Candidate Reference Measurement Procedure for the Quantification of Total Serum Cortisol with LC-MS/MS

James M. Hawley,1* Laura J. Owen,1,2 Finlay MacKenzie,3 Chris Mussell,4 Simon Cowen,4 and Brian G. Keevil1,2

BACKGROUND: Accurate measurement of serum cortisol is required to diagnose and treat adrenal disorders. Although certified reference materials (CRMs) are available to standardize cortisol measurements, External Quality Assessment (EQA) schemes still demonstrate a wide dispersion of results. We present a serum cortisol candidate reference measurement procedure that, through analysis of a Joint Committee for Traceability in Laboratory Medicine–listed panel of higher-order CRMs, provides metrologically traceable results.

METHOD: Isotope-labeled internal standard was added to samples before supported liquid extraction. Extracts were analyzed with LC-MS/MS in positive electrospray ionization mode. Multiple reaction monitoring was used to detect cortisol and its corresponding internal standard transitions. We measured samples in triplicate over 3 days and calculated the mean result.

RESULTS: Mean intra- and interassay imprecision were 1.3% and 1.5%, respectively, for concentrations of 154, 510, and 769 nmol/L. Ionization efficiency studies and structural analog analysis proved the method to be robust against interferences. Through analysis of 34 CRMs (83–764 nmol/L), expanded measurement uncertainty was calculated to be 5% (95% CI). The mean bias between the measured and target CRM concentrations was statistically insignificant at −0.08%.

CONCLUSIONS: The accuracy and low measurement uncertainty of this method qualify it as a CRM procedure. Metrological traceability has been achieved through the analysis of higher-order CRMs. This method could be used to underpin serum cortisol EQA schemes to provide samples with a traceable target value, enabling participating laboratories to determine the accuracy and measurement uncertainty of their assays.

A standard reference material (SRM)5 for the measurement of cortisol has been available from NIST since 1973. Since its release, SRM 921 has been used to develop several reference measurement procedures (RMPs) for the quantification of total serum cortisol (1–6). These methods have subsequently been used to assign target values to certified reference materials (CRMs) to promote the standardization of routine cortisol assays. Agreement across RMPs is ensured through participation in the IFCC’s External Quality Assessment Scheme for Reference Laboratories in Laboratory Medicine (7). Furthermore, work has been conducted by NIST in collaboration with the Bureau International des Poids et Mesures to assess the agreement of serum cortisol RMPs across National Metrology Institutes. Recent results were excellent, with an interlaboratory relative SD across the 6 participating National Metrology Institutes of 1.1% at a mean cortisol concentration of 100.5 ng/g (8).

Despite the outstanding performance and concurrence of existing RMPs, the performance of routine laboratory immunoassays remains highly variable and prone to matrix effects and interferences (9–11). This variability is particularly relevant in the differential diagnosis of hypo- or hypercortisolism, for which measurement of serum cortisol is integral to dynamic function test interpretation. Dynamic function tests are designed to assess the hypothalamic-pituitary-adrenal axis through an exogenous challenge that either stimulates or inhibits cortisol production. Test interpretation is traditionally based on single-value cutoffs; however, these cutoffs have been shown to be variable across assays (12). In addition to their diagnostic utility, serum cortisol measurements...
are used to assist in the management of patients with hypercortisolism, for which treatment with the 11β-hydroxylase inhibitor metyrapone may be titrated against the cortisol concentration. Recent work has demonstrated that routine assays suffer from variability in measuring cortisol in serum obtained from patients on metyrapone (13). Hence, accurate and precise serum cortisol measurements are fundamental to correct patient diagnosis and management. The variability observed across current immunoassays therefore represents a potential risk to patient care.

