1. Title page:

**Title:** Determination of the Method-Specific Normal Serum and Salivary Cortisol Response to the Short Synacthen Test

**Name of investigational product:** Synacthen (Tetracosactide)

**Indications studied:** Healthy volunteers, Patients with adrenal insufficiency, nephrotic syndrome or liver cirrhosis

**Sponsor:** Cardiff University

**Protocol identification:** EudraCT No: 2007-000056-14
Ethics reference No: 08/WSE04/13
R+D No: 06/DTD/3791E

**Development phase:** Phase IV

**Study initiation date (first patient visit):** 01.10.2008

**Study completion date (last patient visit):** 28.07.2012

**Principal investigator:** Dr Aled Rees, Centre for Endocrine and Diabetes Studies, School of Medicine, Cardiff University, Heath Park, Cardiff, CF14 4XN, United Kingdom.
Email: reesda@cf.ac.uk. Telephone: (029) 20742341. Fax: (029) 20744671

**Sponsor signatory:** Dr K J Pittard-Davies, Research and Commercial division, Cardiff University, 7th floor, 30-36 Newport Road, Cardiff, CF24 0DE.
Email: davieskp2@cf.ac.uk. Tel: (029) 20874189.

This study was performed in compliance with Good Clinical Practice (GCP) including the archiving of essential documents.
Date of Report: 23rd May, 2013

2. Synopsis:

**Name of Sponsor:** Cardiff University

**Name of finished product:** Tetracosactide (Synacthen)

**Name of active ingredient:** Synthetic Adrenocorticotrophin (ACTH)

**Title of study:** Determination of the Method-Specific Normal Serum and Salivary Cortisol Response to the Short Synacthen Test

**Investigators:** Dr Aled Rees (Principal investigator), Dr Carol Evans (Co-investigator), Dr Nadia El-Farhan (Co-investigator).

**Study Centre:** Clinical Research Facility, University Hospital of Wales, Heath Park, Cardiff

**Publication:** Method-specific serum cortisol responses to the adrenocorticotrophin test: comparison of gas chromatography-mass spectrometry and five automated immunoassays. El-Farhan N, Pickett A, Ducroq D, Bailey C, Mitchem K, Morgan N, Armston A, Jones L, Evans C, Aled Rees D. Clin Endocrinol (Oxf). 2012 Sep 20. [Epub ahead of print]

**Studied period:** First enrolment 01.10.2008 Last completed 28.07.2012

**Phase of development:** phase IV

**Objectives:** 1. To define the serum cortisol response to Synacthen in normal volunteers using the gold-standard gas chromatography- mass spectrometry (GC-MS) method and five commercially available cortisol immunoassays, and to compare the performance of cortisol immunoassays to the GC-MS assay. 2. To investigate the effect of oestradiol-containing oral contraceptives and low protein status on the cortisol response to Synacthen.

**Methodology:** An ACTH test (250 micrograms iv $ACTH_{1-24}$) was undertaken in healthy volunteers, patients with adrenal insufficiency and patients with low serum albumin. Serum cortisol in the samples collected from healthy volunteers was measured using GC-MS, Advia Centaur (Siemens), Architect (Abbott), Modular Analytics E170 (Roche), Immulite 2000 (Siemens) and Access (Beckman) automated immunoassays. The estimated lower reference limit for the 30 min cortisol response to ACTH was derived from the 2.5th percentile of log-transformed concentrations in this healthy population.

**Number of subjects:** 206 (165 healthy volunteers; 30 potential adrenal insufficiency; 11 hypoalbuminaemia – nephrotic syndrome or liver cirrhosis)

**Diagnosis and main criteria for inclusion:** To be eligible to take part the healthy volunteers had to be in self-proclaimed good health, free of illness on the day of testing and not taking any drug therapy. The 30 patients in the adrenal insufficiency group were recruited from the Endocrine Investigation unit after referral by an Endocrinologist for a short Synacthen test either to exclude adrenal insufficiency on the basis of symptoms or to evaluate adrenal function in patients with known pituitary disease on steroid replacement therapy. Patients in the hypoalbuminaemia group had a clinical diagnosis of nephrotic syndrome or liver cirrhosis and were recruited by consultants in Nephrology and Hepatology respectively. The only inclusion criterion was an albumin concentration < 35 g/L.

**Test product, dose and mode of administration:** Synacthen 250 ug, administered intravenously

**Duration of treatment:** Single patient visit; duration of Synacthen test – 30 minutes.

**Statistical methods:** Data analysis was performed using SPSS version 16.0 (SPSS Inc., Chicago, Illinois). All data were log-transformed to create a Gaussian distribution for analysis. Differences between method means and bias ratios were compared using the paired t-test. Bland-Altman plots and correlation graphs were also used to demonstrate assay differences. Gender differences and differences between non-OCP females and OCP-females were evaluated using the unpaired t-test or Mann-Witney U test in cases of non-parametric data. In all cases, differences were considered to be significant when $P < 0.05$.

**Summary – Conclusions:**

**Safety Results:** Overall the study medication was well tolerated. Eleven adverse events were reported during the course of the study, 6 of which were expected (nausea, rash, dizziness, facial flushing) and 5 unexpected (headache, tiredness, tearfulness, paranoia & anxiety, pallor & clamminess, swollen ankle). Most adverse events were mild in intensity and all had resolved by the time of the follow-up phone call a week after the Synacthen test.

