E170. All five immunoassays included in this study differed significantly from GC-MS either in the distributional form of the cortisol response in women not taking the OCP or by virtue of a gender difference in concentration or both. As expected, mean cortisol concentration was higher for all immunoassays than with GC-MS, and each immunoassay exhibited its own unique cortisol response to ACTH stimulation. These findings may be explained by differences in assay calibrants or in the specificity of assay antibodies, but irrespective of the origin, there is the potential for this to lead to misclassification if assay differences are not recognised. By including five of the most commonly used modern immunoassays, we provide clinicians with access to gender-specific, normative ranges for the cortisol response to

ACTH. Furthermore, the inclusion of the GC-MS method provides a reference point from which normative cortisol responses for future immunoassays, and current assays not studied here, can be derived by the estimation of assay bias relative to GC-MS.

In contrast to mean responses, the lower limit of the normative range, defined as the 2·5th percentile value of the log-transformed data, showed much less of a gender difference, ranging from 2 to 14 nM for the Centaur, Architect, Immulite (2000) and Access assays. As these differences are too small to be of any clinical significance and fall within the precision limits of their assays, gender-related reference limits are not required. For the E170 assay, the difference between the male and female lower limits was significant (50 nM); hence, for this assay gender-specific lower limits are required. However, we emphasise that these lower reference limits should not be confused with diagnostic cut-offs for diseases of the HPA axis; such decision limits would need to be established by comparing these results with samples from patients with primary or secondary adrenal insufficiency using ROC curve analysis. The validity of using a single cut-off to diagnose adrenal insufficiency should also be questioned because disease of the HPA axis is a continuum and cortisol values lying just above the lower reference limit may well represent impaired HPA axis function in patients where the clinical features are suggestive. In such cases, a single cut-off is no replacement for good clinical judgement. Nevertheless, our findings are likely to be helpful for endocrinologists because application of the widely used cortisol cut-offs of 500[19,20] or 550 nM[2,21,22] would result in misclassification of a significant number of healthy individuals with adrenal insufficiency (12%, 19%, 4% and 9% for the Centaur, Architect, Immulite (2000) and Access assays, respectively, at 500 nM; 27%, 42%, 16% and 21% for the respective assays at 550 nM). Our findings thus have the potential to reduce inappropriate and costly follow-up investigations for patients referred with a low pretest probability of HPA axis disease.

Our study has several limitations. We did not set out to test the influence of other physiological factors such as fasting, exercise, posture or BMI on post-ACTH cortisol. Although some reports have shown that food may result in an increase in cortisol,[23,24] Klose *et al.*[9] did not find a difference in 30-min cortisol values between the fasting and nonfasting state. They were also unable to demonstrate an effect of intermittent, light exercise on cortisol responses but did find an independent effect of central obesity on the 30-min cortisol response in men. However, as others have shown no effect of obesity on cortisol responses to ACTH,[25] it is unclear whether waist circumference-specific reference limits are needed. Dhillo *et al.*[22] showed that CBG and total cortisol concentrations fall significantly within 30 min of adopting a supine from a standing position. We did not specifically test the influence of posture but standardised our protocol to collect samples in the sitting position, adopted at least 15 min in advance of the baseline collection. We thus recommend that clinicians adopt a similar standardised approach if they are to translate the results from our study to their own practice.

In conclusion, we have shown that cortisol responses to the adrenocorticotrophin test are influenced significantly by assay and treatment with oestrogens. We also report a negative assay bias in women on the OCP which further complicates interpretation of the adrenocorticotrophin test in this group. Endocrinologists investigating patients with suspected adrenal insufficiency should be aware of these limitations and should clarify which immunoassay method is in use in their laboratories before interpreting post-adrenocorticotrophin cortisol results. In the light of our findings, we also recommend that normative responses to the insulin stress test may need to be re-established.

Acknowledgements

We thank Professor Robert Newcombe for his advice on statistical analysis, the staff of the Quality Laboratory, WEQAS, Cardiff; Department of Medical Biochemistry and Immunology, University Hospital of Wales, Cardiff; Biochemistry Department, Royal Gwent Hospital, Newport; Biochemistry Department, Prince Charles Hospital, Merthyr Tydfil; Biochemistry Department, Bristol Royal Infirmary, Bristol and Biochemistry Department, Southampton General Hospital, Southampton for their help with analysis of cortisol samples, and Siemens, Abbott, Roche and Beckman for kindly supporting this work. No grants or fellowships were received supporting the writing of this paper.

Disclosure

The authors have no conflicts of interest to declare.

References

1 Gleeson, H.K., Walker, B.R., Seckl, J.R. *et al.* (2003) Ten years on: safety of short synacthen tests in assessing adrenocorticotropin deficiency in clinical practice. *Journal of Clinical Endocrinology and Metabolism*, **88**, 2106–2111.

2 Reynolds, R.M., Stewart, P.M., Seckl, J.R. *et al.* (2006) Assessing the HPA axis in patients with pituitary disease: a UK survey. *Clinical Endocrinology (Oxford)*, **64**, 82–85.

3 Dickstein, G. (2001) Hypothalamo-pituitary-adrenal axis testing: nothing is sacred and caution in interpretation is needed. *Clinical Endocrinology (Oxford)*, **54**, 15–16.

4 Courtney, C.H., McAllister, A.S., Bell, P.M. *et al.* (2004) Low- and standard-dose corticotrophin and insulin hypoglycaemia testing in the assessment of hypothalamic-pituitary-adrenal function after pituitary surgery. *Journal of Clinical Endocrinology and Metabolism*, **89**, 1712–1717.

5 Klose, M., Lange, M., Kosteljanetz, M. *et al.* (2005) Adrenocortical insufficiency after pituitary surgery: an audit of the reliability of the conventional short synacthen test. *Clinical Endocrinology (Oxford)*, **63**, 499–505.

6 Abdu, T.A.M., Elhadd, T.A., Neary, R. *et al.* (1999) Comparison of the low dose short synacthen test (1ug), the conventional dose short synacthen test (250 ug), and the insulin tolerance test for assessment of the hypothalamic-pituitary-adrenal axis in patients with pituitary disease. *Journal of Clinical Endocrinology and Metabolism*, **84**, 838–843.

7 Shah, A., Stanhope, R. & Matthew, D. (1992) Hazards of pharmacological tests of growth hormone secretion in childhood. *British Medical Journal*, **304**, 173–174.

8 Clark, P.M., Neylon, I., Raggatt, P.R. *et al.* (1998) Defining the normal cortisol response to the short synacthen test: implications for the investigation of hypothalamic-pituitary disorders. *Clinical Endocrinology (Oxford)*, **49**, 287–292.

9 Klose, M., Lange, M., Rasmussen, A.K. *et al.* (2007) Factors influencing the adrenocorticotropin test: role of contemporary cortisol assays, body composition and oral contraceptive agents. *Journal of Clinical Endocrinology and Metabolism*, **92**, 1326–1333.

10 Mattingly, D. (1962) A simple fluorimetric method for the estimation of free 11-hydroxycorticoids in human plasma. *Journal of Clinical Pathology*, **15**, 374–379.

11 Tai, S.S. & Welch, M.J. (2004) Development and evaluation of a candidate reference method for the determination of total cortisol in human serum using isotope dilution liquid chromatography/mass spectrometry and liquid chromatography/tandem mass spectrometry. *Analytical Chemistry*, **76**, 1008–1014.

12 Woods, L., Ducroq, D.H., Fraser, H.L. *et al.* (2008) Measurement of urinary free cortisol by tandem mass spectrometry and comparison with results obtained by gas chromatography-mass spectrometry and two commercial immunoassays. *Annals of Clinical Biochemistry*, **45**, 380–388.

13 Kirschbaum, C., Kudielka, B.M., Gaab, J. *et al.* (1999) Impact of gender, menstrual cycle phase and oral contraceptives on the activity of the hypothalamus-pituitary-adrenal axis. *Psychosomatic Medicine*, **61**, 154–162.

14 Stewart, P.M., Penn, R., Holder, R. *et al.* (1993) The hypothalamo-pituitary-adrenal axis across the normal menstrual cycle and in polycystic ovary syndrome. *Clinical Endocrinology (Oxford)*, **38**, 387–391.

15 Horrocks, P.M., Jones, A.F., Ratcliffe, W.A. *et al.* (1990) Patterns of ACTH and cortisol pulsatility over twenty-four hours in normal males and females. *Clinical Endocrinology (Oxford)*, **32**, 127–134.

16 Roelfsema, F., van den Berg, G., Frőlich, M. *et al.* (1993) Sex-dependent alteration in cortisol response to endogenous adrenocorticotropin. *Journal of Clinical Endocrinology and Metabolism*, **77**, 234–240.

17 Coolens, J.-L., van Baelen, H. & Heyns, W. (1987) Clinical use of unbound plasma cortisol as calculated from total cortisol and corticosteroid-binding globulin. *Journal of Steroid Biochemistry*, **26**, 197–202.

18 Jung, C., Ho, J.T., Torpy, D.J. *et al.* (2011) A longitudinal study of plasma and urinary cortisol in pregnancy and postpartum. *Journal of Clinical Endocrinology and Metabolism*, **96**, 1533–1540.

19 Tordjman, K., Jaffe, A., Trostanetsky, Y. *et al.* (2000) Low-dose (1 microgram) adrenocorticotrophin (ACTH) stimulation as a screening test for impaired hypothalamo-pituitary-adrenal axis function: sensitivity, specificity and accuracy in comparison with the high-dose (250 microgram) test. *Clinical Endocrinology (Oxford)*, **52**, 633–640.

20 Suliman, A.M., Smith, T.P., Labib, M. *et al.* (2002) The low-dose ACTH test does not provide a useful assessment of the hypothalamic-pituitary-adrenal axis in secondary adrenal insufficiency. *Clinical Endocrinology (Oxford)*, **56**, 533–539.

21 Orme, S.M., Peacey, S.R., Barth, J.H. *et al.* (1996) Comparison of tests of stress-released cortisol secretion in pituitary disease. *Clinical Endocrinology (Oxford)*, **45**, 135–140.

22 Dhillo, W.S., Kong, W.M., Le Roux, C.W. *et al.* (2002) Cortisol-binding globulin is important in the interpretation of dynamic tests of the hypothalamic-pituitary-adrenal axis. *European Journal of Endocrinology*, **146**, 231–235.

23 Quigley, M.E. & Yen, S.S. (1979) A mid-day surge in cortisol levels. *Journal of Clinical Endocrinology and Metabolism*, **49**, 945–947.

24 Pasquali, R., Biscotti, D., Spinucci, G. *et al.* (1998) Pulsatile secretion of ACTH and cortisol in premenopausal women: effect of obesity and body fat distribution. *Clinical Endocrinology (Oxford)*, **48**, 603–612.

25 Maccario, M., Grottoli, S., Divito, L. *et al.* (2000) Adrenal responsiveness to high, low and very low ACTH 1–24 doses in obesity. *Clinical Endocrinology (Oxford)*, **53**, 437–444.

Original Article

The Association for Clinical Biochemistry & Laboratory Medicine

Annals of Clinical Biochemistry
50(4) 345–352

DOI: 10.1177/0004563212473448
acb.sagepub.com
SAGE

Development of a rapid assay for the analysis of serum cortisol and its implementation into a routine service laboratory

Laura J Owen[1,2], Joanne E Adaway[1,2], Sarah Davies[2], Sarah Neale[3], Nadia El-Farhan[3,4], David Ducroq[5], Carol Evans[3], D Aled Rees[4], Finlay MacKenzie[6] and Brian G Keevil[1,2]

Abstract

Background: LC-MS/MS is rapidly becoming the technology of choice for measuring steroid hormones. We have developed a rapid LC-MS/MS assay for the routine analysis of serum cortisol. We have used this assay to investigate the effects of gender and exogenous steroid interference on the immunoassay measurement of serum cortisol.

Methods: Zinc sulphate (40 μL) was added to 20 μL of sample. This was vortexed for 10 s followed by the addition of 100 μL of internal standard in methanol. Following mixing and centrifugation, 10 μL of sample was injected into an Acquity LC system coupled to a Quattro Premier tandem mass spectrometer.

Serum samples ($n = 149$) were analysed by LC-MS/MS and two commercial immunoassays. Results were then compared for all samples and for gender differences. A further set of serum samples ($n = 171$) was analysed by the LC-MS/MS assay and a GC-MS assay.

Results: Cortisol had a retention time of 0.98 min and the assay had an injection-to-injection time of 2.6 min per sample. Mean recovery was 99% and mean CV was 8%. The immunoassays gave comparisons of: Roche = 1.23 × LC-MS/MS −1.12 nmol/L and Abbott = 0.94 × LC-MS/MS + 11.97. The comparison with GC-MS showed LC-MS/MS = 1.11 × GC-MS – 22.90.

Discussion: We have developed an LC-MS/MS assay for serum cortisol analysis that is suitable for routine clinical use and has been in use in our laboratory for 12 months. The availability of this assay will give more reliable results in patients receiving exogenous steroid therapy.

Keyword
Serum cortisol; LC-MS/MS

Accepted: 23rd November 2012

Introduction

Liquid chromatography tandem mass spectrometry (LC-MS/MS) is increasingly becoming the method of choice for steroid hormone measurements due to the poor specificity of immunoassays. Serum cortisol immunoassays are particularly susceptible to interferences due to exogenous steroids since these are structurally similar to cortisol and are widely used therapies

[1]Manchester Academic Health Science Centre, University Hospital of South Manchester, The University of Manchester, Manchester, UK
[2]Biochemistry Department, University Hospital of South Manchester, Manchester, UK
[3]Department of Medical Biochemistry and Immunology, University Hospital of Wales, Cardiff, UK
[4]Centre for Endocrine and Diabetes Sciences, Institute for Molecular and Experimental Medicine, School of Medicine, Cardiff University, Cardiff, UK
[5]Quality Laboratory, WEQAS, Cardiff, UK
[6]UK NEQAS, University Hospitals Birmingham, Birmingham, UK

Corresponding author:
Laura J Owen, Manchester Academic Health Science Centre, University Hospital of South Manchester, The University of Manchester, Manchester, UK.
Email: laura.owen@uhsm.nhs.uk

for many disorders. UK NEQAS data show that a wide range of concentrations can be obtained by immunoassays when measuring a sample. Bimodal distributions can also be observed in some female samples, which suggest that binding proteins may influence the measurement of cortisol by these assays.

Due to their structural similarities, exogenous steroids may also present a challenge for mass spectrometry-based assays. Prednisolone, in particular, has been reported to cause interferences in both immunoassays and LC-MS/MS assays,[1] and long LC runs have been used to circumvent this, resulting in assays not suitable for routine clinical use.

We developed a rapid LC-MS/MS assay for serum cortisol for routine use and share our experiences over 12 months in a large teaching hospital which provides clinical services in respiratory and transplant medicine and therefore has a large population of patients on exogenous steroids. We also investigated the effects of gender on the measurement of cortisol.

Materials and methods

Sample preparation

Stock solutions were prepared by dissolving hydrocortisone (Sigma, Poole, UK) in methanol (Sigma, Poole, UK) and were stored at −20°C for up to six months. Separate stock solutions were prepared for standards and quality control samples (QCs), both 10 mg/mL. Working standards and weighed-in QCs were prepared by diluting the stock solution with phosphate buffered saline (PBS) pH 7.4 (Sigma, Poole, UK) containing 0.1% (w/v) bovine serum albumin (BSA, Sigma, Poole, UK) to give concentrations of 0–2000 nmol/L and 100, 400 and 800 nmol/L. Aliquots (200 µL) of these were stored at −30°C for up to 6 months. In addition, commercial QCs were used (Lyphochecks, BioRad, Hemel Hempstead, UK) according to the manufacturers' instructions. Deuterated cortisol-9, 11, 12, 12 (D4) was used as an internal standard (CDN Isotopes, Quebec, Canada) at a working concentration of 0.1 mg/L (276 nmol/L) in methanol.

Standard or QC sample (20 µL) was manually pipetted directly into the well of a 96-deep well block (Thermo, Hemel Hempstead, UK). To this, 40 µL of 0.1 mol/L zinc sulphate was added. This was vortexed for 10 s followed by the addition of 100 µL of internal standard. The block was heat-sealed (Thermo, Hemel Hempstead, UK) and vortexed for 1 min, then centrifuged at 8000 *g* for 5 min. Following centrifugation, the plate was transferred directly to the autosampler for analysis; 10 µL of sample was injected into the liquid chromatography (LC) system using partial loop mode.

In addition, to investigate a method for the prompt analysis of urgent requests, samples ($n = 39$) were prepared as described above following the completion of their initial analysis and processed using the calibration curve from the original analysis. The results from the repeat analysis were then compared to the original analysis to identify if this method of adding samples to an existing calibration would provide fast reliable results for urgent samples.

Chromatography

A Waters Acquity UPLC liquid chromatograph (Waters, Manchester, UK) was used for chromatography. The mobile phases utilized were: (A) distilled water with 2 mmol/L ammonium acetate (Sigma, Poole, UK) and 0.1% formic acid (VWR International Ltd, Leicestershire, UK) and (B) methanol (LC-MS grade, Sigma, Poole, UK) with 2 mmol/L ammonium acetate and 0.1% formic acid. The sample was injected onto a Phenomenex C8 4 × 2 mm guard cartridge (Phenomenex, Macclesfield, UK) connected to a Phenomenex 30 × 2.1 mm C8 Kinetex analytical column using 30% mobile phase B at a flow rate of 0.6 mL/min. Cortisol is eluted with a 1 min linear gradient to 50% mobile phase B at 0.6 mL/min. Following this, the mobile phase composition was stepped up to 98% B at a flow rate of 0.6 mL/min. The flow was maintained at 98% B for 0.5 min before returning to 30% B for the final 0.4 min. Both guard and analytical columns were maintained at 50°C.

Mass spectrometry

The eluate was injected from the LC directly into a Quattro Premier™ Tandem mass spectrometer (Waters, Manchester, UK) with a Z spray ion source. MassLynx NT 4.1 software was used for system control and the MassLynx QuanLynx programme allowed data processing. This software used the area of the detected peaks, 1/x weighting and linear least squares regression to produce a standard curve. The 1/x weighting gives greater accuracy at lower concentrations. The mass spectrometer was operated in electrospray positive mode, the capillary was maintained at 0.3 kV and the source temperature was 140°C. The desolvation temperature and gas flow were 380°C and 800 L/h, respectively. The transitions identified were m/z 363.2> 97.15 and 363.2>121.1 for cortisol and 367.1>97.1 for D4 cortisol. The cone and collision energies were 28 V and 26 eV for cortisol and D4 cortisol, respectively. Transitions were monitored in multiple reaction monitoring (MRM) mode, with a dwell time of 0.15 s.

Validation

The assay was validated against published acceptance criteria for linearity, precision, recovery and sample stability.[2]

Ion suppression

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 1000 nmol/L solution of D4 cortisol in 50% (v/v) methanol/water directly into the mass spectrometer to give a constant background signal. Serum samples ($n = 6$) were injected simultaneously via the autosampler. A reduction in the background signal is observed when ion suppression occurrs and ion suppression is deemed significant if a reduction in signal of >10% is observed where the compound of interest elutes. Serum samples ($n = 5$) were serially diluted with PBS containing 0.1% BSA to assess the linearity of the dilutions and as a further test on ion suppression since cortisol concentrations would dilute in a non-linear manner in the presence of ion suppression.

Accuracy

The accuracy of the assay was determined by the analysis of European Reference Materials DA 192 and 193. Accurate measurement of this matrixed reference material would also confirm the lack of matrix effects given the use of PBS-based calibrators. To further confirm this, the method of standard additions was used with serum and PBS-based spiked samples with a target of within 5% of each other.

Recovery

Three different concentrations of cortisol were spiked into six different serum samples with cortisol concentrations ranging from 122 to 384 nmol/L. The concentrations of cortisol added to each sample were 97, 207 and 433 nmol/L. The recovery was calculated from measured compared to expected concentrations and absolute ion counts were compared for any variances between matrices.

Imprecision

Imprecision of the method was assessed against a range of concentrations using PBS-based and commercial QC samples. These samples were analysed daily for 12 days to calculate inter-assay imprecision. To determine intra-assay imprecision, the same samples were analysed 12 times within one batch. Percentage deviation was calculated from the difference between mean observed and nominal concentrations to assess bias.

Linearity

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 cortisol concentration in nmol/L. Calibration curves were judged linear if the correlation coefficient (R^2) was better than 0.9900 as calculated by weighted linear regression.

