Christopher Kyriacou, Wei Yang, Shikha Kapur, Shanuja Maheetharan, Margaret Pikovsky, Nina Parker, Jennifer Barcroft, Shabnam Bobdiwala, Shyamaly Sur, Catriona Stalder, Deborah Gould, Dede Ofili-Yebovi, Andrea Day, Nick Unsworth, Edmund H. Wilkes, Tricia Tan and Tom Bourne*

Ambulatory human chorionic gonadotrophin (hCG) testing: a verification of two hCG point of care devices

https://doi.org/10.1515/cclm-2023-0703
Received July 5, 2023; accepted October 10, 2023;
published online October 30, 2023

Abstract

Objectives: Quantitative human chorionic gonadotropin (hCG) measurements are used to manage women classified with a pregnancy of unknown location (PUL). Two point of care testing (POCT) devices that quantify hCG are commercially available. We verified the i-STAT 1 (Abbott) and the AQT 90 FLEX (Radiometer) prior to use in PUL triage.

Methods: Tests for precision, external quality assurance (EQA), correlation, hook effect and recovery were undertaken alongside a POCT usability assessment during this prospective multi-center verification.

Results: Coefficients of variation ranged between 4.0 and 5.1 % for the three i-STAT 1 internal quality control (IQC) solutions and between 6.8 and 7.3 % for the two AQT IQC solutions. Symmetric differences in POCT EQA results when compared with laboratory and EQA stock values ranged between 3.2 and 24.5 % for the i-STAT 1 and between 3.3 and 36.9 % for the AQT. Correlation coefficients (i-STAT 1: 0.96, AQT: 0.99) and goodness of fit curves (i-STAT 1: 0.92, AQT: 0.99) were excellent when using suitable whole blood samples. An hCG hook effect was noted with the i-STAT 1 between 572,194 and 799,089 IU/L, lower than the hook effect noted with the AQT, which was between 799,089 and 1,619,309 IU/L. When hematocrit concentration was considered in sample types validated for use with each device, hCG recovery was 108 % with the i-STAT 1 and 98 % with the AQT. The i-STAT 1 scored lower on usability overall (90/130) than the AQT (121/130, $p<0.001$, Mann-Whitney).

Conclusions: Both hCG POCT devices were verified for use in clinical practice. Practical factors must also be considered when choosing which device to use in each unit.

Keywords: point of care; hCG; pregnancy of unknown location (PUL)

Christopher Kyriacou and Wei Yang share first authorship.

***Corresponding author: Professor Tom Bourne**, Tommy's National Centre for Miscarriage Research, Early Pregnancy and Acute Gynaecology Unit, Department of Obstetrics and Gynaecology, Queen Charlotte's and Chelsea Hospital, Imperial College London, Du Cane Road, London, W12 0HS, UK; and Department of Development and Regeneration, KU Leuven, Leuven, Belgium, Phone: +44 203 313 5131,
E-mail: t.bourne@imperial.ac.uk. https://orcid.org/0000-0003-1421-6059
Christopher Kyriacou, **Margaret Pikovsky**, **Nina Parker**, **Jennifer Barcroft**, **Shabnam Bobdiwala**, **Shyamaly Sur and Catriona Stalder**, Tommy's National Centre for Miscarriage Research, Early Pregnancy and Acute Gynaecology Unit, Department of Obstetrics and Gynaecology, Queen Charlotte's and Chelsea Hospital, Imperial College London, London, UK. https://orcid.org/0000-0001-9001-5545 (C. Kyriacou). https://orcid.org/0000-0003-2538-5851 (M. Pikovsky). https://orcid.org/0000-0002-7962-4451 (N. Parker). https://orcid.org/0000-0001-5258-312X (J. Barcroft). https://orcid.org/0000-0003-0540-2191 (S. Bobdiwala)
Wei Yang, **Shanuja Maheetharan**, **Deborah Gould**, **Nick Unsworth**, **Edmund H. Wilkes and Tricia Tan**, Biochemistry Unit, Department of Laboratory Diagnostics, Hammersmith Hospital, Imperial College London, London, UK. https://orcid.org/0000-0003-1166-6281 (W. Yang). https://orcid.org/0000-0001-5873-3432 (T. Tan)
Shikha Kapur, Gynaecology Emergency Unit, Department of Obstetrics and Gynaecology, St Mary's Hospital, Imperial College London, London, UK
Dede Ofili-Yebovi, Early Pregnancy Assessment Unit, Chelsea and Westminster Hospital, London, UK. https://orcid.org/0000-0002-2910-7050
Andrea Day, Early Pregnancy Assessment Unit, West Middlesex University Hospital, London, UK. https://orcid.org/0000-0002-9486-7734

Introduction

Transvaginal ultrasonography is used to assess pregnancy location and viability in women during the first trimester of pregnancy [1]. However, in up to 42 % of cases, a pregnancy is not visualized [2, 3]. These women are classified as having a pregnancy of unknown location (PUL) pending a confirmed outcome of either [1]: intrauterine pregnancy (IUP: pregnancy correctly located within the endometrial cavity) [2]; ectopic pregnancy (EP: pregnancy incorrectly located external to the endometrial cavity) [3]; persistent PUL (PPUL: PUL with ≥3 serial hCG values that vary by <15 %); or [4] failed PUL (FPUL: PUL with a negative urine pregnancy test two weeks from initial review) [3]. Serial quantitative human chorionic

gonadotropin (hCG) levels are measured and used to triage women with a PUL classification as being at a low (IUP/FPUL) or high risk (EP/PPUL) of clinical complications [4].

