AACC Guidance Document on the Use of Point-of-Care Testing in Fertility and Reproduction

Edited by James H. Nichols, PhD, DABCC, FAACC, Vanderbilt University Medical Center, Nashville, TN

Expert Panel:
Mahesheema Ali, MSc, PhD, NRCC, FAACC, The MetroHealth System, Cleveland, OH
John Anetor, PhD, FIMLS (Nig.), FIBMS (UK), ERT, FACN, FRSC(UK), FAACC, University of Ibadan, Nigeria
Li-Sheng Chen, PhD, DABCC, FAACC, Silver Spring, MD
Yu Chen, MD, PhD, FAACC, FCACB, DABCC, Dr. Everett Chalmers Regional Hospital, Horizon Health Network, and Dalhousie University, Fredericton, NB, Canada
Sean Collins, MD, Vanderbilt University Medical Center, Nashville, TN
Saswati Das, MD, Dr. Ram Manohar Lohia Hospital & Atal Bihari Vajpayee Institute of Medical Sciences, New Delhi, India
Sridevi Devaraj, PhD, DABCC, FAACC, Texas Children’s Hospital and Baylor College of Medicine, Houston, TX
Lei Fu, PhD, DABCC, FAACC, FCACB, Sunnybrook Health Sciences Center, Toronto, ON Canada
Dina Nicole Greene, PhD, DABCC, University of Washington, Seattle, WA
Brad S. Karon, MD, PhD, Mayo Clinic, Rochester, MN
Heba Kary, MD, FRCPC, King Fahd Armed Forces Hospital, Jeddah, Saudi Arabia
Robert D. Nerenz, PhD, Dartmouth-Hitchcock Medical Center, Lebanon, NH
Alex J. Rai, PhD, DABCC, FAACC, Columbia University College of Physicians and Surgeons and New York Presbyterian Hospital, New York, NY
Zahra Shajani-Yi, PhD, DABCC, FAACC, Laboratory Corporation of America (LabCorp), San Diego, CA
Vinita Thakur, PhD, FAACC, Eastern Health Authority, Health Science Center and Memorial University, St. John’s, NL, Canada
Sihe Wang, PhD, DABCC, FAACC, Akron Children’s Hospital, Akron, OH
Hoi-Ying Elsie Yu, PhD, DABCC, FAACC, Geisinger Health System, Danville, PA
Lindsey E. Zamora, MD, MPH, Vanderbilt University Medical Center, Nashville, TN

Abstract

Background: The American Association for Clinical Chemistry (AACC) Academy revised the reproductive testing section of the Laboratory Medicine Practice Guidelines: Evidence-Based Practice for Point-of-Care Testing (POCT) published in 2007.

Methods: A panel of AACC Academy members with expertise in POCT and laboratory medicine was formed to develop guidance for the use of POCT in reproductive health, specifically ovulation, pregnancy, premature rupture of membranes (PROM) and high-risk deliveries. The committee was supplemented with clinicians having Emergency Medicine and Obstetrics/Gynecology training.

Results: Key recommendations include 1) Ovulation POCT - urine luteinizing hormone (LH) tests are accurate and reliable predictors of ovulation. The use of home-based monitors are popular for identifying the fertile window of the menstrual cycle. Urinary LH POCT is used as a part of the procedures in various assisted reproductive technologies to facilitate fertility treatment. 2) Pregnancy POCT - should be considered in clinical situations where rapid diagnosis of pregnancy is needed for treatment decisions, and laboratory analysis cannot meet the required turnaround time (TAT). Antigen-excess, proteinuria, biotin, dilution effects and other sources of interference should be considered with pregnancy tests. 3) Premature rupture of membranes (PROM) – PROM testing using commercial kits alone is not recommended without clinical signs of rupture of membranes (ROM), such as leakage of amniotic fluid from the cervical opening. 4) Delivery - fetal scalp lactate testing is recommended over fetal scalp pH testing for the management of intrapartum pregnancy with abnormal fetal heart rate.
Conclusions: This revision of the AACC Academy POCT Guidelines provides recommendations for best practice use of POCT in fertility and reproduction.

Key Words: Point-of-Care Testing (POCT), Reproduction, Ovulation, Pregnancy, Premature Rupture of Membranes and High-Risk Delivery.

Impact Statement

Point-of-care testing (POCT) is growing in popularity as a means of delivering faster turnaround time of test results closer to the patient. Guidance is needed for optimizing the implementation of POCT in patient care. This guidance document revises previous recommendations and offers best practice for the use of POCT in fertility and reproductive health.

Introduction

POCT is defined as clinical laboratory testing conducted close to the site of patient care, typically by clinical personnel whose primary training is not in the clinical laboratory sciences or by patients (self-testing).(1) POCT refers to any laboratory tests performed outside of the traditional, core or central laboratory. POCT is an increasingly popular means of bringing laboratory testing closer to the patient. POCT can provide faster turnaround of results compared to core laboratory testing, because POCT eliminates transportation and considerably reduces processing and preanalytical steps required for laboratory tests. Faster test results offer the potential for rapid medical decisions. The convenience, ease-of-use, and speed are driving POCT growth. POCT described in this guideline can be performed in the hospital or clinic setting or even purchased over the counter and performed at home.

Despite the apparent simplicity, many factors can negatively affect POCT quality. Exposure to environmental extremes of heat, cold, and humidity during shipping or storage can degrade reagents and controls. Some operators of POCT systems may not have training on par with that of laboratorians. Failure to follow manufacturer’s instructions, such as over-timing or under-timing test development, can lead to false-positives and false-negatives. Failure to analyze controls or troubleshoot when controls are out-of-range can impact accuracy. Seemingly minor interruptions, such as moving test devices during analysis, can affect POCT results. POCT is methodologically distinct, with different interferences and limitations compared to laboratory methods. Thus, POCT results are not necessarily harmonized with laboratory methods and test results may not agree.

Although POCT has great potential to improve patient care, when inappropriately utilized or incorrectly performed, POCT can lead to unnecessary follow-up tests and procedures with considerable clinical and financial implications for the patient. Using POCT to improve the speed of care delivery in time-sensitive settings may also put the patient at risk if errors occur. Clinicians/POCT operators need guidance in utilizing POCT and in implementing good laboratory practices to obtain reliable results. The American Association for Clinical Chemistry (AACC) Academy (formerly known as the National Academy of Clinical Biochemistry, NACB) developed best practice recommendations for use of POCT in patient care. The Laboratory Medicine Practice Guidelines (LMPG): Evidence-Based Practice for Point-of-Care Testing were published in 2007.(1) Those guidelines critically reviewed the peer literature, graded the evidence that links POCT to clinical outcomes, and provided recommendations for optimizing POCT utilization. They offer a comprehensive, systematic review of the POCT literature. However, recent POCT studies have been published, and the original 2007 LMPG consequently needs updating. Instead of revising the entire LMPG, sections of the document were prioritized for revision to narrow the focus of the revision process and expedite publication of updated recommendations. The first section to be revised was published last year as the AACC Guidance Document on Management of POCT.(2) This manuscript addresses revision of the “Reproductive Testing” section of the 2007 POCT guidelines.

Methods
The AACC Academy formed a committee of experts with interest and experience in POCT and laboratory testing for fertility and reproduction. The expert committee was composed of AACC Academy members and was supplemented with clinicians having Emergency Medicine and Obstetrics/Gynecology (OB/GYN) training. The committee divided into several subgroups to formulate clinical questions related to the use of POCT in the assessment of ovulation, pregnancy, premature rupture of membranes and evaluation of fetal distress. Literature searches were conducted for peer literature that could address each clinical question. Common search engines were utilized such as PubMed, Cochrane, Embase and Web of Science. Searches were limited to manuscripts published in the English language with emphasis on recent literature published since the previous 2007 guidelines. Publications on test performance, sensitivity, and potential interferences were also included. Guidance documents from other professional organizations such as the American College of Obstetrics and Gynecology (ACOG) were reviewed to not duplicate recommendations. Once drafted by the committee, prospective guidelines were distributed amongst the AACC membership and to the wider laboratory community for public comment. Revisions were made to address each comment, and the final guidance document was approved by the AACC Academy Council prior to publication. This guidance document represents a consensus opinion of the expert committee and provides best practice recommendations for use of reproductive POCT in patient management.

GUIDANCE RECOMMENDATIONS

POINT-OF-CARE OVULATION TESTING

The menstrual cycle refers to the cyclic changes in the ovaries and endometrium caused by the interplay of female hormones in women of reproductive age to prepare for the possibility of pregnancy. Ovulation is the process of releasing the mature oocyte from the ovary into the fallopian tube, typically near the midpoint of the menstrual cycle. In the fallopian tube, a sperm may fertilize the released egg. The fertility window refers to the period of the menstrual cycle during which pregnancy is most likely to be achieved. The fertility window can vary and is dependent on the length of the menstrual cycle; it usually begins approximately 3–5 days before ovulation and lasts until 24 hours after ovulation.

Ultrasonography, luteinizing hormone (LH), estradiol, and progesterone testing are the main methods employed to detect, predict and monitor ovulation in clinical practice. They are also valuable tools to evaluate ovarian function and optimize the timing of fertilization in assisted reproductive procedures. The development of point-of-care testing such as basal body temperature measurement, salivary and vaginal mucus ferning, and urinary LH/estrone-3-glucuronide testing provides simple and convenient ways for women to identify their fertile windows.

Is the diagnostic accuracy of urine LH tests sufficient for detecting and predicting ovulation when compared to ultrasound as the gold standard for confirming ovulation? Multiple studies have demonstrated that urine LH tests are accurate and reliable predictors of ovulation with the urinary LH surge preceding ovulation by approximately 1 day.

Follicle-stimulating hormone (FSH) is secreted by the pituitary and promotes the growth of multiple follicles inside the ovary. These follicles produce estrogen to induce the abrupt increase of luteinizing hormone (LH) which in turn triggers the dominant follicle(s) to rupture and release the mature oocyte(s). The menstrual cycle is divided into two phases—the follicle phase starts from the first day of menstruation until ovulation, after ovulation the luteal phase begins until the onset of next menstruation. The average length of a woman’s menstrual cycle is considered to be 28 days; however, it has been long known that there exists considerable inter-individual and intra-individual menstrual cycle variability.
The urinary LH POC test were first performed in clinics by professional staffs, but now, over-the-counter urinary LH testing for fertility monitor are available for at home testing without prescription.