**Conclusion:** Normal cortisol responses to the ACTH test are influenced significantly by assay and oestrogen treatment. This study provides immunoassay users with method-specific estimated lower cortisol limits for the SST. It also identifies the need for separate reference limits in premenopausal women on the OCP, and the potential for assay interference in cortisol measurements in this subgroup.

**Date of report:** 23rd May, 2013

**3. Table of contents:** **Page number**

**4. List of abbreviations:**

| | |
|---|---|
| ACTH | Adrenocorticotrophin |
| CBG | Cortisol binding globulin |
| CV | Coefficient of variation |
| EQA | External quality assurance |
| GC-MS | Gas chromatography – Mass spectrometry |
| HPA | Hypothalamic-pituitary-axis |
| ITT | Insulin tolerance test |
| LC – MS | Liquid chromatography – tandem mass spectrometry |
| OCP | Oestradiol-containing oral contraceptive pill |
| RIA | Radioimmunoassay |
| SD | Standard deviation |
| SPSS | Statistical Package for the Social Sciences |
| SST | Short Synacthen test |
| UHW | University Hospital of Wales |
| UKNEQAS | United Kingdom National External Quality Assurance Scheme |
| WEQAS | Welsh External Quality Assurance Scheme |

**5. Ethics:**

***5.1 Independent Ethics Committee***

The study and all amendments were approved by the Joint Cardiff and Vale NHS Trust/Cardiff University Peer and Risk Review Committee, the South East Wales Research Ethics Committee and the Medicines and Healthcare Products Regulatory Authority. The study was registered on the Clinical Trials.Gov website and ascribed the registration number NCT00851942 (http://clinicaltrials.gov/).

***5.2 Ethical Conduct of the Study***

The study was conducted in accordance with the ethical principles that have their origins in the Declaration of Helsinki

### *5.3 Patient information and consent*

Potential participants were recruited from staff at the University Hospital of Wales (UHW) and Cardiff University and their friends. Posters 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 to pass on to colleagues and friends. Potential participants were also identified as they passed through the Endocrine, Renal and Hepatology clinic systems at UHW (Dr D A Rees, Dr S Riley and Dr L Sunderraj). Eligible subjects were invited to participate in these clinics by the clinician responsible for their care, and a patient information sheet was provided.

This explained 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. Each subject was informed that participation in the study was voluntary and he/she may withdraw from the study at any time and that withdrawal of consent would not affect his/her subsequent medical treatment or relationship with the treating physician. Subjects were given adequate time to review the patient information sheet and ask questions about any aspects of the study. Consent was obtained in the Clinical Research Facility or B7 Endocrine Unit at UHW by one of the following individuals:

Dr Nadia El-Farhan (Specialist Registrar)
Dr Aled Rees (Principal Investigator)
Sister Laila Jones (Research Nurse)
Sister Janet Lewis (Endocrine Nurse)
Sister Nikki Davies (Endocrine Nurse)

Informed consent was given by means of a standard written statement, written in non-technical language. The subject read and considered the statement before signing and dating it, and was given a copy of the signed document. No patient could enter the study before his/her informed consent had been obtained.

## 6. Investigators

a) Investigators

Dr Aled Rees MB BCh (Hons), PhD, MRCP, Senior Lecturer, Centre for Endocrine and Diabetes Sciences, School of Medicine, Cardiff University, Heath Park, Cardiff (Principal Investigator)

Dr Carol Evans PhD, FRCPath, Consultant Clinical Scientist, Department of Medical Biochemistry and Immunology, UHW, Cardiff (Co-investigator)

Dr Nadia El-Farhan, MB ChB, MRCP, FRCPath, Consultant Chemical Pathologist, Royal Gwent Hospital, Newport

Sister Laila Jones, Research Nurse, UHW, Cardiff

b) Author of the report: Dr Aled Rees

Responsible statisticians:

Dr Nadia El-Farhan and Dr Aled Rees with advice from Professor Robert G Newcombe PhD, Department of Primary Care and Public Health, School of Medicine, Cardiff University, Cardiff

## 7. Introduction

The adrenocorticotropin (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 non-specific, 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 criteria 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 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 estrogens, 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 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.

## 8. Study Objectives

- To define the cortisol response to Synacthen in normal volunteers using GC-MS.
- To define the method-specific cortisol response to Synacthen in normal volunteers using five commercially available cortisol immunoassays.
- To compare cortisol assay performance against GC-MS.
- To investigate the effect of oestradiol-containing oral contraceptives on the cortisol response to Synacthen.

## 9. Investigational Plan

### *9.1 Overall study design*

This was a prospective study of the normal cortisol response to ACTH stimulation in healthy volunteers. Subjects' involvement in the trial lasted one week, with each subject attending for a single visit, lasting 90 minutes, during which a 30 minute Synacthen test was performed. The primary endpoint was to establish the lower reference limits for the cortisol response to Synacthen stimulation in healthy subjects; therefore, it was necessary that the number of participants was sufficient to allow determination of useful ranges (see below for *determination of sample size*).