HCG is a glycoprotein hormone secreted from pregnancy trophoblast and is quantifiable in maternal circulation shortly after conception via the immunologically specific beta-hCG subunit [5]. HCG plays many roles, including maintenance of corpus luteum progesterone production and angiogenesis [6]. Differences between longitudinal serial hCG levels observed with different pregnancy outcomes has allowed the development of PUL management protocols [4]. Whilst UK National Institute for Health and Care Excellence guidance currently advises use of an hCG ratio (hCG at 48 h divided by hCG at 0 h), other methods of biochemical PUL triage also incorporate progesterone into their algorithm, and/or provide accurate statistical prediction via logistic regression modeling (e.g., the clinically implemented and validated M6) [1, 4, 7–12].

Measuring serial hCG levels with a recognized laboratory method can take hours. This delay, particularly at the 48-h clinical assessment, compromises the ability to effectively triage women and promptly direct resources to those at high-risk of complications, as well as impacting management planning, postponing senior clinical input if after hours, and forcing women to anxiously wait at home for advice throughout the day or overnight. However, there are now commercially available point of care testing (POCT), quantitative hCG devices that may improve PUL workflow [13].

Devices such as the i-STAT 1 (Abbott), AQT 90 FLEX (Radiometer), i-CHROMA II (Boditech), CS (Stratus®) and Easy Reader (VEDA.LAB) can provide hCG measurements in an outpatient setting [13–21]. Whilst hCG data on the i-CHROMA II, CS and Easy Reader are limited, data from both i-STAT 1 and AQT 90 FLEX suggest wider use of both systems, with acceptable levels of accuracy, precision, and linearity [13, 16, 17, 19, 20].

Our aim was to verify the i-STAT 1 and the AQT devices as POCT systems that safely measure hCG levels in the early pregnancy setting.

Materials and methods

Study design, setting and participants

This prospective multi-center hCG POCT verification was performed as an audit at four London hospitals between January and December 2021 following review and support from the North–West London POCT committee. Samples were either collected by the biochemistry team, or prospectively collected under the remit of two ethically approved studies, references 14/NS/1078 (ClinicalTrials.gov ID: NCT04738370) and 20/LO/0477 (ClinicalTrials.gov ID: NCT04739956) between September 2018 and November 2021. Informed consent for these studies was confirmed in writing.

The i-STAT 1 (Abbott, Chicago, USA) and the AQT 90 FLEX (Radiometer, Copenhagen, Denmark) POCT devices were verified by 22 healthcare professionals alongside POCT device usability questionnaires. Quality control (QC), plasma, serum, and whole blood samples were stored and/or processed in accordance with local operating and safety protocols. HCG levels obtained by the i-STAT 1 and the AQT analyzers were verified against the Architect or Alinity (Abbott, Chicago, USA) and the Beckman Access (Beckman, Brea, USA) hCG assays, run by laboratory technicians. Tests for precision, external quality assurance (EQA), correlation, hook effect and recovery were undertaken.

Data collection

Point of care testing hCG verification: The i-STAT 1 is validated to process untreated or lithium-heparin (LH)-stabilized whole blood and plasma samples for total beta-hCG (total: measures both intact hCG (joined alpha and beta-hCG subunits) and free beta-hCG subunits). Samples were applied to single-use cassettes, which functioned via a two-site enzyme-linked immunosorbent assay (ELISA) method in 10–12 min, with a quantitative hCG range of 5–2,000 IU/L. Calibrators and controls for the device are traceable to target hCG concentrations defined using the World Health Organization (WHO) 5th International Standard [22]. The AQT 90 FLEX is validated to process EDTA or LH-stabilized whole blood and plasma samples for total beta-hCG. These were inserted into the device in original (stabilized whole blood) or custom manufacturer (plasma) tubes, where an all-in-one dry chemistry system delivered ELISA driven results in 18–20 min, with a quantitative hCG range of 2–5,000 IU/L. Calibrators for the device are traceable to target hCG concentrations defined using WHO 4th International Standard [17]. Please see Supplementary Table S1 for a summary of the assay characteristics, storage, calibration, and QC.

Laboratory hCG reference standard: The Architect and Alinity total beta-hCG chemiluminescent microparticle immunoassays can process LH, sodium-heparin or EDTA-stabilized plasma, or serum specimens. Each assay requires 20 min but can take longer when dilution is required (>15,000 IU/L). Within run and within laboratory/total coefficients of variation (CV) for pooled serum samples ranged from 1.2 % (hCG concentration: 5,052.1 IU/L) to 4.9 % (hCG concentration: 21.1 IU/L) when using the Architect [23]. CV when using the Alinity ranged from 2.1 % (hCG concentration: 9,421.1 IU/L) to 7.6 % (hCG concentration: 5.3 IU/L) [24]. The Beckman total beta-hCG chemiluminescent immunoassay can process LH-stabilized plasma or serum samples in 20 min, unless dilution is required (>1,350 IU/L). CV when using the Beckman ranged from 1.8 % (hCG concentration: 106.7 IU/L) to 21.7 % (hCG concentration: 0.6 IU/L) [25]. Six calibrators run every 28–30 days, with QC performed daily. Please see Supplementary Table S2 for a summary of the assay characteristics, storage, calibration, and QC.

External quality assurance schemes: Each hospital laboratory is subscribed to a national EQA scheme to ensure a high standard for obtaining reproducible results and confirm compliance with international standards (ISO15189:2012). All devices were subject to EQA: National External Quality Assurance Scheme (NEQAS, Sheffield, UK) is the EQA provider for the AQT 90 FLEX, Architect and Beckman, and Wales External Quality Assurance Scheme (WEQAS, Cardiff, UK) is the EQA provider for the i-STAT 1.

Outcome and statistical analysis

Precision: Internal QC solutions were provided by the POCT device manufacturers. For the i-STAT 1, QC processing was performed five times per day over five days at one unit by one user, and five times per day over one day at a second unit by a different user. The data was combined (n=30). Three i-STAT 1 internal QCs were used with LOT A20261 cassettes: QC1 (LOT 351126), QC2 (LOT 361129); and QC3 (LOT 371134). For the AQT, two internal QCs were performed five times per day for five days at one unit by one user (n=25): QC1 (LOT 18456) and QC2 (LOT 18457). Mean with standard deviation (SD) were calculated. CV were compared to assigned targets, and results were compared to Westgard hCG target values (TV) of ±18 % [26].