Limit of quantitation and detection

The lower limit of quantitation (LLOQ) was defined as the concentration for which 10 replicates of PBS-based samples prepared with low concentrations of cortisol gave a CV of less than 20% and bias of less than 20%. The lower limit of detection (LLOD) was defined as the concentration which gave a peak with a 3 to 1 signal to noise ratio.

Specificty

Solutions of various related steroids, both natural and synthetic, were prepared in 50% (w/v) methanol/water and injected directly into the mass spectrometer via the autosampler, without internal standard. Steroids tested to a final concentration of 1000 nmol/L were testosterone, dehydroepiandrosterone sulphate, dehydroepiandrosterone, oestradiol, progesterone, aldosterone, cortisone, corticosterone, 17 hydroxyprogesterone, 11 deoxycortisol, 21 deoxycortisol, pregnenolone, 19 nortestosterone, epitestosterone, dihydrotestosterone and androstenedione. Others tested to 1 mg/L were fludrocortisone, dexamethasone, methylprednisolone, budesonide, prednisone, beclometasone, cyproterone, ethinyl oestradiol, norethisterone, triamcinolone and prednisolone.This excess concentration was chosen to allow easy identification of any potentially interfering peaks. The retention times of any compounds found to give a signal in the specific channels for cortisol or D4 cortisol were subsequently assessed to determine if they would cause interference. Previous reports have shown the M + 2 isotope of prednisolone may cause an isobaric interference in the cortisol transition[1]; however, we have not found this to be a problem in a previously published salivary cortisol assay using similar equipment.[3]

Stability

The stability of cortisol has been well documented.[4] Therefore, we investigated the stability of cortisol

once the sample preparation has taken place and also the stability of the detector response upon multiple injections of an extracted sample over 16 h.

Comaparative assays and Short Synacthen® Testing

Serum samples were analysed by the LC-MS/MS method and compared to two commercially available immunoassays. These were the Roche E170 electrochemiluminescent immunoassay (ECLIA, $n = 149$) and the Abbott Architect chemiluminescent microparticle immunoassay (CMIA, $n = 148$). Following this, the data from unstimulated (pre-Synacthen®, $n = 103$) samples were investigated for gender-related differences by assessing the difference between males and females in a method comparison.

Furthermore, to evaluate the performance of our LC-MS/MS method pre-and post-synacthen, we analysed serum samples from healthy subjects which had been collected as part of another study that aimed to investigate the effects of different assay systems in the measurement of cortisol during Short Synacthen® Testing (SST).[11] The samples were collected from individuals aged 20 to 66 years and were stored at −20°C and transported frozen. SST samples ($n = 171$) from this normal population were analysed by the LC-MS/MS assay and were compared to the results of a gas chromatography mass spectrometry (GC-MS) assay at the Welsh External Quality Assurance Scheme (WEQAS) laboratory. The GC-MS assay measured total cortisol using a modified version of their isotope dilution gas chromatography-mass spectrometry (ID-GC-MS) method.[5] Interassay CVs for this assay were 5.4%, 6.2% and 6.8% at concentrations of 120, 363 and 657 nmol/L, respectively.[11]

This assay used NIST 921 for calibration and DA 192 and 193 for quality control. The measured percentage deviation from the target concentrations was 0.35% for DA 192 and 0.87% for DA 193 for the GC/MS assay. Data analysis was performed for all data, males and females, with and without oral contraceptive (OCP) use. The lower reporting limit for the 30 min SST cortisol response was estimated using the 2.5th centile of the log-transformed concentrations. A morning reference range for cortisol using the LC-MS/MS assay was derived using the 2.5th and 97.5th centiles of the 0 min data excluding samples from women on the OCP.

Data review

Chromatograms over a two-month period were inspected closely for the presence of exogenous steroids which was suggested by peaks in the chromatogram at a different retention time to cortisol in either transition. These samples were collected when available and analysed on the Abbott CMIA to investigate the degree of interference in the immunoassay. An arbitrary cut-off of immunoassay results >20% higher than the LC-MS/MS results were identified as having interference.

Statistical analysis

All statistical analyses were performed using Analyse-it software (Analyse-it Software Ltd, Leeds, UK).

Results

Both cortisol and D4 cortisol had retention times of 1.02 min (Figure 1) and a cycle time, injection-to-injection, of 2.6 min. The time to prepare and run 10

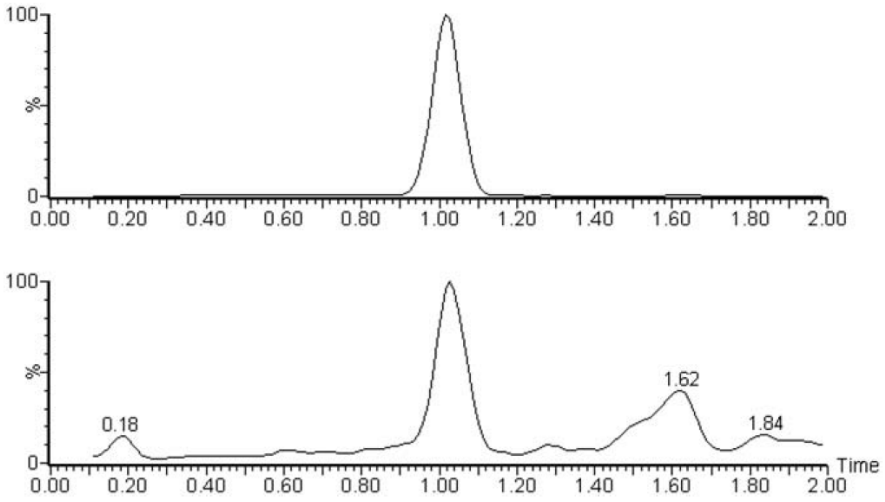


Figure 1. Chromatogram of a serum sample containing 47 nmol/L of cortisol. The upper panel shows the internal standard and the lower panel shows the cortisol quantifier transition. The *y*-axis shows relative ion count normalised to the same intensity.

patients' samples was approximately 75 min. The 363.2 to 97 transition for cortisol was used as the quantifier as the 363.2 to 121 transition displayed some interference in samples from patients taking prednisolone. This interference manifested as a shoulder on the peak at an earlier retention time to cortisol and prednisolone (both eluting at 1.02 min) and was therefore likely to be a drug metabolite. There was no significant suppression of ionisation in the region of the chromatogram where cortisol elutes. The mean recoveries following the addition of cortisol to six serum samples were 99% (range 85–123%).

Analysis of matrixed reference materials gave concentrations within 3% of the target concentrations. The method of standard additions showed that the PBS containing 0.1% (w/v) BSA and serum spiked with 0–2000 nmol/L of cortisol were within the 5% target.

The inter-assay imprecision (%CV) for the PBS-based QCs was 13%, 9% and 5% at concentrations of 100, 400 and 800 nmol/L, respectively. The inter-assay imprecision for commercial QCs was 8%, 7% and 6% at concentrations of 80, 480 and 842 nmol/L, respectively. The intra-assay imprecision was 12%, 7% and 9% for the PBS-based QCs and 9%, 6% and 5% for the commercial QCs. All means were within 4% of the PBS-based QC targets. The reference material was measured within 3% of the target values.

Standard curves were made by plotting cortisol concentrations on the x-axis and cortisol/D4 cortisol 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. Dilutional linearity experiments showed that serum can be diluted in PBS containing 0.1% BSA mean bias +1% (range −11% to +13%).

The lower limits of detection and quantitation were 5 and 12.5 nmol/L, respectively. The CV at 12.5 nmol/L concentration was 17% and the bias was +2%. None of the steroids, endogenous or exogenous, gave a response above the LLOQ at the retention time for cortisol or D4 cortisol.

A continuous series of injections over a 16-h period did not show any systematic decrease in the area counts of either cortisol or D4 cortisol and therefore did not cause any loss of sensitivity. The CV of the response (cortisol peak area/D4 cortisol peak area) was 8.1%. Extracts were stable at 4°C for 24 h as shown by a paired sample t-test result of $P = 0.22$.

The re-analysis of samples prepared independently to the calibration curve to simulate a potential mechanism for the handling of urgent requests showed a bias of −1.1 nmol/L (Figure 2). The maximum CV of the pairs was 16.5% and the difference between the two sets of results was not significant (paired sample t-test, $P = 0.89$).

A method comparison of serum samples analysed by this assay ($n = 149$) and a Roche immunoassay gave a Bland-Altman bias plot (Figure 3) with a mean bias of −101 nmol/L. Passing–Bablock analysis showed an equation of: Roche = 1.23 × LC-MS/MS − 1.12 nmol/L. The correlation coefficient (R^2) was 0.94. Four samples were excluded from the data analysis due to a large discrepancy between the results. The immunoassay cortisol concentrations were considerably higher than the LC-MS/MS assay. All these samples were taken from patients taking prednisolone. When the data (males $n = 47$, females $n = 57$) were examined for gender-related

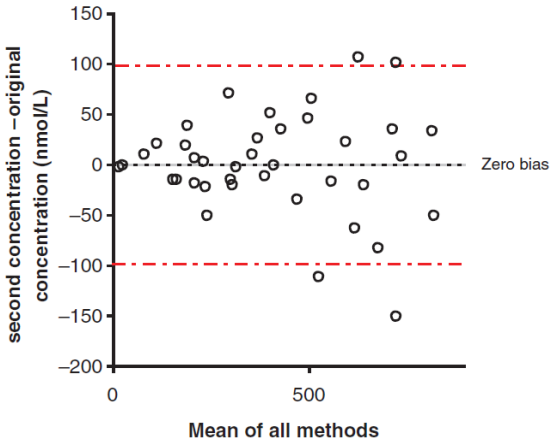


Figure 2. Bland–Altman plot comparing the results of 39 serum samples that were re-extracted and analysed for cortisol following the completion of their original analysis. The calibration of the original sample preparation was used to calculate the concentrations of the re-extracted samples.

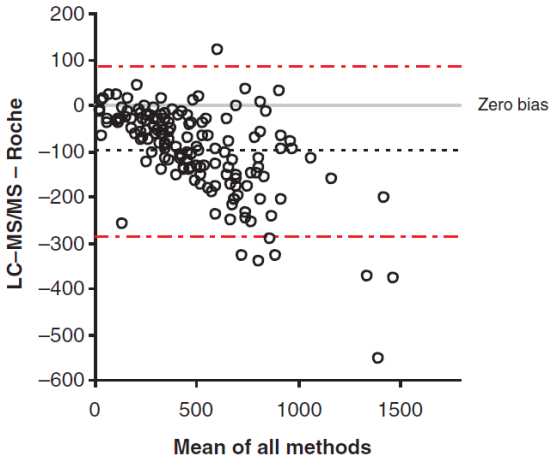


Figure 3. Bland–Altman plot comparing the LC-MS/MS assay and Roche immunoassay ($n = 149$), mean bias of LC-MS/MS assay was −101 nmol/L.

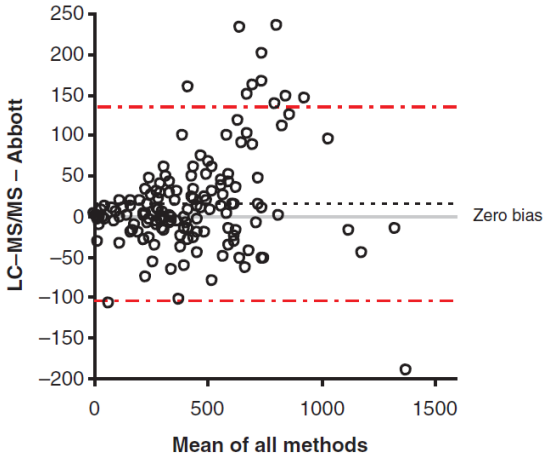


Figure 4. Bland–Altman plot comparing the LC-MS/MS assay and Abbott immunoassay ($n = 148$), mean bias of the LC-MS/MS assay was +14.5 nmol/L.

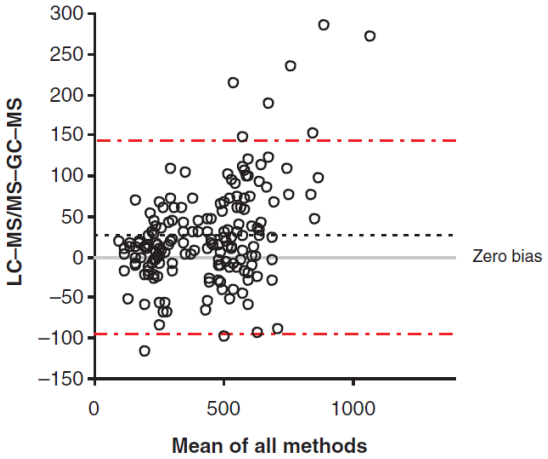


Figure 5. Bland–Altman plot comparing the LC-MS/MS assay and WEQAS GC/MS ($n = 171$), mean bias of the LCMS/MS assay was +23.6 nmol/L.

differences, following the exclusion of any samples post Synacthen®, none were found. The Passing–Bablock method comparison for males was: Roche = 1.25 × LC-MS/MS −9.82 and for females was: Roche = 1.24 × LC-MS/MS + 7.43.

A method comparison using the Abbott immunoassay gave a Bland-Altman bias plot (Figure 4) with an average bias of +15 nmol/L. Passing–Bablock regression analysis gave the following equation; Abbott = 0.94 × LC-MS/MS + 11.97. The correlation co-efficient (R^2) was 0.96. The same four samples showed discrepancy between LC-MS/MS and the Abbott immunoassay, although to a smaller extent, and were excluded from the data analysis. When the data (males $n = 44$, females $n = 59$) were examined for gender-related differences, following the exclusion of any samples post Synacthen®, both constant and proportional biases were observed with the assay giving slightly lower results in most female samples. The Passing–Bablock method comparison for males was: Roche = 0.98 × LC-MS/MS + 6.45 and for females was: Roche = 0.94 × LC-MS/MS + 13.13.

A comparison with a GC-MS assay ($n = 171$) for cortisol available at the WEQAS laboratory in Cardiff gave a Bland-Altman bias plot (Figure 5) with a mean bias of 24 nmol/L for all samples pre- and post-Synacthen®. The Passing and Bablock analysis gave the following equation: LC-MS/MS = 1.11 × GC-MS – 22.90 ($R^2 = 0.92$). Examining the data for SST samples showed the LC-MS/MS assay had a mean bias of 10.5 nmol/L and Passing–Bablock equation of LC-MS/MS = 1.08 × GC-MS–10.6 nmol/L for the pre-Synacthen® samples ($n = 87$). For post- Synacthen® (30 min) samples, the LC-MS/MS had a mean bias of 37.2 nmol/L and a Passing–Bablock equation of LC-MS/MS = 1.18 × GC-MS – 64.9. When examining the data for gender related differences, none were observed (males $n = 74$, females $n = 97$). Furthermore, similar comparisons were observed when comparing male samples ($n = 74$) with female samples, both on ($n = 13$) and off the OCP pill ($n = 84$).

The 30 min lower reporting limit for cortisol concentration post-Synacthen® for males and females who were not taking the OCP was 417 nmol/L ($n = 83$, 45 females and 38 males). There was an insufficient number of samples ($n = 14$) from women taking the OCP to enable a specific cut-off to be used in this group. The reference range of the 2.5th and 97.5th centiles for morning samples pre-Synacthen® excluding samples that were taken from women on the OCP was 108–491 nmol/L.

Following the review of two months of requests for cortisol measurement, 128 samples (of the total requests, i.e. 33%) were found to show evidence of exogenous steroid use. These were identified as having extra peaks on the chromatogram at a different retention time from the cortisol peak in either the quantifier or the qualifier transitions or shoulder on the qualifier peak, as is the case for prednisolone metabolites. No samples exhibited shoulders on the quantifier peaks. Of these, 105 samples were available for further investigation. The Bland–Altman analysis of the LC-MS/MS assay compared to the Abbott immunoassay (Figure 6). Of the 105 samples, 24 samples displayed a cortisol concentration determined by the immunoassay as >20% higher than the concentration determined by LC-MS/MS. These samples represented 6% of all cortisol requests.

Discussion

We have developed a rapid LC-MS/MS assay for serum cortisol analysis that is suitable for routine clinical use.

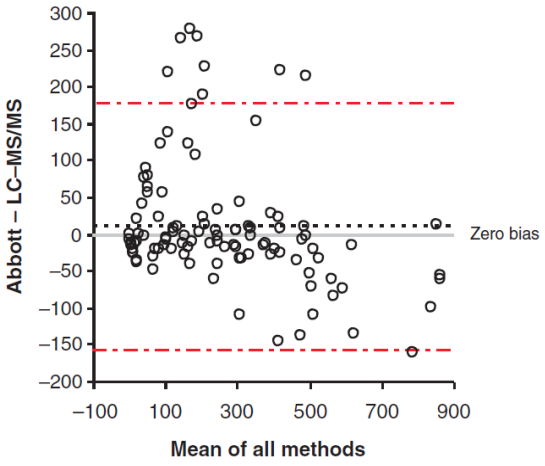


Figure 6. Bland–Altman bias plot comparing results of the LC-MS/MS and Abbott immunoassays of samples ($n = 105$) from patients samples that showed evidence of exogenous steroid use.

There have been LC-MS/MS assays for serum cortisol reported[1,6,7] but these are not suitable for routine use due to the long run times and extensive sample preparation. As a result, the uptake of LC-MS/MS assay has been poor compared to other steroids and consequently there is only one routine LC-MS/MS assay registered with UK NEQAS and none are registered with the WEQAS. The sample preparation in our LC-MS/MS assay involves a simple protein precipitation technique which takes approximately 10 min to prepare a run of six standards, quality controls and 12 samples, a typical daily run size for our large teaching hospital laboratory. However, we tend to prepare two smaller batches to minimize the turnaround time of these samples. Extracting samples and analysing them after a run has finished with the previous existing calibration still gave acceptable results. This approach could be used for handling urgent samples out of hours as this would take less than 10 min hands on time to prepare the samples and report the results. There would be also a small amount of 'walk away time' during the centrifugation and analysis times. We estimate this to be less than 10 min for one sample with the addition of 2.6 min for subsequent samples.

In 12 months, we have analysed 2860 samples for cortisol requiring three guard and analytical columns. The use of a C8-fused core column gave adequate separation without extensive run times as it allows rapid re-equilibration. The sample volume of only 20 µL is advantageous for small samples, particularly paediatric samples. The assay demonstrated excellent performance during the method validation process and had CVs comparable or less than those observed for the Abbott immunoassay despite manual sample preparation. The LLOQ of 12.5 nmol/L also rivals those obtained by immunoassays despite the fact that the LC-MS/MS assay was developed for speed of analysis rather than sensitivity.

The high degree of specificity conferred by the use of tandem mass spectrometry coupled with chromatography on a C8 column ensured that none of the endogenous or exogenous steroids tested resulted in any observed interferences. Prednisolone co-elutes with cortisol in this assay, however, due to different fragmentation patterns of prednisolone and cortisol, high concentrations (exceeding those found in individuals on therapy) of prednisolone do not cause a measurable peak in either cortisol transitions. A metabolite of prednisolone can be observed as an unresolved peak in the 363.2 to121 transition of cortisol but not in the 363.2 to 97 transition, and hence this was used for quantitation. This promotes the use of this LC-MS/MS assay in patients receiving prednisolone therapy and has previously been shown to be essential in patients with altered steroid metabolism such as those patients receiving metyrapone therapy.[8,9]

The two routine immunoassays for cortisol demonstrated good comparisons with the LC-MS/MS assay; however, the Roche assay tended to measure cortisol 23% higher than the LC-MS/MS assay, but no gender-related differences were observed using this assay. However, these data may include samples from women taking the OCP and therefore if these are excluded, differences may become apparent. The Abbott immunoassay gave the best overall comparison of the two immunoassays but appeared to underestimate the cortisol concentrations in female samples. This supports the differences observed between these two immunoassays after the analysis of UK NEQAS samples from females. This may be due to the inability of some immunoassays to completely remove cortisol from its binding protein for measurement. This effect would be more pronounced in females due to the influence of oestrogens on binding protein concentrations.[10]

The LC-MS/MS assay compared well to the WEQAS reference ID-GC-MS assay.

No differences were observed for males, females or females taking the OCP. This assay may therefore also be used in females on the OCP to give a more accurate result than some immunoassays and some samples have been analysed from referral laboratories for this purpose. A slight difference was observed when comparing samples pre- and post-synacthen®.

