

Opinion Paper

Catharine Sturgeon, Stephen A. Butler, Fiona Gould, Sarah Johnson*, Sam Rowlands, Ulf-Håkan Stenman and David G. Grenache

Recommendations for validation testing of home pregnancy tests (HPTs) in Europe

<https://doi.org/10.1515/cclm-2020-1523>

Received October 13, 2020; accepted December 3, 2020;
published online January 28, 2021

Abstract: Home pregnancy tests (HPTs) available in Europe include accuracy and other performance claims listed on their packaging. Due to the lack of guidance on the standardisation of such products, it is often difficult to replicate these claims when tested on a clinical sample, whether in a laboratory setting or by lay users. The *In Vitro* Diagnostic Regulation is a set of requirements that mandate comprehensive validation data on human pregnancy tests and other *in vitro* devices. It is due to replace the current European Directive (98/79/EC) and fully implemented in Europe by 2022. In June 2019, a panel of seven experts convened to discuss the validation studies required to provide the information needed to meet the new regulation for HPTs in Europe and proposed 15 recommendations for best practice. Defining best practice at all stages of validation of these important tests may ensure that tests marketed in Europe are fit for purpose, enabling lay users to be confident of the high quality of the HPT results they obtain. The panelists believe that the recommendations proposed here for the validation of HPTs may constructively contribute to improved standardisation of validation procedures in Europe.

Keywords: European Directive; home pregnancy tests; human chorionic gonadotrophin; *in vitro* diagnostic regulation; validation testing.

Introduction

Home pregnancy tests (HPTs) became widely available in 1976, enabling women to determine their pregnancy status in the privacy and comfort of their own homes [1–3]. Significant technical innovations have resulted in the development of the tests used today. These have simple, one-step direct sampling formats with either colourimetric or digital outputs that facilitate interpretation as positive, negative or equivocal [2, 3]. The basis of all pregnancy tests is detection of the hormone human chorionic gonadotrophin (hCG) [2]. hCG is a glycoprotein that has a structure consisting of two non-covalently linked dissimilar α - (91 amino acids) and β - (145 amino acids) subunits [1, 2]. Following implantation, the placenta begins to develop, gradually producing increasing amounts of hCG until weeks 10–12 of pregnancy, after which the hCG concentration decreases moderately and stabilises until late pregnancy, with a further decrease during the last month [2]. hCG is not present in measurable concentrations in non-pregnant, pre-menopausal women except in relatively rare malignancies [4]. Its measurement, therefore, has enabled development of both laboratory tests and HPTs for detection of pregnancy [2].

At least 65 million HPTs are sold in Europe each year [5]. These are available as strips, cassettes and mid-stream dipsticks. Mid-stream tests either have results displayed as a line or cross for visual interpretation by the user (visual tests, also known as line tests) or the results are displayed digitally to the user on an LCD screen (digital tests) [5]. The characteristics of typical devices are shown in Table 1. Many HPTs are sold over the counter for home use, others are intended for professional use only [6].

HPTs are designed for single use for qualitative detection of hCG in a urine specimen collected at any time of the day, with a positive result indicating possible pregnancy [7].

***Corresponding author: Dr. Sarah Johnson**, Clinical Research Director, SPD Development Company Limited Bedford, Bedford, MK44 3UP, UK, Tel: +44 (0)1234 835486, Fax: +44 (0)1234 835 006, E-mail: sarah.johnson@spdspark.com

Catharine Sturgeon, Department of Laboratory Medicine, UK National External Quality Assessment Service (UK NEQAS), Edinburgh, UK

Stephen A. Butler, SAB Scientific Consultancy Ltd, Flitton, Bedfordshire, UK





Fiona Gould, Alere International, Bedford, UK

Sam Rowlands, Department of Medical Sciences & Public Health, Bournemouth University, Bournemouth, UK

Ulf-Håkan Stenman, Department of Clinical Chemistry and Haematology, University of Helsinki, Helsinki, Finland

David G. Grenache, TriCore Reference Laboratories, Albuquerque, NM, USA

Table 1: Broad characteristics of HPTs commonly available in Europe.^a

Device format	Strip	Cassette	Mid-stream visual	Mid-stream digital
				
Mode of specimen collection	Sample collected in a container (usually not supplied with the test).	Sample collected in a container (usually not supplied with the test). A sample pipette is usually provided with the kit for sample application.	Sample by holding directly into the urine stream. Alternatively, most mid-stream tests also provide an option for dipping the sampling wick into a collected sample.	Sample by holding directly into the urine stream. Alternatively, most mid-stream tests also provide an option for dipping the sampling wick into a collected sample.
Sampling requirements	Strip dipped into specimen up to the “dip line” for the sampling time (usually 3–10 s).	Urine applied by pipetting a set number of drops to sample well (usually 3–5 drops).	Held in the urine stream for a prescribed time period (usually 3–5 s). Dipping time may differ from in-stream sampling time.	Held in the urine stream for a prescribed time period (usually 3–5 s). Dipping time may differ from in-stream sampling time.
Signal detection / read-out	Colourimetric	Colourimetric	Colourimetric	Digital
Range of time to read-out	1–5 min	1–5 min	1–5 min	1–5 min
Do not read generally after	10 min.	10 min.	10 min.	Digital results remain valid until display screen goes blank (days–months).

^aAll require use of urine as the specimen. Exact sampling times or requirements should be clearly defined in each products' instructions for use. The sampling times, amount of urine to be applied, and reading times provided in the table above are typical, but some tests may vary from these instructions. HPT, home pregnancy test.

However, detection limits may vary considerably, with one study reporting detection limits ranging from 6.3 to 50 IU/L [8]. There is currently no universal standardisation guidance for medical devices such as HPTs; thus, their analytical and clinical performance inevitably vary considerably [9]. The US Food and Drug Administration (FDA) has published summaries of HPTs cleared for sale in the USA, thereby providing some indication of how validation is being conducted. Details of the tests performed and the type of hCG samples used are included in these summaries [10, 11]. Table 2 illustrates the variability in testing of analytical precision for four HPTs recently cleared by the FDA, with the number of tests performed during evaluation ranging from 540 to 6,000 [10]. It is highly desirable that validation of HPTs includes assessment of their clinical accuracy when used by non-laboratory personnel, but how rigorously this is assessed also varies markedly as presented in Table 2 [11].