Cortisol is not the only steroid hormone that has been well characterized yet suffers from poor performance on routine platforms. This problem has been widely acknowledged, and recent Endocrine Society guidelines for the minimum performance of steroid assays have recognized that metrological traceability of an assay to a certified standard is not currently available for all steroid hormones, yet remains a necessary goal (14–16). This important work has already commenced, with initiatives such as the Hormone Standardization Program for testosterone (17, 18) and the Vitamin D Standardization Certification Program (19). These programs have culminated in the formation of the broader Partnership for the Accurate Testing of Hormones initiative, and central to this is External Quality Assessment (EQA) performance. EQA schemes are prioritized the standardization of estradiol and thyroid hormone assays (20). Crucially, the standardization of routine cortisol assays has yet to be considered, despite their poor routine performance relative to the clinical importance of cortisol.

Verification of end-user performance is a vital component of the Partnership for the Accurate Testing of Hormones initiative, and central to this is External Quality Assessment (EQA) performance. EQA schemes are designed to assess the performance of routine methods, with both commutability of the material and accuracy of the participating methods contributing to the variability of results. Whereas EQA scheme organizers can minimize the effects of sample noncommutability by adhering to consensus guidelines (21), the true assessment of accuracy can be made only through underpinning the scheme with a reference system that is metrologically traceable to SI units. This can be achieved through the use of an RMP.

Current serum cortisol RMPs are not routinely available and typically require large sample volumes, thus limiting their use to underpin EQA schemes. These methods primarily use isotope dilution (ID)-GC-MS (1–4). However, advances in sample preparation, liquid chromatography, and mass spectrometry have resulted in the ability to accurately and precisely measure serum cortisol with ID-LC-MS/MS, thereby enabling the development of LC-MS/MS reference systems (5, 6).

The aim of this work was therefore to develop and validate an LC-MS/MS serum cortisol RMP that could provide metrologically traceable target values for the UK National External Quality Assessment Scheme (NEQAS) serum cortisol EQA scheme.

Materials and Methods

We used cortisol CRM at a concentration of 1 g/L as a calibration standard (Cerilliant). Stable deuterium-labeled cortisol-d4 (9,11,12,12-d4) with an isotopic purity of 97% (CDN Isotopes) was used as an internal standard at a working concentration of 350 nmol/L (128 μg/L) in 50% methanol/water (vol/vol) (LC-MS grade, Sigma).

PREPARATION OF CALIBRATORS AND QC MATERIALS

We prepared calibrators and QC materials in PBS 0.1% (wt/vol) BSA (Sigma). Calibrators ranged from 0 to 1000 nmol/L (0–362 μg/L), and QC materials were prepared at concentrations of 154, 510, and 769 nmol/L (56, 185, and 279 μg/L).

PREPARATION OF SAMPLES

We prepared samples in triplicate over 3 separate days. Each calibrator, QC material, and sample was manually pipetted with a calibrated positive-displacement pipette (Gilson Microman®) in 50-μL aliquots into wells of a 96-well plate (Porvair Sciences). To this, 20 μL internal standard was added, vortex-mix mixed for 30 s, centrifuged at 880 g for 5 min, and allowed to equilibrate for 30 min at ambient temperature. After equilibration, 150 μL deionized water was added to each well, vortex-mixed for 30 s, and centrifuged at 880 g for 5 min. A 200-μL aliquot was then transferred to the corresponding well of a 200-μL supported liquid extraction (SLE) plate (Biotage®), drawn on to diatomaceous earth with a vacuum pump for 30 s (KNF Neuberger), and left for 20 min. Samples were subsequently extracted with 1000 μL methyl tert-butyl ether (Fisher Scientific) into a second 96-well plate. The resulting eluate was dried down under nitrogen (Biotage SPE Dry), reconstituted in 100 μL 40% methanol/water (vol/vol), vortex-mixed, and centrifuged at 880 g for 5 min.