For SST, a lower reporting limit of 417 nmol/L is proposed when using this assay as this was the 2.5th centile of the log transformed 30-min cortisol concentrations during the synacthen® testing of normal individuals. This is similar to what was achieved using the ID-GC-MS assay. This is considerably lower than what is currently quoted for immunoassays (500–550 nmol/L). However, as the immunoassays compare variably to this

assay, method-specific reference ranges should be employed.[11]

A review of two months chromatograms at the end of the first year of routine use of this assay resulted in a large proportion of samples showing evidence of exogenous steroid use. Of these samples, many gave a falsely elevated result when measured by immunoassay which could cause in misclassification, misdiagnosis or incorrect treatment of patients.

Declaration of conflicting interests

None.

Funding

None.

Ethical approval

The synacthen testing study protocol was approved by the South East Wales Research Ethics Committee (clinical trial registration number: NCT00851942).

Guarantor

LJO.

Contributorship

The developmental and validation of the assay described was performed by LJO and JEA. SD supported the interference studies. FM contributed to the conception of part of the work. BGK supervised the work and made the collaboration. SN, NA-F, DD, CE and DAR contributed to the method comparisons. All authors have contributed to the manuscript.

References

1. Ionita DM and Fast F Akhlaghi. Development of a sensitive and selective method for the quantitative analysis of cortisol, cortisone, prednisolone and prednisone in human plasma. *J Chromatogr B Analyt Technol Biomed Life Sci* 2009; 877: 765.
2. *Guidance for Industry. Bioanalytical Method Validation*. US Department of Health and Human Services. Food and Drug Administration, Center for Drug Evaluation and Research (CDER) and Center for Vetinary Medicine (CVM). May 2001.
3. Owen LJ, Haslam S, Adaway JE, Wood P, Glenn C and Keevil BG. A simplified liquid chromatography tandem mass spectrometry assay, using on-line solid-phase extraction, for the quantitation of cortisol in saliva and comparison with aroutine DELFIA method. *Ann Clin Biochem* 2010; 47: 131–6.
4. Diver MJ, Hughes JG, Hutton JL, West CR and Hipkin LJ. The long-term stability in whole blood of 14 commonly-requested hormone analytes. *Ann Clin Biochem* 1994; 31: 561–5.
5. Woods L, Ducroq DH, Fraser HL, et al. Measurement of urinary free cortisol by tandem mass spectrometry and comparison with results obtained by gas chromatography-mass spectrometry and two commercial immunoassays. *Ann Clin Biochem* 2008; 45: 380–8.
6. Ruiter AF, Teeninga N, Nauta J, Endert E and Ackermans MT. Determination of unbound prednisolone, prednisone and cortisol in human serum and saliva by on-line solid-phase extraction liquid chromatography tandem mass spectrometry and potential implications for drug monitoring of prednisolone and prednisone in saliva. *Biomed Chromatogr* 2012; 26: 789–96.
7. Fanelli F, Belluomo I, Di Lallo VD, et al. Serum steroid profiling by isotopic dilution-liquid chromatography-mass spectrometry: comparison with current immunoassays and reference intervals in healthy adults. *Steroids* 2011; 76: 244–53.
8. Owen LJ, Halsall DJ and Keevil BG. Cortisol measurement in patients receiving metyrapone therapy. *Ann Clin Biochem* 2010; 47: 573–5.
9. Monaghan PJ, Owen LJ, Trainer PJ, Brabant G, Keevil BG and Darby D. Comparison of serum cortisol measurement by immunoassay and liquid chromatography–tandem mass spectrometry in patients receiving the 11β–hydroxylase inhibitor metyrapone. *Ann Clin Biochem* 2011; 48: 441–6.
10. Klose M, Lange M, Rasmussen AK, et al. Factors influencing the adrenocorticotropin test: role of contemporary cortisol assays, body composition, and oral contraceptive agents. *J Clin Endocrinol Metab* 2007; 92: 1326–33.
11. El–Farhan N, Pickett A, Ducroq D, *et al*. Method–specific serum cortisol responses to the adrenocorticotrophin test: comparison of gas chromatography–mass spectrometry and five automated immunoassays. *Clin Endocrinol* 2013; 78: 673–680.

Clinical Endocrinology (2014) **80**, 766–771

LETTERS TO THE EDITOR

Response to 'Determining the utility of the 60-min cortisol measurement in the short Synacthen test'

Dear Sir,

We write with respect to the recent article by Chitale *et al.*[1] Determining the utility of the 60-min cortisol measurement in the short Synacthen test (SST). The authors suggest that a significant proportion of people undergoing SSTs may be inappropriately diagnosed as having adrenal insufficiency if a 30-min cut-off of 550 nM is used and that a 60-min sample is required to improve the specificity of the SST. However, the authors have not considered the effect of cortisol immunoassay bias when interpreting their results.

We[2] and others[3,4] have shown that cortisol immunoassay bias has a significant effect on the cortisol response to Synacthen; whilst the historical 550 nM cut-off continues to be widely used by clinicians within the UK,[5] we would suggest that adopting a method-specific lower reference limit (Table 1) would improve the specificity of the SST, potentially avoiding the need for a further sample at 60 min.

Furthermore, we believe there is an error in the publication relating to the quoted bias of the cortisol assays. In fact, the UKNEQAS 2011 steroid annual review reports the method bias for the Roche Elecsys as +7·8% and for the Siemens Advia Centaur as +0·2%, derived by comparison with the all laboratory trimmed mean. These biases appear to have been transposed in the manuscript. It is also important to note that the percentage bias does not represent an individual laboratory's assay bias, nor is it a once and for all measure of assay bias. The all laboratory trimmed mean is continually updated over time and so method bias as compared to this can change as assays become relatively more or less popular due the relative weightings, or are re-formulated either actively or passively. There is also no mention of the reformulated Siemens assay released in 2010, with uptake by the majority of laboratories by mid-to-late summer 2010. This makes it likely that analysis of cortisol samples from the SSTs carried out at KCH between January and October 2010 was undertaken using a mixture of both assays.

We have previously investigated the bias of all three cortisol immunoassays likely to have been included in this study when compared to the gold standard GC-MS. This shows a significant gender-dependent positive bias for the Roche assay (mean bias ratio 1·36 in male samples and 1·25 in female samples). The original Siemens assay had a lesser positive bias (mean bias ratio 1·20 in male samples and 1·16 in female samples), which decreased further for the reformulated assay (mean bias ratio 1·08 in male samples and 1·05 in female samples).

In summary, we advocate the use of appropriate method and gender (in the case of the Roche assay)-specific lower limits for the cortisol response to Synacthen to ensure its correct interpretation. We believe that the use of such cut-offs would improve the specificity of the SST without the need for a 60-min sample, although further evidence for this is needed. We suggest it may be valuable for the authors to revisit the 30-min cortisol responses using our quoted reference limits and evaluate the impact of this on their findings.

Table 1. Assay-specific estimated lower reference limits for post-ACTH cortisol in healthy volunteers

Assay	Males	Females
GC-MS	418	421
Centaur (post mid 2010)	448	446
Centaur (premid 2010)	496	502
Roche	574	524

The estimated lower reference limit was determined by back transformation of the 2·5th percentile value (mean − 1·97*SD) of the log-transformed data (81 female subjects and 60 male). Results are expressed in nM.

Disclosure statement

The authors have no conflicts of interest to declare. No grants or fellowships were received supporting the writing of this paper.

Carol Evans*, Nadia El-Farhan†, David Ducroq‡ and Aled Rees§

**Department of Medical Biochemistry, University Hospital of Wales Heath Park, Cardiff, †Biochemistry Department, Royal Gwent Hospital, Newport, ‡Quality Laboratory, WEQAS, Cardiff and §Centre for Endocrine and Diabetes Sciences, Institute for Molecular and Experimental Medicine, School of Medicine, Cardiff University, Cardiff, UK*

E-mail: carol.evans9@wales.nhs.uk

[Correction added on 26 July 2013, after first online publication: Finlay MacKenzie and his related affiliation was removed from the author list.]

doi: 10.1111/cen.12272

References

1 Chitale, A., Musonda, P., McGregor, A.M. *et al.* (2013) Determining the utility of the 60 min cortisol measurement in the short synacthen test. *Clinical Endocrinology*, **79**, 14–19.

2 El-Farhan, N., Pickett, A., Ducroq, D. *et al.* (2013) Method-specific serum cortisol responses to the adrenocorticotropin

test: comparison of gas chromatography-mass spectrometry and five automated immunoassays. *Clinical Endocrinology*, **78**, 673–680.
3 Clark, P.M., Neylon, I., Raggatt, P.R. *et al.* (1998) Defining the normal cortisol response to the short synacthen test: implications for the investigation of hypothalamic-pituitary disorders. *Clinical Endocrinology*, **49**, 287–292.
4 Klose, M., Lange, M., Rasmussen, A.K. *et al.* (2007) Factors influencing the adrenocorticotropin test: role of contemporary cortisol assays, body composition and oral contraceptive agents. *Journal of Clinical Endocrinology and Metabolism*, **92**, 1326–1333.
5 Chatha, K.K., Middle, J.G. & Kilpatrick, E.S. (2010) National UK audit of the short Synacthen(r) test. *Annals of Clinical Biochemistry*, **47**, 158–164.

Response to: 'Determining the utility of the 60 min cortisol measurement in the short synacthen test'

Dear Sir,

With respect to the recent article by Chitale *et al.* – Determining the utility of the 60 min cortisol measurement in the short synacthen test,[1] I agree that despite many years of clinical use, the interpretation of the short synacthen test (SST) is still debated. The adrenal response to synacthen is a continuous variable which is dichotomized into adequate/inadequate by the decision limit used in the SST. As the authors rightly point out the only time point that has been validated to a 'gold standard' is the 30 min cortisol sample.

In this study, the authors fail to verify the status of the patients included in the study, they have arbitrarily classified those patients who recorded levels of cortisol <550 nmol/l at 30 min and >550 nmol/l at 60 min as false positives, failing the synacthen test but having 'normal' adrenal reserve; however, this classification is based solely on the index test without mention of a reference test.

The SST is a screening test for adrenal insufficiency. The decision limit chosen is a balance between sensitivity and specificity. This balance depends on the relative importance of missing a diagnosis, false negatives, versus the burden of false positives. Screening tests are generally biased towards sensitivity, as it is deemed more important not to miss a diagnosis, at the cost of accepting lower specificity. Raising the decision limit typically increases the specificity of a test at the expense of lower sensitivity, some patients with the condition will be missed.[2] In those cases where the clinical findings and laboratory testing do not lever enough evidence to make a diagnosis, a second tier test with higher specificity should be performed – in the case of suspected adrenal insufficiency either a metyrapone test or insulin tolerance test (ITT).

I would caution against calculating the sensitivity and specificity of the cortisol thresholds using the other time point as the 'gold standard' as these values are not independent of each other.[3] As the authors point out no patient who passed the test at 30 min, failed at 60 min and this is reflected in the statistics presented in Tables 3 and 4.

Guidelines are available for the reporting of diagnostic accuracy studies, and the STARD initiative provides a checklist for authors.[4]

Richard I. King
Department of Clinical Biochemistry, Mater Pathology, South Brisbane, Qld, Australia
E-mail: richard.kingnz@gmail.com

doi: 10.1111/cen.12289

References

1 Chitale, A., Musonda, P., McGregor, A.M. *et al.* (2013) Determining the utility of the 60 min cortisol measurement in the short synacthen test. *Clinical Endocrinology*, **79**, 14–19.
2 Florkowski, C.M. (2008) Sensitivity, specificity, receiver-operating characteristic (ROC) curves and likelihood ratios: communicating the performance of diagnostic tests. *Clinical Biochemist Reviews*, **29** (Suppl. 1), S83–S87.
3 Greenhalgh, T. (1997) How to read a paper: papers that report diagnostic or screening tests. *British Medical Journal*, **315**, 540–543.
4 Bossuyt, P.M., Reitsma, J.B., Bruns, D.E. *et al.* (2003) Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *British Medical Journal*, **326**, 41–44.

The Authors' Reply

Dear Prof. Bevan,

We are very grateful to be given the opportunity to reply to the letters by Evans et al. and King about this study.

Evans *et al.* mention that there are several potential sources for error and misinterpretation when using cortisol readings, unadjusted for either gender or assay type. However, in clinical practice, cortisol measurements are often reported as only 'time zero', 'time 30 min' and 'time 60 min', with no additional data on how these should be adjusted according to gender and assay, and thus, how these values should be interpreted without these sources of bias being known. Most decisions would be based only on the clinical interpretation of the three provided values. Therefore, whilst we accept that there are methodological influences on the actual values provided by the test; to the jobbing clinician, these may play a minor role in influencing individual patient management. We would welcome a debate as to whether the additional information on gender and assay differences should be provided by different laboratories for every short synacthen test report to see if this changes treatment decisions. This, however, would necessitate a prospective study to address the validity of the proposed upper limits of normal for the different assays.

We acknowledge that our study was also limited by the fact it was a retrospective analysis. As Evans *et al.* describe, there was a reformulation of the Siemens assay during the time, the short synacthen tests were carried out, and this could have influenced

Original Article

The Association for Clinical Biochemistry & Laboratory Medicine

Annals of Clinical Biochemistry
2014, Vol. 51(3) 379–385

DOI: 10.1177/0004563213514567
acb.sagepub.com
SAGE

The effect of serum matrix and gender on cortisol measurement by commonly used immunoassays

AJ Dodd[1], DH Ducroq[2], SM Neale[1], MP Wise[3], KL Mitchem[4], A Armston[5], JH Barth[6], N El-Farhan[7], DA Rees[8] and C Evans[1]

Abstract

Background: Considerable intermethod bias has been observed between cortisol immunoassays, with some also displaying a gender difference. Cortisol immunoassay performance is affected by serum matrix effects such as changes in steroid binding proteins and presence of interfering steroids which can be altered in various clinical settings. This study investigates cortisol immunoassay bias in pregnancy, renal failure and intensive care patients.

Methods: Serum remaining after routine analysis from pregnant patients, patients on the intensive care unit and patients with renal failure were obtained prior to disposal and used to create 20 anonymous samples per group. A male and female serum pool was prepared and spiked with cortisol. Samples were aliquoted and distributed to four hospitals for cortisol analysis by immunoassays from four different manufacturers. Cortisol was also measured by an isotope dilution-gas chromatography–mass spectrometry method for comparison of assay bias.

Results: Differences in cortisol immunoassay bias were observed across the different patient groups. A negative bias compared to pooled serum samples was observed for pregnancy serum, whilst a more positive bias was seen in renal failure and intensive care patients. Variation in bias was greatest in renal failure with the Roche E170 the most affected and the Abbott architect the least (interquartile ranges 44% and 14%, respectively).

Conclusions: Cortisol immunoassay bias may be affected by gender and differences in serum matrix from patients with various clinical conditions. Users of cortisol assays should be aware of differing matrix effects on their assay and the relevance of these for the interpretation of clinical results.

Keywords

Steroid hormones, endocrinology, immunoassay

Accepted: 15th October 2013

Introduction

National External Quality Assurance schemes demonstrate significant bias between cortisol immunoassays[1,2] with a higher degree of interlaboratory variability apparent for pooled female samples largely due to the behaviour of the Roche Elecsys/E170 assay.[3] Whilst studies evaluating the cortisol response to adrenocorticotropic hormone (ACTH) stimulation have all

[1]Department of Medical Biochemistry and Immunology, University Hospital of Wales, Cardiff, UK
[2]Quality Laboratory, WEQAS, Cardiff, UK
[3]Adult Critical Care, University Hospital of Wales, Cardiff, UK
[4]Department of Clinical Biochemistry, Prince Charles Hospital, Merthyr Tydfil, UK
[5]Department of Lab Medicine, University Hospital Southampton, Southampton, UK
[6]Blood Sciences, Old Medical School, Leeds Teaching Hospitals Trust, Leeds, UK
[7]Department of Medicine, Royal Gwent Hospital, Newport, UK
[8]Centre for Endocrine and Diabetes Sciences, School of Medicine, Cardiff University, Cardiff, UK
Corresponding author:
Carol Evans, Department of Medical Biochemistry and Immunology, University Hospital of Wales, Heath Park, Cardiff CF14 4XN, UK.
Email: Carol.evans9@wales.nhs.uk

demonstrated assay-dependent differences in the stimulated cortisol response,[4–6] the findings on gender-related differences have been conflicting.[5,6] Clark et al.[6] found that the response of serum cortisol to ACTH was gender-dependent although the extent of this varied between methods, whilst Klose et al.[5] did not observe any effect of gender. Previously, we were able to confirm that mean cortisol concentrations at baseline and post-ACTH did not differ between men and women using a gold standard gas chromatography–mass spectrometry (GC-MS) method.[4] However, gender-specific differences were found when cortisol was measured utilizing some immunoassays, with a clinically significant effect being found for the Roche E170 assay. This suggests that analytical factors specific to the immunoassays rather than genuine gender-related physiological differences are the cause of these observations. Differences in immunoassay bias relative to GC-MS have been reported between females taking or not taking oestrogen containing oral contraceptives (OCPs).[4] Matrix effects on cortisol immunoassay have also been demonstrated in patients in intensive care. Briegel et al.[7] demonstrated variation in cortisol immunoassay bias relative to GC-MS in patients with septic shock compared to those attending outpatients. Similarly, Barnes and Swaminathan[8] noted under-recovery of cortisol in a patient group with hypoalbuminaemia using the Bayer Advia Centaur assay.

It has long been recognized[9] that the specificity of a steroid immunoassay is determined not only by cross reactivity to steroids in serum but also the presence of steroid binding proteins and the affinity of the antibody used. There is increasing awareness that assay bias should be considered when interpreting cortisol measurements,[4–6] but less information is available on the effect of variability in the serum matrix on cortisol measurement by immunoassay.

The aims of the current study were (1) to investigate measurement of cortisol by immunoassay in male and female serum and (2) to use relative bias as compared to GC-MS to investigate the effect of different serum matrices on immunoassay performance.

Methods

Sample preparation

Surplus serum was selected prior to routine disposal from specimens previously collected from adult patients attending the University Hospital of Wales. A total of 70 samples were prepared for analysis: five spiked samples from both a male and a female serum pool, and 20 each from pregnancy, renal failure and intensive care patients. All samples were mixed thoroughly after preparation, aliquoted and stored at −20°C until analysis. All samples were anonymised prior to analysis.

Patient pools

Electronic patient records were searched for evidence of steroid use. Serum from any patient taking steroids was excluded from use in pooled samples. Two pools were created, one female and one male. For females, only samples where follicle stimulating hormone and luteinising hormone results were available were included. Pregnancy or oral contraceptive pill use were excluded by only including specimens with gonadotrophins within the range 1–10 U/L from women under 40 years of age, or a follicle stimulating hormone greater than 30 U/L in a woman over 50 years of age.

Each pooled sample was mixed by inversion for 30 min and split into five aliquots. Cortisol solution for spiking was prepared by dissolving Sigma hydrocortisone (H4001 1G, lot number 061M1142V) in methanol at a concentration of 0.1 mg/mL. This was diluted 1/10 with phosphate-buffered saline (PBS) to spike cortisol concentration in aliquots by 0, 100, 200 and 400 nmol/L. The final aliquot was spiked with cortisol by 800 nmol/L using 0.1 mg/mL in methanol, to ensure all spiking volumes were of 1% of pool volume or less. All volumes added were confirmed gravimetrically. Samples were then mixed by inversion for a further hour. Recoveries of cortisol by GC-MS in spiked patient pools were all within 6% of target, confirming accuracy of spike concentrations.

Pregnancy

Samples from second trimester Down's syndrome screening were collected immediately prior to routine disposal. Samples were pooled pairwise to provide sufficient volume for the study. Twenty of these two patient pools were produced.

Intensive care

Samples from 20 patients in the intensive care unit (ICU) at the University Hospital of Wales, with albumin <20 g/L at the time of collection, were collected immediately prior to routine disposal. Ten male and 10 female patients were included. Due to reduced haematocrit, many samples were of sufficient volume without pooling; where this was not the case, multiple samples taken from the same patient within 48 h were pooled to provide sufficient volume. The patients' notes were examined on the ward to ensure no patients were included who had recently received steroids.

Renal

Samples from patients who had a creatinine >300 µmol/L for at least three months were identified and collected immediately prior to routine disposal. Multiple samples from the same patient within 48 h were pooled as available; otherwise, samples were pooled pairwise as necessary for volume. The patients' electronic records were examined for evidence of steroid use and samples from any patients on steroids excluded from use.

Welsh External Quality Assessment Scheme data

Bias information for each assay were taken from data provided by Welsh External Quality Assessment Scheme (WEQAS) for distributions 171, 172, 173, 176, 184 and 185. All samples included ($n = 10$) were single patient samples with no exogenous steroids.