Earlier reports established that commercially available urinary LH test kits have excellent diagnostic sensitivity (85-100%, median 100%) and predictive value for ovulation (85-100%, median 93%) with ovulation occurring within approximately 1 day of the LH surge. As the performance of urinary LH surge detection kits has been established, more recent studies have focused on understanding the underlying population variability in order to improve ovulation prediction, including 1) developing optimal methods and algorithms for detecting the surge, and 2) establishing guidance on the cycle-length tailored testing protocol in the menstrual cycle to increase the predictive value of the fertility monitoring methods.

Alliende et al. showed that hormone trajectories in fertile women differed from mean hormone curves. Park et al. characterized the LH surge and noted a high degree of variability in configuration (single peak vs. double/multiple peaks), amplitude (range in the peak level of LH relative to baseline), and duration of the LH surge. Direito et al. further characterized and analyzed the relationship of the LH surge relative to profiles of the other reproductive hormones. The study results showed that there exists a high variation in the range of amplitudes (2.1x-61x) of the LH surge, and only 48% of the study women had a single LH peak, while 33-41% of participants had two or more LH peaks or exhibited an LH plateau that lasted >5 days (11%). Godbert et al. determined that the ability to correctly identify the LH surge in urine varied widely depending on the statistical algorithm used. This study reviewed the published predictive methods and classified them into 3 broad categories based on how the baseline LH is estimated: “fixed days”, “peak LH day”, and “estimated LH surge” (TABLE 1), and defined LH surge as a sustained rise in LH concentration above the baseline LH. The primary difference between the methods is how the LH baseline assessment is established.

The “fixed days” method required no previous cycle information; whereas the other two methods (“peak LH day” and “estimated LH surge”) needed complete cycle data and could only be evaluated retrospectively. The “peak LH day” identified the cycle peak day and established a reference LH concentration based on setting a set number of days (ranging between 3-6) before the peak. The “estimated LH surge” estimated the LH surge based on analysis of the measured concentrations and set the baseline concentration LH day 1-2 days prior to the surge. Performance evaluation was conducted using AutoDELFIA (Perkin Elmer, Waltham, MA, USA) LH assay with a sensitivity of 0.1 mIU/ml and <5% for both inter-and intra-assay coefficients of variation. Samples were collected from 227 premenopausal women (age 18 – 45 years) and the complete profiles of LH were determined using the above three methods to calculate the LH baseline. The study results show that the “estimated LH surge” method delivers the most accurate performance as shown in TABLE 1. The use of the second day before the estimated surge as reference plus the previous 4/5 days to calculate the baseline LH gave the most reliable results.

The Menstrual Cycle Monitoring (MeMO) was a prospective study of normally menstruating women between the ages of 18- 40 (n=40) that examined the inter-individual variation of urine and serum reproductive hormones and their relationship to ovulation, as confirmed by ultrasound measurements. The investigation found that, among the study participants, the menstrual cycle length ranged from 22-37 days, the length of the luteal phase ranged from 3-15.5 days, and the day of ovulation ranged from day 8 to day 26. More recent large-scale studies further confirmed the wide variations in the cycle length and date of ovulation at the population level. The MeMO study also found that the LH surge preceded ovulation in all participants by a mean of 0.81 days. In a separate report of the same study, the urinary ranges of LH, E3G and FSH reference to the actual ovulation day (determined by ultrasonography) were established. Daily urine samples of a complete cycle were collected and batch-analyzed; LH concentrations were evaluated side-by-side using AutoDELFIA platform and an in-house developed assay. AutoDELFIA assay detects intact LH,
free β LH and the dominant urinary metabolite LH β core fragment (LH-βcf), while the in-house assay only recognizes intact LH. The study showed that timing of the urinary LH peak was assay-dependent and did not always precede ovulation. The profiling using AutoDELFIA assay gave relatively variable patterns of LH profile and also recorded the LH peak about 1-day lag behind the corresponding peak using intact LH-specific assay. However, both assays worked equitably well in defining the day of the LH surge.(16)

Leiva et al. recently conducted a study using receiver operating characteristic (ROC) analysis to determine the optimal urinary LH threshold for the LH peak method in predicting ovulation within 24 hours of LH surge. The best performance was at a threshold of 25 mIU/mL, which yielded a sensitivity of 54%, specificity of 97%, PPV of 50%, and NPV of 98%. Therefore, the timing of the LH peak alone should not be used for predicting or determining ovulation status. A complimentary biomarker may be used concurrently to improve specificity and predictive value.(9)

Each of these studies adds to our understanding of the individual variability in menstrual cycle physiology. More studies are needed to further clarify the mechanistic basis of this variability and address the complex interplay between female reproductive hormones and other important factors known to affect the female menstrual cycle such as body mass index (BMI), race, age, and environmental pollutants.

Does the use of ovulation predicting kits (measuring urine luteinizing hormone and estrone-3-glucuronide) for predicting ovulation in women not treated in a fertility clinic improve outcomes (i.e., increase conception rates, decrease the number of clinic visits/the numbers of unwanted pregnancies) compared to not using prediction tests?

The use of home-based monitors to measure both urinary LH and estrone-3-glucuronide is an established and popular method for identifying the fertile window of the menstrual cycle. Studies have shown that the use of ovulation predicting kits (OPKs) may improve the likelihood of conception among healthy fertile women seeking pregnancy.

Infertility refers to the inability to conceive within one year of unprotected intercourse. Infertility can occur in both men and women. Worldwide infertility is estimated at 15-25% of women at reproductive age, and there is an upward trend of infertility prevalence in the last two decades, especially in developing countries.(17, 18) In resource-limited settings, the options of medically assisted reproduction are either unaffordable or inaccessible. The use of ovulation predictor kits (OPKs) to identify the optimal timing for intercourse in the fertile window is convenient and economical, with the potential of increasing the conception rates, or conversely, for preventing unwanted pregnancies.

Home-use urinary LH testing kits, targeting women who were not under fertility treatment, were developed and became commercially available in the 1980s, and have now become a commonly used tool for evaluating the LH surge to time intercourse.(3) Our literature search revealed several recent publications examining the clinical impacts of home-use urinary LH testing as an ovulation predictor for women who are not under fertility treatment, and the findings are summarized below:

**Impact of OPK use on the conception rate**

The ClearBlue fertility monitor, which measures both LH and estrone-3-glucuronide to signal the onset and the closure of the fertility window respectively, is the most studied device among the current commercially available urinary LH POCT. In a prospective randomized control trial (RCT) with 653 healthy participants, the study results showed that the cumulative pregnancy rate (for two cycles) was significantly higher in the test group using Clearblue Easy Fertility Monitor compared with the control group (22.7% versus 14.4%).(19) The improved version, Bluetooth-compatible Connected Ovulation Test System, was evaluated similarly in a more recent study involving 844 volunteers.(20) The study results revealed the conception rates of the test versus the control group were 25.4% vs. 14.7% (p <0.001) after one cycle and 36.2% vs. 26.8% (p = 0.026) after two cycles.
In a systematic review conducted to inform World Health Organization (WHO) guidelines on self-care interventions, Yeh et al. performed a meta-analysis that included four studies (three RCTs and one observational study conducted between 1996-2013) with a total of 1,487 participants from high-income countries. The analysis showed a higher self-reported pregnancy rate in those using OPKs in all three randomized control trials. The result from the pooled data analysis found that the ratio of pregnancy among participants with OPK use compared to pregnancy without OPK use (i.e. the relative risk, RR) is 1.36 (95%CI = 1.07-1.73).(21)

In addition, a recent study supported the adoption of online tracking systems and fertility monitoring apps as a simple, economical, and effective way to help couples achieve pregnancy. In a 24-cycle prospective effectiveness study (n=256), the fertile window was identified with Clearblue fertility monitor or cervical mucus monitoring, and an online tool was deployed to primary care clinic for participants to record and monitor their observations and use the data to time intercourse. The study results showed that the pregnancy rate reached 100% at 24 cycles for those women using the hormonal (LH and estrone-3-glucuronide) fertility monitor.(22)

Impact of OPK use on pregnancy avoidance

Fertility awareness is a collective term describing several different contraceptive methods based on the collection and guided interpretation of various personal fertility signs/symptoms (including basal body temperature, cervical fluid, menstrual cycle length, and urinary assays for reproductive hormones) to predict or identify the fertile window of the menstrual cycle, so an individual can use the information to reduce the probability of pregnancy. All fertility awareness methods meet the WHO criteria of modern methods with “a sound basis in reproductive biology, a precise protocol for correct use and the evidence of efficacy under various conditions based on appropriately designed studies”.(23)

The Marquette Method involves users both observing their cervical fluid and measuring urinary hormone levels (both LH and estrone-3-glucuronide) with the recommended home-use POCT, the Clearblue Fertility monitor. The effectiveness of this method was supported by a 12-month retrospective study with 204 couples, who were taught to correctly follow the protocol of the method for avoiding pregnancy. The study results showed that the 12-month pregnancy rate of a “correct use” group was 0.6 % (i.e., 99.4% effective) and of a “typical use” group was 10.6 (i.e., 89.4% effective) per 100 users.(24) Subsequently, two comparative efficacy studies targeting perimenopausal (n= 160) and breastfeeding women (n=816) suggest this method can be effective for pregnancy prevention in older and nursing women, especially with the adoption of online tracking/monitoring systems.(25, 26)

Compared to other more effective contraceptive methods (contraceptive pill and physical barriers), the adoption of fertility awareness methods in the US is low, but the demand has increased 3 fold from 2008 to 2015 as the result of the wider use of the app- or tech-based fertility trackers as contraception.(23) More education, training, and research are needed to properly use hormonal fertility monitors as a means to empower women with awareness and good decision-making for their own reproductive wellbeing.