Currently, tests available in Europe include accuracy and other performance claims on their packaging. However, as illustrated above, the lack of standardised criteria against which performance can be assessed makes it difficult to compare results for analytical accuracy [2]. This also applies to claims for clinical accuracy, i.e. the correct determination of “pregnant” (positive) and “non-pregnant” (negative) results for specimens for which the

true clinical status is known. Many HPTs claim to be more than 99% clinically accurate when used from the day of the missed menstrual period, but it is not always clear how this was determined – a problem highlighted in the USA almost two decades ago [8]. Different approaches can yield differing results depending on the types of specimens used for testing, sample size and other factors, as well as the selection of lay users contributing to validation studies.

A recent study reported that only three of seven HPTs available on the European market met accuracy and reliability claims of >99% of the products' claimed analytical sensitivity (i.e. the stated concentration of hCG at which the product is expected to return pregnant results), while the other four tests had analytical accuracies of <99% (81.6–95.9%) [12]. In this context, it is important to note that if such assessments are performed by trained laboratory technicians under ideal conditions [2, 13] they may not reflect real-life usage. As the main aim of analytical validation is to demonstrate that the procedure is suitable for its intended purpose [14], for HPTs it is essential that there is clear evidence that the test can be successfully performed by lay users.

Manufacturers who market HPTs in the USA are now required by the FDA to provide evidence of product performance that includes data for testing analytical and

Table 2: Validation studies conducted to assess the precision of four FDA-cleared HPTs with claimed analytical sensitivities of 10 U/L.^a

Device	Mid-stream test K123436	Three formats in one submission: Strip, cassette and mid-stream test K150022	Two formats in one submission: Cassette and mid-stream test K172256	Three formats in one submission: Strip, cassette and mid-stream test K181551
Number of hCG specimens in the testing panel	6	10	8	8
Calibrated against	WHO 4th International Standard for hCG	WHO 4th International Standard for hCG	WHO 4th International Standard for hCG	WHO 4th International Standard for hCG
hCG concentrations added to “non-pregnant” urine, IU/L	0, 3, 6, 10, 15, 25	0, 2.5, 3, 5, 7.5, 8, 10, 15, 20, 25, 50	0, 3, 5, 8.5, 10, 25, 50, 100	0, 3, 6, 7.5, 9, 10, 15, 25
Number of replicates per condition	10	10	10	10
Number of consecutive days on which tests performed	3	5	5	5
Number of lots tested	5	3	3	3
Number of operators	3 (different operator on each different day)	Three operators at three sites (total of nine operators)	5	Not stated, three sites used (a different test format at each site)
Total number of tests done	900	Three formats of tests examined 1,500 per format	Two formats of tests examined 6,000 per format	Three formats of tests examined 1,200 per format

^aSource of the hCG used for these studies not specified. The K number is the FDA device registration number that can be used as a unique identifier for the product. FDA, US Food and Drug Administration; hCG, human chorionic gonadotrophin; HPT, home pregnancy test; WHO, World Health Organisation.

clinical sensitivity, equivalency of results when using different sampling methods (e.g. sampling directly in urine stream and dip, or dipping into a collected sample) and early pregnancy detection rate before the menstrual period is missed [2]. The current guidance document from the FDA (Guidance for over-the-counter [OTC] human chorionic gonadotrophin [hCG] 510[k]s) is out of date and awaiting revision [15], but recent information is available from the FDA detailing current FDA testing requirements, and Clinical Laboratory Standards Institute (CLSI) guidelines can also be used to guide testing. However, manufacturers are not required to meet these FDA requirements for the European market and in the absence of other recommendations, there is a notable lack of transparency and detailed information regarding HPTs available in Europe.

This lack of guidance will in part be addressed by the new *In Vitro* Diagnostic Regulation (IVDR), a set of requirements due to replace the current European Directive (98/79/EC). Transition to the IVDR commenced in May 2017, with the intention of full enforcement in Europe by 2022 [16, 17]. The requirements include the necessity for comprehensive validation data (including data on customer usability) on HPTs and other *in vitro* devices [17, 18].

An expert panel (with the following members: Dr Stephen Butler [Clinical Research Scientist]; Dr Fiona Gould [Industry, Regulatory Expert]; Professor David Grenache [Director of Clinical Reference Laboratory]; Dr Sarah Johnson [Industry, Diagnostics Expert]; Dr Sam Rowlands [Specialist in Sexual and Reproductive Health]; Professor Ulf-Håkan Stenman [Emeritus Professor of Clinical Chemistry] and Dr Catharine Sturgeon [Director of UK National External Quality Assessment Service for Pregnancy Testing]) convened in London (UK) in June 2019 to consider these issues and the validation studies required for HPTs available in Europe as these new regulations come into force. While HPTs were specifically considered at that meeting, the conclusions are equally valid for pregnancy tests intended for use by health professionals in hospitals and other healthcare settings. As there are international standards relating to the validation of shelf life and stability of HPTs, these were not considered by the panel. Key literature references used to guide the discussion are listed in Supplementary Table 1.

The panel’s recommendations for best practice are described below and key points are highlighted in Table 3. The panel encourage manufacturers to apply these recommendations during implementation of the new IVDR and for notified bodies to use this guidance when examining the data underpinning HPT performance.

Table 3: Recommendations for validation of HPTs in Europe.

-
- (1) **Nomenclature of HPTs, description of hCG isoforms recognised and units for reporting**
- (a) Achieving agreement on the broad descriptions of different types of device formats is a priority
 - (b) Data regarding the relative recognition of hCG isoforms expressed in molar units should be included in technical data sheets
 - (c) hCG results should be expressed using the same units
- (2) **Analytical requirements for validation of HPTs**
- (a) Pre-analytical requirements for urine collection and storage prior to validation studies should be clearly defined
 - (b) The minimum number of urine specimens containing added hCG and the appropriate concentration range for validation of analytical accuracy should be defined
 - (c) Analytical sensitivity should be expressed as the lowest concentration of hCG that the HPT detects $\geq 99\%$ of the time
 - (d) Analytical precision should be assessed by: (a) repeating the test at least 20 times per condition, across standards which include hCG concentrations near the detection limit of the method and the “50:50” point; and (b) repeating the testing series on at least three separate days spaced across a minimum five-day time frame, and with a minimum of three different operators and three different lot numbers
 - (e) Cross-reactions with LH, hCG β cf and other potential clinically relevant interferences should be assessed in HPT validation studies and results documented in technical data sheets
- (3) **Clinical requirements for validation of HPTs**
- (a) Data supporting claims for “early pregnancy testing” should be provided and how studies were conducted described in full in the technical data sheet
 - (b) Potential drawbacks of highly sensitive HPTs should be assessed, and relevant information included in technical data sheets and package inserts
 - (c) Statements that any HPT is 100% accurate should not be permitted in data sheets or other promotional material
- (4) **Requirements for validation of HPTs by lay users**
- (a) Sensitivity, specificity, PPV and NPV should be calculated
 - (b) A study where the HPT is used by women representative of the lay user and the volunteer results are compared to clinical pregnancy status should be conducted. The study should also consider ease of use, with results presented in the technical data sheet.
 - (c) Clear and detailed instructions about urine collection and timing between steps in the testing process must be provided in the instructions for use. The clarity of these instructions should be assessed in practice by an appropriately constituted panel of lay users during validation of the HPT
 - (d) Clear information about limitations of the HPT should be included in the instructions for use. The clarity of the information provided should be assessed during validation through questionnaires provided to lay users participating in the analytical validation studies
-

hCG, human chorionic gonadotrophin; hCG β cf, β core fragment of hCG; HPT, home pregnancy test; LH, luteinising hormone; NPV, negative predictive value; PPV, positive predictive value.