CHROMATOGRAPHY

We used an Acquity® ultra-performance liquid chromatography I-Class system (Waters) to achieve chromatographic separation. Mobile phases consisted of (A) deionized water with 2 mmol/L (154.2 mg/L) ammonium acetate (LC-MS grade, Sigma) and 0.1% formic acid (VWR International) and (B) methanol with 2 mmol/L (154.2 mg/L) ammonium acetate and 0.1% formic acid. The prepared sample (10 μL) was injected onto a KrudKatcher® Ultra HPLC 0.5-μm inline filter (Phenome-
mass spectrometer with a Z spray ion source. We used MassLynx NT version 4.1 software for system control and MassLynx TargetLynx for data processing. The mass spectrometer was operated in the positive electrospray ionization mode, the capillary voltage was maintained at 0.34 V, and the source temperature was 150°C. The desolvation temperature and gas flow were 850°C and 30 L/h, respectively. Cortisol transitions were monitored. Endogenous analogs tested to a concentration of 1000 nmol/L were beclomethasone, budesonide, cyproterone, 11-deoxycorticosterone, 21-deoxycorticosterone, cortisol, corticosterone, dehydroepiandrosterone sulfate, dihydrotestosterone, 17-hydroxyprogesterone, pregnenolone, progesterone, and testosterone. Exogenous analogs tested to a concentration of 1 mg/L were beclomethasone, budesonide, cyproterone, desogestrel, dexamethasone, epitestosterone, and any reduction or increase in the background signal was noted.

**RESULT CALCULATION**

To calculate the final result for a given sample, we averaged each set of triplicates to produce 3 results that corresponded to each day’s analysis. We then calculated their mean to produce the final value.

**IONIZATION EFFICIENCY**

Ionization efficiency is defined as the ratio of the number of ions generated to the number of molecules consumed in the source of a mass spectrometer. Ideally, this should remain constant; however, under certain conditions or in the presence of particular species, the ionization efficiency can be either reduced or increased, resulting in the respective phenomena of ion suppression or ion enhancement (22). To investigate ionization efficiency, serum, EDTA, and lithium heparin samples were collected in tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and EDTA, and lithium heparin samples were collected in tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greiner, and tubes supplied by Becton Dickinson, Greener. The respective cone voltages were 120.95 (quantification) and 363.1 (confirmation); 367.0 (367.0/H11022) 120.95. The respective cone and collision energies were 32 and 28 eV for both cortisol and cortisol-d4. All transitions were monitored in the multiple reaction monitoring (MRM) mode with a dwell time of 0.174 s.

**LIMIT OF DETECTION**

The limit of detection was defined as the concentration of cortisol that produced a peak with a signal-to-noise ratio of 3:1.

**SPECIFICITY**

To investigate analytical specificity, solutions of structurally analogous steroids were prepared in 40% methanol/water (vol/vol) and directly injected into the mass spectrometer via the autosampler. All cortisol and cortisol-d4 MRM transitions were monitored. Endogenous analogs tested to a concentration of 1000 nmol/L were aldosterone, androstenedione, corticosterone, cortisone, dehydroepiandrosterone sulfate, dihydrotestosterone, 11-deoxycorticosterone, 21-deoxycorticosterone, 17-hydroxyprogesterone, pregnenolone, progesterone, and testosterone. Exogenous analogs tested to a concentration of 1 mg/L were beclomethasone, budesonide, cyproterone, desogestrel, dexamethasone, epitestosterone,
ethinyl estradiol, fludrocortisone, fluticasone, fluocinolone, levonorgestrel, methylprednisolone, norethisterone, 19-nortestosterone, prednisolone, prednisone, and triamcinolone. The absence of a peak at the same retention time as cortisol or cortisol-d4 was used to confirm that the structural analogs tested did not interfere with quantification of serum cortisol. In addition, we examined the results of the 34 higher-order CRM samples to compare the mean difference between the quantifying ion and confirmatory ion concentrations. For further confirmation, we analyzed 152 anonymized surplus serum samples (50 male, 50 female, 25 prednisolone, and 27 metyrapone) with the candidate RMP and compared the quantifying:confirmatory ion ratio to that of a calibrator.

**MEASUREMENT UNCERTAINTY**

We calculated measurement uncertainty from the dispersion of the measured concentrations of the 34 CRMs, which included intra- and interday variation. Each assay was performed on a separate day. Because each of the 3 days consisted of a repeated instance of the entire analytical process, the measurement uncertainty incorporated all sources of variability present at the intermediate precision level.