Impact of OPK use on other health outcomes

Home-based hormonal fertility monitors also impact other outcome measures, including the general acceptance and satisfaction of end-users to improve fertility awareness and knowledge, and enabling women to know their bodies better. The study findings also supported that the use of OPKs did not cause additional stress/anxiety in general, especially among those who do become pregnant.(21)
The wide adoption of online tracking technology may further promote OPK use as an inexpensive and effective tool of fertility management to benefit women living in low-resource settings. However, more research is required to investigate the values/preferences and social harms/adverse events using a diverse set of OPKs other than the Clearblue Monitor system, especially in resource-limited settings.

The benefits of using mobile computing apps have been revealed in several recent studies: in a study using the data collected from a North American prospective internet-based cohort (n=8363), it was found that the use of cycle apps to monitor and correlate multiple fertility indicators was associated with a 12-20% increase of conception rate after adjusting for possible confounding factors. In another study, analysis of “big data” (total 75,981 cycles) collected from 32,595 users of a connected ovulation monitor confirmed the existence of wide individual variations of cycle length and ovulation dates and the data were used to derive the population range and a probability table of ovulation per cycle day. Meanwhile, the market for fertility apps is rapidly growing, and these apps are gaining wider acceptance for identifying female fertility windows. However, regulatory bodies, academia, and industry should collaboratively develop a standard metric evaluation system to objectively assess the performance and usefulness of these apps based on their intended uses, and to inform consumers by periodically reporting the evaluations so they may selectively use only the evidence-based effective products.

Does the use of urine LH tests for predicting ovulation in women undergoing fertility treatment improve outcomes (i.e., increase conception rates, decrease the number of clinic visits/the number of fertility treatment cycles) compared to not using prediction tests? Urinary LH point-of-care testing demonstrated a comparable performance among other ovulation monitoring methods for timing intrauterine insemination and confirming sufficient ovulation induction before oocyte retrieval during in vitro fertilization (IVF). Limited data supported urinary LH as a cost-effective measure. However, the actual cost saving may vary depending on the other factors such as population demographics, geographic locations, health resource allocations, and insurance coverage; therefore, the cost savings need to be empirically determined.

Infertility, as defined earlier in this document, is a common disease and affects one in six couples in Western countries including the United States. In these countries, if the expectant management of unassisted pregnancy failed, fertility treatment via various assisted reproductive technologies (ARTs) are viable options for achieving a successful pregnancy.

OPK use for timing of insemination in intrauterine insemination

Intrauterine insemination (IUI) with or without ovulation induction is usually the first line of treatment for couples with unexplained infertility. IUI involves procedures to place concentrated sperm directly in the uterus around the time of ovulation. The timing of IUI is critical to increasing the chance of fertilization and subsequent pregnancy. The use of an at-home ovulation prediction kit (OPK) is one of the available methods to predict the optimal timing of IUI. In a meta-analysis including 2,279 infertile couples from 14 RCTs comparing the effectiveness of different synchronization methods for IUI, no differences in pregnancy rate, live birth rates or adverse events were found among the evaluated methods (i.e. serum or urinary LH detection, ultrasound detection of ovulation, basal body temperature, and human chorionic gonadotropin (hCG)/gonadotropin releasing hormone agonist (GnRHa) administration). A more recent cohort study (n=232 normo-ovulatory women) compared the timing of therapeutic donor sperm inseminations using urinary LH versus using ultrasound detection combined with ovulation triggering by hCG. The results of this study also showed no difference in the outcomes of pregnancy rate, live birth rates, and adverse events between these two synchronization methods.
A recent report estimated that 8 million babies have been born worldwide through IVF since its inception in 1978. IVF is the most effective form of assisted reproductive technologies (ARTs). It consists of a complex series of procedures to retrieve the mature oocytes, fertilize the chosen retrieved oocyte in the laboratory setting and then transfer the fertilized oocyte (embryo) back to the uterus to complete conception. IVF is the mainstream treatment for women with tubal disease and men with poor sperm quality but is also increasingly used for unexplained subfertility. IVF is time-consuming, expensive and invasive, but has enabled many infertile couples to become parents.(29)

GnRHa trigger is a recently introduced procedure for stimulating oocyte maturation before egg retrieval to prevent the complication of hyperstimulation syndrome. The measurement of serum LH at 12 hours after GnRHa trigger has been suggested for confirming the sufficiency of oocyte stimulation. Detecting an insufficient LH surge allows a rescue hCG bolus to be administered during the same cycle in order to avoid failed oocyte retrieval. However, the optimal serum LH threshold concentrations after the GnRHa trigger have not yet been established.(32) Urinary LH concentrations are highly correlated with serum LH concentration in the fertile window, and urinary LH has been used for timing insemination in IUI. Thus, self-testing of the LH surge (defined as LH>15 mIU/mL at 12 hours after injection) with home-use OPKs followed by communication via cellular phone has been proposed to replace serum LH testing as a simple, safe and convenient way to confirm the adequacy of GnRHa trigger for oocyte maturation. In a recent multicenter prospective cohort study conducted in Spain, Brazil, and Denmark, urinary LH results were obtained from 359 oocyte donors, 356 participants recorded positive and only 3 had negative urinary LH results, with one false-positive and one false-negative as confirmed by ultrasonography and serum LH results. This gave an overall sensitivity of 99.7% (355/356) and specificity of 66.7% (2/3).(33) In fact, the serum LH measurement for the only false-negative case was 18.6 mIU/mL, which is just above the defined positive trigger cutoff but is below the lower detection limit (25 mIU/mL) for the ovulation kit used. Therefore, a few negative LH results from an at-home ovulation monitoring kit would require reflex confirmation by serum testing to rule out false negative results. A cost analysis (including direct and indirect cost) was also performed, and OPK use was shown to have a significant cost saving. Serum LH would have cost 14,840€ (~$17,680) while urine LH kits cost only 1855€ (~$ 2,010).(33) More empirical cost-benefit studies are needed to provide definitive evidence before integrating OPK testing as a standard procedure for oocyte retrieval protocol.

So far, many RCTs and a few longitudinal studies on fertility treatments are mostly focusing on the impact of a particular ART procedure on the outcome measures of time-to-pregnancy, the number of clinic visits, and the number of fertility treatment cycles. There is scarce evidence to support or refute that the use of OPK alone can significantly improve these outcomes.

TABLE 2 summarizes the research design/methodologies, the study size, and the main results of the studies cited in answering the above two questions.

Is the diagnostic accuracy of non-urine ovulation POCT (including marketed and emerging devices) sufficient to predict ovulation when ultrasound is used as a gold standard? Recently commercialized salivary ferning tests are reusable and easy-to-use but have sub-par reliability and reporting accuracy. The new-generation biosensor-based basal body temperature monitors equipped with algorithms for pattern recognition enhanced their ability to detect ovulation, henceforth may be suitable for family planning. The claimed high accuracy by the device manufacturers needs to be further augmented by post-market performance studies. In addition, new devices equipped with software algorithms for continuously monitoring skin temperature, pulse rate, and ferning patterns are emerging into the ovulation point-of-care market. Preliminary studies of these emerging devices showed promising results. However, but more comprehensive studies are warranted to investigate the actual performance and justify the usability of these emerging devices.
Current commercially available non-urine ovulation tests include basal body temperature monitoring and salivary ferning analysis.

Salivary Ferning Test

The phenomenon of ferning (or crystallization) results from the cyclical increases of sodium and chloride concentrations in body fluids under the influence of estrogen. The ferning appearance due to the crystallization of NaCl is observable by examining the dried saliva or cervical mucus under a microscope. (5)

Two point-of-care devices based on the measurement of salivary electrical resistance have been marketed in the 1990s but did not gain traction, likely due to the inferior performance compared to other competing devices in the marketplace. The reported diagnostic accuracy of both devices varied from 52 to 74%. (34-36)

Recently, two relatively simple and user-friendly salivary ferning test kits were developed and became commercially available; both include a pocket-sized microscope for evaluating the ferning patterns. Know when Ovulation Monitoring System has a hand-held mini microscope and an accompanying smartphone app for recording the results to predict the fertile window. In a small open-label prospective study, compared with actual ovulation assessed by transvaginal ultrasound, salivary ferning was observed in 29 of 30 ovulatory cycles, while false-positive results occurred twice in 10 anovulatory cycles. (37) A subsequent prospective observational study involving 107 healthy women for a total of 114 cycles reported a sensitivity of 88.6% and a specificity of 80% with a positive predictive value of 93.3% and a negative predictive value of 69% compared to the results using ultrasoundography. (38) Another salivary ferning device, Geratherm® ovu Control, has a mini-microscopic lens with a switchable light source fixed on a lipstick-sized stand. When assessed in women under IVF treatment, this device showed a specificity of 78% and a sensitivity of 80% when compared to the rise of estradiol (E2) after follicle-stimulating hormone/hCG stimulation. The same study also showed an agreement of 89.4% between the evaluations of the ferning patterns performed by patients and by laboratory staff. (39) In another comparison study, 74 healthy women performed Geratherm® ovu Control and a urinary LH test side-by-side, the paired-results showed a high level of conformity, from the 5th (100%) until the 14th (84%) cycle day and from the 18th (80%) until the 22nd (96%) cycle day, corresponding to the pre-and post-ovulatory period, however the lack of ultrasonography to confirm the ovulation makes the study results hard to interpret. (40)

The most attractive features of the salivary ferning ovulation test are its reusability and relatively low cost, but the impact of physiological condition, disease state, and certain medications on the ferning patterns significantly limits the reliability of its results.

Basal Body Temperature (BBT) Monitors

The pattern of dip-rise-return of BBT reflects the changing level of thermogenic progesterone during the menstrual cycle; BBT reaches nadir approximately one day before ovulation, rises 0.5–1.0 °F after ovulation, plateaus, and then returns to a lower range around the time of menstrual bleeding. BBT monitoring has been and still is widely used to estimate the time of ovulation retrospectively. Modern digital thermometers and the available online services/smartphone apps make the recording of results more convenient. Innovative biosensor-based devices were introduced into the market to facilitate the pattern recognition of BBT and thus improve the predictability of ovulation.