External quality assurance: 14 (distribution 384) NEQAS and six (distributions 77 and 78) WEQAS EQA materials were processed once on both the i-STAT 1 and the AQT at one unit. POCT results were compared with the findings of laboratory processing of NEQAS and WEQAS samples. When possible, they were also compared against the NEQAS grouped laboratory trimmed mean (GLTM) and the NEQAS AQT average values. The group means of logarithm transformed data (natural log units) were presented alongside symmetric percentage (%) differences between the groups of interest.

Correlating POCT hCG values with laboratory hCG values using serum and EDTA-stabilized plasma samples: Although the i-STAT 1 and AQT devices are validated for LH-stabilized plasma, they are not for serum. The i-STAT 1 is also not validated for EDTA-stabilized plasma. Nonetheless, 80 paired serum samples and 30 paired EDTA-stabilized plasma samples were processed by both POCT devices at one unit. Comparison data were presented as group means of logarithm transformed data alongside symmetric percentage (%) differences between the groups of interest. In addition, for the serum samples, coefficients of correlation (R) were identified following Pearson correlation analyses and presented with 95 % confidence intervals (95 % CI). Coefficients of determination/goodness of fit (R^2), regression coefficients (95 % CI) and linear equations were identified following simple linear regression. Bland-Altman plots of differences in hCG levels between methodologies described device bias with SD and 95 % limits of agreement (95 % LOA – defined as the mean of the paired differences in hCG levels ± 1.96 standard deviations).

Correlating whole blood sample POCT hCG values with serum sample laboratory hCG values: The majority of the available 742 fresh whole blood samples (untreated or LH-stabilized) were processed using the i-STAT 1 at two units. Ninety-four EDTA-stabilized whole blood samples were processed using the AQT at two different units. Each of the paired units used the same laboratory methods. The sample types selected were validated for use on each POCT device. R (95 % CI) were again obtained following Pearson correlation analyses, with R^2, regression coefficients (95 % CI) and linear equations identified following simple linear regression. Bland-Altman plots of differences in hCG levels between methodologies were again used to describe device bias with SD and 95 % LOA.

Hook effect: One serum sample with known high hCG concentration (>1,000,000 IU/L) following laboratory processing was tested in duplicate on the i-STAT 1 and AQT at one unit. Serial dilutions were prepared using hCG radioimmunoassay dilution buffer by the laboratory, and each of these samples were tested until the hook effect was no longer apparent, confirmed when the devices presented qualitative high hCG results (i-STAT 1: >2,000 IU/L, AQT: >5,000 IU/L). POCT results were compared with the findings of the laboratory to determine hook effect thresholds. The details of the hCG radioimmunoassay dilution buffer are as follows: 50 mM phosphate (di-sodium hydrogen phosphate dodecahydrate – $Na_2HPO_4{\cdot}12H_2O$ and potassium di-hydrogen phosphate – KH_2PO_4) containing 50 mM EDTA, 0.075 % bovine serum albumin, 0.05 % polysorbate 20 and 0.01 % sodium azide.

Recovery: 186 IU purified WHO 5th International hCG standard was prepared in 10 mL hCG radioimmunoassay dilution buffer (as described above) for a stock concentration of 18,600 IU/L. This was used to produce reducing concentrations: 9,300 IU/L; 4,650 IU/L; 2,325 IU/L; 1,162.5 IU/L; 581.3 IU/L; 290.6 IU/L; and 0 IU/L. EDTA-stabilized whole blood, LH-stabilized whole blood, and serum was collected at one unit from hCG negative participants. The samples were divided, and serial hCG standard dilutions were spiked with a 1:10 dilution factor: one part standard to nine parts negative pool. The whole blood and serum samples were thus spiked to the following hCG concentrations: 1,860 IU/L, 930 IU/L, 465 IU/L, 232.5 IU/L, 116.3 IU/L, 58.1 IU/L, 29.1 IU/L and 0 IU/L. All samples were processed in duplicate or quadruplicate depending on sample and consumable availability, to confirm the findings of both POCT devices. Recovery was calculated in percent (%).

Usability: Twenty-six questions were developed to assess POCT device usability. Each question was designed to be answered using a Likert scale, from zero (strongly disagree) to five (strongly agree) and were filled in by healthcare professionals 35 times across three units. Please see Supplementary Table S3 for a copy of this questionnaire. Average scores for each POCT device were compared by device, by unit and by usability assessment subgroup (usefulness, ease of use, ease of learning and satisfaction). Shapiro-Wilk testing was used prior to deciding on the most appropriate parametric or non-parametric statistical test: unpaired t test for normal population distribution, or the Mann-Whitney test for non-normal population distribution. Statistical significance was defined as a $p<0.05$.

Statistical analysis was performed using Excel (Microsoft, Redmond, USA) and PRISM (GraphPad, San Diego, USA).

Results

Precision

CV across the three i-STAT 1 internal quality control solutions ranged between 4.0 and 5.1 %, with all 30 values within their assigned ranges and meeting manufacturer CV targets (Table 1). When i-STAT 1 internal quality control levels were compared to Westgard TV (±18 %), one of 30 was beyond TV [26]. The CV of the two AQT internal quality control solutions ranged between 6.8 and 7.3 %, with two of the 25 values reported out of the assigned range at each level. There was no manufacturer assigned CV for the AQT. When AQT

Table 1: Point of care testing device precision assessment when compared to internal quality control solutions.