**Results**

After SLE purification and LC-MS/MS analysis, cortisol and cortisol-d4 eluted after 3.00 and 2.97 min, respectively. Throughout the validation, m/z 363.1>96.95 was used as the quantifier in preference to the m/z 363.1>120.95 confirmatory transition. Although analytically more sensitive, the confirmatory transition suffered from incomplete baseline resolution from 2 adjacent peaks in samples prepared from patients taking prednisolone (Fig. 1, A–D).

Less than 10% ion suppression or enhancement was observed in the region of the chromatogram in which cortisol or its internal standard eluted in any of the collection tubes investigated. The structural analogs investigated were either chromatographically separated from both cortisol and cortisol-d4 or did not produce a notable signal in any of the MRM channels. The mean difference between the concentrations obtained for the 34 CRMs according to the quantifying ion and confirmatory ion transitions was 0.06% (95% CI, −0.17 to 0.28). The quantifying:confirmatory ion ratios of the 152 serum samples analyzed by the candidate RMP were examined and compared to the mean quantifying:confirmatory ion ratio of a calibrator (Table 1).

The recoveries ranged from 99% to 103%. Analysis of the 34 CRMs yielded a Passing–Bablok equation of LC-MS/MS = 0.9967 × ID-GC/MS + 0.4125 (95% CIs were (slope) 0.986–1.005 and (intercept) −1.485 to 3.108), with a correlation coefficient ($R^2$) of 1.0 (Fig. 2). Comparison of the results obtained from this assay against the certified target concentrations gave a Bland–Altman mean bias of −0.87 nmol/L (95% CI, −2.52 to 0.78) (Fig. 3).
Multiple injections of PBS 0.1% BSA QC pools yielded intraassay imprecisions of 1.7%, 0.8%, and 1.5% at concentrations of 154, 510, and 769 nmol/L (56, 185, and 279 μg/L), respectively. The interassay imprecision for these same PBS 0.1% BSA QC pools was 1.3%, 2.0%, and 1.1%. The interassay imprecision of the 4 matrix-matched samples was 1.4%, 1.1%, 1.3%, and 1.2% at concentrations of 174, 293, 376, and 702 nmol/L (63, 106, 136, and 254 μg/L), respectively.

The limit of detection was obtained from analysis of surplus anonymized prednisolone serum samples. Cortisol concentrations of 2 nmol/L (0.7 μg/L) produced a signal-to-noise ratio of 6:1. Extrapolating from this, the limit of detection was 1 nmol/L (0.4 μg/L).

Standard curves were constructed by plotting cortisol concentrations on the x axis and cortisol/cortisol-d4 peak height ratios on the y axis. The observed increase in variance with concentration was accounted for with weighted 1/x linear regression. Examination of the residuals and the fitting of a higher-order quadratic regression model provided no evidence that a further term was required. The curve was linear (R² > 0.999) over the calibration range of 0–1000 nmol/L (0–362 μg/L). Examination of the curve over 20 consecutive batches produced a mean (SD) R² value of 0.99967 (0.00017) (95% CI, 0.99960–0.99974) and a mean (SD) intercept of 0.0020 (0.0014) (95% CI, 0.0013–0.0026), showing that the calibration line was reproducible between batches.

An expanded measurement of uncertainty was calculated to be 5% (95% CI) across the range of 83–764 nmol/L (30–277 μg/L) (Fig. 4). The measured concentration was observed to follow the expected log-normal distribution with a constant relative SD. A log transformation produced a data set with a normal distribution and constant variance to which standard statistical methods could be applied. Details pertaining to the statistical analysis used to calculate the measurement of uncertainty, including fitted model coefficients, can be found in Supplemental Data, which accompanies the online version of this article at http://www.clinchem.org/content/vol62/issue1.