Because Addison's disease is such a rare disease it was deemed impractical to recruit a large enough group to allow case-control comparisons. Thus, target recruitment for the patient groups was for a minimum of 10 participants, to allow broad differences between the groups to be identified. Establishing a reference cut-off for patients with hypoadrenalism or hypoalbuminaemia was beyond the scope of this study.

### *9.2 Selection of study population*

#### *9.2.1 Inclusion criteria*

To be eligible to take part 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.

#### *9.2.2 Exclusion 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. Of 172 healthy subjects showing interest a total of 7 were excluded: due to asthma (3), topical steroid use (1), history of allergy to aspirin and walnuts (1), autoimmune hypothyroidism (1) and difficulty in venepuncture (1).

## 9.4 Treatments

### *9.4.1 Treatments administered*

All study participants received intravenous synthetic $ACTH_{1-24}$ (Synacthen, Alliance Pharmaceuticals Ltd, Wiltshire, UK).

### *9.4.2 Selection of dose and timing of dose*

The standard Synacthen dose of 250 micrograms was used. The test was performed in the morning between 08.30 and 11.30 h, as cortisol secretion is diurnal – peaking in the early morning and falling throughout the day.

### *9.4.3 Prior and concomitant therapy*

One of the inclusion criteria for healthy volunteer participation in the trial was that subjects should not be taking drug therapy. For the two patient groups (Adrenal insufficiency and Hypoalbuminaemia) a list of all medication was recorded at the time of the Synacthen test. Patients taking exogenous steroids were excluded from taking part.

## 9.5 Efficacy and safety variables

Participants attended the Clinical Research Facility or Endocrine unit once for the Synacthen test to be performed. Pre-menopausal women underwent a pregnancy test and were excluded from participation if this was positive. A follow-up phone call, to exclude any late-onset adverse events, was made one week after the participant had attended for the Synacthen test. Four investigators (Nadia El-Farhan, Laila Jones, Janet Lewis and Nikki Davies) were responsible for performing Synacthen tests and completing the follow-up phone call.

One aliquot of sample from each time point was analysed upon completion of the Synacthen test and any results that failed to reach the 550 nmol/L cut-off in use at the time were referred for assessment by an Endocrinologist, to exclude genuine disease of the HPA axis.

### *9.5.1 Synacthen Test standardisation*

The short Synacthen tests were carried out in the Clinical Research Facility at the UHW 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 to enable interpretation of 17-OHP measurements in this cohort, but in men and older women there were no particular timing requirements.

Once appropriate consent had been obtained, subjects were asked to collect a 5ml saliva sample by passive drooling into a Universal container. A 21-gauge butterfly needle was inserted into a superficial antecubital vein and a 20ml serum sample was collected followed by intravenous administration of a 250mcg bolus of synthetic $ACTH_{1-24}$ (Tetracosactide) (Synacthen, Alliance Pharmaceuticals Ltd, Wiltshire). Thirty minutes after the administration of Synacthen a further 20ml serum sample was collected and subjects were asked to collect a second 5ml saliva sample.

### *9.5.2 Sample handling*

Baseline and 30 minute serum samples were split into 9 aliquots; one was analysed directly on the Siemens Centaur assay, while the remainder were stored frozen with the saliva samples at -20°C. When samples were later removed for analysis, one aliquot was re-analysed on the Centaur to exclude sample degradation. The remaining aliquots were sent to external laboratories for analysis in batches, comprising a maximum of 17 samples per batch, once a week over a period of 10 weeks.

### *9.5.3 Analytical methods*

Total cortisol 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[12]. Interassay CVs for this assay were 5.4%, 6.2% and 6.8% at concentrations of 120, 363 and 657 nmol/L respectively. Total cortisol was also analysed by five automated immunoassays: the Advia Centaur (Siemens AG, Erlangen, Germany) with inter assay CVs of 5.2, 4.5 and 2.9% at concentrations of 105,

571 and 784 nmol/L 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 nmol/L respectively, Immulite 2000 (Siemens AG, Erlangen, Germany) with interassay CVs of 10.2, 7.7 and 7.2% at concentrations of 117, 695 and 978 nmol/L respectively, Access (Beckman Coulter, Brea, CA) with interassay CVs of 9.9, 5.0 and 2.5% at concentrations of 124, 620 and 828 nmol/L respectively, and Architect (Abbott Laboratories, Illinois) with interassay CVs of 10.4%, 5.4% and 6.8% at concentrations of 97, 549 and 840 nmol/L respectively. All assays were solid-phase competitive binding immunoassays using chemiluminescent detection, except the E170 which uses electrochemiluminescent detection.

Aliquots of 3 different serum pools with GC-MS assigned target concentrations (76.1, 527.7 and 696.4nmol/L respectively) were sent with each batch of samples. Analysis of these serum pools over the 10 week period provided the data for determining the interassay (between-batch) precision for each assay.

All participating laboratories were enrolled in the UK National External Quality Assurance Scheme (NEQAS) at the time of analysis and their cortisol assays were performing well within their method groups.

### *9.5.4 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, was collected and recorded. An adverse event was defined as any undesirable sign, symptom or medical condition occurring during, immediately after, or within one week of the Synacthen test.