Requirements for validation of HPTs in Europe

Achieving consistency for the validation of these important tests requires agreement regarding the pre-analytical, analytical, clinical and user assessments that should be undertaken, as well as how this information should be presented to enable ready comparison of data from different sources. Panel recommendations are presented below.

Considerations relating to nomenclature

Description of device formats

HPTs are available in a number of formats (Table 1). Reaching broad agreement about how these tests should be described and adoption of consistent nomenclature for the different types of tests would assist individual lay users to decide which type of test they would find most convenient to use.

Recommendation: Agreement on the broad descriptions of different types of device formats is a priority.

Description of hCG isoforms recognised

Recognition of isoforms of hCG

Numerous isoforms of hCG exist in urine, and not all assays detect the various forms to the same extent. Intact hCG (in which the α - and β -subunits remain associated) is the predominant urinary form in early pregnancy [19]. In early pregnancy the free β -subunit of hCG (hCG β) is present at a concentration below 10% of that of intact hCG and is not detected in many HPTs [20–24]. The β core fragment of hCG (hCG β cf), a metabolically degraded fragment of hCG β , is undetectable in early-pregnancy urine, but becomes the predominant form by 10 weeks of gestation [21, 25]. hCG and the nicked form are less stable than hCG β [26]. In quantitative assays, these variations can lead to different “total” hCG results for clinical specimens, depending on the concentrations of the hCG isoforms present [26]. International standards available for most forms of hCG enable characterisation of quantitative assays in terms of their detection of the different forms [26–31]. The panel considered whether HPTs should undergo similar characterisation.

HPTs provide qualitative results only – they indicate whether the hCG concentration is above or below a certain claimed detection limit. The panel agreed that information relating to recognition of different hCG isoforms by HPTs is not relevant for lay users, as actual clinical performance in

relation to pregnancy status is the key metric. However, it is important that manufacturers assess which isoforms of hCG are recognised by each HPT during early characterisation of the method (Section “Units of measurement”). This can be conveniently achieved using existing International Reference Reagents for these isoforms and expressing comparative results in substance concentrations (mol/L) assigned to these Reference Reagents [26].

Recommendation: Manufacturers should include data regarding the relative recognition of hCG isoforms as expressed in molar units (mol/L) in technical data sheets.

Units of measurement

Achieving agreement regarding the units in which hCG concentrations are expressed for HPTs would be desirable. In an FDA guidance document for OTC hCG tests, the units used throughout are mIU/mL [15]; these units are established by the World Health Organisation Expert Committee on Biological Standardization. Originally based on hCG bioactivity measurements, these somewhat arbitrary units were assigned by bioassay to each successive International Standard for hCG to maintain continuity of clinical hCG results (and pharmacological hCG preparations). Adoption of molar units (substance concentrations) enables direct comparison of recognition of different forms of hCG as described above and facilitates analytical standardisation of both qualitative and quantitative hCG methods. The International Federation of Clinical Chemistry and Laboratory Medicine recommends the use of substance concentrations (mol/L) wherever possible; however, it was accepted by the panel that as clinicians and lay users are familiar with hCG results expressed in mIU/mL, IU/L or U/L, changing to molar or mass units would cause considerable confusion. It would be desirable to express results in terms of U/L, the units for hCG used in most clinical laboratories. However, when comparing concentrations of hCG β and hCG β cf with hCG, molar units must be used. Earlier established International Units (IUs) for hCG β are not at all comparable with IUs of hCG and the standards for hCG β and hCG β cf have not been assigned a value in IU.

Recommendation: hCG results should be expressed using the same units, ideally U/L.

Analytical requirements for validation of HPTs

Validation studies for HPTs differ from those for many diagnostic tests performed in clinical laboratories as they include studies conducted both by the diagnostic

companies manufacturing the HPT and by lay users representative of those likely to use the product. The aims of these two types of validation differ, as the first focuses primarily on the analytical characteristics of the HPT, while the second assesses whether the HPT is fit for purpose in the hands of the lay user. The urgent need for international guidance on validation of HPTs is clear from the variation in practice shown in Table 2 and Supplementary Table 1. During laboratory testing, technicians should handle the tests according to the instructions for use, for example, lie on a flat surface and replace cap if this is detailed in the instructions. If a different testing procedure is used, there should be statistically robust studies showing equivalence of methods.

Specimen requirements for validation studies

It is essential that HPT validation studies of both types are performed using urine. Validation studies for methods appropriate for use with blood or serum specimens must also be performed in these matrices. hCG standards used to assess detection limits should be prepared in the relevant matrices from pre-menopausal non-pregnant women. Samples should have urinary hCG <1 U/L, measured using a quantitative assay validated for use in urine samples. Alternatively, pregnancy status could be determined using serum hCG concentration taken on the same day as the urine sample. Given that it is relatively easy to source female urine, there should be no need to use male urine, which may not completely reflect the matrix in which tests are ordinarily performed. When preparing standards, specific gravity should be recorded and highly dilute urine should not be used.

Validation panels should include specimens from individual women, both pregnant and non-pregnant. This is in accordance with FDA guidance, which states that at least 100 “fresh” authentic human urine specimens should be used in validation testing of HPTs [15].

The conditions of collection for urine and other specimens for HPT validation studies should be agreed, as current definitions of what constitutes a “fresh” urine specimen are vague and inadequate. Whether specimens should be collected at a specified time of day must also be considered. The time delay between collection of urine specimens and their use in validation studies is particularly important for manufacturers as it is unlikely to be feasible to test specimens immediately after collection. Conditions of storage, including temperature and the maximum time of storage from when specimens are collected until they are used in validation studies, need to be specified. This is because studies have shown that

storage conditions can affect the measured concentration of hCG in urine samples [19, 25, 32–34]. For example, significant loss of immunoreactivity has been reported for samples stored at -20°C , but not at $+4^{\circ}\text{C}$ or -80°C [35]. Inclusion of agents to retard bacterial growth is advisable for refrigerated samples (e.g. sodium azide) because bacterial enzymes have been reported to cleave the peptide chains of hCG. Manufacturers should therefore validate their storage and collection conditions if using banked samples, and clearly describe these in the relevant technical data sheets.