### Table 1. Quantifying: confirmatory ion ratios of calibrator solutions and patient samples.

<table>
<thead>
<tr>
<th>Sample</th>
<th>n</th>
<th>Concentration</th>
<th>Quantifying: confirmatory ion ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>nmol/L, μg/L</td>
<td>Mean (95% CI) Median (minimum-maximum)</td>
</tr>
<tr>
<td>Calibrator</td>
<td>21</td>
<td>25, 9.1</td>
<td>0.331 (0.329–0.332) 0.331 (0.323–0.335)</td>
</tr>
<tr>
<td>Male</td>
<td>50</td>
<td>108–400, 39.1–145.0</td>
<td>0.332 (0.331–0.332) 0.332 (0.325–0.338)</td>
</tr>
<tr>
<td>Female</td>
<td>50</td>
<td>96–866, 34.8–113.8</td>
<td>0.332 (0.331–0.332) 0.331 (0.324–0.339)</td>
</tr>
<tr>
<td>Metyrapone</td>
<td>27</td>
<td>91–535, 33.0–193.9</td>
<td>0.330 (0.329–0.330) 0.330 (0.324–0.335)</td>
</tr>
<tr>
<td>Prednisolone</td>
<td>25</td>
<td>2–99, 0.7–35.9</td>
<td>0.307 (0.300–0.314) 0.319 (0.220–0.353)</td>
</tr>
</tbody>
</table>

Fig. 2. Passing–Bablok plot comparing the candidate RMP result against the ID-GC/MS certified target concentration for the 34 CRMs.

Fig. 3. Bland–Altman plot comparing the candidate RMP results of the 34 CRMs against the ID-GCMS certified target values.

The mean bias of the candidate RMP was −0.87 nmol/L (95% CI, −2.52 to 0.78).
Discussion

There remains a clinically significant bias across routine cortisol assays, emphasizing a gap between the translation of reference measurements and the implementation of metrologically traceable results into clinical practice (9–13). This has potential ramifications on patient care, as the poor standardization and nonspecificity of routine assays can directly impact result interpretation and therefore patient management. In an attempt to close this gap and improve the standardization of routine assays, we developed an LC-MS/MS candidate RMP for the quantification of total cortisol in serum.

The application of Cerilliant CRM as a reference compound was verified through the analysis of 34 higher-order CRM samples, therefore completing the traceability chain. The Cerilliant material is manufactured and certified under International Organization for Standardization (ISO) Guide 34, ISO/International Electrotechnical Commission 17025, ISO 9001, and ISO 13485 and therefore qualifies as a CRM as defined by ISO. The use of Cerilliant CRM is preferable, since it is provided as a methanolic stock solution that requires only 1 weighing-in procedure to be diluted into the desirable working concentration range. Conversely, because NIST SRM 921 is a powder, additional preparation would be required to produce calibrators within the desirable working concentration range and therefore systematically increase measurement uncertainty.

Current RMPs typically use derivatization to alter the physiochemical properties of cortisol, making it amenable to GC-MS (1–4). Although derivatization can improve analytical sensitivity, there are potential disadvantages to this approach. Aside from the added expense and associated health and safety issues, the process of derivatizing can affect the assay itself by extending sample preparation time and increasing total analytical error. Furthermore, it has been suggested that the conditions required to derivatize can hydrolyse metabolites and conjugates to generate the analyte of interest, therefore reducing assay specificity (26).

The use of SLE plates for sample purification is advantageous, as they provide a reproducible and efficient 96-well support to ensure complete recovery of cortisol while minimizing the retention of undesirable compounds such as salts, proteins, and phospholipids. This is reflected by the negligible ion suppression or enhancement observed in multiple specimen collection tubes and

Fig. 4. Difference plot between the mean candidate RMP result and the certified concentration (nmol/L) for each of the 34 CRMs analyzed.

The error bars applied are the candidate RMPs’ expanded measurement uncertainty at the 95% CI.
the small level of noise in samples with low cortisol concentrations, allowing cortisol to be detected at concentrations of 1 nmol/L (0.4 μg/L).