Analytical methods

Total cortisol was measured by GC-MS at the WEQAS Reference Laboratory using a modified version of their GC-MS reference method.[10] The bracketed standard curve used in the reference method was replaced with a conventional six-point standard curve. Total cortisol was also measured by four automated immunoassays in four separate hospitals: Abbott Architect (Abbott Laboratories, Illinois), Beckman Access (Beckman Coulter, Brea, CA), Advia Centaur (Siemens AG, Erlangen, Germany) and the Modular Analytics E170 (Roche, Mannheim, Germany). All laboratories performing analysis were enrolled in the UK NEQAS scheme for cortisol and were performing within their method group at the time of analysis. Inter-assay coefficients of variation for quality control data around the time of analysis were 4.7%, 3.3% and 5.7% at cortisol concentrations of 107.6, 355.7 and 634.8 nmol/L for GC-MS; 10.8%, 5.8% and 4.5% at 66.5, 540.3 and 915.1 nmol/L for the Roche E170; 5.1%, 3.6% and 3.1% at 76.7, 327.7 and 720.8 nmol/L for the Abbott architect; 6.1% and 6.0% at 145.5 and 954.4 nmol/L for the Siemens Centaur and 6.7%, 5.6% and 3.9% at 119.4, 432.7 and 973.4 nmol/L for the Beckman Access, respectively.

Statistics

All statistical analysis was performed with IBM SPSS Statistics version 21. Assay-specific bias ratios were calculated by dividing the cortisol immunoassay result by the matched GC-MS result. Variance in bias relative to GC-MS across groups was compared with the Levene statistic for each assay. Due to differences in sample size and variance between groups, bias between all patient groups were compared for each method with the Kruskal–Wallis non-parametric test for independent samples. The sources of difference were determined *post hoc* with the Mann–Whitney U test.

Results

Patient pools

Recovery of cortisol in male and female pooled serum is shown in Figure 1. Increased inter-assay variation in cortisol recovery is seen in the female compared to the male pool. Average percentage recovery of cortisol in female and male serum was 126% and 109% for Roche E170, 87% and 110% for Abbott architect, 106% and 100% for Siemens Centaur and 93% and 99% for Beckman Access cortisol immunoassays, respectively. When actual measurements were compared to GC-MS, the closest agreement with GC-MS was found for the male pool using the Beckman Access immunoassay and the largest (positive) bias observed for the male pool using the Roche E170 method.

Mean assay bias relative to GC-MS in the spiked patient pools is shown in Table 1. Differences in assay bias compared to previously reported values[4] were obtained. For the Abbott architect assay, there was a decrease in assay bias between 2010 when the previous samples were assayed and the current study with analysis in 2013.

Patient serum samples

Differences in cortisol assay bias compared to GC-MS across the groups studied are shown in Figure 2 and Table 1. Pregnancy serum samples all showed a greater negative bias for all assays compared to pooled serum (Figure 2). This was particularly marked for the Beckman Access, which had either no bias or a slight positive bias in all groups except pregnancy, where a strong negative bias was seen. This bias in pregnant patients appeared to be greater at higher concentrations of cortisol for all assays except the Roche E170, which had a relatively constant bias (Figure 3).

Variation in bias was greatest in serum from patients with end-stage renal disease, and to a lesser degree in the ICU patient serum, than in the relatively healthy patient groups (patient pools and pregnancy; also WEQAS distributions) for all assays (Figure 2). The proportional bias did not show a clear trend at different concentrations of cortisol for either ICU or renal patients (Figure 3). In general, assays over-recovered in these patient groups relative to GC-MS; the least affected by this was the Abbott architect, and the most affected was the Roche E170, although all assays appeared to show more variation, and greater

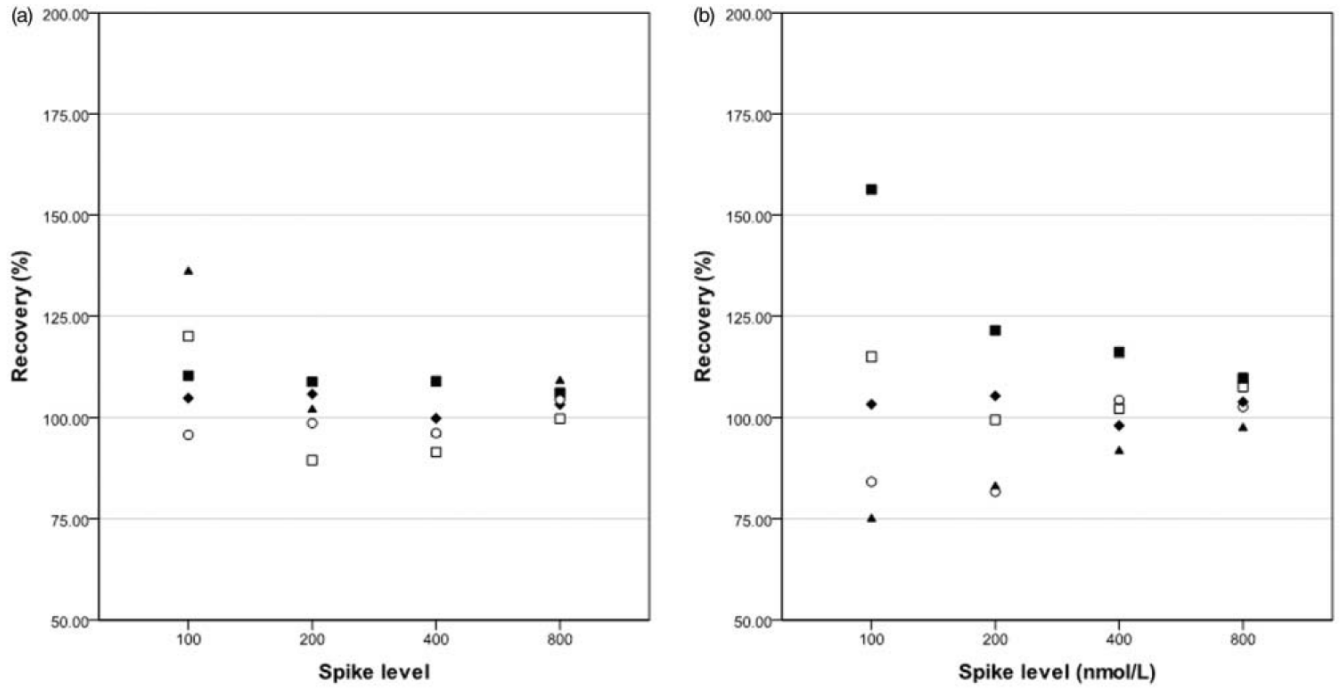


Figure 1. Percentage recovery of cortisol spiked into pools of (a) male and (b) female serum. ◆: GC-MS; ■: Roche E170; ▲: Abbott architect; □: Siemens centaur; ○: Beckman access. Recoveries are calculated relative to cortisol measured in unspiked pool, as determined by each assay.

Table 1. Mean bias ratios for cortisol immunoassays compared to GC-MS in different patient groups.

	Roche E170	Abbott architect	Siemens centaur	Beckman access
WEQAS distributions	1.19	0.95	1.23	1.02
Spiked patient pools (Total)	1.26	0.94	1.17	0.98
Male	1.28	0.95	1.19	1.00
Female	1.23	0.93	1.14	0.95
Pregnant patients	1.14 (0.002)	0.81 (0.000)	1.05 (0.004)	0.74 (0.000)
ICU patients	1.45 (0.16)	0.96 (0.312)	1.19 (0.538)	1.12 (0.008)
Renal patients	2.03 (0.000)	1.08 (0.008)	1.57 (0.000)	1.29 (0.000)

positive bias, in renal patients compared to the pools or pregnant patients. *Post hoc* Mann–Whitney statistics did not indicate a significant difference in bias between the spiked pools and ICU groups for the Roche E170, Abbott architect and Siemens Centaur ($p = 0.16$, 0.31, 0.54, respectively).

Discussion

There is increasing recognition that interpretation of cortisol measurements requires knowledge of the assay bias[4–6] and for some immunoassays there will be significant gender effects to consider as well.[3–5] This study and others[4–8] clearly demonstrate that physiological and pathological conditions which affect serum matrix can have an effect on cortisol assay bias. For all immunoassays studied, cortisol bias compared to the gold standard GC-MS was more negative in pregnant individuals than in non-pregnant healthy volunteers or pooled normal female serum. Previous studies[4,6] have demonstrated a similar trend in the serum of women taking the OCP pill compared to other female volunteers. In pregnancy and women taking the OCP, this is likely to be an effect of increased concentrations of cortisol binding globulin (CBG). Cortisol immunoassays measure total cortisol (free cortisol and cortisol bound to CBG and albumin). To do this, the assay manufacturers must include a step to displace cortisol from its binding proteins. In most cases, this is commercially sensitive and not included in the assay

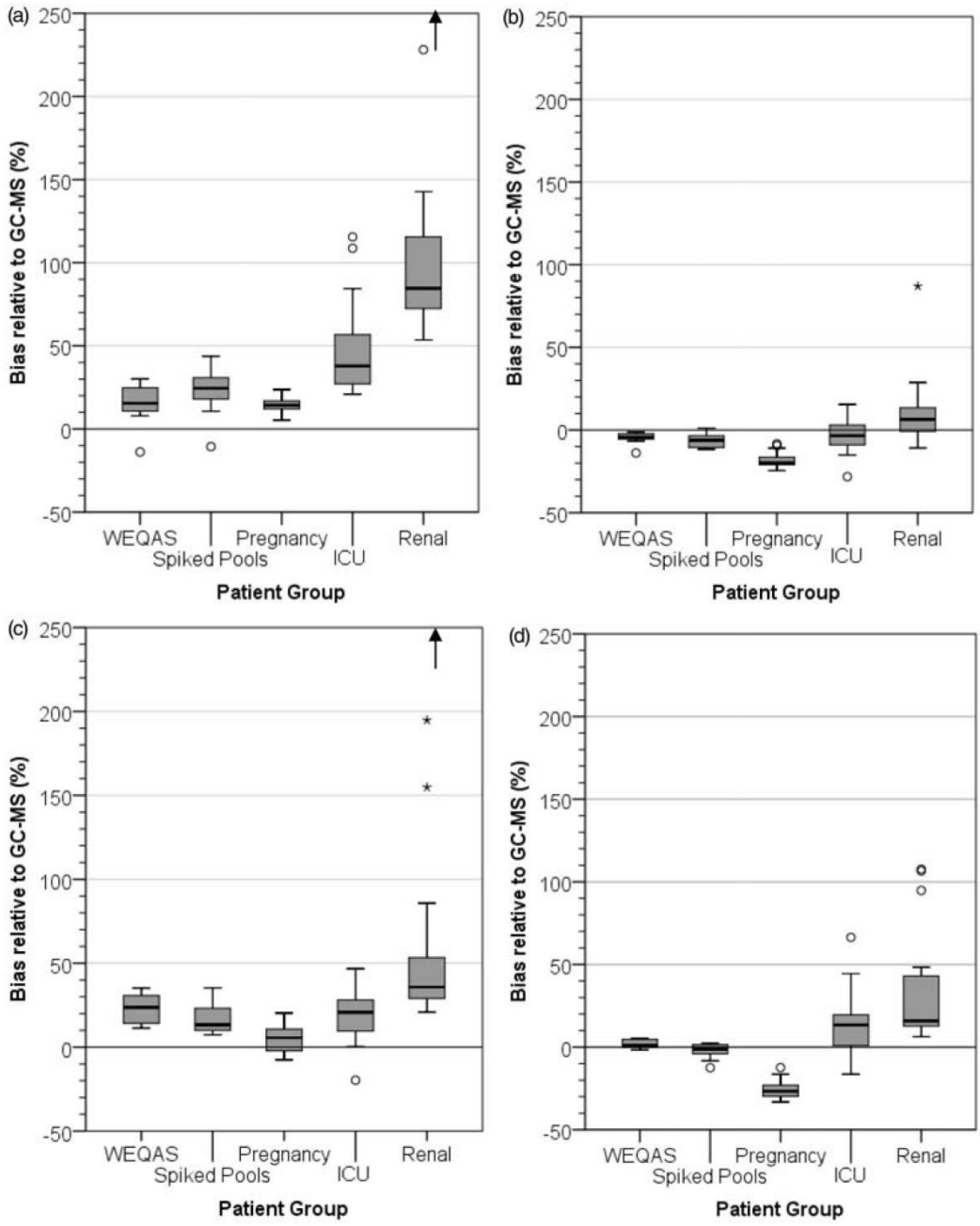


Figure 2. Box and whisker plots of bias between immunoassay and GC-MS showing marked effect due to the clinical source of patients' samples. Each plot shows a different assay: (a) Roche E170, (b) Abbott architect, (c) Siemens centaur and (d) Beckman access. Minor and major outliers as determined by SPSS displayed by ○ and *, respectively. One renal outlier for both (a) and (c) are not shown on plot for clarity (indicated by arrow, 389% for (a) and 456% for (c)).

information available for customers. Use of another steroid, a change in pH[9] or use of a high affinity antibody to compete for binding is a speculative approach. Our results, however, may suggest that when CBG is increased the methods in use are not totally efficient. We speculate that the negative bias compared to normal female serum is due to some cortisol remaining CBG bound. Where cortisol binding proteins are decreased, e.g. in the hypoalbuminaemic ICU subjects included in this study this will result in a decrease in bound cortisol. This may also affect the efficiency of any cortisol displacement step included in the assay.

We have observed increased intermethod variation in the serum from both renal and ICU patients, compared to pooled serum samples. This is most significant for the Roche E170 method. We speculate that this is a function of the specificity of this immunoassay. During critical illness, cortisol production is increased, and

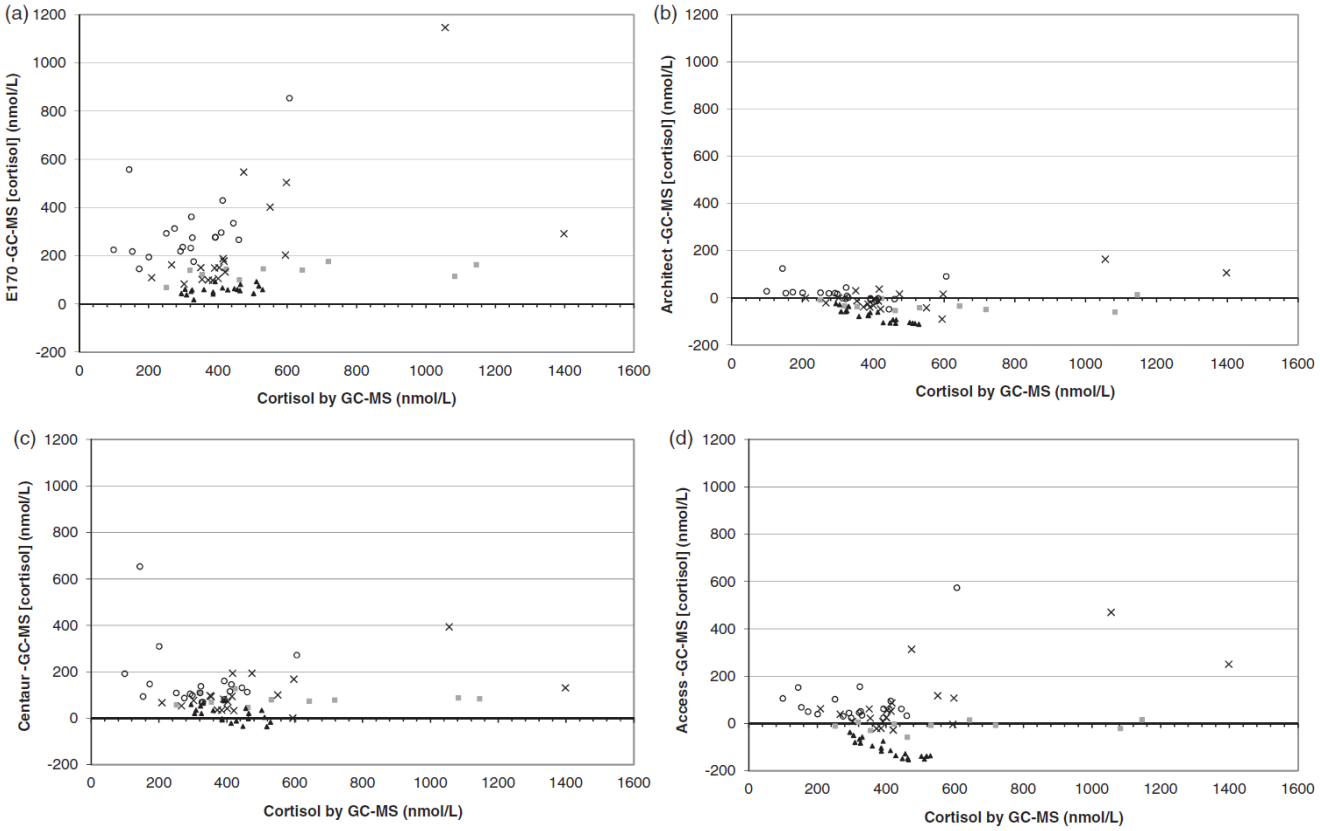


Figure 3. Difference plots (Bland–Altman) for (a) Roche E170, (b) Abbott architect, (c) Siemens centaur and (d) Beckman access versus GC-MS cortisol concentrations. ■: spiked pools; ▲: second trimester of pregnancy; x: intensive care unit; ○: patients with end-stage renal disease. One intensive care patient not shown on plots for clarity (GC-MS 2379 nmol/L: Roche E170 +581 nmol/L, Abbott architect −672 nmol/L, Siemens centaur −469 nmol/L and Beckman access −389 nmol/L).

cortisol breakdown is also reduced as a consequence of suppressed expression and activity of cortisol metabolizing enzymes.[11] Decreased clearance also occurs in patients with renal impairment. These changes in cortisol metabolism may lead to an accumulation of other steroids, which because of the similarities to the structure of cortisol are recognized by the immunoassay. Others have also observed differences in bias of the Roche E170 assay compared to GC-MS in the serum of sepsis patients compared to patients routinely attending outpatients.[7] Increased 'apparent' cortisol measurement has also been reported in patients with congenital adrenal hyperplasia using the Roche E170 method[12] which was attributed to recognition of other steroids which are increased in the serum of patients with this condition.

Comparison of cortisol immunoassay bias against GC-MS in this study compared to previous work[4] and cortisol spiking studies reported annually by UKNEQAS over time[1,13] demonstrates that immunoassay bias is not constant over long periods. Laboratories will be familiar with changes in assay bias between lots of immunoassay reagents or following assay reformulations by the manufacturer. But clinicians may not be cognisant to this fact. Where specific cut-offs are used for interpretation of results, e.g. the short Synacthen test, this becomes a significant factor to consider.

Owen et al.[14] have recently described development of a cortisol assay suitable for clinical use using tandem mass spectrometry. This method offers improved specificity over immunoassays and overcomes the limitations reported here for immunoassays. However, tandem mass spectrometry technology is still relatively new to clinical biochemistry and for the time being is less robust than automated immunoassay. Accordingly, the use of tandem mass spectrometry is likely to be limited to laboratories with specialist expertise in this technology. Whilst the use of immunoassay to measure cortisol predominates, it is important that laboratories and clinicians interpreting cortisol results are aware that serum matrix as well as assay bias and gender must be considered. Specific knowledge concerning the behaviour of the cortisol immunoassay used by the laboratory is essential, as it is clear that for some immunoassays there is a significant difference between true and 'apparent' measured cortisol in some clinical settings.

Acknowledgements

The authors thank staff at WEQAS and Department of Medical Biochemistry and Immunology, University Hospital of Wales, for their assistance.

Declaration of conflicting interests

None.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Ethical approval

This study was approved as a service evaluation project by Cardiff and Vale University Health Board.

Guarantor

CE.

Contributorship

CE, NE-F and DAR conceived the study. AJD undertook laboratory studies with assistance from DHD and SMN. AJD analysed the data. KLM, AA and JHB supervised analysis in their laboratories, AJD and CE wrote the first draft of the manuscript. All authors contributed to the analysis and interpretation of data, drafting and revising the manuscript and approval of the final version.