OvuSense and OvuRing are vaginal biosensors. An open-label clinical study (n=148) showed that OvuRing has 99% accuracy for ovulation date detection and an 89% accuracy rate in predicting the ovulation date when referenced to ultrasound assessment, while OvuSense claimed to have equivalent performance. (41, 42)

Rather than its apparent rise during ovulation, BBT monitors rely on a dip of BBT to predict the onset of ovulation, thus it is less accurate in predicting than in detecting ovulation retrospectively. (43) Besides, BBT may
fluctuate with the changes in climate, room temperature, alcohol use, medications, many physiological conditions, and stress.

Emerging Wearable Biosensors

Various forms of wearable sensors, such as armbands, bracelets, and earbuds, are designed to measure the nocturnal skin temperature; some of them are still at the stage of pilot or proof-of-concept studies, while a few has emerged into market and offer a new way for monitoring ovulation in menstrual cycle.

The Oura ring is a commercially available wearable sleep and activity tracker and has a negative temperature coefficient (NTC) thermistor as its temperature sensor. In a clinical study, several algorithms for tracking the start of menstruation and ovulation were developed and tested involving 22 women for whom ovulations were predicted by recording and analyzing the changes of intravaginal temperature using proprietary software. The best-performing algorithm was shown to have the sensitivity of 83.3% and with a similar positive predictive value. A wrist-worn armband was shown to detect a sustained 3-day BBT shift pattern in 357/437 cycles (82%) in an observational study with 136 healthy non-pregnant participants. Another non-invasive wearable device is a thermometer that consists of an earpiece, which measures the ear canal temperature every 5 minutes during night sleep hours, and a base station that transmits the data to a smartphone application for analysis. Results from a feasibility study of 34 users yielded detection accuracy with a sensitivity of 92.31% when compared with data obtained from an ovulation test kit.

The changing levels of estrogen and progesterone are known to affect the cardiovascular system, and studies have demonstrated that pulse rate significantly increases during the fertile window compared to the menstrual phase. Ava bracelet (Ava AG, Zurich, Switzerland) detects the significant, concurrent phase-based shifts in wrist skin temperature, heart rate, and respiratory rate, followed by analyzing the data collected by the bracelet using an accompanying machine learning algorithm. A prospective longitudinal study involving 237 conception-seeking Swiss women showed Ava bracelet detected the fertile windows with 90% accuracy when compared with ovulation date predicted by the Clear Blue Ovulation Monitor. In another prospective observational clinical trial with 91 healthy non-pregnant women, the pulse rates (PR) were measured by a photoplethysmography-based wrist sensor for fertility monitoring. A significant increase of PR (2.1 beats per minute, p<0.01) during the fertile window was observed compared to the PR in menstrual phase.

So far none of the emerging devices has yet published randomized prospective clinical studies using the gold standard ultrasonography to confirm pregnancy result. Demonstration of correlation between the use of the wearable BBT devices and the improved pregnancy outcome data is required to support the efficacy of these device as reliably accurate ovulation predictor.

Overall, we have seen the continuous improvements and innovations of ovulation point-of-care devices in the past decade, especially the biosensor-based BBT monitors and the emerging wearable biosensors, which are especially popular for tech-savvy generation and those who may benefit for those who have abnormal ovulation pattern, such as those with polycystic ovary syndrome. Non-urine ovulation tests provide convenience and privacy for the device users to predict the fertile window at home either for increasing the possibility of conception or to avoid pregnancy.

POINT-OF-CARE PREGNANCY (BETA-HUMAN CHORIONIC GONADOTROPIN) TESTING

When should POCT hCG testing be considered in place of laboratory hCG analysis? POCT should be considered in clinical situations where rapid diagnosis of pregnancy is needed for treatment decisions, and laboratory analysis cannot meet the required turnaround time (TAT). This would include situations where patients present in the acute setting with
unstable vital signs and symptoms raising concern for ruptured ectopic pregnancy. A positive urine or whole blood bedside POCT hCG test combined with an ultrasound suggesting free fluid in the abdomen would be helpful when considering early operative intervention.

Beta-human chorionic gonadotropin (hCG) is a sialoglycoprotein that is initially secreted by the trophoblastic cells of the placenta shortly after implantation of the fertilized ovum into the uterine wall. (50, 51) The rapid rise in hCG serum levels after conception allows for the use of hCG as an early biomarker of pregnancy.

Physiologically, ovarian hCG maintains the corpus luteum, thereby allowing synthesis of progesterone and estrogens that support the endometrium. As uncomplicated pregnancies progress, the placenta assumes the production of these hormones. Serum hCG levels increase to a peak concentration, then decrease and plateau. hCG circulates as the intact molecule in the serum of normal women with uncomplicated pregnancy. Several hCG variants are differentially expressed over the course of pregnancy. Ovarian hCG promotes progesterone production to the endometrium in the first 3-4 weeks after embryo implantation. Trophoblasts produce predominantly hyperglycosylated hCG, which contains extended sialoglycan sidechains and acts as a crucial paracrine factor for promoting implantation and placental development. Hyperglycosylated hCG accounts for about 80% and 50% of circulating total hCG in week 3 and week 4 respectively. Intact hCG production picks up rapidly after and reaches 80% of total circulating hCG in 2nd trimester. (52) Detectable amounts (>5mIU/mL) are present in the serum 8 – 11 days after conception, which is the third week of pregnancy as measured from the last menstrual period. (53) hCG becomes detectable in urine 1 to 3 days later, although the timing of urine hCG detection is highly variable. (54) hCG concentrations in approximately half of pregnant women reach 25 mIU/mL on the first day of their missed period. In normal pregnancy, hCG concentration rapidly increases, doubling about every 2 days during the first four weeks of pregnancy, reaching a peak at 8 – 10 weeks around 100,000 mIU/mL. (55) hCG levels are about 30 – 50% higher for twins compared to single pregnancy. (56) circulating hCG is degraded and cleared by the kidney and appears in urine mainly as β-core fragment (hCGβcf). (57) Extremely wide hCG levels are found during pregnancy, especially in the first trimester due to inter-individual variations, differences in methods used for dating gestational age as well as the sensitivity and reactivity to hCG variants of the analytical method used. (58)

hCG belongs to a glycoprotein superfamily whose members all consist of two subunits, alpha and beta. The alpha subunit is identical among the superfamily, including luteinizing hormone (LH), follicle stimulating hormone (FSH), and thyroid stimulating hormone (TSH), while the beta subunit determines the biological and immunochemical specificity of these hormones. (59) Since the immunogenicity and epitope specificities of anti-hCG antibodies have been thoroughly investigated and mapped to the hCG molecule in detail, current FDA-cleared hCG tests are designed so that the cross-reactivity with the analogue hormones described above should be minimal, however, there still exists a test-to-test variability. (57, 58, 60)

POCT has advantages and limitations when compared to laboratory tests. (TABLE 3) POCT urine hCG tests take 3 – 5 minutes to develop a positive result, and 3 – 15 minutes for a negative result. Whole blood (WB) POCT hCG takes 10 – 14 minutes, while the central laboratory has analytical times of 9 – 25 minutes. Laboratory turnaround time is longer than this because it is a combination of analytical time and processing time, including time for the sample to clot and time for centrifugation. This is the inherent advantage of POCT which utilizes native urine or whole blood samples that do not require processing. Faster stat laboratory hCG methods for plasma/serum samples are available on some laboratory analyzers (ranging between 9 - 10 minutes). These turnaround times are comparable to the whole blood POCT hCG analysis times without considering the added processing time required for laboratory samples. In a recent prospective observational study (n=265), the use of WB to replace urine as a specimen for a rapid hCG bedside pregnancy test resulted in a mean time reduction of 21 minutes in the emergency department (ED) (95% CI 16 – 25 min, p<0.001). There was a 99.6% concordance between WB and urine test results. (61) However, WB is not an approved
specimen for all rapid hCG test kits and off-label use as a modified method with a WB specimen is not recommended. (62)

One advantage of POCT is its portability. POCT kits can be carried to patient rooms in a clinic, by visiting nurses to a patient’s home, and even deployed in military field hospitals and other community settings. This is contrasted with the dedicated space required for laboratory instrumentation. Although, some POCT devices and readers require access to a power source and may have downloaders that require fixed space on a countertop.

Laboratory methods are now mostly automated, with walk-away capabilities. POCT kits without readers are manual, visually interpreted tests that require timing and manual recording of results in the patient’s medical record. While some POCT devices and readers have electronic interface capabilities, manual documentation risks transcriptional errors. POCT devices and readers require minimal maintenance beyond cleaning and disinfection while laboratory instrumentation requires ongoing preventive maintenance and repairs. Urine POCT hCG methods, however, only provide qualitative results with cutoff concentrations for positivity in the 20 – 25 mIU/mL range depending on the manufacturer, some with even higher cutoffs. There are a few whole blood hCG POCT methods that report quantitative results with positive cutoffs of 5 -10 mIU/mL, although these methods have limited upper reportable ranges of 1250 -2000 mIU/mL and POCT whole blood samples cannot be diluted to extend the range. Laboratory methods, on the other hand, report quantitative results, have wide analytical measurement ranges that can be extended by sample dilution to higher reportable results. Clinicians can follow patient trends over time when laboratory results are analyzed by the same method. POCT methods, however, do not necessarily match laboratory results, because the two methods are different. These steps in the core laboratory give more confidence and reliability in the ensuing results.

Urine POCT hCG is considered waived testing under the Clinical Laboratory Improvement Amendments of 1988 (CLIA) law. CLIA waived testing has minimum documentation requirements. Staff only need to follow manufacturer’s directions, pay a biennial CLIA certificate fee, and agree to be inspected. This allows the test to be conducted in a wide variety of settings under physician or nursing supervision without laboratory involvement. Laboratory hCG and whole blood POCT hCG are CLIA moderate complexity which requires method validation, operator training and competency, daily quality control, participation in a proficiency testing program, biennial inspection, and other quality requirements.