Specificity has been ensured through use of an XSelect High Strength Silica pentafluorophenyl column. The pentafluorophenyl phase and isocratic elution allow excellent separation of structurally analog steroids. The mean concentration difference of 0.06% for the quantifying ion and confirmatory ion results of the 34 CRMs provides evidence that both transitions are accurate. In routine analysis, the application of quantifying:confirmatory ion ratios provides a means of detecting additional interfering compounds. Here, analysis of the male, female, and metyrapone serum samples all met the criterion of a quantifying:confirmatory ion ratio within 5% of the calibrator’s quantifying:confirmatory ion ratio (Table 1). The 95% CIs for these samples also closely agree with those of the calibrator, thereby further supporting the absence of interfering compounds in these samples. Analysis of the prednisolone samples yielded a mean quantifying:confirmatory ion that was outside the 5% acceptance criteria, indicating the presence of interference. This interference manifests as 2 coeluting peaks in the confirmatory ion transition and is easily identifiable by examining the chromatogram (Fig. 1B). However, because the range of the prednisolone quantifying:confirmatory ion ratios spans that of the calibrators, ion ratios alone would not detect all prednisolone samples. This underlines the importance of examining the chromatograms and further validates the use of the m/z 363.1>96.95 transition for quantification (Fig. 1A). The presence of peaks not completely resolved from cortisol in the m/z 363.1>120.95 channel has been reported previously in patients taking prednisolone (27).

Importantly, the method is accurate and possesses a statistically insignificant bias of 0.08% compared with existing RMPs, as verified through the analysis of JCTLM-listed higher-order CRMs (Fig. 3). All 34 CRM samples analyzed encompass the certified target value when the 5% (95% CI) expanded measurement uncertainty is applied (Fig. 4). This confirms the method’s accuracy and further validates the model used to calculate the measurement uncertainty.

It has been suggested that to qualify as an RMP, the candidate method must demonstrate a bias that is one-third the limit of routine assays, with an imprecision one-half the limit of routine assays (28, 29). Current Westgard recommendations for routine serum cortisol assays specify a bias of 10.3% and an imprecision of 7.6% (30). The analytical performance parameters of this method therefore satisfy those proposed for an RMP.

The measurement uncertainty calculation required estimates of the intra- and interassay variance. These variances were obtained by fitting a linear mixed-effects model with maximum likelihood estimation. This approach is advantageous because analyzing all of the data at once effectively pools the residuals, increases the number of degrees of freedom, and improves the estimates. In addition, the fact that 1 data set was discarded owing to unacceptable intraassay variation does not appreciably increase the uncertainty of its estimated concentration, as would be the case if the CRM compounds were treated separately. Further information regarding the model fitted is available in the online Supplemental Data. On the basis of this statistical analysis, when the candidate RMP is applied in the same manner as for analysis of the CRM samples, an expanded measurement uncertainty of 5% (95% CI) can be applied to samples between the concentrations of 83 and 764 nmol/L (30 and 277 μg/L). Outside of this range, the 5% (95% CI) expanded measurement uncertainty can be applied provided interassay QC specifications are satisfied.

In summary, this method demonstrates excellent accuracy and imprecision, is robust against interferences, and has low measurement uncertainty. It therefore qualifies as an RMP for total serum cortisol quantification. Our goal is to use this assay to underpin the UK NEQAS serum cortisol EQA scheme, thereby allowing >200 laboratories across 20 different countries to compare their method to a certified target from which their accuracy and measurement uncertainty can be calculated.

Author Contributions: All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.

Authors’ Disclosures or Potential Conflicts of Interest: Upon manuscript submission, all authors completed the author disclosure form. Disclosures and/or potential conflicts of interest:

Employment or Leadership: S. Cowen, LGC Ltd.
Consultant or Advisory Role: None declared.
Stock Ownership: None declared.
Honoraria: None declared.
Research Funding: None declared.
Expert Testimony: None declared.
Patents: None declared.
Role of Sponsor: No sponsor was declared.