References

1. MacKenzie F. UKNEQAS for steroid hormones Annual review 2012. Participants version. https://results.ukneqas.org.uk/output/QEHSTER/[1]SteroidAnnualReview_2011.pdf (Wolfson EQA Laboratory, UKNEQAS, Birmingham, UK).
2. Ducroq DH and Thomas MA. Traceability of cortisol assays in the UK: comparison with JCTLM listed reference method. *Ann Clin Biochem* 2013; 50(supplement): 56.
3. Middle JG. UKNEQAS for steroid hormones Annual Review 2009. Participants version. https://results.ukneqas.org.uk/output/QEHSTER/06_C_2009_part.pdf (Wolfson EQA Laboratory, UKNEQAS, Birmingham, UK).
4. El-Farhan N, Pickett A, Ducroq D, et al. Method-specific serum cortisol responses to the adrenocorticotrophin test: comparison of gas chromatography-mass spectrometry and five automated immunoassays. *Clin Endocrinol (Oxf)* 2013; 78: 673–680.
5. Klose M, Lange M, Rasmussen AK, et al. Factors influencing the adrenocorticotropin test: role of contemporary cortisol assays, body composition, and oral contraceptive agents. *J Clin Endocrinol Metab* 2007; 92: 1326–1333.
6. Clark PM, Neylon I, Raggatt PR, et al. Defining the normal cortisol response to the short Synacthen test: implications for the investigation of hypothalamic-pituitary disorders. *Clin Endocrinol (Oxf)* 1998; 49: 287–292.
7. Briegel J, Sprung CL, Annane D, et al. Multicenter comparison of cortisol as measured by different methods in samples of patients with septic shock. *Intensive Care Med* 2009; 35: 2151–2156.
8. Barnes SC and Swaminathan R. Effect of albumin concentration on serum cortisol measured by the Bayer Advia centaur assay. *Ann Clin Biochem* 2007; 44: 79–82.
9. Pratt JJ. Steroid immunoassay in clinical chemistry. *Clin Chem* 1978; 24: 1869–1890.
10. Woods L, Ducroq DH, Fraser HL, et al. Measurement of urinary free cortisol by tandem mass spectrometry and comparison with results obtained by gas chromatography mass spectrometry and two commercial immunoassays. *Ann Clin Biochem* 2008; 45: 380–388.
11. Boonen E, Vervenne H, Meersseman P, et al. Reduced cortisol metabolism during critical illness. *N Engl J Med* 2013; 368: 1477–1488.
12. Curtis GE and Stutchfield PR. A case to illustrate diagnostic pitfalls in a district general hospital: analytical and interpretative. *Ann Clin Biochem* 2009; 46(supplement 1): 130.
13. MacKenzie F. UKNEQAS for steroid hormones Annual review 2010. Participants version. https://results.ukneqas.org.uk/output/QEHSTER/[1]SteroidAnnualReview_2012_participants.pdf (Wolfson EQA Laboratory, UKNEQAS, Birmingham, UK).
14. Owen LJ, Adaway JE, Davies S, et al. Development of a rapid assay for the analysis of serum cortisol and its implementation into a routine service laboratory. *Ann Clin Biochem* 2013; 50: 345–352.

Review Article

The Association for Clinical Biochemistry & Laboratory Medicine
Better Science, Better Testing, Better Care

Annals of Clinical Biochemistry
2017, Vol. 54(3) 308–322

DOI: 10.1177/0004563216687335
journals.sagepub.com/home/acb
SAGE

Measuring cortisol in serum, urine and saliva – are our assays good enough?

Nadia El-Farhan[1], D Aled Rees[2] and Carol Evans[3]

Abstract

Cortisol is a steroid hormone produced in response to stress. It is essential for maintaining health and wellbeing and leads to significant morbidity when deficient or present in excess. It is lipophilic and is transported bound to cortisol-binding globulin (CBG) and albumin; a small fraction (~10%) of total serum cortisol is unbound and biologically active. Serum cortisol assays measure total cortisol and their results can be misleading in patients with altered serum protein concentrations. Automated immunoassays are used to measure cortisol but lack specificity and show significant inter-assay differences. Liquid chromatography – tandem mass spectrometry (LC-MS/MS) offers improved specificity and sensitivity; however, cortisol cut-offs used in the short Synacthen and Dexamethasone suppression tests are yet to be validated for these assays. Urine free cortisol is used to screen for Cushing's syndrome. Unbound cortisol is excreted unchanged in the urine and 24-h urine free cortisol correlates well with mean serum-free cortisol in conditions of cortisol excess. Urine free cortisol is measured predominantly by immunoassay or LC-MS/MS. Salivary cortisol also reflects changes in unbound serum cortisol and offers a reliable alternative to measuring free cortisol in serum. LC-MS/MS is the method of choice for measuring salivary cortisol; however, its use is limited by the lack of a single, validated reference range and poorly standardized assays. This review examines the methods available for measuring cortisol in serum, urine and saliva, explores cortisol in disease and considers the difficulties of measuring cortisol in acutely unwell patients and in neonates.

Keywords

Endocrinology, steroid hormones

Accepted: 12th December 2016

Introduction

Cortisol

Cortisol is a steroid hormone produced by the zona fasciculata of the adrenal cortex. Its secretion is directly controlled by the release of adrenocorticotrophin (ACTH) from the anterior pituitary which is regulated by corticotrophin-releasing hormone (CRH), and less directly, arginine vasopressin (AVP), secreted by the hypothalamus. Cortisol exerts negative feedback on both ACTH and CRH secretion, while stress acts directly on the hypothalamus to stimulate CRH release.

ACTH secretion is pulsatile and follows a circadian pattern in line with CRH. An unstressed adult produces 18–25 pulses in a 24-h period, starting with

[1]Biochemistry Department, Royal Gwent Hospital, Newport, UK
[2]Neuroscience and Mental Health Research Institute, Cardiff University, Cardiff, UK
[3]Department of Medical Biochemistry and Immunology, University Hospital of Wales, Cardiff, UK

Corresponding author:
Nadia El-Farhan, Biochemistry Department, Royal Gwent Hospital, Newport NP20 2UP, UK.
Email: Nadia.El-Farhan@wales.nhs.uk

high-amplitude pulses released at roughly 90-min intervals between 3 and 9 a.m., which peak between 7 and 11 a.m. and reduce in frequency between 6 p.m. and midnight. Rapid blood sampling is, therefore, required to fully appreciate the dynamic nature of ACTH concentrations in blood.[1] Cortisol mirrors ACTH's circadian rhythm and pulsatile pattern, with highest peak in the early morning and nadir between midnight and 3 am.[1,2] Its rhythm is further affected by multiple factors, including sleep deprivation/shift-working,[3] ethnicity,[4] gender, age, BMI[5] and menstrual cycle phase.[6,7]

Cortisol is lipophilic and requires a protein carrier for transport around the body. Approximately 80% is bound to cortisol-binding globulin (CBG), 10% is albumin-bound and 10% is free (unbound) and biologically active. Cortisol is involved in energy metabolism through direct effects on the synthesis and breakdown of carbohydrates, protein and lipids. It increases gluconeogenesis by activating key enzymes, e.g. phosphoenolpyruvate carboxykinase and glucose-6-phosphatase and by increasing muscle breakdown and inhibiting amino acid uptake and protein synthesis to provide additional substrate. It reduces peripheral glucose uptake into cells, activates glycogen synthase and inhibits glycogen phosphorylase leading to increased glycogen stores. It activates lipolysis, resulting in increased circulating free fatty acids and enhances adipogenesis through activation of lipoprotein lipase and increased leptin concentration.[8,9] The net result is increased blood glucose concentrations mediated by the breakdown of protein and lipids; and cortisol, along with glucagon, adrenaline and growth hormone, is released in response to hypoglycaemia.

Cortisol is involved in water and electrolyte balance and blood pressure control. It increases angiotensinogen synthesis, enhances the vasopressor effects of catecholamines and angiotensin II and reduces nitric oxide-mediated endothelial dilatation. In the kidneys, it increases glomerular filtration rate and acts on the distal renal tubule to increase sodium retention and potassium loss; it also increases sodium transport and free water loss through antagonism of AVP.[9] Its renal effects are mediated through the mineralocorticoid receptor to which it binds with equal affinity to aldosterone. The magnitude of cortisol binding to the receptor is reduced by the action of 11β-hydroxysteroid dehydrogenase type 2 (11β-HSD2) which oxidizes cortisol to cortisone, which has much weaker affinity for the receptor.

Cortisol has both anti-inflammatory and immunosuppressive effects. It reduces circulating lymphocyte counts by redistribution to the spleen, lymph nodes and bone marrow. It inhibits immunoglobulin synthesis, induces lymphocyte apoptosis and prevents inflammatory cytokine (e.g. interleukin-1, 2, 6, tumour necrosis factor [TNF], interferon-γ) production through inhibition of nuclear factor kappa B (NF-κB). Monocyte differentiation into macrophages and subsequent macrophage function are inhibited by cortisol, and local inflammation is suppressed by reduced histamine production and impaired prostaglandin synthesis.

Cortisol is also involved in maintaining healthy pregnancy and in the initiation of birth, although its exact role is still unclear;[10] further study being hindered by differences between humans and animal models. Increased placental production of CRH has been noted in the run up to human labour, resulting in increased foetal cortisol production. Raised serum cortisol and placental CRH, both believed to be manifestations of maternal stress, have been shown to increase the risk of obesity and diabetes in later life.[11–13]

Measuring cortisol

Acute cortisol deficiency presents clinically as an Addisonian crisis, characterized by catastrophic dehydration and salt wasting. This is potentially fatal so early diagnosis is essential. However, as mild adrenal insufficiency is difficult to detect clinically due to non-specific symptoms such as fatigue and weakness, diagnosis relies on appropriate biochemical testing. Similarly, the early symptoms of cortisol excess (Cushing's syndrome), which include weight gain, hypertension and impaired glucose tolerance, are easily attributable to other causes, such as simple obesity, leading to delays in diagnosis.

A random serum cortisol concentration rarely gives enough information for a diagnosis of adrenal disease to be made.[14] A recent study using the Siemens Centaur assay demonstrated the extent of inter-individual variation in morning cortisol, with concentrations ranging between 96 nmol/L and 722 nmol/L in samples collected from healthy volunteers between 8.00 and 11.00 am.[15] This wide variation can be partly explained by the timing of the sample relative to a pulse of cortisol.

An early morning cortisol measurement, although rarely diagnostic, can contribute some useful information, with concentrations above 450–550 nmol/L virtually excluding hypoadrenalism, concentrations below this indicating the need for more definitive testing[14,16,17] and concentrations below 140 nmol/L suggesting adrenal insufficiency,[18] although requiring confirmation (these quoted concentrations will vary according to the accuracy and bias of the assays used). Plasma ACTH can aid in the diagnosis of adrenal insufficiency; with concentrations >66 pmol/L being highly predictive for disease when combined with low early morning cortisol (<140 nmol/L) or an impaired response to Synacthen (these quoted concentrations will vary

according to the accuracy and bias of the assays used).[14] The most useful tests of adrenal function, therefore, are those that stimulate or suppress the HPA axis.

Serum/plasma cortisol assays

Cortisol can be measured in either serum or plasma,[19] with direct comparison between the two showing they are interchangeable (unpublished data). The first plasma cortisol assays were restricted to research laboratories as they required several labour-intensive steps, given the low concentrations being measured and the presence in plasma of numerous other steroid hormones. In 1962, a simple fluorimetric assay utilizing the natural fluorescence of 11-hydroxycorticoids (cortisol and corticosterone) was described, marking the beginning of clinical cortisol assays.[20] However, this method was limited by poor specificity for cortisol and low sample throughput, and radioimmunoassay (RIA) soon replaced it as the method of choice.[21] Further assay development saw the elimination of pre-analytical sample extraction steps and the introduction of enzyme and chemiluminescent detection,[22,23] and the first automated cortisol immunoassay was described in 1992.[24] These have now become the method of choice in many clinical laboratories despite their limitations which include differences in antibody specificity and affinity for cortisol and other steroid hormones and the need to release the steroid molecule from its carrier protein before it can be measured.

Immunoassays

Endogenous and exogenous steroids. Many steroids are structurally similar to cortisol (Figure 1), which makes the generation of entirely specific antibodies difficult.[25] Cortisol immunoassays are thus hampered by varying degrees of antibody cross-reactivity with other steroids, endogenous and exogenous,[26] and can be unreliable in certain clinical settings such as congenital adrenal hyperplasia[27] (CAH) and in patients treated with synthetic glucocorticoids.

Information provided by assay manufacturers can sometimes help laboratories identify possible interference; however, this information is not exhaustive, and labs are not always aware of patients' medication. This was highlighted by the recent implementation of a liquid chromatography-tandem mass spectrometry (LC-MS/MS) serum cortisol assay in a laboratory serving a large respiratory unit. Exogenous steroids were found in the serum of up to 50% of patients tested.[28] Synthetic steroids were also found in the urine of 5% of patients investigated for Cushing's syndrome when an LC-MS/MS assay was used.[29]

Binding proteins. Current assays measure total serum cortisol (bound and unbound fractions) and as a result are affected by changes in CBG and albumin concentration. The cortisol response to ACTH in a group of patients with the nephrotic syndrome was shown to be lower, by between 63 and 179 nmol/L, depending on the assay used, than that in healthy volunteers.[30] This translated into a 50% failure rate in interpretation of the response for two of the assays, when a 500 nmol/L cut-off was used. Similar findings were reported in patients with liver disease and mean albumin concentration of 21 g/L, with 46% of patients failing to achieve a total cortisol of 550 nmol/L post-stimulation.[31] Low protein concentration also affects the interpretation of cortisol results in acutely unwell patients.

Conversely, increased binding protein concentrations, which occur with oestrogen excess, e.g. pregnancy and treatment with oestrogen-containing oral contraceptives (OCP), can cause higher concentrations of cortisol. Thus, mean post-ACTH cortisol concentrations up to 484 nmol/L higher, depending on the assay used, are seen in women taking exogenous oestrogens, compared with their untreated counterparts[15] (Table 1). Furthermore, the normal relationship between immunoassay and gas chromatography-mass spectrometry (GC-MS) or LC-MS/MS cortisol measurement is reversed, with most immunoassays showing negative bias relative to the mass method result.[15,32]

This difference can be eliminated by heat-treatment[32] and has been attributed to reduced cortisol availability for measurement due to increased protein binding. Most current immunoassays make some attempt to displace cortisol from CBG prior to measurement, using either steroid substitutes, e.g. Danazol (Roche, personal communication) or pH change (Abbott, personal communication), although these measures are not always sufficient when excess CBG is present.

Assay standardization. Poor standardization, partly due to differences in antibody specificity and partly the lack of a single reference material or method for cortisol, is another factor contributing to the variability of cortisol immunoassays. There are at least six certified reference materials for cortisol – ERM-DA192 and 193 (1985), ERM-DA451/IFCC (1999), SRM 921 (1993), SRM 971 (2011) and the U.S. Pharmacopeia (USP) Hydrocortisone Reference Standard,[33] and information from manufacturers of current immunoassays reveals traceability to several of these (personal correspondence, Table 2) and to both the isotope dilution-GC-MS reference method[34] and several different LC-MS/MS reference methods.[35]

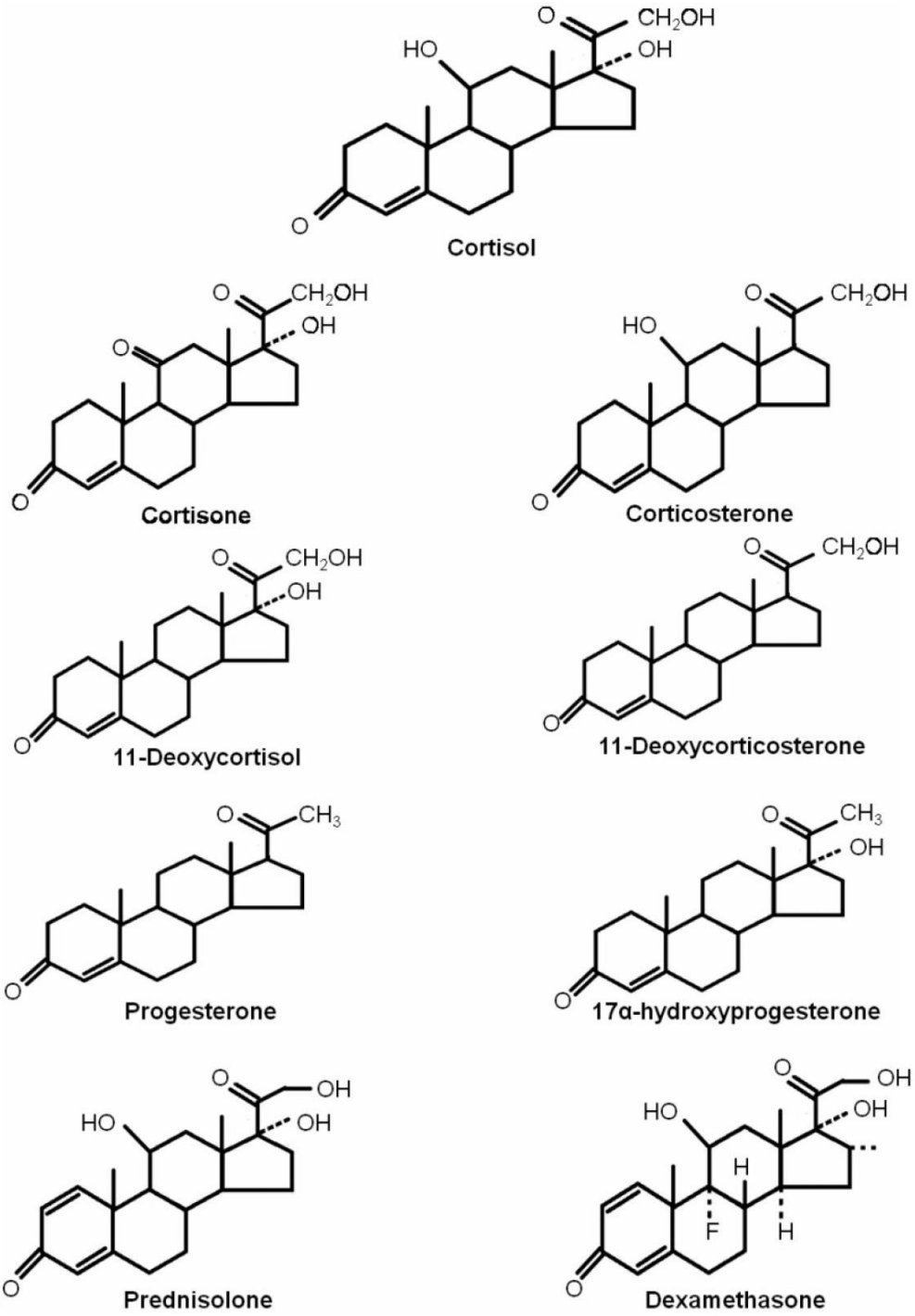


Figure 1. Structural homology of cortisol and various endogenous and exogenous steroid molecules.

The effect of gender and sample matrix. A recent study of the effect of gender and matrix on immunoassay cortisol measurement[36] demonstrated increased inter-assay variation in patients with significant renal disease and those who are critically ill. This was presumed to be a matrix effect due to decreased cortisol clearance in the former group and its increased production and reduced breakdown in the latter. Inter-assay variation in cortisol recovery was also noted to be higher in spiked samples from female patients than in males, presumably due to differences in CBG concentrations.

Divergence in external quality assurance. The majority of UK participants in the UKNEQAS cortisol scheme measure cortisol using one of the following immunoassays: Roche Elecsys (44%), Siemens ADVIA Centaur

Table 1. Mean post-ACTH cortisol concentrations in non-OCP females and OCP females by assay.

	Post-stimulation cortisol (nmol/L)		
Assay	Non-OCP females $n = 79$	OCP females $n = 24$	P
GC-MS	555	870	0.001
Centaur	578	763	0.001
Abbott	542	747	0.001
Roche	712	1026	0.001
Immulite	628	850	0.001
Beckman	594	757	0.001

Non-OCP females: women not using an oestrogen-containing oral contraceptive pill; OCP females: women using an oestrogen-containing oral contraceptive pill.
P-value < 0.05 indicates a significant difference between means.
Reproduced with permission of authors and publishers.[15]

Table 2. Information provided by assay manufacturers regarding the reference material and method to which each assay was traceable.

Assay	Traceability
Beckman Access	Traceable to United States Pharmacopea reference material.
Abbott Architect	Assay is designed to have a slope of 1.0 ± 0.1 and a correlation coefficient (r) of $\geqslant 0.95$ for serum samples when compared with Liquid Chromatography Mass Spectrometry/Mass Spectrometry (LC-MS/MS).
Siemens Centaur	GC-MS
Roche E170	Standardized against another method which is standardized against ID-MS. IFCC-451 Panel ID/GC/MS gives 89–111% recovery.
DPC Immulite	Correlated to Coat-A-Count cortisol assay. Traceable to an internal standard manufactured using qualified materials and measurement procedures.