Each adverse event was also described by:

1. Its duration
2. The severity grade (mild, moderate, severe)
3. Its relationship to the Synacthen test (suspected, non suspected)
4. The action(s) taken

*Serious adverse events*

No serious adverse events occurred in this study.

### 9.5 Data quality assurance

Dr Nadia El-Farhan (Co-investigator) was responsible for the day-to-day conduct of the study. She was also responsible for co-ordinating recruitment and ensuring the study protocol was adhered to throughout. Obtaining consent, performing the Synacthen tests and completing the follow-up call were carried out by Dr El-Farhan, Sisters Laila Jones, Nikki Davies and Janet Lewis. Sample handling, storage and analysis were co-ordinated by Mr Alan Pickett (Senior Biomedical Scientist, Biochemistry Department, UHW). Dr El-Farhan was responsible for ensuring all data was entered correctly and samples were stored appropriately. At the end of the study, she conducted the statistical analyses, with advice from Dr Aled Rees and Professor Robert Newcombe (Department of Primary Care and Public Health, Cardiff University).

Dr Aled Rees, the Principal Investigator and Dr Carol Evans, acted as Dr El-Farhan's supervisors providing advice and guidance for the clinical and laboratory aspects of the study, respectively. A formal meeting took place on a monthly basis to discuss study progress and future direction.

Source data verification was undertaken by Dr Rees at regular intervals.

### 9.7 Statistical Methods

#### *9.7.1 Statistical plan*

Data analysis was performed by Dr Nadia El-Farhan with support from Dr Aled Rees and Dr Carol Evans. Professor Robert Newcombe was consulted regarding the non-Gaussian distribution of some data and advised on the most appropriate method for transforming these. Subsequently, all data were log-transformed to create a Gaussian distribution for analysis. Differences between method means and bias ratios were compared using the paired t-test. Bland-Altman plots and correlation graphs were also used to demonstrate assay differences. Gender differences and differences between

non-OCP females and OCP-females were evaluated using the unpaired t-test or Mann-Witney U test in cases of non-parametric data. In all cases, differences were considered to be significant when $P < 0.05$.

Data analysis was performed using SPSS version 16.0 (SPSS Inc., Chicago, Illinois).

### *9.7.2 Determination of 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, our data confirmed the findings by Clark et al[8] that the distributional properties are independent of age. Thus, univariate reference ranges are appropriate.

The International Federation of Clinical Chemists (IFCC) and US Food and Drug Administration (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.

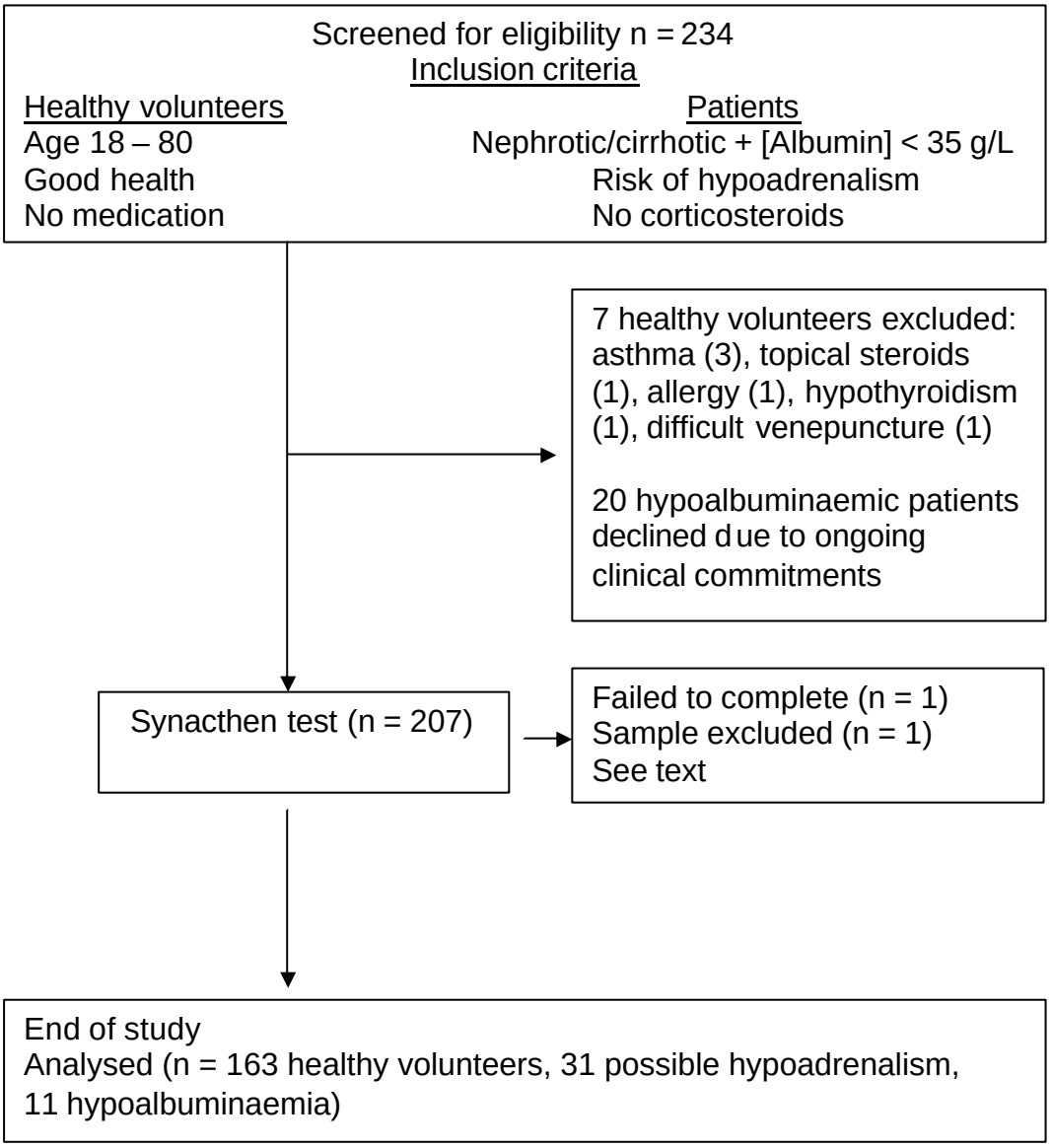
Our data was non-Gaussian, but transformed logarithmically into a Gaussian distribution. Thus, our sample size of 60 male and 81 female participants (excluding the group of 24 women taking an oestrogen-containing oral contraceptive pill) was large enough to allow gender-specific reference ranges to be established.