The same considerations apply to lay users who are testing specimens provided by the manufacturer, particularly if this is being performed at home (Table 4). This is likely to be less relevant when lay users are testing their own urine, assuming they will test at the time of urination. However, lay users participating in validation studies should be reminded of this requirement and also advised of the ideal time of day at which to test their urine if specific timing is recommended in the instructions with the product (e.g. sampling the first morning urine).

Recommendation: Pre-analytical requirements for urine collection and storage prior to validation studies should be clearly defined.

Assessing analytical accuracy, sensitivity and reproducibility of HPTs

It is very important to differentiate analytical and clinical accuracy and sensitivity of HPTs, an issue that the panel considered problematic, particularly in advertising claims made for some HPTs. These distinctions should be explained and the terms clearly defined in kit inserts and/or technical data sheets.

Analytical accuracy

Analytical accuracy must be assessed by the manufacturer, and reflected in kit inserts and/or technical data sheets in the claimed detection limit or analytical sensitivity of the method. These are most often 10 or 25 U/L of hCG but may be as high as 50 U/L or as low as 2 U/L. Analytical accuracy is assessed using urine specimens containing known amounts of added purified intact hCG, which has been calibrated against the International Standard for hCG. Manufacturers should document the version of International Standard used for calibration. The concentration ranges of these specimens should include specimens below and above the claimed detection limit of the method – concentrations that should give consistently negative (not pregnant) or positive (pregnant) results. Specimens of intermediate concentrations near the “transition” or “50:50” concentration (at which on

Table 4: Assessment of performance of four FDA-cleared HPTs with claimed analytical sensitivities of 25 U/L when used by non-laboratory personnel (lay users).

Type of device	Three formats in one submission: Strip, cassette and mid-stream K043443	Three formats in one submission: Strip, cassette and mid-stream K183097	Mid-stream digital test K173229	Cassette K172627
Total number of lay users	189 (87 strip, 54 cassette, 48 mid-stream)	360 (90 strip, 90 cassette, 90 dip tested mid-stream, 90 in-stream tested mid-stream)	200	360
Testing specimens	Own urine	Own urine	Own urine	Own urine
Number of pregnant volunteers	110	185	86	153
Number of non-pregnant volunteers	79	175	114	207
Method for assessing ease of use and leaflet comprehension	Questionnaire to assess readability of labelling.	Ease of use questionnaire.	Ease of use and leaflet comprehension questionnaire, and Flesch–Kincaid analysis of leaflet to show reading age of Grade 7.	Ease of use questionnaire and Flesch–Kincaid analysis of leaflet to show reading age of Grade 8.

FDA, US Food and Drug Administration; HPT, home pregnancy test.

half the testing occasions, the HPT will give a positive result and on the other half will give a negative result) should also be included (See further discussion below). Test performance at these concentrations is most likely to be affected by factors such as different lot number or operator, thus providing additional valuable information. Some panel members recommended a detection limit of 25 U/L, but there was no consensus.

Recommendation: The minimum number of urine specimens containing added hCG and the appropriate concentration range for validation of analytical accuracy should be defined.

Analytical sensitivity

As HPTs are threshold-based tests rather than quantitative tests, standard definitions of Limit of Detection (LoD) and Limit of Quantification (LoQ) have not historically been applied. Rather, tests have been described in terms of “sensitivity”, meaning analytical sensitivity. The definition of analytical sensitivity varies and in some data sheets is “the lowest concentration of hCG which it is possible to detect” rather than, as is appropriate, the lowest concentration which the HPT detects reliably and reproducibly. The panel agreed that the analytical sensitivity of HPTs should be defined as the lowest concentration of hCG that can be reliably detected more than 99% of the time. Although it would be preferable to use LoD, the term “sensitivity” is now being used by consumers and health-care professionals to understand the performance of tests; thus, a new term in consumer labelling would be unhelpful. Details of how analytical sensitivity was assessed should be included in technical data sheets.

Recommendation: Analytical sensitivity should be expressed as the lowest concentration that the HPT detects $\geq 99\%$ the time.

Analytical precision

The precision of HPTs, which reflects their repeatability and reproducibility [14], should be assessed during validation. Repeatability is a measurement of precision in which replicate measurements using the same procedure, operators, system, operating conditions location (laboratory) are conducted on the same or similar specimens over a short period of time. Reproducibility aims to identify the parameters that contribute to variability of the assay by combining factors known to influence assay precision, such as operators, lots and batches, in a matrix to provide sufficient data points for statistical analysis. As presented in Table 2, considerable variation in how the precision of HPTs is currently assessed was reported. The panel agreed that in accordance with recommendations for other laboratory tests, repeatability should be assessed by obtaining 20 measurements from the same specimen (ideally a pooled urine standard) at different hCG concentrations representative of the working range of the assay. By including specimens of appropriately low concentration, the limit of detection can be determined similarly. Testing the same samples across five non-consecutive days by three different operators and for three different lot numbers provides additional evidence of precision and reproducibility. Guidance regarding experimental design for reproducibility is available from CLSI Evaluating Quantitative Measurement Precision-A3 [36].

Recommendation: Analytical precision should be assessed by: (1) repeating the test at least 20 times per condition, across standards which include hCG concentrations near the detection limit of the method and the “50:50” point; and (2) repeating the testing series on at least three separate days spaced across a minimum five-day time frame, and with a minimum of three different operators and three different lot numbers.

Analytical robustness to potential interferences

HPT validation studies should include evaluation of the effect of potential clinically relevant cross-reactants, including luteinising hormone (LH) at concentrations that may be present in urine (e.g. 500 U/L). The FDA requests that manufacturers also test follicle stimulating hormone (e.g. 1,000 U/L) and thyroid stimulating hormone (e.g. 1 U/L) [10, 11]. Dietary supplements such as biotin may suppress signal production [37] causing erroneous results. Vulnerability to such interference should be assessed during validation. The urine matrix used in spiking experiments should have an appropriate pH (5–7) and specific gravity (1.010–1.030).