IQC levels	Assigned range, IU/L	Assigned mean, IU/L	Values OOR, n	Mean, IU/L	SD, IU/L	Assigned CV, %	CV, %	Westgard TV, %	Values >TV, n
i-STAT 1 (n=30)									
1	19.3–35.8	27.5	0	26.5	1.4	<10	5.1	±18	0
2	779.6–1,447.8	1,113.7	0	1,184.7	51.3	<10	4.3	±18	1
3	982.2–1,824	1,403.1	0	1,511.2	60.5	<10	4.0	±18	0
AQT 90 FLEX (n=25)									
1	14.3–22.2	17.3	2	15.7	1.1	NA	6.8	±18	2
2	267–354	310	2	300.5	22.1	NA	7.3	±18	1

IQC, internal quality control; OOR, out of range; SD, standard deviation; CV, coefficient of variation; TV, target value; NA, not applicable.

internal quality control levels were compared to Westgard TV (±18 %), three of 25 were beyond TV [26].

External quality assurance

For the i-STAT 1, five of the 14 NEQAS samples were below the quantitative limits of detection (<5 IU/L) and excluded from analysis. This was correctly detected in four cases. In one case, the device measured <5 IU/L, whilst the laboratory and GLTM values were 5.6 and 6.1 IU/L respectively. Analysis of the remaining nine NEQAS samples identified a symmetric difference of 3.2 % (POCT vs. laboratory) and 4.0 % (POCT vs. GLTM) (Table 2). There were no NEQAS i-STAT 1 average values provided for comparison. One of six WEQAS samples was within quantitative detection limits, with a symmetric difference of 24.5 % when i-STAT 1 results were compared with laboratory processing.

Four NEQAS samples were below the quantitative limits of the AQT device (<2 IU/L) and were excluded, the device correctly detecting each one. Analysis of the remaining 10 NEQAS samples identified a symmetric difference of 3.3 % (POCT vs. laboratory), 3.9 % (POCT vs. GLTM) and 7.0 % (POCT vs. EQA AQT average values). Two of the six WEQAS samples were within quantitative detection limits, with a symmetric difference of 36.9 % when compared with laboratory processing.

Correlating POCT hCG values with laboratory hCG values using serum and EDTA-stabilized plasma samples

54 of the 80 serum samples were above (>2,000 IU/L) or below (<5 IU/L) the quantitative detection limits of the i-STAT 1 and correctly detected in all cases other than one, which had succumbed to the hook effect (1,215.4 IU/L i-STAT 1 vs. 1,516,158 IU/L laboratory). Analysis of the remaining 26 samples identified a 1.0 % symmetric difference between POCT and laboratory hCG processing (Supplementary

Table 2: Point of care testing (POCT) device performance using external quality assurance (EQA) materials when compared with laboratory processing of EQA materials, EQA grouped laboratory trimmed mean, and EQA grouped point of care testing results.

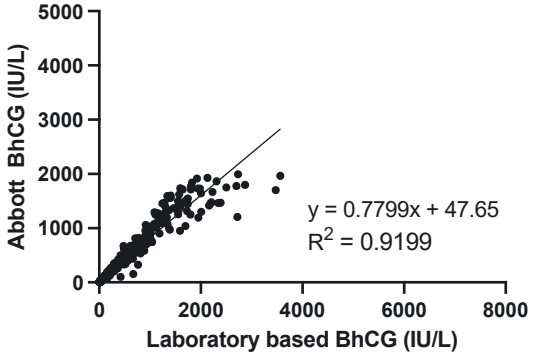
EQA service	NEQAS[b]				WEQAS[a]	
Processing of EQA materials	**POCT**	**Local lab**	**GLTM**	**EQA POCT av**	**POCT**	**Local lab**
i-STAT 1, n		9			1	
Group means of log transformed data, natural log units	5.4[b]	5.4	5.5	NA[b]	5.7	6.0
Symmetric difference with POCT, %		3.2[b]	4.0[b]	NA[b]		24.5
AQT, n		10			2	
Group means of log transformed data, natural log units	5.1	5.0	5.1	5.1	6.8[a]	7.2
Symmetric difference with POCT, %		3.3	3.9	7.0		36.9[a]

NEQAS, National External Quality Assurance Scheme; WEQAS, Wales External Quality Assurance Scheme; GLTM, grouped laboratory trimmed mean; lab, laboratory; av, average; log, logarithm; NA, not applicable. [a]The AQT device is validated against NEQAS. [b]The i-STAT 1 device is validated against WEQAS. No NEQAS EQA POCT av values provided for comparison with i-STAT 1 device.

Table S4). Following correlation analysis and simple linear regression, R was 0.99 (95 % CI 0.98–>0.99) and R^2 was 0.98 (Supplementary Figure S1).

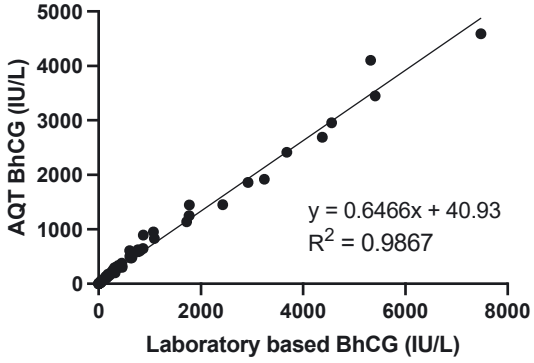
Thirty of the 80 serum samples were above (>5,000 IU/L) the quantitative AQT detection limit and correctly detected in all cases other than one (>5,000 IU/L AQT vs. 4,723.2 IU/L laboratory). Analysis of the remaining 50 samples identified a 9.0 % symmetric difference between POCT and laboratory hCG processing (Supplementary Table S4). Following correlation analysis and simple linear regression, R was >0.99 (95 % CI 0.99–>0.99) and R^2 was 0.99 (Supplementary Figure S1).