## 10. Study Patients

### *10.1 Disposition of patients*

Figure 1 summarises the flow of participants through the study. Two-hundred and seven participants completed the study. No subjects withdrew from the study, but in one case the Synacthen test could not be completed due to difficulty obtaining blood after Synacthen had been administered. Samples from a second volunteer were withdrawn from analysis due to underlying Hepatitis B infection.

Figure 1: Flow of participants through the study:



## 11. Efficacy evaluation

### *11.1 Baseline cortisol*

When measured by GC-MS, baseline cortisol was normally distributed in male but not non-OCP female volunteers, with no significant gender difference in mean values (table 1) or any effect of age ($R^2 = 0.003$, 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.

**Table 1: Geometric mean of baseline cortisol concentrations in male, non-OCP female and OCP-female subjects.**

| Assay | Males n = 60 | Non-OCP Females n = 79 | p-value[c] | OCP Females n = 24 | p-value[d] |
|---|---|---|---|---|---|
| **GC-MS** | 274 (131 - 575) | 254 (139 - 463) | 0.193 | 542 (318 - 922) | < 0.001 |
| **Centaur** | 298 [a] (158 - 565) | 257 [a] (138 - 477) | 0.023 | 488 (323 - 738) | < 0.001 |
| **Architect** | 289 [a] (151 - 556) | 247 [b] (134 - 455) | 0.018 | 465 (301 - 718) | < 0.001 |
| **E170** | 370 [a] (182 - 750) | 292 [a] (147 - 581) | 0.001 | 646 (383 - 1090) | < 0.001 |
| **Immulite (2000)** | 316 [a] (165 - 604) | 267 [a] (144 - 495) | 0.003 | 510 (330 - 788) | < 0.001 |
| **Access** | 293 [a] (160 - 538) | 252 [b] (143 - 444) | 0.011 | 429 (286 - 643) | < 0.001 |

Results are expressed as geometric mean ($2.5^{th}$ – $97.5^{th}$ percentile) in nmol/L.

[a] p-value for immunoassay vs GC-MS <0.005

[b] p-value 0.95 and 0.21 for Architect and Access assays vs GC-MS respectively.

[c] p-value for difference between genders.

[d] p-value for difference between women taking an oral contraceptive pill and those who were not.

### *11.2 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 post-stimulation 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).

**Table 2: Geometric mean of post-ACTH stimulation cortisol concentrations in male, non-OCP female and OCP-female subjects.**

| Assay | Males n = 60 | Non-OCP Females n = 79 | p-value [b] | OCP Females n = 24 | p-value [c] |
|---|---|---|---|---|---|
| **GC-MS** | 563 (418 - 757) | 555 (421 - 731) | 0.594 | 870 (643 - 1177) | < 0.001 |
| **Centaur** | 599 [a] (448 - 802) | 578 [a] (446 - 750) | 0.138 | 763 (619 – 940) | < 0.001 |
| **Architect** | 577 [a] (430 - 773) | 542 [a] (416 - 707) | 0.012 | 747 (577 - 967) | < 0.001 |
| **E170** | 772 [a] (574 - 1039) | 712 [a] (524 - 967) | 0.003 | 1026 (791 - 1330) | < 0.001 |
| **Immulite (2000)** | 641 [a] (469 - 874) | 628 [a] (478 - 826) | 0.449 | 850 (688 - 1051) | < 0.001 |
| **Access** | 625 [a] (459 - 852) | 594 [a] (455 - 777) | 0.045 | 757 (604 - 948) | < 0.001 |

Results are expressed as geometric mean ($2.5^{th}$ – $97.5^{th}$ percentile) in nmol/L.

[a] p-value for immunoassay vs GC-MS <0.02

[b] p-value for difference between genders.