High urinary concentrations of hCG in late pregnancy, and of hCG β cf at the end of the first trimester and through later pregnancy may saturate antibody binding sites in the HPT, producing a false negative test result (i.e. the “hook effect”) [2, 21, 38–41]. Validation of HPTs should include testing to demonstrate that the test still returns “Pregnant” results at high concentrations of hCG (e.g. 500,000 U/L). For hCG β cf, validation studies should demonstrate that “hooking” that results in false negative results does not occur in clinically possible scenarios. This has to be done in a matrix with high concentrations of intact hCG to mimic the actual clinical conditions under which this phenomenon has been observed to occur (e.g. hCG β cf diluted to achieve a final concentration of 1 μ mol/L into late first trimester pooled urine) [2, 20, 21, 41–44].

Results of such validation studies should be included in technical data sheets, which should also clearly state whether concentrations above which the method is affected are likely to be encountered in clinical specimens. hCG β cf is likely to be found at the highest concentration in urine at 10–12 weeks of gestation [43].

Recommendation: Cross-reactions with LH, hCG β cf and other potential clinically relevant interferences should be assessed in HPT validation studies and results documented in technical data sheets.

Robustness to potential errors when using the HPT

Although challenging to address, assessing the vulnerability of the HPT to small errors in specimen application

(e.g. holding a mid-stream test in the urine stream for too long, or adding too many urine drops to a cassette test) or errors in timing prior to reading the tests should be included in validation studies. Mid-stream tests often provide two options for sample application: in-stream and dipping; equivalency of these methods should be demonstrated. If the HPT is insufficiently robust, it may be necessary to consider re-designing method.

Clinical requirements for validation of HPTs

Assessment of clinical accuracy should provide users with information about how accurately the test will identify pregnancy status at the time the sample is taken, and should include robust data with a statistically appropriate sample size and clear definition of how pregnancy status is defined. Particular attention is required during validation studies to address the following issues.

Validation of claims for “early pregnancy testing”

Manufacturers making claims relating to “early pregnancy testing” (often described as “testing before the missed period”) should be fully transparent regarding the proportion of pregnancies detected when testing so early and provide data supporting their claim. It is not clear whether many of the tests available in Europe have validated their early claim performance on clinical samples and if so, how this was achieved or how the specimens were collected. It is likely that rather than conducting studies examining early pregnancy detection rate, some may base the claim purely on the analytical sensitivity of the test.

The language surrounding early testing claims appears to have arrived at a common terminology, “testing before the missed period”, and as consumers are familiar with seeing data presented using the missed period as reference, moving to alternative, more precise terms would be confusing. The timing of the missed period is notably inaccurate, especially for women with irregular periods, but unless the woman is monitoring ovulation using home ovulation tests or undergoing *in vitro* fertilisation, there is currently no better method for assessing when to time a test. Claims on currently marketed HPTs tend to be derived from calculating day of expected period as 15 days following ovulation, where ovulation day can be determined by a variety of methods (pre-conception day of LH surge [urine or serum], or use of urinary oestrone-3-glucuronide:pregnanediol-3-glucuronide ratio or ultrasound observed ovulation). Although this may not necessarily mirror the day a consumer would calculate, it

does enable direct comparison between performance characteristics of different tests.

Claims such as “Results over 99% accurate from the day of your expected period” should be supported by referenced clinical studies.

Recommendation: Data supporting claims for early pregnancy testing should be provided and how studies were conducted described in full in the technical data sheet.

Potential disadvantages of highly sensitive HPTs

HPT methods should be designed so that analytical sensitivity is not so high that it generates “apparently false” positive results in women experiencing early pregnancy loss, which depending on the circumstances can be very distressing. Approximately one in four pregnancies end in early pregnancy loss, many of which will cause a positive HPT result at the time of the first missed period [43, 45, 46]. In addition, hCG concentrations in non-pregnant peri- and post-menopausal women can be slightly raised relative to the pre-menopausal reference interval [47]; therefore, HPTs should be assessed for specificity when used by women in this age group. This issue is increasingly relevant as more women delay pregnancy until a later age when irregular periods, often associated with the perimenopause, may be mistaken for pregnancy, prompting a pregnancy test. It has been suggested that an upper reference limit of 14 U/L should be used when interpreting serum hCG results in women >55 years of age [47]. Manufacturers should quantify this risk for each HPT and consider including this information in the package insert if the risk is relatively high.

Recommendation: Potential drawbacks of highly sensitive HPTs should be assessed and relevant information included in technical data sheets and package inserts.

Misleading claims of 100% accuracy

The panel was particularly concerned that claims of “100% accuracy” made for some HPTs can be misleading to health professionals and lay users as it is often not clear whether “100%” relates to analytical or clinical accuracy. Lay users in particular are likely to assume such claims mean that an HPT is 100% clinically accurate for the detection of pregnancy. This is never the case due to potentially confounding factors, including specimens that are too dilute or the possibility of early pregnancy loss, as well as the nature of all immunoassays. Such claims could be exploited by manufacturers to imply that their test is clinically superior. In agreement with current FDA guidance, the panel agreed that such claims should not be permitted in Europe.

Recommendation: Statements that HPTs are 100% accurate should not be permitted in data sheets or other promotional material.

Requirements for validation of HPTs by lay users

HPTs are developed primarily for lay users who generally will have little practical experience of laboratory testing. It is important to ensure that HPTs include clear instructions for use, that the tests can be readily performed outside a laboratory environment, that they are robust and that interpretation of the results is straightforward. Incorporating lay user testing during validation of HPTs is therefore essential to assess ease of use and confirm whether intended users are readily able to follow the product instructions, as well as to ensure that test results are interpreted correctly. The lay user panel should include a reasonable number of representative women from a range of age groups and its composition should be described in the technical data sheet. Panels would ideally include women seeking to become pregnant and women who suspect pregnancy (whether wanted or unwanted) because their period is late. In practice, however, recruiting women in the latter group is likely to be more difficult.

The results obtained by lay users testing their own sample should be compared to their true clinical status, as determined by an independent method. This is especially important as lay users may be less able than laboratory staff to read fainter lines at the test sensitivity. The study should be sufficiently sized to enable clinical accuracy, sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) to be calculated robustly, as required by the IVDR.

Recommendation: Sensitivity, PPV and NPV should be calculated.

Ease of use

Both practical and questionnaire assessments are likely to be required. The latter could take the form of Likert-style questionnaires, where participants provide responses to questions such as “How easy was it to read the result?” on a 7-point scale ranging from 1 (e.g. “Extremely easy”) to 7 (e.g. “Extremely difficult”). Criteria for acceptability should be established before questionnaires are issued.

Recommendation: A study where the HPT is used by women representative of the lay user and the volunteer results are compared to clinical pregnancy status should be

conducted. The study should also consider ease of use, with results presented in the technical data sheet.