References

2. Jonckheere JA, De Leenheer AP. Candidate definitive method for the determination of cortisol in human se-
er JS, Sampson EJ, Bayse DD. Determination of steroid
homones in a human-serum reference material by iso-
tope dilution-mass spectrometry: a candidate definitive
AP. Candidate reference method for determining serum
cortisol based on isotope dilution-gas chromatography/ 
mass spectrometry using heptafluorobutyric acid as 
5. Tai SS, Welch MJ. Development and evaluation of a can-
didate reference method for the determination of total
cortisol in human serum using isotope dilution-liquid 
chro-
matography/mass spectrometry and liquid chroma-
tography/tandem mass spectrometry. Anal Chem 
6. Jung PG, Kim B, Park SR, So HY, Shi LH, Kim Y. Determina-
tion of serum cortisol using isotope dilution-liquid 
chro-
matography/mass spectrometry as a candidate refer-
7. RELA-IFCC. External Quality assessment scheme for 
reference laboratories in laboratory medicine. Proce-
8. Tai SS-C, Duerwer DL. Final report on CCQM-K63.a,b: 
7
8. RELA-IFCC External Quality assessment scheme for
6. RELA-IFCC. External Quality assessment scheme for 
reference laboratories in laboratory medicine. Proce-
10. WEQAS. Traceability of cortisol assays in the UK: com-
www.weqas.com/download/traceability-of-cortisol-
assays-in-the-uk-comparison-with-jctlm-listed-
11. Dodd AJ, Ducrocq DH, Neale SM, Wise MP, Mitchem KL,
on cortisol measurement by commonly used immuno-
12. El-Farhan N, Pickett A, Ducrocq D, Bailey C, Mitchem K,
Morgan N, et al. Method-specific serum cortisol re-
sponses to the adrenocorticotrophin test: comparison 
of gas chromatography-mass spectrometry and five au-

tomated immunoassays. Clin Endocrinol (Oxf) 2013;
BG, Darby D. Comparison of serum cortisol measure-
ment by immunoassay and liquid chromatography-
tandem mass spectrometry in patients receiving the 
11β-hydroxylase inhibitor metyrapone. Ann Clin 
14. Handelsman DJ, Wantofsky J. Requirement for mass 
spectrometry sex steroid assays in the Journal of Clini-
cal Endocrinology and Metabolism. J Clin Endocrinol
15. Monaghan PJ, Kevil BG, Stewart PM, Trainer PJ. Case 
for the wider adoption of mass spectrometry-based ad-
renal steroid testing, and beyond. J Clin Endocrinol
16. Wierman ME, Aarhus R, Haisenleder DJ, Hall JE,
Handelsman D, Hankinson S, et al. New instructions to 
authors for the reporting of steroid hormone 
measurements [Editorial]. Mol Endocrinol 2014;28:
1917.
17. Vesper HW, Botelho JC. Standardization of testoster-
one measurements in humans. J Steroid Biochem Mol 
18. Botelho JC, Shacklady C, Cooper HC, Tai SS, Van Uyt-
fanghe K, Thienpont LM, Vesper HW. Isotope dilution 
liquid chromatography-tandem mass spectrometry 
candidate reference method for total testosterone 
20. PATH. Partnership for the Accurate Testing of Hor-
2015).
21. CLSI. Characterisation and qualification of commutable 
reference materials for laboratory medicine, approved 
guideline. CLSI document C53-A. Wayne, PA: CLSI; 
2010.
Definition of terms relating to mass spectrometry 
lsu.edu/msterms/index.php/Ionization_efficiency (Ac-
cessed May 2015).
23. Annesley TM. Ion suppression in mass spectrometry. 
24. IRMM. Certification report: The characterisation of 
cortisol concentrations in a reference serum panel-
ec.europa.eu/jrc/sites/default/files/erm/
25. Bureau International des Poids et Mesures. JCTLM 
database: Laboratory Medicine and In Vitro Diagnos-
26. Soldin SJ, Soldin OP. Steroid hormone analysis by 
1061–6.
27. Owen LJ, Adaway JE, Davies S, Neale S, El-Farhan N, 
Ducrocq D, 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:
345–52.
28. Stöckl D, Sluss PM, Thienpont LM. Specifications for 
trueness and precision of a reference measurement 
system for serum/plasma 25-hydroxyvitamin D analy-
29. Petersen PH, Fraser CG, Callner A, Kenny D. Strategies 
to set global quality specifications in laboratory medi-
30. Westgard QC. Desirable Biological Variation Database 
html (Accessed April 2015).