[c] p-value for difference between women taking an oral contraceptive pill and those who were not.

### *11.3 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 nmol/L), although this relationship varied based on gender and exogenous oestrogen use. Figure 2 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 which 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 (Figure 2). This OCP-dependent difference in assay behaviour was further explored in Bland-Altman plots (Figure 3) 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 3: Bias ratios for cortisol immunoassay compared to GC-MS.**

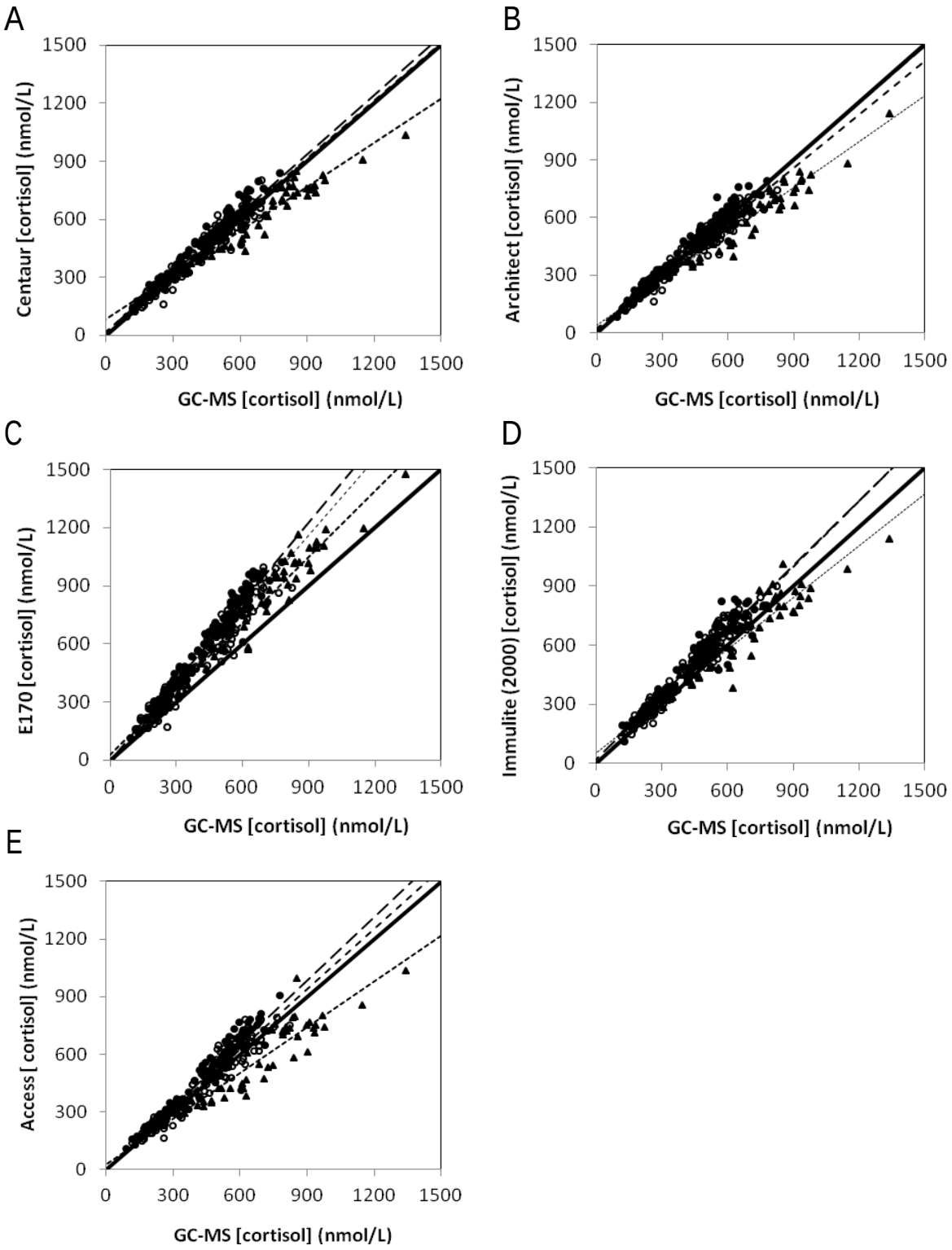
| Assay | Males | Non-OCP Females | p-value [a] | OCP Females | p-value [b] |
|---|---|---|---|---|---|
| **Centaur** | 1.08 | 1.05 | 0.012 | 0.88 | <0.001 |
| **Architect** | 1.04 | 1.00 | <0.001 | 0.85 | <0.001 |
| **E170** | 1.36 | 1.25 | <0.001 | 1.18 | <0.001 |
| **Immulite (2000)** | 1.15 | 1.11 | 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 minute) and post-ACTH (30 minute) cortisol concentrations.

[a] p-value for difference between genders.