Many variables should be evaluated when considering POCT hCG testing. Operators can perform the tests reliably with minimal training, while laboratory instrumentation requires technical staff with laboratory experience to operate, troubleshoot and maintain the equipment. POCT kits can be stored at room temperature, depending on the manufacturer, while laboratory reagents and controls require refrigeration and freezers. POCT can be utilized in areas with low test volumes because the tests are unit-use, single packaged kits. This is an advantage for those settings with occasional testing requirements but can become a limitation as the demand for testing grows. The POCT operator must perform one test at a time which can slow turnaround of results if multiple samples need to be tested at the same time, unless the optical reader can process multiple tests at the same time. Laboratory analyzers, on the other hand, have higher sample throughput and can analyze multiple tests simultaneously. Laboratory analyzers utilize bulk reagents that can produce more tests at lower cost compared to POCT that are singly packaged. This appears to raise the unit cost of POCT tests.

Thus, POCT hCG has a number of features which allow for rapid test results in a variety of clinical settings. However, urine POCT is limited to qualitative results and the available quantitative whole blood POCT has limited reportable ranges.

In what clinical scenarios should hCG POCT be considered? Is whole blood hCG POCT preferred over urine hCG POCT in patient care? Compared to laboratory-performed hCG testing, routine urinary hCG bedside testing with POCT devices may reduce the test TAT. hCG POCT should be considered for the rapid determination of pregnancy status in conjunction
with ultrasound, symptoms and medical history. However, location of pregnancy, intrauterine versus ectopic, cannot be
determined from POCT alone. The wait time for urine collection often offsets the gain in the testing turnaround time
reduction. WB hCG POCT is more invasive than urine but is not affected by urine dilution or dependent upon the ability to
spontaneously void. Thus, WB hCG may be useful when the patient is unwilling to provide a urine sample or unable to
urinate due to trauma, an acute presentation or painful condition making urine collection problematic. Weighing the
costs associated with potential harms due to an undiagnosed pregnancy into consideration, accurate perioperative and
ED hCG POCT can be advantageous.

Females of childbearing age frequently present to EDs, urgent care centers, physician offices, clinics and other
care providers with symptoms of pregnancy or other clinical conditions that may be pregnancy related, such as
abdominal pain, vaginal bleeding, syncope or shock. Qualitative and quantitative analysis of hCG in urine or blood
(whole blood, serum and plasma) is used for the early detection of pregnancy.(63-65) Qualitative (yes/no) results are
useful for detecting pregnancy prior to exposing the patient to radiation, teratogenic medications, and surgical
procedures that can harm the fetus. Clinicians need to determine the pregnancy status of their patients in order to
make timely treatment decisions, for teratogenic medications are contraindicated and diagnostic imaging that exposes a
developing fetus to radiation should be avoided during pregnancy. However, the menstrual history alone is not reliable,
for this reason, a rapid hCG test at the point-of-care may be deployed. (66)

Improving patient flow can shorten ED length of stay (LOS), and shorter LOS is associated with higher patient
satisfaction. One way to reduce ED LOS is the adoption of POCT, enabling rapid delivery of test results through
immediate bedside testing, to reduce test turnaround time and potentially lead to shorter time to decision-making, a
critical operational benchmark in urgent care settings.(67) The significant delay in awaiting the urine collection often
compromises the time-saving advantage of beside urinary POC hCG testing. Blood hCG should be considered when the
patient is unable or unwilling to urinate. Collection of blood samples is more invasive than urine samples, but
phlebotomy may be preferred over patient catheterization when delays in testing and treatment/procedures are
pending the collection of a urine sample. Spontaneous voiding can be complicated by clinical states where the patient
can’t urinate (dehydration, pelvic trauma, and unstable clinical presentation), and blood hCG may be required in these
situations. Few published studies report the actual turnaround time improvement of POC hCG testing in ED settings. One
prospective study (n=498) found that compared to laboratory-performed hCG testing, the adoption of urinary POC hCG
testing in the ED significantly reduced the time to initial report and time to availability on the chart, with mean
differences of 25 minutes (95% CI, 22–27) and 60 minutes (95% CI, 56–64) respectively.(68) White et al. conducted a
prospective analysis of the laboratory process improvement intervention using LEAN methodologies in a teaching
hospital ED with an on-site POC “kiosk” laboratory. The results showed that the median laboratory turnaround time
(TAT) decreased in five of the six monitored tests, including the shortening of urine hCG TAT by 10 minutes following the
Lean intervention.(69)

The reduction in TAT of a given test does not always translate to an improved patient flow in ED. Only if the
shortened TAT also leads to a significant improvement of result time to clinical decision-making, will POCT become a
cost-saving measure.(67) Nonetheless, hCG bedside testing can rapidly rule out pregnancy to prevent unnecessary
exposure to medical treatments associated with harmful effects on the fetuses and/or pregnant women.(70) To this end,
the accuracy of rapid hCG testing is of utmost importance to realize the potential cost savings.

Results from hCG or any other diagnostic test should be used and interpreted only in the context of the overall
clinical picture.(71, 72) hCG test results should be interpreted in conjunction with other data (history, symptoms, clinical
impressions, and results of other tests). If blood hCG is inconsistent with clinical evidence, results should be confirmed
by an alternative method – this may include a qualitative urine hCG test. If urine and blood hCG results are discrepant, further workup is required (for example, ultrasound or laboratory evaluation of interfering substances).

Ectopic pregnancy is a common and serious problem, with significant morbidity and high risk of maternal death. hCG POCT (or quantitative hCG) should be used to screen for pregnancy in every sexually active, reproductive-aged woman who presents to the ED or outpatient settings with abdominal pain or vaginal bleeding, regardless of whether she is currently using contraception. Transvaginal ultrasound is to follow for the confirmation of pregnancy in patients with stable condition. Indeterminate ultrasound results may be clarified by quantitative measurement (single or serial) of the serum β-hCG concentration. There has been increasing recognition that the previously accepted threshold for hCG in evaluation of ectopic pregnancy led to misdiagnosis. This has led to the discriminatory threshold for hCG in evaluation of ectopic pregnancy increasing to 3500 mIU/mL in order to be “conservatively high” to avoid treating a potentially viable intrauterine pregnancy with methotrexate. The discriminatory threshold is the hCG level that distinguishes patients with intrauterine pregnancies in whom a gestational sac can be seen from those in whom it cannot. However, the upper detection limits may differ among instruments, and dilution may be needed for platforms with reportable ranges < 3500 mIU/mL. Ectopic pregnancy cannot be distinguished from normal pregnancy by hCG measurements alone. A discriminatory zone based on hCG alone cannot exclude an ectopic pregnancy without further testing (serial hCG or confirm with transvaginal ultrasound) in order to safely send a patient home versus giving methotrexate.

Demonstration of the normal doubling of hCG levels over 48 hours suggests a diagnosis of fetal viability but it does not rule out ectopic pregnancy, and a rate of rising hCG concentration that fails to reach a 50% increase over initial hCG level or plateaus within 2 days, suggests a failing or ectopic pregnancy. Falling levels confirm nonviability of the fetus but do not rule out ectopic pregnancy. Ectopic pregnancy may resolve spontaneously through tubal abortion. The risk of tubal rupture is similar across a wide range of hCG values (10 – 190,000 mIU/mL). Quantitative hCG levels are used not only to diagnose ectopic pregnancies but to guide treatment, as a quantitative hCG level greater than 5000 mIU/mL is considered a relative contraindication to use of methotrexate for treatment of ectopic pregnancies due to its higher risk of failure in these patients. Once methotrexate is used, the quantitative hCG needs to be measured serially to ensure the pregnancy is resolving and guide further treatment either through continued methotrexate administration or surgery. Thus, interpretation of hCG results in the evaluation of normal and abnormal pregnancy should be done in conjunction with clinical and sonographic findings to arrive at the correct diagnosis.

Though POCT is commonly deployed in the diagnosis of pregnancy, a number of conditions other than pregnancy, including trophoblastic disease and certain non-trophoblastic neoplasms such as testicular tumors, prostate cancer, breast cancer and lung cancer, are known to cause highly elevated levels of hCG. Therefore, hCG can also be considered a tumor marker for monitoring the treatment outcomes of hCG-secreting malignancies. However, the FDA has not approved hCG for use as a tumor marker. hCG alone should not be used for the diagnosis of these neoplastic states. There have been reports of patients receiving unnecessary medical treatment and surgery, including chemotherapy and hysterectomy, when hCG results were used alone in the diagnosis of abnormal conditions.

The presence of hCG in a serum or urine specimen should not be used independent of clinical correlation to diagnose pregnancy unless these conditions have been ruled out.

Low levels of hCG can occur in healthy, nonpregnant subjects. Post-menopausal patients may have low levels of hCG unrelated to pregnancy. In postmenopausal women, hCG is produced alongside the structurally similar luteinizing hormone (LH) by the anterior pituitary. LH shares a common alpha subunit with hCG and has 50% of the biological homology of hCG produced in pregnant women. Clinicians and laboratorians need to be aware of normal pituitary secretion of hCG in postmenopausal women (<14 IU/L). Because hCG levels double every 48 hours in a normal
pregnancy, patients with very low levels of hCG should be resampled and retested after 48 hours or tested with an alternate hCG method (that utilizes a different antibody). (63, 65, 84)

Other sources of hCG, such as an hCG injection, can result in detectable levels of hCG in blood or urine, which can cause diagnostic confusion in patients being screened for pregnancy. (85) Chronic renal failure (due to decreased clearance) (86) and passive transfer of hCG to plasma recipients from unknowingly pregnant blood donors (87) have also been reported as sources of elevated hCG in nonpregnant women. There have been rare cases of individuals with familial hCG syndrome—an idiopathic elevation of hCG that has been shown to be inherited in an autosomal manner. The genetic basis of this rare syndrome has not been elucidated but is characterized by persistent low hCG levels in multiple family members with no other underlying etiology. (88)

**Do quantitative hCG results enhance clinical interpretation compared to qualitative hCG at the point-of-care?**

Quantitative hCG results at the point-of-care are only available from a few WB methods. WB POCT hCG is of limited clinical value in serial testing and trending hCG results early after conception due to the narrow reportable range of POCT methods and the rapid rise of hCG levels after implantation. If hCG trends are being monitored, POCT results must be followed by the same analytic methodology. hCG results will vary between methods due to differences in antibodies and test affinity to various forms of hCG.