Presentation of practical information on how to use the HPT

Clear instructions about specimen collection

The clarity of instructions regarding how to collect urine specimens and apply the specimen to the HPT should be assessed as how the specimen is applied to the device may influence the results. For example, some tests require uniform flow of urine through the test strip. If this is perturbed because too much or too little urine is applied, erroneous results may be obtained. Validation studies should demonstrate the results from lay user testing match those for similar urine specimens as tested by laboratory professionals [12].

Clear instructions about timing

Clear instructions about timing, both from the time of sample collection to the time of use of the HPT, and between application of urine to the device and the read time, should be included in the information leaflet. Additionally, the maximum time period for which the result remains valid from the time of running the test should be noted. This is because when lateral flow-based tests dry out, line intensity can be altered to no longer reflect the true result of the test, so there is usually a maximum reading time.

Recommendation: Information provided with the HPT should include clear and detailed instructions about urine collection and timing between steps in the testing process. The clarity of these instructions should be assessed in practice by an appropriately constituted panel of lay users during validation of the HPT.

Presentation of information on interpretation of HPT results

General format of information provided

Information should be clearly presented using wording readily understood by the general public, including young teenagers, and should avoid use of scientific terms wherever possible. The font size on all package labelling and information leaflets should be large enough to be readily legible.

Shelf life

HPTs have a defined shelf life as determined by stability studies, which should be conducted by the manufacturer (usually 2–3 years). The “use by” date should be clearly printed on the test carton.

Error results

As with any immunoassay, HPTs can occasionally fail to function properly, whether due to over- or under-sampling, or misuse by the lay user, or due to a faulty test (e.g. damaged in shipping). Therefore, it is important that a lay user is able to understand when a test result is an error, and not misinterpret the result as “Not pregnant” (e.g. no lines on the test) or “Pregnant” (e.g. just the test line visible). The instructions for use should clearly provide information on what error results present like.

Information about test interpretation

Information leaflets should include a statement that HPTs are not 100% accurate for confirmation of pregnancy and that negative tests should be repeated if there is a strong suspicion of pregnancy. Clear images of the different results should be included to aid interpretation.

Inappropriately negative results may be obtained in early pregnancy if hCG concentrations are not high enough for detection or if urine is very dilute (early morning urine is often requested to reduce this risk). Repeating the test 3–5 days later is generally desirable. Measurement of serum hCG when requested by the woman’s healthcare provider is another option that should provide a definitive answer. This is essential if the woman is experiencing abdominal pain or bleeding as these may indicate an ectopic pregnancy, which can produce very low concentrations of hCG [41, 48, 49].

Recommendation: Information leaflets provided with the HPT should include clear and detailed information about the limitations of HPTs. The clarity of the information provided should be assessed through questionnaires provided to lay users participating in the analytical validation studies.

Conclusions

The authors trust that the recommendations proposed here for validation of HPTs may constructively contribute to improved standardisation of validation procedures for HPTs in Europe. Defining best practice at all stages of validation of these important tests should help to ensure that tests marketed in Europe are fit for purpose and that lay users can be confident of the high quality of the HPT results they obtain.

Acknowledgments: The authors would like to thank Abisola Akinbowale from integrated medhealth

communication for providing medical writing assistance, which was funded by SPD Development Company Limited.

Research funding: The study was funded by SPD Development Company Limited (Bedford, UK), a wholly owned subsidiary of SPD Swiss Precision Diagnostics GmbH (Geneva, Switzerland).

Author contributions: All authors attended the panel meeting, provided expert recommendations, and were involved in manuscript preparation and review. All authors have accepted responsibility for the entire content of this manuscript and approved its submission.

Competing interests: Sarah Johnson – Employee of SPD Development Company Ltd., a fully owned subsidiary of SPD Swiss Precision Diagnostics GmbH, a manufacturer of home pregnancy and ovulation tests. Fiona Gould – Contracted to support Abbott Rapid Diagnostics, a manufacturer of home pregnancy and ovulation tests.

References

- Haarburger D, Pillay TS. Historical perspectives in diagnostic clinical pathology: development of the pregnancy test. *J Clin Pathol* 2011;64:546–8.
- Gnoth C, Johnson S. Strips of hope: accuracy of home pregnancy tests and new developments. *Geburtshilfe Frauenheilkd* 2014;74:661–9.
- Eichner SF, Timpe EM. Urinary-based ovulation and pregnancy: point-of-care testing. *Ann Pharmacother* 2004;38:325–31.
- Stenman UH, Alfthan H. Determination of human chorionic gonadotropin. *Best Pract Res Clin Endocrinol Metab* 2013;27:783–93.
- Johnson S. 100 years of human chorionic gonadotropin: reviews and new perspectives. Cole LA, Butler SA, editors. Elsevier; 2020. Chapter 2.14:107–21.
- Nickmans S, Vermeersch P, Van Eldere J, Billen J. Performance of qualitative urinary hCG assays. *Acta Clin Belg* 2014;69:277–9.
- Grenache DG. Variable accuracy of home pregnancy tests: truth in advertising? *Clin Chem Lab Med* 2015;53:339–41.
- Butler SA, Khanlian SA, Cole LA. Detection of early pregnancy forms of human chorionic gonadotropin by home pregnancy test devices. *Clin Chem* 2001;47:2131–6.
- Tomlinson C, Marshall J, Ellis JE. Comparison of accuracy and certainty of results of six home pregnancy tests available over-the-counter. *Curr Med Res Opin* 2008;24:1645–9.
- FDA. Summary of last 10mIU/ml pregnancy tests cleared by FDA for sale in the USA (can be used 6 days before the missed period).
- FDA. Summary of recent 25mIU/ml pregnancy tests cleared by FDA for sale in the USA (can be used 5 days before the missed period).
- Boxer J, Weddell S, Broomhead D, Hogg C, Johnson S. Home pregnancy tests in the hands of the intended user. *J Immunoassay Immunochem* 2019;40:642–52.
- Scolaro KL, Lloyd KB, Helms KL. Devices for home evaluation of women's health concerns. *Am J Health Syst Pharm* 2008;65:299–314.
- ICH Harmonised Tripartite Guideline. Validation of analytical procedures: Text and methodology Q2(R1); 2005.
- Guidance for over-the-counter (OTC) human chorionic gonadotropin (hCG) 510(K)s. 2000.
- Official Journal of the European Union. Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017. Available from: <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32017R0745&from=EN.%20Available%20at:%20https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32017R0745&from=EN> [Accessed August 2020].
- GOV.UK. An introductory guide to the medical device regulation (MDR) and the in vitro diagnostic medical device regulation (IVDR). Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/640404/MDR_IVDR_guidance_Print_13.pdf [Accessed August 2020].
- TÜV SÜD. FAQs: in vitro diagnostic medical device regulation (IVDR). Available from: <https://www.Tuvsud.Com/en-gb/industries/healthcare-and-medical-devices/medical-devices-and-ivd/medical-device-market-approval-and-certification/eu-in-vitro-diagnostic-medical-device-regulation/faqs-in-vitro-diagnostic-medical-device-regulation-ivdr> [Accessed August 2020].
- Kovalevskaia G, Kakuma T, Schlatterer J, O'Connor JF. Hyperglycosylated hCG expression in pregnancy: cellular origin and clinical applications. *Mol Cell Endocrinol* 2007;260–262:237–43.
- Johnson S, Eapen S, Smith P, Warren G, Zinaman M. Significance of pregnancy test false negative results due to elevated levels of beta-core fragment hCG. *J Immunoassay Immunochem* 2017;38:449–55.
- Grenache DG, Greene DN, Dighe AS, Fantz CR, Hoefner D, McCudden C, et al. Falsely decreased human chorionic gonadotropin (hCG) results due to increased concentrations of the free beta subunit and the beta core fragment in quantitative hCG assays. *Clin Chem* 2010;56:1839–44.
- Cervinski MA, Lockwood CM, Ferguson AM, Odem RR, Stenman UH, Alfthan H, et al. Qualitative point-of-care and over-the-counter urine hCG devices differentially detect the hCG variants of early pregnancy. *Clin Chim Acta* 2009;406:81–5.
- Stenman UH, Bidart JM, Birken S, Mann K, Nisula B, O'Connor J. Standardization of protein immunoprocures. Choriogonadotropin (CG). *Scand J Clin Lab Invest Suppl* 1993;216:42–78.
- Alfthan H, Schroder J, Fraser R, Koskimies A, Halila H, Stenman UH. Choriogonadotropin and its beta subunit separated by hydrophobic-interaction chromatography and quantified in serum during pregnancy by time-resolved immunofluorometric assays. *Clin Chem* 1988;34:1758–62.
- McChesney R, Wilcox AJ, O'Connor JF, Weinberg CR, Baird DD, Schlatterer JP, et al. Intact hCG, free hCG beta subunit and hCG beta core fragment: longitudinal patterns in urine during early pregnancy. *Hum Reprod* 2005;20:928–35.
- Sturgeon CM, Berger P, Bidart JM, Birken S, Burns C, Norman RJ, et al. Differences in recognition of the 1st WHO international reference reagents for hCG-related isoforms by diagnostic immunoassays for human chorionic gonadotropin. *Clin Chem* 2009;55:1484–91.
- Birken S, Berger P, Bidart JM, Weber M, Bristow A, Norman R, et al. Preparation and characterization of new WHO reference reagents