A i-STAT 1: Correlation curve and summary data (N=588)*



Summary data	
Symmetric difference, %	7.3
Correlation coefficient - R (95% CI)	0.96 (0.95-0.97)
	Slope 0.8 (0.8-0.8)
Regression coefficients (95% CI)	Y-intercept 47.7 (33.9-61.4)
	X-intercept -61.1 (-80.0--42.8)
Goodness of fit - R^2	0.92

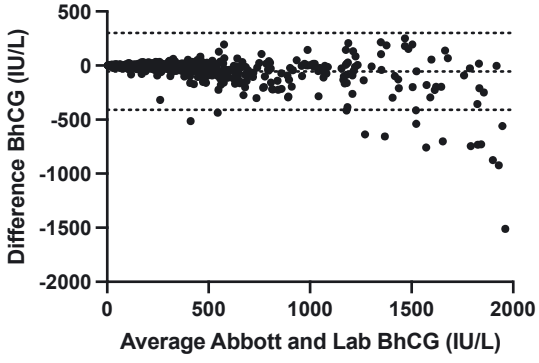
B AQT 90 FLEX: Correlation curve and summary data (N=82)*



Summary data	
Symmetric difference, %	25.2
Correlation coefficient - R (95% CI)	0.99 (0.99->0.99)
	Slope 0.6 (0.6-0.7)
Regression coefficients (95% CI)	Y-intercept 40.9 (14.1-67.8)
	X-intercept -63.3 (-106--21.5)
Goodness of fit - R^2	0.99

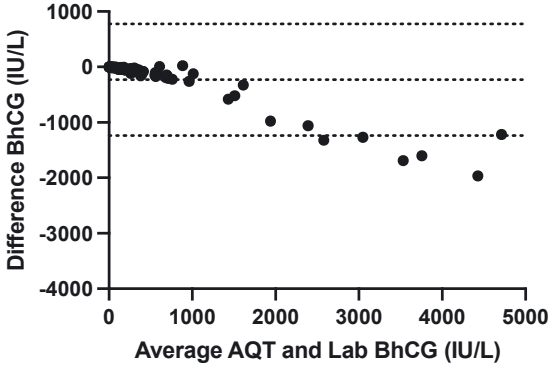
C Bland-Altman Bias

i. i-STAT 1 (N=588)*



Bias, IU/L (SD)	-55.2 (181.5)
95% LOA, IU/L	-410.9 to 300.5

ii. AQT 90 FLEX (N=82)*



Bias, IU/L (SD)	-231.5 (513)
95% LOA, IU/L	-1237 to 774

Figure 1: Correlating point of care testing hCG values within quantitative device detection ranges following whole blood sample processing, with paired serum sample laboratory hCG values. hCG, human chorionic gonadotrophin; SD, standard deviation; 95 % CI: 95 % confidence interval; 95 % LOA: 95 % limits of agreement. *All point of care testing and laboratory devices measure total beta-hCG (i.e., both intact hCG and free beta-hCG subunits).

Upon performing a Bland-Altman analysis, bias was 21.7 IU/L (SD 93.9) with 95 % LOA −162.2 to 205.7 IU/L using the i-STAT 1, compared with 175.2 IU/L (SD 160.5) with 95 % LOA −139.3 to 489.6 IU/L using the AQT. When comparing both POCT devices with one another, bias was −56.2 IU/L (SD 106.6) with 95 % LOA −265 to 152.7 IU/L (Supplementary Figure S1).

EDTA-stabilized plasma comparisons were highly inaccurate with the i-STAT 1, secondary to this sample type not being validated for use with this device (Supplementary Table S4). With the AQT, one of the 30 plasma samples was incorrectly above the quantitative detection limit (>5,000 IU/L AQT vs. 4,768.5 IU/L laboratory). Analysis of the remaining 29 samples showed a 2.7 % symmetric difference between POCT and laboratory hCG processing. When directly comparing matched AQT serum and plasma hCG results, the symmetric difference was 3.2 % (Supplementary Table S4).

Table 3: Point of care testing device hook effect. Samples were evaluated in duplicate on the i-STAT 1 and the AQT (n=1, diluted).

Laboratory hCG, IU/L	i-STAT 1 hCG, IU/L		AQT hCG, IU/L	
1,619,309	1,264.3	1,261.4	4,343	3,168
799,089	1,826.2	1,886.2	>5,000	>5,000
572,194	>2,000	>2,000	NA	NA

NA, not applicable.

Correlating whole blood sample POCT hCG values with serum sample laboratory hCG values

154 of the 742 whole blood samples for i-STAT 1 processing were above or below the quantitative detection limit of the device. Analysis of the remaining 588 samples identified a 7.3 % symmetric difference between POCT and laboratory hCG processing (Figure 1). R was 0.96 (95 % CI 0.95–0.97) and R^2 was 0.92.

Twelve of the 94 whole blood samples for AQT processing were above or below the quantitative detection limit of the device. Analysis of the remaining 82 samples identified a 25.2 % symmetric difference between POCT and the laboratory (Figure 1). R was 0.99 (95 % CI 0.99–>0.99) and R^2 was 0.99.

On Bland-Altman analysis, bias was −55.2 IU/L (SD 181.5) with 95 % LOA −410.9 to 300.5 IU/L using the i-STAT 1,

Table 4: Recovery of 5th international hCG standard from whole blood using laboratory and point of care testing methods once hematocrit concentration (42 %) was considered within the analysis.