Urine hCG tests have reported 99% accuracy for the detection of pregnancy as early as the same day of a missed menstrual period with published cutoffs of 20 – 50 mIU/mL. However, the actual performance is lower due to variability among tests to recognize the different forms of hCG present in urine. (89) Urine dilution is also a concern, and all urine hCG tests provide qualitative results.

Blood hCG concentrations at or around the time of the missed menstrual period (the 4th completed week since the last menstrual period) report a median serum concentration ranging from 205 mIU/mL (3 – 7340 mIU/mL) (90) to 560 mIU/mL (6 – 19,950 mIU/mL). (91) Cut-off concentrations discriminating premenopausal, perimenopausal, postmenopausal, early normal pregnancy, pregnancy in the 2nd-3rd trimester, or gestational trophoblastic neoplasia may vary depending on the clinical category and manufacturer assay. Quantitative WB POCT hCG methods with an upper reportable range of 1500 – 2000 mIU/mL would be of limited value when trending hCG levels over time, depending on the patient’s initial hCG level, since blood hCG concentration rises rapidly in the first several days of pregnancy.

Early pregnancy loss or spontaneous abortion occurs in about 22% of clinically unrecognized pregnancies and 31% of pregnancies overall. (92, 93) Because of the high sensitivity of hCG assays, women with serum/plasma hCG results above the upper reference limit during the initial days after conception may generate negative results in subsequent samples due to natural termination of pregnancy.

Most hCG assays utilize calibrators that are traceable to the World Health Organization’s standard. However, the use of quantitative results to follow patients over time must use the same methodology, since POCT and laboratory hCG immunoassays may generate variable results due to differences in antibody affinity for various forms and fragments of hCG in circulation. (89, 94) Contradictory results between POCT and laboratory testing should be closely investigated for the cause of the discrepancies to mitigate possible clinical ramifications. In summary, quantitative WB hCG is of limited clinical value due to the narrow reportable range of POCT methods and the variability of results between different methods, so quantitative POCT has not been shown to be superior to qualitative POCT hCG methods.

**What is the stability of hCG in urine?**
The stability of hCG in urine at different temperatures has showed significant variation.

Fátima Reis et al. and some earlier studies have reported urinary hCG is stable at temperatures of 2-8 degrees C for up to 48 hours and temperatures around -20 degrees C for longer periods, extending to over 3 months.(95) For point-of-care urinary hCG testing, long-term storage of urine specimens is rarely needed. And it was reported that hCG immunoreactivity in urine was not significantly affected when stored below 10 degrees C for up to 5 days.(96) However, the urine samples should not stand at ambient temperature for an extended period before testing to prevent the deterioration of urinary hCG.

On the other hand, quality control materials and population study specimens may need long-term storage. Lempiäinen et al. found variability in urinary hCG stability for samples stored at -20°C from 3-10 months, and the magnitude of hCG loss was correlated with high sample urea while adding 5-10% glycerol or storing at -80°C mitigated the activity loss. It is noteworthy that the stability of different isoforms of urinary hCG vary and may account for the variations of observed hCG stability.(97)

When should measurement of urine specific gravity be considered?

Specific gravity measurement should be considered on urine samples when measuring hCG to determine viable pregnancy.

False-negative urine hCG results can occur due to dilute urine, very early pregnancy stage, or very high hCG concentrations (hook-effect). A dilute urine specimen (specific gravity of less than 1.007) may not contain a high enough hCG concentration to produce a positive result despite the presence of a viable, intrauterine pregnancy. To mitigate the risk of a false negative urine hCG result, manufacturers recommend analyzing the urine specific gravity when conducting qualitative or quantitative urine hCG tests. For dilute urine samples, another sample should be collected, ideally the first-morning void, or the urine hCG results should be confirmed using a blood sample.(98-100)

Does proteinuria and other interferents affect urine pregnancy tests?

There are a few case reports in the literature of false-positive urine pregnancy tests as a result of proteinuria/other interferents. Heterophile antibodies and other interferences are the primary causes for false positive/negative urine hCG tests. However, urine tests for hCG may not give appropriate results based on the alkalinity or acidity of the urine specimen following medication/drug use or if there are very high levels of hCG such as in trophoblastic diseases.(101, 102) In septic shock, transient passage of interferents into the urine have been shown to result in falsely elevated urine hCG.(103) False-positive urine pregnancy tests have been seen in patients with nephrotic range proteinuria, systemic lupus erythematosus, ovarian failure, elevated gonadotrophins and tubo-ovarian abscess. In addition, a false-positive urine pregnancy test has been reported in a 28-year-old woman with a history of tubal ligation, who had a delayed diagnosis of obstructive pyelonephritis due to renal calculus. Marzinke, et al., reported a false positive urine pregnancy test in a patient with membranoproliferative glomerulonephritis type I.(104-106) Proteinuria and rheumatoid factor have been shown to produce false positive urine hCG.(106)

There are anecdotal reports of false positive urine hCG results in patients with UTI, in addition to warnings on POC hCG kit package inserts, and inaccurate statements on health websites.(107, 108) However, there is scant data in the scientific literature that supports such claims. Mitchell et al. performed a study using 95 urine samples testing positive on urine culture with Gram-positive bacteria.(108) Only five of the 95 samples tested positive for hCG. Of these, two were from pregnant women and three were from cancer patients. The authors concluded that UTIs do not
cause false positive results. This study had some limitations in that data were collected from a single site, used a single test kit, and the study was not designed to evaluate the potential for false negative results. Nevertheless, the results seem to be compelling and point towards the inability of bacterial isolates to produce false positive results with these assay kits.

There are also reports in the literature dating back several decades of an hCG-like molecule expressed on multiple bacterial strains. The authors conclude that “cancer-associated” bacteria can synthesize a protein that mimics hCG. Mitchell et al. refer to these studies and suggest that the bacterial hCG molecule is likely an hCG variant derived from neoplastic disease, rather than from bacteria.

Overall, there are limited published studies on the role of interference by UTIs. The evidence does not suggest that UTIs can cause false positive results on urine based hCG POCT assays. One should be cautious when laboratory findings are inconsistent with the clinical scenario, and exercise prudence by utilizing different methods/assays as test platforms and/or other sample types, e.g. serum, for confirmation of ambiguous results.

Are qualitative POC hCG devices susceptible to false negative results due to antigen excess?

False negative qualitative POC hCG results may occur when testing urine specimens containing elevated concentrations of intact hCG or hCG variants.

Urine hCG devices use a sandwich immunoassay format consisting of an immobilized capture antibody and a soluble, labeled detector antibody. In the presence of hCG, antibody-hCG-antibody “sandwich” accumulates at the test line, forming a visible band. Excess labeled antibody flows past the test line and accumulates at the control line, confirming validity of the test result. When hCG is absent, labeled antibody accumulates at the control line only, indicating a negative result. Any substance that prevents formation of the test line when hCG is present, creates a test line when hCG is absent or prevents the formation of a control line may lead to misclassification of pregnancy status and adverse patient outcomes.

Qualitative POC hCG devices are designed to accommodate intact hCG concentrations typically observed in normal pregnancy (peak around 10 – 12 weeks with levels between 25,700 – 288,000 mIU/mL) and are generally not susceptible to the hook effect when used to aid in the detection of pregnancy. Device manufacturers often evaluate this threshold and indicate in the package insert the highest intact hCG concentration confirmed to generate positive results. Two commercially available devices indicate resistance to the hook effect up to 500,000 mIU/mL, although it is important to note that this threshold varies by manufacturer and a careful reading of the package insert is recommended.

Several types of tumors produce an elevated level of the hCG, the high-level free β-subunit is especially related to the aggressiveness of the malignancies. hCG level can reach > 3,000,000 mIU/ml in patients with the complete hydatidiform mole (advanced molar pregnancy) or disseminated gestational choriocarcinoma. The amounts of hCG in the specimens of these patients greatly exceed and thus saturate the antibodies present in POC hCG devices and cause false-negative results due to the Hook effect. While gestational choriocarcinoma is rare, hydatidiform mole occurs in approximately 1 per 1,000 pregnancies in the United States. Hydatidiform mole can cause the death of the embryo as a result of delayed diagnosis. Choriocarcinoma can evolve very rapidly and result in death if left without treatment.

Because the levels of elevation of hCG in hCG-secreting tumors vary widely, hCG testing should not be used solely in the initial assessment of the suspected cases. In fact, the false-negative urine/serum hCG results caused by the Hook effect is a well-known problem that complicates the diagnosis of these two types of gestational trophoblastic diseases, as evidence by the sporadic case reports in the literature. (111-113)
False negative results may also occur when variant forms of hCG are present at sufficiently high concentrations to saturate the hCG antibodies in the test device and the antibodies demonstrate some binding affinity for the hCG variant(s). Excess hCG variant prevents the device antibodies from binding to intact hCG present in the sample, even though the intact hCG concentration may be well below the device's indicated upper limit of detection. The variant hook effect is most commonly caused by hCG beta core fragment (hCGβcf), a product of proteolytic digestion of the hCGβ subunit formed during renal excretion and present only in urine. While the variant hook effect due to hCGβcf may be observed at any point in pregnancy, it is most likely when urine hCGβcf concentrations are highest, typically coinciding with the peak of plasma hCG concentrations observed between weeks 8-12.(114) hCGβcf is present only in urine and has not been demonstrated to interfere with hCG measurement in serum, plasma or whole blood specimens. Although urine hCGβcf concentrations are typically highest between 10 - 16 weeks gestation, excretion patterns are variable and false negative results due to excess hCGβcf may be encountered at any point in pregnancy.(115, 116) Commonly used qualitative POC hCG devices vary in their susceptibility to hCGβcf interference as some devices generated strong positive results, others generated weak positive results and some generated clear negative results when used to evaluate a series of urine specimens with known intact hCG and hCGβcf concentrations.(117) Some devices have been reformulated, resulting in improved performance in specimens containing high concentrations of hCGβcf.(118, 119)

Recommendations to reduce false negative hCG hook effect:

- Maximal intact hCG concentrations listed in the device package insert can be verified by clinical laboratory personnel
- If a false negative result due to the hook effect is suspected in the clinical conditions discussed above. The urine specimen should be diluted using an hCG-free material and retested. A positive result upon dilution confirms the hook effect.
- When clinically appropriate, quantitative hCG measurement should be performed in serum or plasma specimens to avoid hCGβcf interference. If urine testing is required, a device with minimal susceptibility to hCGβcf variant effect should be used.