(20%), Abbott Architect (19%) and Beckman Access (10%).[37] The scheme organizers calculate a B score, or specimen percentage bias, for each laboratory to demonstrate how far its result deviates from the target concentration – the all laboratory trimmed mean (ALTM). An arbitrary acceptable limit of performance of 10% above or below the target mean has been set, with the percentage of laboratories failing to meet this target shown in Figure 2.

This increasing divergence between cortisol assays[37–39] is unlikely to be explained by worsening laboratory performance, but indicates that the differences between current cortisol immunoassays are too great for a single, meaningful mean to be calculated. UKNEQAS plans to address this by replacing the ALTM with the LC-MS/MS method mean (the field method-mass spectrometry (FM-MS) mean) as of April 2017.[40] Occasional comparisons will also be made to an LC-MS/MS reference method (RM-MS) to ensure the validity of the FM-MS mean.

Differences between immunoassays. The significance of differences between automated cortisol immunoassays was not widely considered until a study by Clark et al.[41] Cortisol concentration pre- and post-Synacthen was measured using four widely available immunoassays. Median cortisol concentration 30 min post-Synacthen ranged from 707 nmol/L to 866 nmol/L, with a lower limit (defined as the 5th percentile) between 510 nmol/L and 626 nmol/L. This demonstrated the assay dependence of cortisol cut-offs and the risk of misdiagnosing adrenal insufficiency if the wrong cut-off was used. Nevertheless, most laboratories continued to use the fluorimetrically derived 550 nmol/L cut-off.[42]

Diagnosing adrenal insufficiency. Prompted by increasing clinical concern over possible overdiagnosis of adrenal insufficiency, the cortisol response to Synacthen stimulation was recently defined for five contemporary assays, with direct comparison to GC-MS[15] (Table 3). The extent of variation seen confirmed the need for assay-specific interpretation of the Synacthen test and demonstrated the impact even minor assay reformulations could have on these measurements (Table 4).

Ongoing assay changes. Dodd et al.[36] demonstrated changes in immunoassay performance over a three-year period, despite no major changes in assay formulation (Table 5), and the increasing divergence in external quality assurance (EQA) performance suggests these assay changes are ongoing. It is, therefore, not surprising that some laboratories and clinicians are reluctant to adopt method-specific cortisol cut-offs, as these soon become obsolete. However, if assay differences continue to be overlooked when interpreting cortisol concentrations, particularly in the context of dynamic tests, the clinical validity of these tests will be lost. Laboratories must, therefore, act to ensure their cut-offs are valid by regularly reviewing the bias of their assay relative to the gold standard – either GC-MS or LC-MS/MS, a process which could be facilitated by an EQA provider.

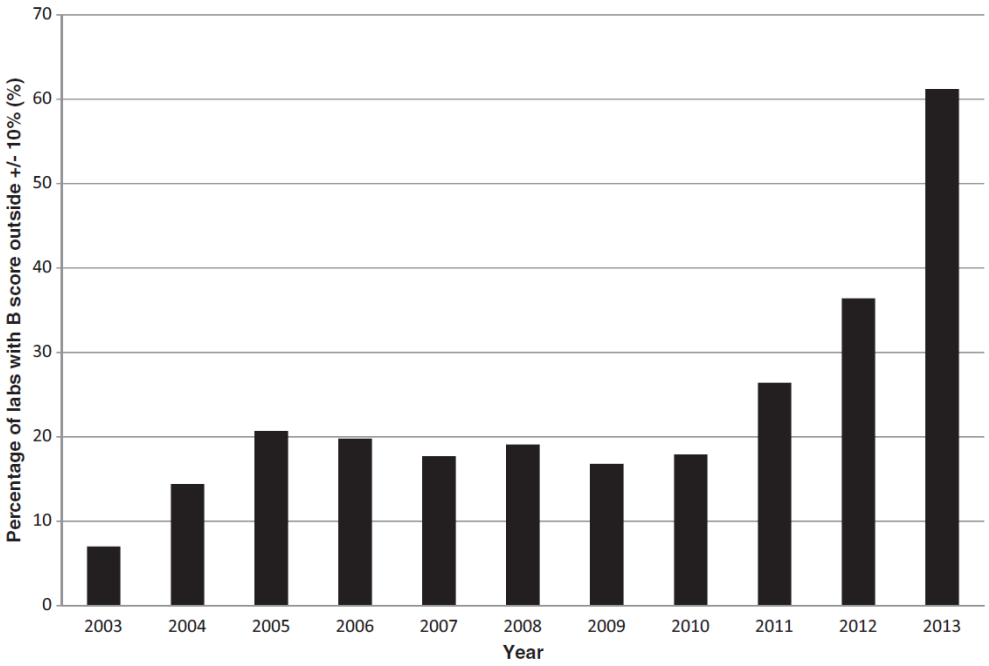


Figure 2. Proportion of laboratories participating in UKNEQAS cortisol EQA scheme with bias scores outside the acceptable limits of performance.
The B score is the average %bias ([result−target]/[target] × 100%) of 30 specimens distributed over a six month time window.
Reproduced with permission from F. MacKenzie, UKNEQAS.

Table 3. Assay-specific estimated lower reference limits for post-ACTH cortisol according to gender and OCP-status.

Assay	Males	Non-OCP females	OCP-females
GC-MS	418	421	643
Siemens Centaur	448	446	619
Abbott Architect	430	416	577
Roche E170	574	524	791
Siemens Immulite (2000)	469	478	688
Beckman access	459	455	604

Note: The estimated lower reference limit was determined by back transformation of the 2.5th percentile value (mean − 1.96 × SD) of the log-transformed data. Results are expressed in nmol/L.
Non-OCP females: women not using an oestrogen-containing oral contraceptive pill; OCP females: women using an oestrogen-containing oral contraceptive pill; NA: not applicable.
Reproduced with permission of authors and publisher.[15]

Table 4. Estimated lower reference limits for post-ACTH cortisol according to gender and OCP-status for Siemens Centaur assay pre- and post-assay reformulation.

Assay	Males	Non-OCP females	OCP	Mean bias ratio
Old Centaur	496	502	682	1.177
New Centaur	448	446	619	1.060

Note: The estimated lower reference limit was determined by back transformation of the 2.5th percentile value (mean: 1.96 × SD) of the log-transformed data. Results are expressed in nmol/L. The bias ratio was calculated by dividing each sample's immunoassay cortisol concentration by the corresponding GC-MS concentration. The mean ratio for all subjects at all time points was then calculated.
Non-OCP females: women not using an oestrogen-containing oral contraceptive pill; OCP: women using an oestrogen-containing oral contraceptive pill.

The recent adjustment of the Roche cortisol assay, to align its performance with LC-MS/MS cortisol, has resulted in substantially lower concentrations and will provide some indication of how well laboratories adapt to ongoing changes. Continuing with old cortisol cut-offs for the Synacthen test will result in overdiagnosis of adrenal insufficiency, but simply reducing the cut-off in line with the observed bias of the new assay will not provide a clinically validated cut-off. The best approach would, therefore, be to directly compare post-Synacthen results across the two assays, which would be easier done by the kit manufacturer or an EQA provider than by each kit user individually.

Mass methods

In 1975, the first reference method for measuring cortisol using GC-MS was described.[34] Cortisol was derivatized to its heptafluorobutyrate ester to improve

Table 5. Mean bias ratios for cortisol immunoassays compared with GCMS in 2010 and 2013.

	Mean bias ratio 2010			Mean bias ratio 2013		
Assay	Total	Males	Females	Total	Males	Females
Roche E170	1.30	1.36	1.25	1.26	1.28	1.23
Siemens Centaur	1.06	1.08	1.05	1.17	1.19	1.14
Abbott Architect	1.02	1.04	1.00	0.94	0.95	0.93
Beckman access	1.07	1.09	1.05	0.98	1.00	0.95

Note: The bias ratio was calculated by dividing each sample's immunoassay cortisol concentration by the corresponding GC-MS concentration. The mean ratio was then calculated.
Total: mean bias ratio for all subjects at all time points.
Reproduced with permission of authors and publishers.[36]

assay sensitivity and tritiated cortisol was added as an internal standard to correct for losses due to steroid adsorption to the chromatography column. Today, GC-MS cortisol assays continue to outperform immunoassays in both specificity and sensitivity, but they remain confined to reference and research laboratories as they are labour intensive and require specialist equipment and highly skilled laboratory staff.

In 2001, an LC-MS/MS method for measurement of multiple steroid hormones in patients with 11 β-hydroxylase and 21-hydroxylase deficiencies and Addison's disease was described.[43] This combined the improved specificity and definitive analyte identification of mass spectrometry with the ability to perform multiple tests simultaneously in a cost-effective way and offered the potential for mass methods to spread to clinical laboratories. In 2004, an LC-MS/MS reference method, which avoided the need for preanalytical sample derivatization was described as an alternative to GC-MS for routine method comparisons,[35] and in 2016, an LC-MS/MS candidate reference measurement procedure that could be used to assign traceable target concentrations to EQA samples was described.[44]

Currently, there are 12 LC-MS/MS cortisol assays registered with the UKNEQAS cortisol scheme[40] and tandem-mass spectrometry is being used globally for routine steroid hormone measurement, either in isolation[28] or as part of a panel of related molecules.[45–50] However, due to the greater specificity of LC-MS/MS, the cortisol response to ACTH stimulation and dexamethasone suppression may need to be redefined to ensure correct interpretation. Concerted efforts to standardize LC-MS/MS assays, including traceability to a single reference method and material, and rigorous validation[51] of these assays are also needed to avoid similar inter-assay differences to those affecting immunoassays.

Such differences have already been reported with LC-MS/MS vitamin D and testosterone assays, due to the lack of common calibrators,[52] differences in sample extraction techniques and differences in the type of LC column and/or mass spectrometer used.[53] There have been no direct comparisons between cortisol LC-MS/MS assays as yet, although UKNEQAS data from the past year shows that the variability between these assays (CVs: 3.1–15.8%) is less than differences between methods (CVs: 11.6–24.9%) and not too dissimilar to the performance of a single immunoassay.[54]

Sample matrix has also been shown to affect the reproducibility and accuracy of LC-MS/MS assays. Endogenous impurities, e.g. salts (typically in urine) and co-eluting substances, e.g. analyte metabolites, can cause ion suppression or enhancement – a reduction or increase in the efficiency of ion formation.[55–57] Careful evaluation of this effect is, therefore, essential before a new LC-MS/MS cortisol assay can be introduced into clinical practice. Careful selection of the internal standard is also necessary, with stable isotopically labelled analogues preferred over structurally similar, but non-identical, analogues[58] and other compounds which have the same retention time but a different structure.[59] The use of isotopically labelled standards ensures co-elution of analyte and internal standard and their subsequent ionization under identical conditions,[59] thus improving analytical accuracy.

Exogenous interferences in LC-MS/MS cortisol assays are few and recent developments have managed to eliminate the interference from prednisolone and fenofibrate which had previously been problematic.[60,61] For a more detailed discussion of LC-MS/MS, cortisol assays readers are directed to a recent review by Hawley and Keevil.[58]

Free cortisol assays

Direct measurement of serum-free cortisol has been possible since the 1950s, although available methods – equilibrium dialysis, gel filtration and ultrafiltration – are too labour intensive and time consuming for routine clinical use. Attempts to develop a semi-routine assay for free cortisol measurement are ongoing, with ultrafiltration offering the most likely option.[62]

Equilibrium dialysis is particularly time consuming. A dialysis membrane separates a small volume of plasma from a buffer solution and the two are incubated for up to 24 h.[63] Free cortisol crosses the membrane until equilibrium is reached and can be measured in the dialysis buffer. However, the accuracy of this method is affected by serum leakage across the dialysis membrane and by changes in cortisol binding equilibrium within the plasma sample due to dilution with

dialysis buffer. Nevertheless, an equilibrium dialysis isotope-dilution LC-MS/MS method using a commercially available dialysis cell (Amicon) has been described and offers potential for a clinical assay.[62]

Gel filtration relies on unbound cortisol moving down a glass column packed with Sephadex G-25 more slowly than its protein-bound counterpart.[64] As the latter passes through the column, the cortisol dissociates from its binding protein until the concentration of dissociated cortisol in the column equals the unbound fraction in the original sample and no further dissociation occurs. As the protein-bound cortisol elutes first the concentration of serum-free cortisol can be measured in the later fraction. This method requires large sample volumes to create a steady state in the gel column and, although analysis is possible at any temperature, 37°C is preferred as temperatures below this may affect cortisol-CBG binding equilibrium.

Ultrafiltration is a relatively simple method and benefits from the availability of commercial filters, e.g. Amicon ultracentrifugal filters with a 30 KDa cut-off (Millipore, UK). Plasma, separated from a collection chamber by a dialysis membrane, is centrifuged for 10 min at 37°C and the cortisol concentration in the resulting filtrate is measured.[65] However, substantial variations in filtrate yield can be seen, presumably due to the forced filtration disturbing cortisol binding equilibrium at the membrane surface. Precise temperature control can also be difficult to achieve, particularly in a routine laboratory setting.

Routine measurement of serum-free cortisol is further hampered by the lack of a single, robust reference range. Most studies are small and have used in-house assays to measure cortisol, resulting in significant differences in the ranges quoted. Low substrate concentration and lack of stable quality control material have also contributed to these differences, although good correlation between serum-free cortisol measured by ultrafiltration and equilibrium dialysis has been reported.[66] In view of the difficulty measuring serum-free cortisol directly, interest has grown in deriving calculated estimates as an alternative.

Coolens et al.[67] derived an equation for calculating serum-free cortisol (CFC), taking into account the binding properties of both CBG and albumin, and demonstrated excellent correlation between this estimate and serum-free cortisol measured by ultrafiltration. The free cortisol index (FCI) is another estimate, defined as serum total cortisol divided by CBG, and has been shown to correlate well with serum free cortisol measured by gel filtration and with Coolens' CFC.[68] However, calculated free cortisol estimates have not yet been validated in patients with HPA axis disease and Coolens' CFC has been shown to underestimate measured free cortisol in patients undergoing a short Synacthen test.[69] Substantial bias between calculated and measured free cortisol in critically ill patients has also cast doubt on its role in this population.[70,71] Better understanding of how serum-free cortisol behaves in health and disease is, therefore, essential if it is to become a practical alternative to total cortisol.

Urinary and salivary cortisol assays

Urine

Under normal conditions, the kidneys filter, and partially reabsorb, unbound cortisol, and only a very small fraction (~2%) is excreted unchanged in the urine. Urinary cortisol is unaffected by changes in hepatic metabolism, unlike urinary cortisol metabolites, and better reflects endogenous cortisol concentration.

Hypercortisolaemic patients have significantly higher renal cortisol clearance than patients with normal serum cortisol due to the rapid rise in serum-free cortisol once the binding capacity of CBG has been exceeded (at around 500 nmol/L), and due to continued cortisol excretion overnight. Thus, 24-h urine free cortisol correlates well with mean serum-free cortisol in conditions of cortisol excess in all but severe renal impairment, and is widely used as a screening test.[72] In some instances, e.g. in patients with adrenal incidentalomas or cyclical Cushing's syndrome, cortisol metabolites perform better than urinary cortisol,[73] despite their variable 24-h excretion pattern.[74]

Urinary cortisol was first measured fluorimetrically, with subsequent development of radioimmunoassay (RIA), enzyme-linked immunoassay (EIA) and chemiluminescent assays (CLIA).[75] However, the presence of conjugated cortisol metabolites in urine which can cross-react with immunoassay antibodies,[76] e.g. allotetrahydrocortisol glucuronide, necessitated extraction of free cortisol prior to analysis. By mixing the urine with an organic solvent, e.g. dichloromethane and removing the aqueous layer, cortisol, which is less water-soluble than most of its metabolites, can be separated from the latter prior to analysis.[8,25] To further improve the specificity of urine cortisol assays, chromatography was introduced – initially as a purification step prior to immunoassay[77] – and has now become the method of choice.

Early HPLC assays were not ideal due to the long run times needed to ensure adequate separation; and by coupling HPLC or GC to mass-spectrometry for detection, run times were significantly reduced, although low throughput and high costs precluded these assays from widespread use.[78,79] More recently, LC-MS/MS has provided highly sensitive and specific assays which are easy to use, with high sample throughput and which compare well to the gold standard GC-MS method.[78]

These assays, however, can be more susceptible than serum assays to ion suppression due to the variable salt content and numerous steroid metabolites present in urine or sample extract.

Saliva

Cortisol enters saliva by diffusion, independently of salivary flow rate.[80] It reflects cortisol's circadian rhythm and early morning peak, and responds to changes in plasma cortisol concentration quickly and reliably.[81] Interest in measuring salivary cortisol began in the 1960s but was hampered by the lack of sensitive cortisol assays as concentrations in saliva are less than one tenth those in serum.[80]

Salivary cortisol was first measured by direct RIA in 1978[82] and, more recently, automated serum cortisol immunoassays have been successfully adapted to measure cortisol in saliva.[83,84] However, these assays are limited by their poor specificity, particularly where there is significant antibody cross-reactivity with cortisone, which is present in high concentration in saliva. Assay sensitivity can also present a problem as late night salivary cortisol concentrations often fall close to, or below, the functional limits of detection.[83,85]

The first LC-MS/MS assay for measuring salivary cortisol was described in 2003 and promised improved specificity over existing assays.[86] Efforts have since concentrated on adapting LC-MS/MS to provide a quick, high-throughput service,[28] with numerous successful assays now available.[87–90] However, widespread application of salivary cortisol measurement is limited by poorly standardized assays[85] and the lack of a single, validated reference range.[89,90]

Thus, mean salivary cortisol concentrations in the early morning of 3.6 nmol/L[87] and 8.3 nmol/L[88] have been reported in healthy laboratory workers in two studies using different in-house LC-MS/MS assays. Similarly, the late night salivary cortisol cut-off for diagnosing Cushing's syndrome has been reported as 2.95 nmol/L and 2.1 nmol/L,[91,92] and although it is likely to lie somewhere below 3 nmol/L, a single cut-off is yet to be determined. A further study of late night salivary cortisol showed significant overlap between concentrations in healthy volunteers and obese subjects, with ranges of <0.1–116 nmol/L and <0.11–17.7 nmol/L respectively.[89] Laboratories will, therefore, need to define their own assay-specific cut-offs until better assay standardization has been achieved and should be aware of the potential overlap between patients and healthy volunteers.

Sample collection technique may also contribute to the differences observed between salivary cortisol assays. Saliva is often collected by passive drooling into a container but can be collected by chewing an absorbent cotton pad, typically the Sarstedt Salivette®, which is later centrifuged to release its liquid content for analysis. Steroid recovery from cotton pads has been shown to differ from other collection devices, e.g. Whatman® foam-tip applicator and blood collection cards,[92] resulting in different concentrations simply due to the collection technique used. Thus, the development of any assay should include an evaluation of the chosen collection device.

Other potential interferences with salivary cortisol measurement include sample contamination with blood during collection and smoking before sample collection; both of which can lead to falsely elevated cortisol concentrations.[89] Avoiding the analysis of visibly pink saliva is, therefore, important and most authors suggest collecting saliva at least 30 min after eating, drinking or teeth-brushing,[84,93] although there is little evidence to support this approach.[94]

Salivary cortisol is now widely used in clinical laboratories: it has been suggested as a convenient alternative to serum for monitoring hydrocortisone replacement,[95] there is interest in it as an alternative to serum total cortisol in the interpretation of the Synacthen test,[96,97] particularly in patients with altered protein concentrations[98] and late night salivary cortisol is a recommended first-line screening test for Cushing's syndrome.[99] Salivary cortisol has also been used to research basal cortisol patterns in neonates and in evaluating their response to stress.[100]

An alternative marker of serum-free cortisol is salivary cortisone, which shows excellent correlation with serum-free cortisol concentration[88,101] and is found in higher concentration than salivary cortisol,[101,102] due to the presence of 11β-HSD2 in parotid tissue.[103] Further work is now needed to establish suitable reference ranges and cut-offs for salivary cortisone and to investigate the effect of altered 11β-HSD2 activity on its diagnostic utility before it can be used in a clinical setting.

Cortisol in disease

Addison's disease

The gold standard test of HPA axis function is the insulin tolerance test (ITT) in which insulin-induced hypoglycaemia stimulates the hypothalamus, resulting in CRH release and an increase in serum cortisol.[14] The short Synacthen test (SST) is a quick and safe alternative to the ITT and has replaced it as the first line test,[104] although as it does not assess the entire HPA axis, it can give false negative results in acute pituitary disease and in the early weeks post-pituitary surgery.[17,105,106] It involves parenteral administration of a supraphysiological dose (250 μg) of a synthetic

peptide consisting of the first 24 amino acids of natural ACTH (Synacthen), followed by measurement of peak cortisol concentration 30 min later.[14] The incremental rise from baseline has been used as an alternative to peak cortisol in the past, but as it is inversely related to baseline cortisol concentration, those with high basal concentrations due to endogenous adrenal gland stimulation often show a poor incremental rise, despite an intact HPA axis,[107–109] thus limiting its diagnostic value.