Diluted standard, IU/L	Recovery EDTA, IU/L	Recovery rate EDTA, %	Recovery LH, IU/L	Recovery rate LH, %	Recovery EDTA + LH, IU/L	Recovery rate EDTA + LH, %
A) i-STAT 1 (n=2, diluted, duplicate/quadruplicate)						
2,990.4	>	NA	>	NA	>	NA
1,495.2	1,388.9	93	1,650.9	110	1,563.6	105
747.6	685.6	92	1,073.9	144	944.5	126
373.8	363.5	97	439.9	118	401.7	107
186.9	167.6	90	176.1	94	171.9	92
93.4	80.9	87	83.6	89	82.2	88
46.7	50.6	108	42.4	91	46.5	99
0	<	NA	<	NA	<	NA
Average recovery rate, %		94[a]		108		103[a]
B) AQT 90 FLEX (n=2, diluted, duplicate/quadruplicate)						
2,990.4	3,155.8	106	3,004.5	100	3,105.3	104
1,495.2	1,476.5	99	1,599	107	1,517.3	101
747.6	743.5	99	627	84	704.7	94
373.8	384.3	103	369	99	379.2	101
186.9	184.3	99	178.5	96	182.3	98
93.4	90.8	97	90.5	97	90.7	97
46.7	49.8	106	44	94	47.8	102
0	<	NA	<	NA	<	NA
Average recovery rate, %		101		97		98

EDTA, EDTA-stabilized whole blood; LH, lithium-heparin-stabilized blood; NA, not applicable. [a]i-STAT 1 device not validated for use with EDTA whole blood samples.

Table 5: Usability assessment of point of care testing devices, divided by unit and question subgroups (n=35).

Usability assessment	i-STAT 1 (n=19)	AQT (n=16)	p-Value[a]
Total average score (/130)	90	121	<0.001[b]
Unit 1 average (/130)	111	119	0.288
Unit 2 average (/130)	78	128	<0.001[b]
Unit 3 average (/130)	51	114	<0.001[b]
Usability assessment subgroups			
Usefulness average (/35)	24	33	<0.001[b]
Ease of use average (/45)	29	40	<0.001[b]
Ease of learning average (/30)	22	28	0.005[b]
Satisfaction average (/20)	14	19	0.002[b]

[a]p-Value corresponds to Mann-Whitney (non-parametric test for continuous, non-normal population distribution), or unpaired t (parametric test for continuous, normal population distribution) for the difference between outcome groups. [b]p<0.05 denotes significance.

compared with −231.5 IU/L (SD 513) and 95 % LOA −1,237 to 774 IU/L using the AQT (Figure 1).

Hook effect

An hCG hook effect was noted with the i-STAT 1 between 572,194 IU/L and 799,089 IU/L, lower than the hook effect noted with the AQT, which was between 799,089 IU/L and 1,619,309 IU/L (Table 3).

Recovery

The average recovery of international hCG standard from serum samples for the laboratory was 91 %, compared with 82 and 84 % using the i-STAT 1 and AQT respectively (Supplementary Table S5).

For whole blood samples, the average recovery of international hCG standard using the i-STAT 1 was 166 % (152 % with EDTA-stabilized whole blood, 173 % with LH-stabilized whole blood), and for the AQT was 160 % (163 % with EDTA-stabilized whole blood and 155 % with LH-stabilized whole blood) (Supplementary Table S6). After troubleshooting for the overestimated data, we then took into account the average concentration of hematocrit in whole blood, estimated as 42 % on review of the literature [27]. With calculations modified to account for the 58 % plasma aspect of the samples, the average recovery rate using the i-STAT 1 was 103 % (94 % with EDTA-stabilized whole blood and 108 % with LH-stabilized whole blood), and for the AQT was 98 % (101 % with EDTA-stabilized whole blood and 97 % with LH-stabilized whole blood) (Table 4).

Usability

The i-STAT 1 score lower on usability overall (90/130) than the AQT (112/130, p<0.001, Mann-Whitney test) (Table 5). When split by unit, scores for both devices were similar at Unit 1 (111 vs. 119/130, p=0.288, unpaired t test), but lower for i-STAT 1 at Units 2 and 3 (78 and 51/130 vs. 128 and 114/130, p<0.001, unpaired t test). When split by usability assessment subgroup, the i-STAT 1 had lower scores than the AQT for usefulness (24 vs. 33/35, p<0.001, Mann-Whitney test), ease of use (29 vs. 40/45, p=0.001, unpaired t test), ease of learning (22 vs. 28/30 p=0.005, Mann-Whitney test) and satisfaction (14 vs. 19/20, p=0.002, Mann-Whitney test).

Discussion

This study verified two POCT devices for clinical practice, prior to using their hCG measurements in early pregnancy to triage women classified with a PUL. Both performed with acceptable precision and EQA, with excellent correlation coefficients and goodness of fit curves using appropriate samples validated for use on each device. Accuracy and recovery percentages were comparable once considering hematocrit. The AQT had a higher hook threshold than the i-STAT 1 and scored higher in the usability assessment. Although serum is not validated for use on either instrument, laboratory values correlated well with both POCT devices.

Strengths

The number of factors considered, the prospective sample collection, the comparison of the two available POCT devices for hCG measurement and the validated sample types evaluated makes this the largest quantitative hCG POCT verification study published. Various issues were investigated, involving clinical and laboratory-based specialists who collaborated to comprehensively test the technology.

Limitations

The disproportionate sample size was secondary to longer use of the i-STAT 1 at one of the units. Samples were collected prospectively from multiple centers, and both devices were compared with one another, adding some generalizability for clinicians who may be considering the introduction of this technology to their practice elsewhere. Whilst some of the sample types investigated were not validated for use

with one or both devices, their assessment was performed for completeness.

Quality control

Both devices have QC limitations. Supplementing internal QC with EQA was advised by manufacturers as the AQT lacks internal QC cover from 500 to 5,000 IU/L. There were no i-STAT 1 NEQAS results published that allowed comparison with data from other laboratories. The AQT is not validated against WEQAS. Many of the EQA samples were beyond the limits of detection for one or both devices. Better EQA materials are required for quantitative hCG POCT devices.

Bias

Correlation and bias findings were similar for both POCT serum and whole blood sample processing when compared to laboratory serum processing. Whilst serum samples are not validated for use with either POCT device, this suggests that POCT result variances were most likely due to matrix differences, rather than differences in hCG isoforms the methods were measuring.