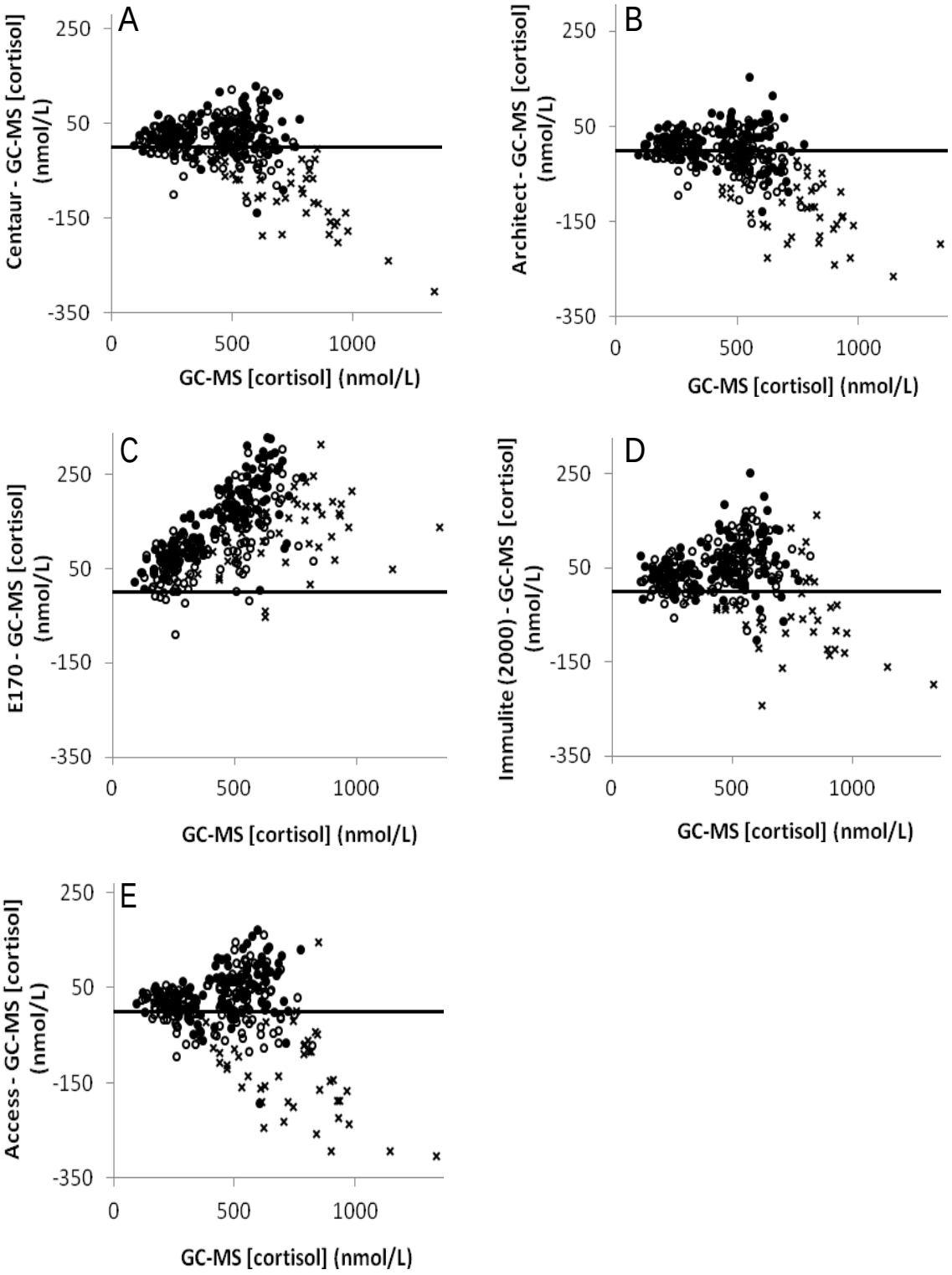
[b] p-value for difference between women taking an oral contraceptive pill and those who were not.

**Figure 2**




Correlation plots demonstrating the relationship between immunoassay and GC-MS cortisol for all baseline and post-ACTH stimulation cortisol measurements combined, for male (●), non-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.

**Figure 3**




Bland-Altman plots showing the difference between immunoassay and GC-MS cortisol plotted against GC-MS cortisol concentration for all baseline and post-ACTH 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-OCP female subjects; ×, OCP-female subjects.

### *11.4 Assay-specific lower reference limits of cortisol post-ACTH*

The lower reference limit of cortisol concentration 30-minutes post-ACTH for GC-MS was 417 nmol/L, 422 nmol/L and 649 nmol/L 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). Since mean cortisol concentration post-ACTH stimulation was not gender-dependent for the GC-MS, Centaur 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 since 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 nmol/L and 4 nmol/L respectively) to be within expected assay variability (assay precision was 5.4% and 5.0% respectively at 549 and 620 nmol/L) 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 nmol/L) making gender related reference ranges necessary.

**Table 4: Assay-specific estimated lower reference limits for post-ACTH cortisol according to gender and OCP-status.**

| Assay | Males | Non-OCP females | Combined male and Non-OCP female subjects[a] | 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 2.5$^{th}$ percentile value (mean – 1.97*SD) of the log-transformed data.

Results are expressed in nmol/L

[a] A single assay-specific combined LRL for males and non-OCP females was calculated by pooling all post-ACTH cortisol results in these subgroups.

NA = not applicable

## 12. Safety evaluation

Overall the study medication was well tolerated.