Defining the most appropriate cortisol cut-off for use with the Synacthen test continues to pose some difficulty.[105,110] Historically, a 30-min cut-off of 550 nmol/L was used, as it correlated well with results from the insulin tolerance test,[111,112] but because of differences between cortisol immunoassays, a single cut-off is no longer applicable and this should be replaced by method-specific alternatives.[15,30,41] These are likely to gain better clinical acceptance if they are shown to correlate well with the cortisol response to insulin-induced hypoglycaemia (which will itself need to be redefined).

Cushing's syndrome

In 2008, the American Endocrine Society's Clinical Guidelines Subcommittee published a guideline on the diagnosis of Cushing's syndrome.[99] The authors collated evidence from studies that had demonstrated significant treatment benefit in patients diagnosed with Cushing's syndrome to identify the tests that best detected the disease. They recommended using 24-h urine free cortisol, late night salivary cortisol, the 1 mg overnight dexamethasone suppression test (DST) or the 48 hour low-dose dexamethasone suppression test (LDDST) (0.5 mg of dexamethasone administered every 6 h for 48 h, beginning at 9 a.m. on day 1, with serum cortisol measured at 9 a.m. on day 3, 6 h after the final dose) for initial screening; the former two on at least two separate occasions. Confirmation of an abnormal screening test using a second test from the list is required and patients with discordant or positive results should be evaluated by an endocrinologist, while those with negative screening tests, reassured. The guideline also identifies population groups at greater risk of diagnostic confusion and makes specific recommendations about the most suitable tests to use.[99]

The effect of immunoassay differences on interpretation of the DST, and hence the diagnosis of Cushing's syndrome, has recently been studied,[113] and the current serum cortisol cut-off of 50 nmol/L, which was first defined in 1989 using a Diagnostic Products Corporation (DPC) cortisol RIA,[114] was shown to perform well with the Abbott assay and LC-MS/MS, with diagnostic sensitivities of 87.5% and 93.8% and specificities of 92.5% and 91.2%, respectively.[113]

Diagnosing Cushing's syndrome can be challenging, as evidenced by the multiple screening tests on offer and cyclical hypercortisolism contributes further diagnostic challenge.[115,116] Cycles of cortisol hypersecretion may be regular or irregular with no discernible pattern, and cortisol concentrations in between each episode may be normal or even low. Repeated testing is, therefore, necessary.[116] Pseudo-Cushing's presents another source of diagnostic confusion. It is a state of cortisol excess caused by increased HPA axis activity due to poorly controlled diabetes, obesity, alcoholism and certain psychiatric disorders.[117,118] Patients present with clinical features of Cushing's syndrome and positive screening tests despite only mild hypercortisolism.[118]

In a study designed to distinguish patients with Cushing's syndrome from those with Pseudo-Cushing's, all recommended first-line tests showed excellent sensitivity but limited specificity,[119] leading several authors[119,120] to recommend the desmopressin stimulation test and combined dexamethasone suppression – CRH stimulation (DST-CRH) test as more discerning alternatives.

The desmopressin stimulation test involves measuring ACTH before and, at intervals, after IV administration of 10 μg of desmopressin (dDAVP). It relies on the presence of vasopressin-2 (V2) receptors on pituitary corticotrophs and results in a rise in ACTH in patients with Cushing's disease, but not other causes of Cushing's syndrome. It is also used to predict relapse of Cushing's disease post-pituitary resection, particularly when preceded by the DST.[121]

The DST-CRH test is recommended as a confirmatory test by the Endocrine society. It consists of the LDDST followed by intravenous (IV) CRH and subsequent measurement of cortisol and ACTH.[122] It relies on the premise that patients with pseudo-Cushing's retain sensitivity to glucocorticoid suppression and fail to respond to CRH stimulation. In contrast, patients with true Cushing's fail to respond to the LDDST, with those with an ACTH-producing tumour subsequently responding to CRH stimulation.

Adrenal incidentalomas (AI) and subclinical Cushing's syndrome (SC) present yet further diagnostic confusion. AI are masses discovered by imaging studies performed for reasons other than the investigation of adrenal pathology and have a reported postmortem prevalence of 1.4–2.9%.[123] Identifying and treating those who require intervention because of hormonal activity or malignancy presents a significant challenge, particularly as the absence of clinical signs does not exclude a secretory mass.[123] SC is autonomous glucocorticoid production without the associated signs and symptoms of Cushing's. It is found in 5 to 20% of patients with AI and carries a risk of progressing to overt Cushing's syndrome.[124,125]

Accurately diagnosing SC remains a challenge. The DST is considered the test of choice,[99,126–128] although there is disagreement over the best cut-off to use, partly due to assay differences and partly a lack of clinical consensus.[127,128] Urinary free cortisol excretion and late night salivary cortisol perform relatively poorly as screening tests on their own[129–131] and most diagnostic algorithms recommend combining the DST with a second measure of hypercortisolism, e.g. 24-h urinary cortisol excretion, failure to respond to CRH stimulation or loss of diurnal rhythm.[128,131]

Acute disease

Cortisol concentrations increase in acute illness via a stress-induced rise in CRH and reduced negative feedback from cortisol.[132] The diurnal pattern of cortisol secretion is lost and CBG concentrations fall, leading to increased circulating free cortisol. These changes are mediated by an increase in circulating cytokines, including tumour necrosis factor α, interleukin-1 and interleukin-6, which also increase cortisol at tissue level by increasing receptor affinity for cortisol and reducing its peripheral metabolism.[132,133] The HPA axis can also be impaired by the disease process, resulting in a functional adrenal insufficiency, i.e. a transient (for the most part) episode of subnormal corticosteroid production.

Evaluating the HPA axis in acutely ill patients is complicated by increased synthesis of endogenous steroids by an activated HPA axis and a high prevalence of heterophilic antibodies.[134] The low protein concentrations typical of these patients affect the serum total cortisol response to Synacthen stimulation, resulting in a diagnosis of adrenal insufficiency in up to 40% of patients with no HPA axis disease.[135] Assay differences are accentuated, with one study showing concordance between the results of the Synacthen test in only four of the nine patients studied when samples were analysed using four different assays.[136]

Similar results were found in an offshoot of the multinational CORTICUS (corticosteroid therapy of septic shock) study by Briegel et al.[134] in which the correlation coefficient for the concordance between serum total cortisol measurements from 12 different cortisol assays and the Roche Elecsys assay was 0.60 and the correlation with LC-MS/MS was as low as 0.43 for some assays. Furthermore, although the correlation between the Roche assay and LC-MS/MS was excellent for a group of outpatients, this fell when samples from critically ill patients were used instead. This has led to concerns that the clinical management of these patients is being determined by the cortisol immunoassay being used by their local laboratory rather than their actual adrenal status, and that serum-free cortisol or total cortisol measured by LC-MS/MS should be used to evaluate their HPA axis.

Cortisol in neonates

Basal cortisol

Early studies of neonatal serum cortisol identified high concentrations in cord blood at birth (880 nmol/L),[137] which started to fall within the first 24 h, plateauing at concentrations between 50 and 400 nmol/L by day 3.[138] More recently, a cortisol range of 4 to 588 nmol/L, with a median concentration of 93 nmol/L, was reported in infants born between 24 and 32 weeks' gestation,[139] with concentrations falling with increasing gestational age.[140] This suggests that early reports of high neonatal cortisol concentration were overestimates, which may have arisen due to the relatively high cortisone concentrations in neonates cross reacting in cortisol assays.

Attempts to define a reference range for salivary cortisol in healthy infants have largely met with failure due to high intraindividual variability.[141] Nevertheless, Ivars et al. reported a median early morning salivary cortisol concentration of 5.1 nmol/L in one-month old infants, increasing to 10.9 nmol/L at 12 months.[142] Strict adherence to sample collection time was essential for meaningful interpretation of results.

Further studies in serum have shown that cortisol's circadian rhythm is absent at birth, although its secretion, and that of ACTH, is pulsatile, and it settles into a circadian rhythm by around three months of age.[143] In contrast, studies in saliva have demonstrated a discernible circadian rhythm in full-term neonates at 1 month of age,[142] albeit with significant intra and interindividual variability in terms of how early the rhythm appears and how persistent it is once it has appeared.[141,142,144] Further study is now needed to differentiate genuine differences from those arising due to non-standardized assays, poor specificity and matrix effects.

Conclusions

Early diagnosis of disorders of the HPA axis is essential and historically, serum total cortisol has been the analyte of choice. Most current cortisol immunoassays lack specificity and are poorly standardized, and in some clinical scenarios, e.g. in patients with altered protein concentrations and the critically ill, their performance is simply not good enough. LC-MS/MS provides a suitable alternative to immunoassay for measurement of serum total cortisol in all clinical scenarios. However, mass spectrometers are not yet ubiquitous in clinical laboratories, and even those that have the equipment may struggle with the extensive validation required to ensure LC-MS/MS methods are fit-for-purpose.[51]

Both salivary cortisol and cortisone show promise as alternatives to serum cortisol, although further work to define reference ranges and their response to dynamic testing is needed before they can be used in routine clinical practice. There has also been interest in steroid hormone profiles for diagnosing adrenal disease, particularly congenital adrenal hyperplasia, although their role in diagnosing adrenal insufficiency and Cushing's syndrome is not yet clear. The role of serum free cortisol measurement in clinical practice remains to be established.

Further improvements to serum and salivary cortisol assays will depend on the development and use of a single cortisol reference material and method to achieve standardized assays. EQA schemes will need to develop a more robust system for evaluating assay performance using samples with MS-assigned target concentrations.

Acknowledgements

We would like to thank Finlay MacKenzie, UKNEQAS for reading and commenting on the manuscript and for permission to use UKNEQAS data. This article was prepared at the invitation of the Clinical Sciences Review Committee of the Association of Clinical Biochemistry and Laboratory Medicine.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Ethical approval

Not applicable.

Guarantor

NE.

Contributorship

NE wrote the first draft of the manuscript. All authors reviewed and edited the manuscript.

References

1. Lightman S and Terry JR. The importance of dynamic signalling for endocrine regulation and drug development: relevance for glucocorticoid hormones. *Lancet Diabetes Endocrinol* 2014; 2: 593–599.
2. Papadimitriou A and Priftis KN. Regulation of the hypothalamic-pituitary-adrenal axis. *Neuroimmunomodulation* 2009; 16: 265–271.
3. Amirian I, Andersen LT, Rosenberg J, et al. Diurnal cortisol rhythm: working night shifts affects surgeons' biological rhythm. *Am J Surg* 2015; 210: 389–395.
4. DeSantis AS, Adam EK, Hawkley LC, et al. Racial and ethnic differences in diurnal cortisol rhythms: are they consistent over time? *Psychosom Med* 2015; 77: 6–15.
5. Veldhuis JD, Roelfsema F, Iranmanesh A, et al. Basal, pulsatile, entropic (patterned), and spiky (staccato-like) properties of ACTH secretion: impact of age, gender, and body mass index. *J Clin Endocrinol Metab* 2009; 94: 4045–4052.
6. Nepomnaschy PA, Altman RM, Watterson R, et al. Is cortisol excretion independent of menstrual cycle day? A longitudinal evaluation of first morning urinary specimens. *PLoS One* 2011; 6: e18242.
7. Boisseau N, Enea C, Diaz V, et al. Oral contraception but not menstrual cycle phase is associated with increased free cortisol levels and low hypothalamo-pituitary-adrenal axis reactivity. *J Endocrinol Invest* 2013; 36: 955–964.
8. Demers LM. The adrenal cortex. In: Burtis CA, Ashwood ER, Bruns DE (eds) *Tietz textbook of clinical chemistry*. 4th ed. St Louis, MO: Elsevier Saunders, 2005, pp.2003–2052.
9. Stewart PM and Krone NP. The adrenal cortex. In: Melmed S, Polonsky KS, Larsen PR, et al. (eds) *Williams textbook of endocrinology*. Saunders: Elsevier Inc., 2011, pp.479–544.
10. Kamel RM. The onset of human parturition. *Arch Gynecol Obstet* 2010; 281: 975–982.
11. Stout SA, Espel EV, Sandman CA, et al. Fetal programming of children's obesity risk. *Psychoneuroendocrinology* 2015; 53: 29–39.
12. Hohwü L, Henriksen TB, Grønborg TK, et al. Maternal salivary cortisol levels during pregnancy are positively associated with overweight children. *Psychoneuroendocrinology* 2015; 52: 143–152.
13. Entringer S. Impact of stress and stress physiology during pregnancy on child metabolic function and obesity risk. *Curr Opin Clin Nutr Metab Care* 2013; 16: 320–327.
14. Bornstein SR, Allolio B, Arlt W, et al. Diagnosis and treatment of primary adrenal insufficiency: an Endocrine Society Clinical Practice Guideline. *J Clin Endocrinol Metab* 2016; 101: 364–389.
15. El-Farhan N, Pickett A, Ducroq D, et al. Method-specific serum cortisol responses to the adrenocorticotrophin test: comparison of gas chromatography-mass spectrometry and five automated immunoassays. *Clin Endocrinol (Oxf)* 2013; 78: 673–680.
16. Jones SL, Trainer PJ, Perry L, et al. An audit of the insulin tolerance test in adult subjects in an acute investigation unit over one year. *Clin Endocrinol (Oxf)* 1994; 41: 123–128.
17. Courtney CH, McAllister AS, McCance DR, et al. Comparison of one week 0900 h serum cortisol, low and standard dose synacthen tests with a 4 to 6 week insulin hypoglycaemia test after pituitary surgery in assessing HPA axis. *Clin Endocrinol (Oxf)* 2000; 53: 431–436.
18. Kazlauskaite R, Evans AT, Villabona CV, et al. Corticotropin tests for hypothalamic-pituitary-adrenal insufficiency: a metaanalysis. *J Clin Endocrinol Metab* 2008; 93: 4245–4253.
19. World Health Organization. Use of anticoagulants in diagnostic laboratory investigations (WHO/DIL/LAB/99.1 Rev.2), http://apps.who.int/iris/bitstream/10665/65957/1/WHO_DIL_LAB_99.1_REV.2.pdf (2002, accessed 16 January 2017).
20. Mattingly D. A simple fluorimetric method for the estimation of free 11-hydroxycorticoids in human plasma. *J Clin Pathol* 1962; 15: 374–379.
21. Campuzano HC, Wilkerson JE, Raven PB, et al. A radioimmunoassay for cortisol in human plasma. *Biochem Med* 1973; 7: 350–362.
22. Kominami G, Fujisaka I, Yamauchi A, et al. A sensitive enzyme immunoassay for plasma cortisol. *Clin Chim Acta* 1980; 103: 381–391.
23. Lindstrom L, Meurling L and Lovgren T. The measurement of serum cortisol by a solid-phase chemiluminescence immunoassay. *J Steroid Biochem* 1982; 16: 577–580.
24. Bacarese-Hamilton T, Cattini R, Shandley C, et al. A fully automated enzyme immunoassay for the measurement of cortisol in biological fluids. *Eur J Clin Chem Clin Biochem* 1992; 30: 531–535.
25. Holder G. Measurement of glucocorticoids in biological fluids. *Methods Mol Biol* 2006; 324: 141–157.
26. Brossaud J, Barat P, Gualde D, et al. Cross reactions elicited by serum 17-OH progesterone and 11-desoxycortisol in cortisol assays. *Clin Chim Acta* 2009; 407: 72–74.
27. Curtis GE and Stutchfield PR. A case to illustrate diagnostic pitfalls in a district general hospital: analytical and interpretative. *Ann Clin Biochem* 2009; 46(Suppl 1): 130.
28. Owen LJ, Adaway JE, Davies S, et al. Development of a rapid assay for the analysis of serum cortisol and its implementation into a routine service laboratory. *Ann Clin Biochem* 2013; 50: 345–352.
29. Djedovic NK and Rainbow SJ. Detection of synthetic glucocorticoids by liquid chromatography tandem mass spectrometry in patients being investigated for Cushing's syndrome. *Ann Clin Biochem* 2011; 48: 542–549.
30. Klose M, Lange M, Rasmussen AK, et al. Factors influencing the adrenocorticotropin test: role of contemporary cortisol assays, body composition and oral contraceptive agents. *J Clin Endocrinol Metab* 2007; 92: 1326–1333.