- for human chorionic gonadotropin and metabolites. *Clin Chem* 2003;49:144–54.
28. Bristow A, Berger P, Bidart JM, Birken S, Norman R, Stenman UH, et al. Establishment, value assignment, and characterization of new WHO reference reagents for six molecular forms of human chorionic gonadotropin. *Clin Chem* 2005;51:177–82.
 29. Berger P, Sturgeon C, Bidart JM, Paus E, Gerth R, Niang M, et al. The ISOBM TD-7 workshop on hCG and related molecules. Towards user-oriented standardization of pregnancy and tumor diagnosis: assignment of epitopes to the three-dimensional structure of diagnostically and commercially relevant monoclonal antibodies directed against human chorionic gonadotropin and derivatives. *Tumour Biol* 2002;23:1–38.
 30. Berger P, Paus E, Hemken PM, Sturgeon C, Stewart WW, Skinner JP, et al. Candidate epitopes for measurement of hCG and related molecules: the second ISOBM TD-7 workshop. *Tumour Biol* 2013;34:4033–57.
 31. Sturgeon CM, Ellis AR. Standardization of FSH, LH and hCG – current position and future prospects. *Mol Cell Endocrinol* 2007;260-262:301–9.
 32. Pettersson K, Alfthan H, Stenman UH, Turpeinen U, Suonpää M, Söderholm J, et al. Simultaneous assay of α -fetoprotein and free β subunit of human chorionic gonadotropin by dual-label time-resolved immunofluorometric assay. *Clin Chem* 1993;39:2084–9.
 33. Cole LA, Kardana A, Park SY, Braunstein GD. The deactivation of hCG by nicking and dissociation. *J Clin Endocrinol Metab* 1993;76:704–10.
 34. Spencer K, Macri JN, Carpenter P, Anderson R, Krantz DA. Stability of intact chorionic gonadotropin (hCG) in serum, liquid whole blood, and dried whole-blood filter-paper spots: impact on screening for Down syndrome by measurement of free β -hCG subunit. *Clin Chem* 1993;39:1064–8.
 35. Lempiäinen A, Hotakainen K, Alfthan H, Stenman UH. Loss of human chorionic gonadotropin in urine during storage at -20°C . *Clin Chim Acta* 2012;413:232–6.
 36. CLSI. Evaluation of precision of quantitative measurement procedures; approved guideline—third edition. CLSI document EPO5-A3. Wayne, PA: Clinical and Laboratory Standards Institute; 2014.
 37. Gifford JL, Sadrzadeh SMH, Naugler C. Biotin interference: underrecognized patient safety risk in laboratory testing. *Can Fam Physician* 2018;64:370.
 38. Nerenz RD, Song H, Gronowski AM. Screening method to evaluate point-of-care human chorionic gonadotropin (hCG) devices for susceptibility to the hook effect by hCG beta core fragment: evaluation of 11 devices. *Clin Chem* 2014;60:667–74.
 39. Anderson Z, Larson E, Khan M, Bell M. False negative urine pregnancy testing with complete molar pregnancy: an example of the hook effect. *S D Med* 2016;69:55–7.
 40. Nerenz RD, Butch AW, Woldemariam GA, Yarbrough ML, Grenache DG, Gronowski AM. Estimating the hCG β cf in urine during pregnancy. *Clin Biochem* 2016;49:282–6.
 41. Greene DN, Grenache DG, Education Committee of the Academy of Clinical Laboratory P, Scientist. Pathology consultation on human chorionic gonadotropin testing for pregnancy assessment. *Am J Clin Pathol* 2015;144:830–6.
 42. Gronowski AM, Cervinski M, Stenman UH, Woodworth A, Ashby L, Scott MG. False-negative results in point-of-care qualitative human chorionic gonadotropin (hCG) devices due to excess hCGbeta core fragment. *Clin Chem* 2009;55:1389–94.
 43. Alfthan H, Stenman UH. Pregnancy serum contains the beta-core fragment of human chorionic gonadotropin. *J Clin Endocrinol Metab* 1990;70:783–7.
 44. Williams GR, Cervinski MA, Nerenz RD. Assessment of biotin interference with qualitative point-of-care hCG test devices. *Clin Biochem* 2018;53:168–70.
 45. Wilcox AJ, Baird DD, Dunson D, McChesney R, Weinberg CR. Natural limits of pregnancy testing in relation to the expected menstrual period. *J Am Med Assoc* 2001;286:1759–61.
 46. Ashra NB, Marriott L, Johnson S, Abrams KR, Crowther MJ. Jointly modelling longitudinally measured urinary human chorionic gonadotrophin and early pregnancy outcomes. *Sci Rep* 2020;10:4589.
 47. Snyder JA, Haymond S, Parvin CA, Gronowski AM, Grenache DG. Diagnostic considerations in the measurement of human chorionic gonadotropin in aging women. *Clin Chem* 2005;51:1830–5.
 48. NHS, UK. Ectopic pregnancy. Diagnosis. Available from: <https://www.nhs.uk/conditions/ectopic-pregnancy/diagnosis/> [Accessed August 2020].
 49. NHS, UK. Ectopic pregnancy. Symptoms. Available from: <https://www.nhs.uk/conditions/ectopic-pregnancy/symptoms/> [Accessed August 2020].
 50. Berger P, Sturgeon C. Human chorionic gonadotropin isoforms and their epitopes: diagnostic utility in pregnancy and cancer. *Expert Opin Med Diagn* 2008;2:1347–64.
 51. Borrelli PT, Butler SA, Docherty SM, Staite EM, Borrelli AL, Iles RK. Human chorionic gonadotropin isoforms in the diagnosis of ectopic pregnancy. *Clin Chem* 2003;49:2045–9.
 52. Butler SA, Cole LA. Evidence for, and associated risks with, the human chorionic gonadotropin supplemented diet. *J Diet Suppl* 2016;13:694–9.
 53. Cole LA, Kardana A, Ying FC, Birken S. The biological and clinical significance of nicks in human chorionic gonadotropin and its free beta-subunit. *Yale J Biol Med* 1991;64:627–37.
 54. Cole LA, Shahabi S, Butler SA, Mitchell H, Newlands ES, Behrman HR, et al. Utility of commonly used commercial human chorionic gonadotropin immunoassays in the diagnosis and management of trophoblastic diseases. *Clin Chem* 2001;47:308–15.
 55. Cole LA, Butler S. Detection of hCG in trophoblastic disease. The USA hCG reference service experience. *J Reprod Med* 2002;47:433–44.
 56. Cole LA, Khanlian SA, Sutton JM, Davies S, Rayburn WF. Accuracy of home pregnancy tests at the time of missed menses. *Am J Obstet Gynecol* 2004;190:100–5.
 57. Cole LA, Butler S. Familial hCG syndrome: production of variable, degraded or mutant forms of hCG. *J Reprod Med* 2014;59:435–42.
 58. Daviaud J, Fournet D, Ballongue C, Guillem GP, Leblanc A, Casellas C, et al. Reliability and feasibility of pregnancy home-use tests: laboratory validation and diagnostic evaluation by 638 volunteers. *Clin Chem* 1993;39:53–9.
 59. Greene DN, Schmidt RL, Kamer SM, Grenache DG, Hoke C, Lorey TS. Limitations in qualitative point of care hCG tests for detecting early pregnancy. *Clin Chim Acta* 2013;415:317–21.
 60. Grenache DG, Moller KA, Groben PM. Endometrial adenocarcinoma associated with elevated serum concentrations of the free β subunit of human chorionic gonadotropin. *Am J Clin Pathol* 2004;121:748–53.
 61. Gronowski AM, Fantz CR, Parvin CA, Sokoll LJ, Wiley CL, Wener MH, et al. Use of serum FSH to identify perimenopausal women with pituitary hCG. *Clin Chem* 2008;54:652–6.