Bland-Altman analyses of whole blood samples indicated that POCT hCG values were consistently lower than laboratory findings. This finding may be due to the differences in hCG isoforms the methods were measuring, as well as differences in reagent antibodies and the material used as calibrators. Whilst the equivalent recovery experiments initially showed overestimation, this was corrected once the average hematocrit proportion of whole blood was considered. The excellent correlation findings suggest that even if hCG values using POCT are lower, intra-technology consistency is high so the pattern of hCG level change is the same. Per patient, serial POCT hCG values could thus be considered in clinical practice for PUL triage when not used interchangeably with laboratory hCG measurements, as the level of agreement when using only one technology (laboratory vs. i-STAT 1 vs. AQT) appears to be high, and the POCT assays showed acceptable levels of imprecision across the concentrations measured. Formal clinical study using PUL triage protocols is already underway.

Usability

Unit 1 had used the i-STAT 1 for longer than the AQT and deemed usability to be similar for the two devices. This was not the case for Units 2 and 3 where both devices were used for same amount of time. Whilst the AQT may therefore be easier to use in a shorter space of time, using either device for long enough may increase overall usability as the unit becomes accustomed to the equipment and protocols.

Other devices and population considerations

Data on hCG POCT using other devices is limited [13, 16, 17, 19–21]. The i-CHROMA II (Boditech) is portable and uses a fluorescence immunoassay on whole blood samples, with a range of 5–50,000 IU/L and an assay time of 15 min [15, 16]. The CS (Stratus®) has a range of 0–1,250 IU/L using whole blood and plasma samples, with a run time of 14 min [14]. The Easy Reader (VEDA.LAB) can process blood and urine for hCG levels up to 1,000 IU/L in 30 s [18].

These devices run similar samples at a similar rate for hCG when compared to the i-STAT 1 and AQT. However, in the context of early pregnancy and PUL triage, most hCG levels will unlikely exceed 1,000–2,000 IU/L, within the quantitative detection ranges for both the i-STAT 1 and the AQT [28]. Levels measured that would be susceptible to the hook effect tend to only apply to molar pregnancies, cases of which are extremely rare and usually suspected following transvaginal ultrasonography. These women will therefore continue to utilize laboratory hCG measurements as they do not require rapid POCT hCG results for clinical assessment and management.

Other unit considerations

Running whole blood samples whilst the patient is in attendance carries considerable advantages, allowing information and advice to be provided promptly and face-to-face. Other considerations for units planning to introduce hCG POCT should include patient throughput (if simultaneous testing is required), portability (handheld or benchtop), setting (hospital or community), system (open or closed), quantitative biochemical ranges, time to process results, training, and cost.

Conclusions

Both devices described in this study were thoroughly verified for use in clinical practice. The next step will be validating their hCG values for use within published PUL triage protocols.

Acknowledgments: We acknowledge the support staff of the early pregnancy units and the laboratories involved with this work.

Research ethics: This prospective multi-center hCG POCT verification was performed as an audit at four London hospitals between January and December 2021 following review and support from the North–West London POCT committee. Samples were either collated by the biochemistry team, or prospectively collected under the remit of two ethically approved studies, references 14/NS/1078 (ClinicalTrials.gov ID: NCT04738370) and 20/LO/0477 (ClinicalTrials.gov ID: NCT04739956) in accordance with the Declaration of Helsinki (as revised in 2013) between September 2018 and November 2021.

Informed consent: Informed consent for individuals under the remit of two ethically approved studies, references 14/NS/1078 (ClinicalTrials.gov ID: NCT04738370) and 20/LO/0477 (ClinicalTrials.gov ID: NCT04739956) between September 2018 and November 2021 was confirmed in writing.

Author contributions: CK, WY, TT and TB participated in the conception and design of the study. CK, WY, SK, SM, MP, NP, JB, SB, SS, CS, DG, DOY, and AD acquired patient data. CK and WY performed the statistical analysis. CK, WY, NU, EW, TT and TB interpreted the results. CK, WY, TT and TB wrote the initial version of the manuscript. The authors have accepted responsibility for the entire content of this manuscript and approved its submission.

Competing interests: CK reports honoraria and non-financial support from Germitec. TT reports research support from the Novo Nordisk (UK) Research Foundation and the Moulton Charitable Research Foundation. TB reports grants, personal fees and non-financial support from Samsung Medison, non-financial support from Roche diagnostics and non-financial support from Abbott Diagnostics. All other authors state no conflict of interest.

Research funding: CK is supported by the Imperial Health Charity, grant number RFPrD1920/116. NP is supported by Tommy's National Centre for Miscarriage Research, Imperial College London. JB is supported by the Imperial Health Charity, grant number RFPR2122_13. SB was supported by NIHR CLAHRC NWL, grant number RDIP0033. TT and TB are supported by the NIHR Imperial Biomedical Research Centre, grant number IS-BRC-1215-20013. Abbott Diagnostics provided i-STAT 1 devices and Radiometer provided AQT 90 FLEX devices for the verification. The views expressed are those of the author(s) and not necessarily those of the National Health Service, the National Institute of Health Research, or the Department of Health. The funders had no role in study design, data collection, analysis and interpretation or reporting.

Data availability: The data that supports the findings of this study are available in the supplementary material of this article.