31. Vincent RP, Etogo-Asse FE, Dew T, et al. Serum total cortisol and free cortisol index give different information regarding the hypothalamus-pituitary-adrenal axis reserve in patients with liver impairment. *Ann Clin Biochem* 2009; 46: 505–507.
32. Jung C, Ho JT, Torpy DJ, et al. A longitudinal study of plasma and urinary cortisol in pregnancy and postpartum. *J Clin Endocrinol Metab* 2011; 96: 1533–1540.
33. www.usp.org/reference-standards (accessed 16 January 2017).
34. Breuer H and Siekmann L. Mass fragmentography as reference method in clinical steroid assay. *J Steroid Biochem* 1975; 6: 685–688.
35. Tai SS-C and Welch MJ. Development and evaluation of a candidate reference method for the determination of total cortisol in human serum using isotope dilution liquid chromatography/mass spectrometry and liquid chromatography/tandem mass spectrometry. *Anal Chem* 2004; 76: 1008–1014.
36. Dodd AD, Ducroq DH, Neale SM, et al. The effect of serum matrix and gender on cortisol measurement by commonly used immunoassays. *Ann Clin Biochem* 2014; 51: 379–385.
37. UK NEQAS for Steroid Hormones. Review for the UK NEQAS Clinical Chemistry Specialist Advisory Group for Endocrinology and Immunoassay: participants' version, https://results.ukneqas.org.uk/output/QEHSTER/%5b1%5d%20Steroid%20Annual%20Review%20Participant_2013.pdf (2013, accessed 8 February 2016).
38. UK NEQAS for Steroid Hormones. Review for the UK NEQAS Clinical Chemistry Specialist Advisory Group for Endocrinology and Immunoassay: participants' version, https://results.ukneqas.org.uk/output/QEHSTER/[1]SteroidAnnualReview_2012_participants.pdf (2012, accessed 8 February 2016).
39. UK NEQAS for Steroid Hormones. Review for the UK NEQAS Clinical Chemistry Specialist Advisory Group for Endocrinology and Immunoassay: participants' version, https://results.ukneqas.org.uk/output/QEHSTER/[1]SteroidAnnualReview_Participants_2010.pdf (2010, accessed 8 February 2016).
40. UK NEQAS report for Steroid Hormones. Distribution 436, https://results.ukneqas.org.uk/scripts/scheme-select.pl (2016, accessed 8 February 2016).
41. Clark PM, Neylon I, Raggatt PR, et al. Defining the normal cortisol response to the Synacthen test: implications for the investigation of hypothalamic-pituitary disorders. *Clin Endocrinol (Oxf)* 1998; 49: 287–292.
42. Chatha KK, Middle JG and Kilpatrick ES. National UK audit of the short synacthen test. *Ann Clin Biochem* 2010; 47: 158–164.
43. Kao PC, Machacek DA, Magera MJ, et al. Diagnosis of adrenal cortical dysfunction by liquid chromatography-tandem mass spectrometry. *Ann Clin Lab Sci* 2001; 31: 199–204.
44. Hawley JM, Owen LJ, MacKenzie F, et al. Candidate reference measurement procedure for the quantification of total serum cortisol with LC-MS/MS. *Clin Chem* 2016; 62: 262–269.
45. Ruiter AFC, Teeninga N, Nauta J, et al. Determination of unbound prednisolone, prednisone and cortisol in human serum and saliva by on-line solid-phase extraction liquid chromatography tandem mass spectrometry and potential implications for drug monitoring of prednisolone and prednisone in saliva. *Biomed Chromatogr* 2012; 26: 789–796.
46. Koal T, Schmiederer D, Pham-Tuan H, et al. Standardized LC-MS/MS based steroid hormone profile-analysis. *J Steroid Biochem Molecular Biol* 2012; 129: 129–138.
47. Kyriakopoulou L, Yazdanpanah M, Colantonio DA, et al. A sensitive and rapid mass spectrometric method for the simultaneous measurement of eight steroid hormones and CALIPER pediatric reference intervals. *Clin Biochem* 2013; 46: 642–651.
48. Keevil BG. Novel liquid chromatography tandem mass spectrometry (LC-MS/MS) methods for measuring steroids. *Best Pract Res Clin Endocrinol Metab* 2013; 27: 663–674.
49. Keefe CC, Goldman MM, Zhang K, et al. Simultaneous measurement of thirteen steroid hormones in women with polycystic ovary syndrome and control women using liquid chromatography-tandem mass spectrometry. *PLoS One* 2014; 9: e93805.
50. Lee S, Lim H-S, Shin H-J, et al. Simultaneous determination of cortisol and cortisone from human serum by liquid chromatography-tandem mass spectrometry. *J Anal Methods Chem* 2014; 2014: 787483.
51. Honour JW. Development and validation of a quantitative assay based on tandem mass spectrometry. *Ann Clin Biochem* 2011; 48: 97–111.
52. Yates AM, Bowron A, Calton L, et al. Interlaboratory variation in 25-hydroxyvitamin D2 and 25-hydroxyvitamin D3 is significantly improved if common calibration material is used. *Clin Chem* 2008; 54: 2082–2084.
53. Owen LJ, MadDonald PR and Keevil BG. Is calibration the cause of variation in liquid chromatography tandem mass spectrometry testosterone measurement? *Ann Clin Biochem* 2013; 50: 368–370.
54. UK NEQAS report for Steroid Hormones. Distributions 415-426; January 2015 – January 2016, https://results.ukneqas.org.uk/scripts/scheme-select.pl (accessed 8 February 2016).
55. Annesley TM. Ion suppression in mass spectrometry. *Clin Chem* 2003; 49: 1041–1044.
56. Allende F, Solari S, Campino C, et al. LC-MS/MS method for the simultaneous determination of free urinary steroids. *Chromatographia* 2014; 77: 637–642.
57. Taylor PJ. Matrix effects: the Achilles heel of quantitative high-performance liquid chromatography-electrospray-tandem mass spectrometry. *Clin Biochem* 2005; 38: 328–334.
58. Hawley JM and Keevil BG. Endogenous glucocorticoid analysis by liquid chromatography-tandem mass spectrometry in routine clinical laboratories. *J Steroid Biochem Mol Biol* 2016; 162: 27–40.
59. Tudela E, Muñoz G and Muñoz-Guerra JA. Matrix effect marker for multianalyte analysis by LC-MS/MS in biological samples. *J Chromatogr B Analyt Technol Biomed Life Sci* 2012; 901: 98–106.
60. Kushnir MM, Rockwood AL, Nelson GJ, et al. Liquid chromatography–tandem mass spectrometry analysis of urinary free cortisol. *Clin Chem* 2003; 49: 965–967.
61. Wear JE, Owen LJ, Duxbury K, et al. A simplified method for the measurement of urinary free cortisol using LC-MS/MS. *J Chromatogr B Analyt Technol Biomed Life Sci* 2007; 858: 27–31.
62. Kirchhoff F, Briegel J and Vogeser M. Quantification of free serum cortisol based on equilibrium dialysis and isotope dilution-liquid chromatography-tandem mass spectrometry. *Clin Biochem* 2011; 44: 894–899.
63. Vogeser M, Briegel J and Zachoval R. Dialyzable free cortisol alter stimulation with Synacthen. *Clin Biochem* 2002; 35: 539–543.
64. Burke CW. Accurate measurement of steroid-protein binding by steady-state gel filtration. *Biochim Biophys Acta* 1969; 176: 403–413.
65. MacMahon W, Thompson J, Bowers W, et al. A simplified ultrafiltration method for determination of serum free cortisol. *Clin Chim Acta* 1983; 131: 171–184.
66. Pretorius CJ, Galligan JP, McWhinney BC, et al. Free cortisol method comparison: ultrafiltration, equilibrium dialysis, tracer dilution, tandem mass spectrometry and calculated free cortisol. *Clin Chim Acta* 2011; 412: 1043–1047.
67. Coolens J-L, Van Baelen H and Heyns W. Clinical use of unbound plasma cortisol as calculated from total cortisol and corticosteroid-binding globulin. *J Steroid Biochem* 1987; 26: 197–202.
68. Le Roux CW, Sivakumaran S, Alaghband-Zadeh J, et al. Free cortisol index as a surrogate marker for serum free cortisol. *Ann Clin Biochem* 2002; 39: 406–408.
69. Barlow NL, Holme J, Stockley RA, et al. An evaluation of measured and calculated serum free cortisol in a group of patients with known adrenal suppression. *Ann Clin Biochem* 2010; 47: 200–204.
70. Molenaar N, Groeneveld AB and de Jong MF. Three calculations of free cortisol versus measured values in the critically ill. *Clin Biochem* 2015; 48: 1053–1058.
71. Cohen J, Venkatesh B and Tan T. Comparison of the diagnostic accuracy of measured and calculated free cortisol in acutely ill patients using the Coolens equation. *Crit Care Resusc* 2013; 15: 39–41.
72. Alexandraki KI and Grossman AB. Is urinary free cortisol of value in the diagnosis of Cushing's syndrome? *Curr Opin Endocrinol Diabetes Obes* 2011; 18: 259–263.
73. Brossaud J, Ducint D and Corcuff JB. Urinary glucocorticoid metabolites: biomarkers to classify adrenal incidentalomas? *Clin Endocrinol (Oxf)* 2016; 84: 236–243.
74. Jerjes WK, Cleare AJ, Wood PJ, et al. Assessment of subtle changes in glucocorticoid negative feedback using prednisolone: Comparison of salivary free cortisol and urinary cortisol metabolites as endpoints. *Clin Chim Acta* 2006; 364: 279–286.
75. Pazzagli M, Kim JB, Messeri G, et al. Luminescent immunoassay (LIA) of cortisol-2. Development and validation of the immunoassay monitored by chemiluminescence. *J Steroid Biochem* 1981; 14: 1181–1187.
76. Krasowski MD, Drees D, Morris CS, et al. Cross-reactivity of steroid hormone immunoassays: clinical significance and two-dimensional molecular similarity prediction. *BMC Clin Pathol* 2014; 14: 33.

77. Schöneshöfer M, Fenner A, Altinok G, et al. Specific and practicable assessment of urinary free cortisol by combination of automatic high-pressure liquid chromatography and radioimmunoassay. *Clin Chim Acta* 1980; 106: 63–73.
78. Taylor RL, Machacek D and Singh RJ. Validation of a high-throughput liquid chromatography-tandem mass spectrometry method for urinary cortisol and cortisone. *Clin Chem* 2002; 48: 1511–1519.
79. Wood L, Ducroq DH, Fraser HL, et al. Measurement of urinary free cortisol by tandem mass spectrometry and comparison with results obtained by gas chromatography-mass spectrometry and two commercial immunoassays. *Ann Clin Biochem* 2008; 45: 380–388.
80. Vining RF, McGinley RA and Symons RG. Hormones in saliva: mode of entry and consequent implications for clinical interpretation. *Clin Chem* 1983; 29: 1752–1756.
81. Vining RF and McGinley RA. The measurement of hormones in saliva: possibilities and pitfalls. *J Steroid Biochem* 1987; 27: 81–94.
82. Walker RF, Riad-Fahmy D and Read GF. Adrenal status assessed by direct radioimmunoassay of cortisol in whole saliva or parotid saliva. *Clin Chem* 1978; 24: 1460–1463.
83. Deutschbein T, Broecker-Preuss M, Flitsch J, et al. Salivary cortisol as a diagnostic tool for Cushing's syndrome and adrenal insufficiency: improved screening by an automatic immunoassay. *Eur J Endocrinol* 2012; 166: 613–618.
84. Sesay AM, Micheli L, Tervo P, et al. Development of a competitive immunoassay for the determination of cortisol in human saliva. *Anal Biochem* 2013; 434: 308–314.
85. Inder WJ, Dimeski G and Russell A. Measurement of salivary cortisol in 2012 – laboratory techniques and clinical indications. *Clin Endocrinol (Oxf)* 2012; 77: 645–651.
86. Jönsson BAG, Malmberg B, Amilon A, et al. Determination of cortisol in human saliva using liquid chromatography-electrospray tandem mass spectrometry. *J Chromatogr B Analyt Technol Biomed Life Sci* 2003; 784: 63–68.
87. Montskó G, Tarjányi Z, Mezó si E, et al. A validated method for measurement of serum total, serum free, and salivary cortisol, using high-performance liquid chromatography coupled with high-resolution ESI-TOF mass spectrometry. *Anal Bioanal Chem* 2014; 406: 2333–2341.
88. Perogamvros I, Owen LF, Newell-Price J, et al. Simultaneous measurement of cortisol and cortisone in human saliva using liquid chromatography-tandem mass spectrometry: application in basal and stimulated conditions. *J Chromatogr B Analyt Technol Biomed Life Sci* 2009; 877: 3771–3775.
89. Baid SK, Sinaii N, Wade M, et al. Radioimmunoassay and tandem mass spectrometry measurement of bedtime salivary cortisol levels: a comparison of assays to establish hypercortisolism. *J Clin Endocrinol Metab* 2007; 92: 3102–3107.
90. Zerikly RK, Amiri L, Faiman C, et al. Diagnostic characteristics of late-night salivary cortisol using liquid chromatography-tandem mass spectrometry. *J Clin Endcrinol Metab* 2010; 95: 4555–4559.
91. Erickson D, Singh RJ, Sathananthan A, et al. Late-night salivary cortisol for diagnosis of Cushing's syndrome by liquid chromatography/tandem mass spectrometry assay. *Clin Endocrinol (Oxf)* 2012; 76: 467–472.
92. Gröschl H and Rauh H. Influence of commercial collection devices for saliva on the reliability of salivary steroid analysis. *Steroids* 2006; 71: 1097–1100.
93. Beko G, Varga I, Glaz E, et al. Cutoff values of midnight salivary cortisol for the diagnosis of overt hypercortisolism are highly influenced by methods. *Clin Chim Acta* 2010; 411: 364–367.
94. Kivlighan KT, Granger DA, Schwartz EB, et al. Quantifying blood leakage into the oral mucosa and its effects on the measurement of cortisol, dehydroepiandrosterone and testosterone in saliva. *Horm Behav* 2004; 46: 39–46.
95. Wong V, Yan T, Donald A, et al. Saliva and bloodspot cortisol: novel sampling methods to assess hydrocortisone replacement therapy in hypoadrenal patients. *Clin Endocrinol (Oxf)* 2004; 61: 131–137.
96. Deutschbein T, Unger N, Mann K, et al. Diagnosis of secondary adrenal insufficiency in patients with hypothalamic-pituitary disease: comparison between serum and salivary cortisol during the high-dose short Synacthen test. *Eur J Endocrinol* 2009; 160: 9–16.
97. Perogamvros I, Owen LJ, Keevil BG, et al. Measurement of salivary cortisol with liquid chromatography-tandem mass spectrometry in patients undergoing dynamic endocrine testing. *Clin Endocrinol (Oxf)* 2010; 72: 17–21.
98. Galbois A, Rudler M, Massard J, et al. Assessment of adrenal function in cirrhotic patients: salivary cortisol should be preferred. *J Hepatol* 2010; 52: 839–845.
99. Nieman LK, Biller BMK, Findling JW, et al. The diagnosis of Cushing's syndrome: an Endocrine Society clinical practice guideline. *J Clin Endocrinol Metab* 2008; 93: 1526–1540.
100. Mehler K, Ulbrich L, Börner S, et al. Multidimensional response to vaccination pain in very preterm, moderate-to-late preterm and full-term infants at age three months. *Early Hum Dev* 2015; 91: 199–204.
101. Perogamvros I, Keevil BG, Ray DW, et al. Salivary cortisone is a potential biomarker for serum free cortisol. *J Clin Endocrinol Metab* 2010; 95: 4951–4958.
102. Raff H and Singh RJ. Measurement of late-night salivary cortisol and cortisone by LC-MS/MS to assess preanalytical sample contamination with topical hydrocortisone (Letter to the editor). *Clin Chem* 2012; 58: 947–948.
103. Smith RE, Maguire JA, Stein-Oakley AN, et al. Localization of 11-hydroxysteroid dehydrogenase type II in human epithelial tissues. *J Clin Endocrinol Metab* 2006; 81: 3244–3248.
104. Reynolds RM, Stewart PM, Seckl JR, et al. Assessing the HPA axis in patients with pituitary disease: a UK survey. *Clin Endocrinol (Oxf)* 2006; 64: 82–85.
105. Dickstein G. Hypothalamo-pituitary-adrenal axis testing: nothing is sacred and caution in interpretation is needed. *Clin Endocrinol (Oxf)* 2001; 54: 15–16.
106. Klose M, Lange M, Kosteljanetz M, et al. Adrenocortical insufficiency after pituitary surgery: an audit of the reliability of the conventional short synacthen test. *Clin Endocrinol (Oxf)* 2005; 63: 499–505.
107. May ME and Carey RM. Rapid adrenocorticotropic hormone test in practice. Retrospective review. *Am J Med* 1985; 79: 679–684.
108. Doi SA, Lasheen I, Al-Humood K, et al. Relationship between cortisol increment and basal cortisol: implications for the low-dose short adrenocorticotropic hormone stimulation test. *Clin Chem* 2006; 52: 746–749.
109. Wallace I, Cunningham S and Lindsay J. The diagnosis and investigation of adrenal insufficiency in adults. *Ann Clin Biochem* 2009; 46: 351–367.
110. Weintrob N, Sprecher E, Josefsberg Z, et al. Standard and low-dose short adrenocorticotropin test compared with insulin-induced hypoglycemia for assessment of the hypothalamic-pituitary-adrenal axis in children with idiopathic multiple pituitary hormone deficiencies. *J Clin Endocrinol Metab* 1998; 83: 88–92.
111. Abdu TA, Elhadd TA, Neary R, et al. Comparison of the low dose short synacthen test (1 microg), the conventional dose short synacthen test (250 microg), and the insulin tolerance test for assessment of the hypothalamo-pituitary-adrenal axis in patients with pituitary disease. *J Clin Endocrinol Metab* 1999; 84: 838–843.
112. Hurel SJ, Thompson CJ, Watson MJ, et al. The short Synacthen and insulin stress tests in the assessment of the hypothalamic-pituitary-adrenal axis. *Clin Endocrinol (Oxf)* 1996; 44: 141–146.
113. Brixey-McCann R, Tennant SM, Geen J, et al. Effect of cortisol assay bias on the overnight dexamethasone suppression test: implications for the investigation of Cushing's syndrome. *Endocrine Abstracts* 2015; 38: 22.
114. Blethen SL and Chasalow FI. Overnight dexamethasone suppression test: normal responses and the diagnosis of Cushing's syndrome. *Steroids* 1989; 54: 185–193.
115. Vance ML. Cyclic Cushing syndrome: definitions and treatment implications. *Neurosurg Focus* 2007; 23: E4, 1–2.
116. Velez DA, Mayberg MR and Ludlam WH. Cyclic Cushing syndrome: definitions and treatment implications. *Neurosurg Focus* 2007; 23: E4, 1–3.
117. Besemer F, Pereira AM and Smit JWA. Alcohol-induced Cushing syndrome. *Neth J Med* 2011; 69: 318–323.
118. Androulakis II, Kaltsas G and Chrousos G. Pseudo-Cushing's states. In: De Groot LJ, Beck-Peccoz P, Chrousos G, et al. (eds) Endotext. South Dartmouth, MA: MDText.com, Inc. Available at: https://www.ncbi.nlm.nih.gov/books/NBK279081/ (2000, accessed 16 January 2017).
119. Pecori Giraldi F, Pivonello R, Ambrogio AG, et al. The dexamethasone-suppressed corticotropin releasing hormone stimulation test and the desmopressin test to distinguish Cushing's syndrome from pseudo-Cushing's states. *Clin Endocrinol (Oxf)* 2007; 66: 251–257.
120. Tirabassi G, Faloia E, Papa R, et al. Use of the desmopressin test in the differential diagnosis of pseudo-Cushing state from Cushing's disease. *J Clin Endocrinol Metab* 2010; 95: 1115–1122.

121. Castinetti F, Martinie M, Morange I, et al. A combined dexamethasone desmopressin test as an early marker of postsurgical recurrence in Cushing's disease. *J Clin Endocrinol Metab* 2009; 94: 1897–1903.
122. Martin NM, Dhillo WS and Meeran K. The dexamethasone-suppressed corticotropin releasing hormone stimulation test and the desmopressin test to distinguish Cushing's syndrome from pseudo-Cushing's states. *Clin Endocrinol (Oxf)* 2007; 67: 476.
123. Aron D, Terzolo M and Cawood TJ. Adrenal incidentalomas. *Best Pract Res Clin Endocrinol Metab* 2012; 26: 69–82.
124. Barzon L, Fallo F, Sonino N, et al. Development of overt Cushing's syndrome in patients with adrenal incidentaloma. *Eur J Endocrinol* 2002; 146: 61–66.
125. Morelli V, Masserini B, Salcuni AS, et al. Subclinical hypercortisolism: correlation between biochemical diagnostic criteria and clinical aspects. *Clin Endocrinol (Oxf)* 2010; 73: 161–166.
126. Mazzuco TL, Bourdeau I and Lacroix A. Adrenal incidentalomas and subclinical Cushing's syndrome: diagnosis and treatment. *Curr Opin Endocrinol Diabetes Obes* 2009; 16: 203–210.
127. Chiodini I. Clinical review: diagnosis and treatment of subclinical hypercortisolism. *J Clin Endocrinol Metab* 2011; 96: 1223–1236.
128. Akehi Y, Kawate H, Murase K, et al. Proposed diagnostic criteria for subclinical Cushing's syndrome associated with adrenal incidentaloma. *Endocr J* 2013; 60: 903–912.
129. Nunes ML, Vattaut S, Corcuff JB, et al. Late-night salivary cortisol for diagnosis of overt and subclinical Cushing's syndrome in hospitalized and ambulatory patients. *J Clin Endocrinol Metab* 2009; 94: 456–462.
130. Masserini B, Morelli V, Bergamaschi S, et al. The limited role of midnight salivary cortisol levels in the diagnosis of subclinical hypercortisolism in patients with adrenal incidentaloma. *Eur J Endocrinol* 2009; 160: 87–92.
131. Palmieri S, Morelli V, Polledri E, et al. The role of salivary cortisol measured by liquid chromatography-tandem mass spectrometry in the diagnosis of subclinical hypercortisolism. *Eur J Endocrinol* 2013; 168: 289–296.
132. Cooper MS and Stewart PM. Corticosteroid insufficiency in acutely ill patients. *N Engl J Med* 2003; 348: 727–734.
133. Chrousos GP. The hypothalamic-pituitary-adrenal axis and immune-mediated inflammation. *N Engl J Med* 1995; 332: 1351–1362.
134. Briegel J, Sprung CL and Annane D. Multicenter comparison of cortisol as measured by different methods in samples of patients with septic shock. *Intensive Care Med* 2009; 35: 2151–2156.
135. Hamrahian AH, Oseni TS and Arafah BM. Measurements of serum free cortisol in critically ill patients. *N Engl J Med* 2004; 350: 1629–1638.
136. Cohen J, Ward G, Prins J, et al. Variability of cortisol assays can confound the diagnosis of adrenal insufficiency in the critically ill population. *Intensive Care Med* 2006; 32: 1901–1905.
137. Stevens JF. Plasma cortisol levels in the neonatal period. *Arch Dis Child* 1970; 45: 592.
138. Weiner DJ, Smith J, Dahlan S, et al. Serum adrenal steroid levels in healthy full-term 3 day old infants. *J Pediatr* 1987; 110: 122–124.
139. Greaves RF, Pitkin J, Ho CS, et al. Hormone modeling in preterm neonates: establishment of pituitary and steroid hormone reference intervals. *J Clin Endocrinol Metab* 2015; 100: 1097–1103.
140. Greaves RF, Zacharin MR, Donath SM, et al. Establishment of hormone reference intervals for infants born < 30 weeks' gestation. *Clin Biochem* 2014; 47: 101–108.
141. Tollenaar MS, Jansen J, Beijers R, et al. Cortisol in the first year of life: normative values and intra-individual variability. *Early Hum Dev* 2010; 86: 13–16.
142. Ivars K, Nelson N, Theodorsson A, et al. Development of salivary cortisol circadian rhythm and reference intervals in full-term infants. *PLoS One* 2015; 10: e0129502.
143. Vermes I, Dohanics J, Toth G, et al. Maturation of the circadian rhythm of the adrenocortical function in human neonates and infants. *Horm Res* 1980; 12: 237–244.
144. De Weerth C, Zijl RH and Buitelaar JK. Development of cortisol circadian rhythm in infancy. *Early Hum Dev* 2003; 73: 39–52.