There were no serious adverse events associated with the use of Synacthen in this trial. A total of 11 subjects out of a study population of 207 experienced adverse events, 7 of which were expected (headache, nausea, “fluttery” stomach, rash, dizziness, facial flushing) and 4 unexpected events (tiredness, tearfulness, paranoia & anxiety/panic attacks, pallor & clamminess, swollen ankle). Most adverse events were mild in intensity and self-limiting; all had resolved by the time of the follow-up phone call a week after the Synacthen test.

The study was designed to minimise risk to recruited subjects. The risk of overdose was reduced by using a standard dose of Synacthen, packaged in individual vials. In healthy volunteers, treatment with any medication other than the oral contraceptive pill was an exclusion criterion, thus reducing the risk of drug interactions. In our patient populations, a full drug history was taken prior to the short Synacthen test being carried out. All subjects are observed for signs of adverse effects for a period of 90 minutes after administration of Synacthen. Pregnant and lactating women were excluded as were subjects under 18 years of age.

There were no drop-outs for safety reasons and no new safety findings arose from the use of Synacthen in this trial. Adverse events occurred infrequently and all reported symptoms resolved quickly.

A summary of the adverse events is provided in table 5.

Table 5: Adverse events recorded during the study

| AE number | Description | Duration | Outcome | Intensity | Expected-ness | Causality | Seriousness |
|---|---|---|---|---|---|---|---|
| 1 | Facial flushing | 30 mins | Resolved | Moderate | Expected | Prob related | Not serious |
| 2 | Headache | 9 hrs | Resolved | Moderate | Unexpected | Poss related | Not serious |
| 3 | Dizziness | 2 mins | Resolved | Mild | Expected | Prob related | Not serious |
| 4 | Pallor, clamminess | 5 mins | Resolved | Moderate | Unexpected | Poss related | Not serious |
| 5 | Itchy rash – arms/lower legs | 5.5 days | Resolved | Mild | Expected | Poss related | Not serious |
| 6 | Tearfulness, paranoia, anxiety | 4 days | Resolved | Mild | Unexpected | Poss related | Not serious |
| 7 | Nausea | 20 mins | Resolved | Mild | Expected | Prob related | Not serious |
| 8 | Swollen L ankle | 3 days | Resolved | Mild | Unexpected | Unrelated | Not serious |
| 9 | Nausea | 5 days | Resolved | Mild | Expected | Poss related | Not serious |
| 10 | Nausea | 5 days | Resolved | Mild | Expected | Poss related | Not serious |
| 11 | Headache, stress | 24 hrs | Resolved | Mild | Unexpected | Poss related | Not serious |

## 13. Discussion and overall conclusions

This is the largest study to examine normative responses to the standard dose ACTH test and the first to compare results 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 estrogens, 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* 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[8]. In contrast, Klose and colleagues, who confirmed clinically significant inter-assay differences in stimulated cortisols, did not find an effect of gender nor evidence for non-normal distribution[9]. 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 estrogen-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 estrogen and consequent effects on 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 estrogen treatment on the cortisol response to ACTH stimulation whether measured by GC-MS or immunoassay. Compared with women not taking

estrogen, 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 minutes respectively. This accords with our understanding of a marked stimulatory influence of estrogen on CBG production[17] and is in agreement with the findings of Klose *et al* who recommended that this test should only be undertaken after the OCP has been discontinued[9]. 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 to 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 estrogen as eucortisolaemic if estrogen status is not taken into account. We thus share Klose and colleagues' views that consideration should be given to stopping estrogen 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 nmol/l in non-OCP females with the Architect assay to 772 nmol/l 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 estimation of assay bias relative to GC-MS.

In contrast to mean responses, the lower limit of the normative range, defined as the 2.5$^{th}$ percentile value of the log transformed data, showed much less of a gender difference, ranging from 2 nmol/L to 14 nmol/L 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 nmol/L) 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 since 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 since application of the widely used cortisol cut-offs of 500[19,20] or 550 nmol/l[2,21,22] would result in misclassification of a significant

number of healthy individuals with adrenal insufficiency (12%, 19%, 4% & 9% for the Centaur, Architect, Immulite (2000) and Access assays respectively at 500 nmol/l; 27%, 42%, 16% & 21% for the respective assays at 550 nmol/l). Our findings thus have the potential to reduce inappropriate and costly follow-up investigations for patients referred with a low pre-test 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* did not find a difference in 30 minute cortisol values between the fasting and non-fasting state[9]. 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 minute cortisol response in men. However, since 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* showed that CBG and total cortisol concentrations fall significantly within 30 minutes of adopting a supine from a standing position[22]. We did not specifically test the influence of posture but standardised our protocol to collect samples in the sitting position, adopted at least 15 minutes 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 ACTH test are influenced significantly by assay and treatment with estrogens. We also report a negative assay bias in women on the OCP which further complicates interpretation of the ACTH 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-ACTH cortisol results. In light of our findings, we also recommend that normative responses to the insulin stress test may need to be re-established.

## 14. Reference list

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., Frolich, 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.

## 15. Appendices

The following appendices are available for the study report on request:

15.1 Study protocol and amendments

15.2 Sample case report form

15.3 Healthy volunteer/patient information sheets, consent forms and GP letter