Does excess Biotin intake affect hCG measurement on POCT device? Interference due to biotin supplementation should be considered if repeat testing on a patient sample generates an invalid test result. A quantitative hCG measurement in plasma/serum or alternatively a qualitative urine hCG device that is not subject to biotin interference should be considered. In non-emergent situations, a second urine specimen may be collected after discontinuation of biotin supplementation for 24 h.

Biotin has recently become a focus of concern as an immunoassay interferent due to its increased popularity as a vitamin supplement, with several case reports and research studies demonstrating that excess biotin in patient samples can interfere with assays reliant on streptavidin-biotin interaction.(120, 121) A recently published AACC guidance document on biotin interference lists immunoassays impacted by biotin; hCG can be falsely decreased in the presence of excess biotin on two commonly used immunoassay instruments.(121)

Approximately 50% of biotin is secreted in urine unchanged. In general, 20% of ingested biotin dose is excreted in urine within 4h. The rate of biotin excretion varies between 3 to 40hrs at low doses and 7.8 to 18.8 hours at high dose.(122, 123) Therefore, qualitative urine hCG testing devices are subject to biotin interference in individuals taking dietary biotin supplements.

In one study the authors investigated potential interference of biotin in qualitative POCT for hCG both in vitro and in vivo and showed that excess biotin produced invalid results on the QuickVue urine hCG device by preventing
streptavidin-biotin interaction at the control line. However, other devices (Alere 20, Alere 25, Icon 20, Osom, QuPID, and SureVue) were not affected. (124) The absence of a control line during patient testing has the potential to delay care due to the generation of an invalid test result and lead to additional unnecessary testing.

If biotin interference is suspected, the ordering clinician should be contacted to inquire about the patient's biotin use. This line of inquiry is limited by the fact that many patients do not include over-the-counter supplements when asked to recite their medication list and clinicians may be unaware of their patients' biotin use. Regardless of whether biotin supplementation has been confirmed, laboratorians should suggest quantitative hCG measurement in plasma or serum, provided the analytical platform is not also subject to biotin interference. (120) In non-emergent situations, a second urine specimen may be collected after discontinuation of biotin supplementation for 24 h, but the timing is assay and dose dependent. As biotin exhibits a short half-life, serum and urine concentrations should be below the threshold required to generate an invalid result within 24 h of supplement discontinuation. (125) As an alternative strategy, laboratorians may choose a qualitative urine hCG device that is not affected by biotin supplementation.

POINT-OF-CARE PREMATURE RUPTURE OF MEMBRANES (PROM) TESTING

What tests are available to predict prelabor rupture of membranes (PROM)?

Conventionally, testing for PROM includes observed pooling of amniotic fluid in vagina / posteria fornix, pH testing of the fluid (pH of vaginal fluid is 3.8-4.5 compares to amniotic fluid is 7.1-7.3), and microscopic examination of dried vaginal fluid (ferning pattern is suggestive of PROM). More recently, multiple commercial test kits are also available to aid the diagnosis of PROM. These include Actim, Amnisure and ROM Plus in the US and / or Canada. Fetal fibronectin (fFN) should not be used for PROM diagnosis. Digital examination should be avoided to reduce risk of infection.

According to the American College of Obstetricians and Gynecologists (126), PROM is defined as membrane ruptures prior to the onset of labor. PROM that occurs before 37 weeks of gestation is considered preterm PROM. It complicates ~2-3% of pregnancies in the US. At term (greater than 37 weeks of gestation), PROM occurs in ~8% of pregnancies in the US and are often followed by prompt delivery. PROM management depends largely on gestational age of the fetus. The most significant complication to the fetus relates to prematurity. As to the mother, intrauterine infection risk increases with duration of membrane rupture. Therefore, the ability to accurately diagnose PROM is crucial in clinical management of mother and fetus.

As noted above, vaginal pooling, pH testing and ferning are often used to confirm the diagnosis of PROM (126). Commercially available test kits (rapid lateral flow immunoassay) such as Actim, Amnisure and ROM Plus are also available. (127-129) Amnisure uses a biomarker called placental-α-microglobulin-1 (PAMG-1). Actim and ROM Plus use a biomarker called placental protein 12 (PP12) / insulin-like growth factor binding protein-1 (IGFBP-1), which are the same protein. (130) These proteins are present in substantially higher concentration in the amniotic fluid than in cervico-vaginal secretions with intact fetal membrane or maternal serum, making them ideal markers to detect ROM. (55, 131) Unlike Actim and Amnisure test kits, the ROM Plus test kit also uses a second biomarker, α-fetoprotein (AFP). AFP level increases in amniotic fluid as gestational age advances but decreases during third trimester.

Historically, transabdominal instillation of indigo carmine dye, followed by the passage of the dye into a vaginal tampon was considered the gold standard. (126) However, given the invasiveness and complexity of the procedure, this is not commonly used in clinical practice.
fFN is generally not recommended to predict ROM due to low specificity. False positive results have been associated with intra-amniotic infection and spontaneous preterm delivery. The sensitivity and specificity of fFN for PROM were 94.5% and 89.1% respectively. Limited studies have shown that fFN has better sensitivity and specificity than ferning or pH testing alone, though not as good as test device like Amnisure. 

When should PROM testing be performed?

Most instances of PROM can be diagnosed based on physical examination, clinical presentation and patient history. PROM testing using commercial kits alone is not recommended without clinical signs of ROM, such as leakage of amniotic fluid from the cervical opening. As with other tests, the test result must be interpreted in the context of the patient's clinical presentation. Unfortunately, there have been reports of misuse resulting in death and health complications for fetus and/or mother.

PROM management varies depending on gestational age and fetal presentation. PROM is typically of minimal risk compared to preterm PROM. Risks associated with preterm PROM are significantly higher with gestational age < 34 weeks. Currently, there is insufficient literature for the committee to recommend whether testing for PROM is clinically beneficial for ≥ 37 weeks pregnancy. In the US, not all the commercial test kits are approved for PROM testing for any gestational age by the U.S. Food and Drug Administration (FDA). For example, Actim PROM is only approved for use with gestational age > 29 weeks, whereas Amnisure and ROM Plus can be used for all gestational age. However, there is little published literature on these tests’ accuracy for gestational age < 20 weeks. Healthcare providers need to be aware of such limitations to assure regulatory compliance and proven test performance for use with all stages of pregnancy.

In 2018, the U.S. FDA released alerts to providers about risks associated with improper use of PROM tests. Briefly, diagnosis of PROM and subsequent management of patient should not be made solely based on test results – over-reliance of test result can cause patient harm. As with other tests, false positives or false negatives can occur. Testing in the absence of clinical signs of PROM can be particularly misleading. In women at term, positive Amnisure test results had been reported without clinical sign of ROM and the positivity rate is higher with patients in active labor. While this is not considered PROM because these women were in labor, the studies suggest potential leakage of amniotic fluid or small amniotic protein as delivery date approaches. Therefore, PROM testing should only be performed if clinically indicated and must be interpreted in conjunction with clinical assessments to diagnose PROM.

How does the performance of commercial test kits (e.g., Actim, Amnisure, ROM Plus) compare to traditional testing (pH, pooling, ferning test)?

Multiple studies have shown that performance of commercial test kits (eg Actim, Amnisure, ROM Plus) is equal to or surpasses the performance of the standard clinical assessment (SCA) which includes traditional testing methods (vaginal pooling, pH testing and ferning). The committee recommends the use of commercial test kits for aiding in the diagnosis of PROM in women with suggestive symptoms.

Multiple prospective, observational studies have compared the performance of rapid commercial test kits to (a) the performance of SCA, and (b) clinical outcomes (48 hours follow up and chart review) and are summarized in TABLE 4. In most of these studies, final PROM diagnosis was adjudicated by a physician who was blinded to the immunoassay test results. Most studies did not define the exact criteria of PROM, so the final diagnosis likely varied depending on the physician. However, some of the following criteria were taken into consideration in many studies: the SCA (some or all three), clinical signs of fetal distress, latency period (varied between 48 hours to 7 days), chorioamnionitis, etc. Esplin et
al.(138) determined a sensitivity, specificity, PPV, and NPV of 91.7%, 97%, 94.8% and 95.1% for the ROM Plus, 93.4%, 95.0%, 91.9% and 96.0% for the Amnisure and 95.0%, 98.5%, 87.4%, 97.1% by SCA compared to the primary outcome. The performance characteristics of the two kits was statistically equivalent and there was no significant difference when compared to the SCA methods.

Albayrak et al.(139) also determined that there was no significant difference in the performance of Actim, Amnisure and the SCA methods compared to the final clinical diagnosis as determined by medical records after delivery. The sensitivities, specificities, PPV, NPV and accuracy as follows: Actim commercial test kit (89.8%, 97.5%, 97.5%, 89.5%, and 93%), Amnisure test kit (94.3%, 97.5%, 97.6%, 93.9%, 95%) and SCA (88.6% 94.9% 95.1% 88.2% and 91%). One notable difference in this study, is that unlike the Esplin study(138), the SCA was considered positive when two of the following three methods (vaginal pooling, pH testing or ferning) yielded a positive result.(139) Similarly designed studies have determined that the Actim, Amnisure and ROM Plus commercial test kits outperform SCA with fewer false negatives, resulting in improved sensitivities and an increase in NPV.(140-142)

Multiple prospective studies have also compared the performance of the rapid commercial tests to just pH testing for amniotic fluid or the ferning test. In these studies, the commercial test kits all outperformed pH testing or ferning in direct comparisons.(140, 142, 143) In a recent study of patients undergoing evaluation for preterm PROM, the Actim and Amnisure test kits outperformed the pH test with sensitivities and specificities and PPV >90% and accuracies ranging from 94-98%.(144)

What might affect the performance of PROM tests?