62. Gronowski AM, Grenache DG. Characterization of the hCG variants recognized by different hCG immunoassays: an important step toward standardization of hCG measurements. *Clin Chem* 2009;55:1447–9.
63. Johnson SR, Miro F, Barrett S, Ellis JE. Levels of urinary human chorionic gonadotrophin (hCG) following conception and variability of menstrual cycle length in a cohort of women attempting to conceive. *Curr Med Res Opin* 2009;25:741–8.
64. Johnson SR, Godbert S, Perry P, Parsons P, Roberts L, Buchanan P, et al. Accuracy of a home-based device for giving an early estimate of pregnancy duration compared with reference methods. *Fertil Steril* 2013;100:1635–41 e1.
65. Johnson S, Cushion M, Bond S, Godbert S, Pike J. Comparison of analytical sensitivity and women's interpretation of home pregnancy tests. *Clin Chem Lab Med* 2015;53:391–402.
66. Johnson S, Stanford JB, Warren G, Bond S, Bench-Capon S, Zinaman MJ. Increased likelihood of pregnancy using an app-connected ovulation test system: a randomized controlled trial. *J Womens Health (Larchmt)* 2020;29:84–90.
67. Larsen J, Buchanan P, Johnson S, Godbert S, Zinaman M. Human chorionic gonadotropin as a measure of pregnancy duration. *Int J Gynaecol Obstet* 2013;123:189–95.
68. Nettleman MD, Ayoola AB, Brewer JR. Utilization of home pregnancy testing among women at risk for unintended pregnancy. *Wom Health Issues* 2009;19:263–7.
69. Pike J, Godbert S, Johnson S. Comparison of volunteers' experience of using, and accuracy of reading, different types of home pregnancy test formats. *Expert Opin Med Diagn* 2013;7:435–41.
70. Ross E. Provisionally pregnant: uncertainty and interpretive work in accounts of home pregnancy testing. *Health (London)* 2018;22:87–105.
71. Snyder ML, Wiley C, Molinaro RJ, Ritchie JC, Fantz CR. Falsely increased hCG in patients with high leukocyte counts. *Clin Chem* 2013;59:1127–9.
72. Sturgeon CM, McAllister EJ. Analysis of hCG: clinical applications and assay requirements. *Ann Clin Biochem* 1998;35(Pt 4):460–91.
73. Wallace LS, Zite NB, Homewood VJ. Making sense of home pregnancy test instructions. *J Womens Health (Larchmt)* 2009;18:363–8.
74. Whittington J, Fantz CR, Gronowski AM, McCudden C, Mullins R, Sokoll L, et al. The analytical specificity of human chorionic gonadotropin assays determined using who international reference reagents. *Clin Chim Acta* 2010;411:81–5.

Supplementary Material: The online version of this article offers supplementary material (<https://doi.org/10.1515/cclm-2020-1523>).