References

1. Webster K, Eadon H, Fishburn S, Kumar G, Committee G. Ectopic pregnancy and miscarriage: diagnosis and initial management: summary of updated NICE guidance. BMJ 2019;367:l6283.
2. Boyraz G, Bozdağ G. Pregnancy of unknown location. J Turk Ger Gynecol Assoc 2013;14:104–8.
3. Bobdiwala S, Al-Memar M, Farren J, Bourne T. Factors to consider in pregnancy of unknown location. Womens Health (Lond) 2017;13:27–33.
4. Bobdiwala S, Saso S, Verbakel JY, Al-Memar M, Van Calster B, Timmerman D, et al. Diagnostic protocols for the management of pregnancy of unknown location: a systematic review and meta-analysis. BJOG 2019;126:190–8.
5. Cole L. Immunoassay of human chorionic gonadotropin, its free subunits, and metabolites. Clin Chem 1997;43:223–43.
6. Cole LA. Biological functions of hCG and hCG-related molecules. Reprod Biol Endocrinol 2010;8:102.
7. Bobdiwala S, Guha S, Van Calster B, Ayim F, Mitchell-Jones N, Al-Memar M, et al. The clinical performance of the M4 decision support model to triage women with a pregnancy of unknown location as at low or high risk of complications. Hum Reprod 2016;31:1425–35.
8. Bobdiwala S, Christodoulou E, Farren J, Mitchell-Jones N, Kyriacou C, Al-Memar M, et al. Triaging women with pregnancy of unknown location using two-step protocol including M6 model: clinical implementation study. Ultrasound Obstet Gynecol 2020;55:105–14.
9. Christodoulou E, Bobdiwala S, Kyriacou C, Farren J, Mitchell-Jones N, Ayim F, et al. External validation of models to predict the outcome of pregnancies of unknown location: a multicentre cohort study. BJOG 2021;128:552–62.
10. Van Calster B, Bobdiwala S, Guha S, Van Hoorde K, Al-Memar M, Harvey R, et al. Managing pregnancy of unknown location based on initial serum progesterone and serial serum hCG levels: development and validation of a two-step triage protocol. Ultrasound Obstet Gynecol 2016;48:642–9.
11. Bourne T, Kyriacou C, Coomarasamy A, Al-Memar M, Leonardi M, Kirk E, et al. ISUOG Consensus Statement on rationalization of early-pregnancy care and provision of ultrasonography in context of SARS-CoV-2. Ultrasound Obstet Gynecol 2020;55:871–8.
12. Bourne T, Kyriacou C, Coomarasamy A, Kirk E, Condous G, Leonardi M, et al. Guidance for rationalising early pregnancy services in the evolving coronavirus (COVID-19) pandemic, Guideline ed London: Royal College of Obstetricians and Gynaecologists; 2020:20 p.
13. Research NIfH. Point-of-care testing for human chorionic gonadotrophin (hCG). In: Co-operative DE, editor. Horizon scan report 0043. Oxford: NIHR Diagnostic Evidence Cooperative Oxford; 2015:6 p.
14. Azzazy HM, Romero LF, Hall L, Ruttle D, Christenson RH. Two-center clinical evaluation of a new automated fluorometric immunoassay for the quantitative analysis of total betaeta-human chorionic gonadotropin. Clin Biochem 2003;36:523–8.
15. Bolodeoku J, Bains S, Pinkney S, Coker O, Fakokunde A. Comparison of the point of care test (POCT), i-CHROMA™ human chorionic gonadotrophin (HCG), leutinzing hormone (LH) and follicle stimulating hormone (FSH) methods with the other laboratory methods in the Randox International Quality Assessment Scheme (RIQAS). Clin Obstet Gynecol Reprod Med 2017;3:2–7.
16. Bolodeoku J, Fakokunde A. Quantitative point of care testing of HCG in early pregnancy units: a review. Integr Gynecol Obstet J 2019;2:1–3.

17. Brun MM, Holloway L, Oleksy A, Dayton J, Estey MP, Goudreau BL, et al. Analytical evaluation of the Radiometer AQT90 FLEX βhCG assay. Pract Lab Med 2019;14:e00116.
18. Reynolds-Wright JJ, Main P, Cameron ST. Accuracy of a point-of-care test for quantifying human chorionic gonadotrophin (hCG) in the management of pregnancy of unknown location in an abortion service. BMJ Sex Reprod Health 2019;45:155–8.
19. Sowder AM, Yarbrough ML, Nerenz RD, Mitsios JV, Mortensen R, Gronowski AM, et al. Analytical performance evaluation of the i-STAT Total β-human chorionic gonadotropin immunoassay. Clin Chim Acta 2015;446:165–70.
20. Wikström AK, Hagmar M, Ronquist G, Larsson A. Evaluation of a plasma hCG method for point of care testing with the aim of shortening test-turnaround-times. Open J Obstet Gynecol 2015;5:341–3.
21. Wilgen U, Pretorius CJ, Gous RS, Martin C, Hale VJ, Ungerer JP. Hook effect in Abbott i-STAT β-human chorionic gonadotropin (β-hCG) point of care assay. Clin Biochem 2014;47:1320–2.
22. Abbott. Total beta-human chorionic gonadotropin (b-hCG) 2019 [10]. Available from: https://www.pointofcare.abbott/int/en/offerings/support/i-stat/cartridge-test-information-sheets.
23. Laboratories A. Architect system: total beta-hCG; 2010. Available from: https://www.ilexmedical.com/files/PDF/TotalBhCG_ARC.pdf.
24. FDA. 510(k) substantial equivalence determination decision summary assay and instrument combination template: total beta human chorionic gonadotropin (β-hCG) in human serum and plasma, Alinity i Total β-hCG Reagent Kit. Available from: https://www.accessdata.fda.gov/cdrh_docs/reviews/K170317.pdf.
25. FDA. 510(k) substantial equivalence determination decision summary assay only template: total beta human chorionic gonadotropin (βhCG) in human serum and plasma, Beckman Coulter, Inc., Access Total βhCG (5 th IS) Assay.
26. QC W. Consolidated comparison of immunoassay performance specifications; 2022. Available from: https://www.westgard.com/consolidated-goals-immunoassay.htm.
27. Billett H, Walker H, Hall W, Hurst J. Clinical methods: the history, physical, and laboratory examinations, 3rd ed. Boston: Butterworths; 1990.
28. Van Calster B, Abdallah Y, Guha S, Kirk E, Van Hoorde K, Condous G, et al. Rationalizing the management of pregnancies of unknown location: temporal and external validation of a risk prediction model on 1962 pregnancies. Hum Reprod 2013;28:609–16.

Supplementary Material: This article contains supplementary material (https://doi.org/10.1515/cclm-2023-0703).