Extended membrane rupture and minimal residual fluid can cause false negative results. The presence of semen, cervical mucus and hepatitis C can cause false positive ferning results. Blood, semen, urine, bacterial vaginitis, cervicitis, trichomonas and anti-septic agents, have also been reported to cause false positive pH/nitrazine results.(140, 145, 146) Blood also affects the performance of different commercially available PROM tests to various degrees.(147, 148)

Ramsaurer et al.(147) investigated the impact of vaginal bleeding on the diagnostic accuracy of the Actim and AmniSure tests in patients with unknown membrane status. Cases that were identified as contaminated with blood had an increased occurrence of false positives and non-evaluable results (2% in Amnisure and 11% in Actim).

Bushman et al.(148) performed an in vitro study using whole blood samples spiked with varying concentrations of amniotic fluid proteins to evaluate the effect of blood contamination on the performance characteristics of three commercially available test kits, Actim, Amnisure and ROM Plus. Blood contamination had no effect on the specificity or PPV value of each assay, which remained 100%. In this study, blood contamination increased the number of false negatives in all three assays, resulting in a reduction of sensitivity, NPV and overall accuracy (Actim: 56.7%, 69.8% and 78.3%; Amnisure: 61.0%, 72.3% and 80.7%; ROM Plus: 96.7%, 96.7% and 97.9%). ROM plus was able to detect amniotic proteins significantly better than the Amnisure or Actim PROM in this study. These discrepancies of the Ramsaurer(147) and the Bushman(148) studies may be due to different cutoffs (analytical sensitivities) of commercial kits for PAMG-1/IGFBP-1/PP12 or AFP levels (TABLE 5), and the different experimental design (real patients vs in vitro spiking samples) utilized.

Is the use of PROM test cost effective?

There are currently limited studies in evaluating the cost effectiveness of PROM testing. Given that prevalence of PROM is ~2-3% preterm and 8% of all term pregnancies in the US(126), concern for PROM is not uncommon for pregnant
women. Timely diagnosis is key in ensuring the health of mother and fetus. Accurate rule out of PROM can reduce unnecessary emergent visits to the hospital/clinic or patient transfer to higher level care facility, and therefore reduce health care cost. *(149, 150)*  

According to Ferro et al.(149), at a large hospital in New Jersey, out of 1250 unscheduled visits, 68 had a primary concern of suspected PROM. Of which, 58 were discharged with PROM rule out. Therefore, if there is a safe and accurate way for pregnant women to determine the presence of PROM, the emergent visits could be avoided. Although there is no at-home device for PROM in the US currently, an at-home pH indicator has been trialed to differentiate amniotic fluid leakage.(151, 152) However, the use of device without proper clinical evaluation can result in false negative diagnosis and presents a significant safety concern.  

In Echebiri et al.(150), the authors use computer modelling and showed that the use of PROM testing can rule out 39-45% of the suspected cases. In resource-limited settings, this can lead to fewer patient transfers to facilities with comprehensive management capabilities for a higher level of care. According to Echebiri et al.(150), a single patient transfer costs $800 to $8800 in the US. This is significantly higher than Medicare reimbursement cost per test (less than $100 USD). Therefore, it presents a significant cost saving, as well as improved patient convenience.  

**POINT-OF-CARE FETAL SCALP LACTATE/CORD BLOOD GAS TESTING**  

**What is the clinical utility of fetal scalp sampling for pH or lactic acid during high-risk deliveries?**  

A fetal blood scalp (FBS) sample taken within one hour prior to birth correlates well with umbilical arterial or venous lactate, pH, and base excess. Lactate values are being used for predicting fetal acidosis and to indicate the need for intervention. Currently available studies indicate that FBS testing has minimal utility in prevention of metabolic acidosis.  

Generally FBS lactate testing is recommended over FBS pH testing for the management of intrapartum pregnancy with abnormal fetal heart rate.(153, 154) An exception was the NICE guideline where both FBS lactate and fetal scalp pH were presented as screening options following an abnormal fetal heart rate, with no preference given to one test over the other.(155) Demaegd H, et. al. recently published a restricted systematic review (156) on whether FBS is useful in detecting and preventing metabolic acidosis when ST-analysis of the fetal ECG (STAN®) is already being used as a second line technique for intrapartum fetal monitoring. According to the authors, only analysis of data coming from one RCT showed some benefit of performing a FBS in the setting of STAN® monitoring when performed in the following specific situations: 1) start of STAN® registration with an abnormal CTG trace, 2) abnormal CTG trace > 60 minutes during the first stage of labor without ST events, 3) poor ECG signal quality in the presence of an intermediary or normal CTG trace.(157-159) FBS identified fetal acidosis in 9.9% of the cases where STAN® monitoring did not indicate intervention. The real benefit of FBS is limited, as only one case of true (neonatal) metabolic acidosis was correctly identified. It was concluded that the introduction of STAN® monitoring resulted in lower-level performance of FBS.  

According to Chandraraharan et. al. a fetus who is exposed to gradually evolving hypoxia would attempt to compensate for the ongoing hypoxic stress by the release of catecholamines (adrenaline and noradrenaline) and consequent peripheral vasoconstriction to divert oxygen from the non-essential organs to the central organs (i.e. centralization) and a progressive increase in the baseline fetal heart rate to obtain fresh oxygenated blood from the placenta.(160) This results in hypoxia for the fetus if this condition remains unchanged for a longer period. Hypoxia induces changes in fetal scalp lactate and pH also. FBS has been found to be unreliable due to multiple variables affecting sample integrity, including alteration of results by the alkaline amniotic fluid(161), and acidic meconium.
containing bile acid. Therefore, the recent NICE guidelines on intrapartum care has concluded that FBS increases the caesarean section rate and operative vaginal delivery rates.

**Does fetal scalp pH or lactic acid above or below a threshold value predict adverse neonatal outcome?**

Fetal scalp lactate identifies the metabolic origin of the fetal acidosis, has a significantly lower failure rate than pH and has a faster turnaround time. The cut-off value for intervention most widely used is a scalp lactate level of 4.8 mmol/L or more. Lactate values between 4.2 and 4.8 mmol/L are considered borderline. A lactate level of 4.2 mmol/L or less is reassuring.

Metabolic acidosis can persist for several hours. It is a predictor of severe neonatal morbidity and mortality. Most clinical trials examining the effect of intrapartum fetal monitoring on neonatal outcome use pH <7.05 and a base deficit >12 mmol/L to define the metabolic acidosis. Incidence of neonatal metabolic acidosis varies between 1.3% and 3.5%.

In the presence of an abnormal fetal heart rate trace, some guidelines recommend fetal scalp blood sampling as a second-line technique to correctly identify hypoxic fetuses. Initially, scalp pH was used to detect fetal acidosis. A pH of >7.25 is normal, whereas a pH of 7.20–7.25 is considered suboptimal and a value of <7.20 is abnormal and requires intervention. Fetal Scalp Lactate identifies more adequately the metabolic origin of the fetal acidosis, has a significantly lower failure rate and the result is known much faster. A fetal scalp blood sample for lactate measurement taken within 60 min prior to birth correlates well with umbilical arterial and venous lactate, pH and base excess measured following birth. However, for the interpretation of scalp lactate values the cut-offs or decision points should be identified. Allen et al. reported that fetal scalp lactate values greater than or equal to 4.2 mmol/L offered the best sensitivity and specificity for detecting clinically important outcomes. In some studies, lactate<5.1 mmol/L was used to indicate absence of acidosis and thus no requirement for intervention; lactate>6.6 mmol/L was used to indicate metabolic acidosis of the fetus and thus requires direct intervention and intermediate lactate concentration between 5.1–6.6 mmol/L indicates reassessment of the fetus and repeat of the lactate measurement after 20–30 minutes. The cut-off value for intervention most widely used is a lactate level of 4.8 mmol/L or more. Lactate values between 4.2 and 4.8 mmol/L are considered borderline. A lactate level of 4.2 mmol/L or less is reassuring.

**Does fetal scalp pH or lactic acid monitoring reduce the rate of caesarian delivery?**

There is insufficient evidence to conclude that use of FBS will reduce the rate of caesarian delivery.

One prospective RCT compared auscultation with CTG and CTG plus FBS of 695 high-risk deliveries where 232 were monitored by auscultation alone, 233 by CTG alone, and 230 by CTG and FBS. When comparing CTG with CTG plus FBS, there was a reduction of CS (18% vs. 11%) in the CTG plus FBS. Data from the systematic review of CTG + FBS trials and observational data from Australian hospitals support a conservative estimate of up to a 40% relative risk reduction in caesarean sections when FBS is added to CTG monitoring. In contrast, a recent Cochrane review concluded that use of FBS as an adjunct to CTG increased instrumented deliveries, reduced fetal acidosis, but did not impact any other fetal outcomes. FBS also poses risks for the fetus included CSF leaks and hemorrhage and sepsis.

**Can fetal scalp pH and fetal scalp lactic acid be used interchangeably to predict fetal acidosis and hypoxia during complicated deliveries?**
Fetal Scalp Lactate is used more than fetal scalp pH for fetal acidosis due to higher success rate and low volume of sample required.

Two randomized clinical trials(169, 170) compared some of the effectiveness and risks of fetal scalp sampling for lactate and pH measurement to assess fetal well-being, after a non-reassuring cardiotocography trace during 3348 deliveries. No statistically significant difference between the two groups were found for neonatal encephalopathy, death, low Apgar score at 5 min, admission to NICU, metabolic acidosis, or mode of birth. Because lactate testing requires less specimen volume than pH, fetal scalp sampling success rate was estimated to be 98.7% and 79.4% for lactate and pH, respectively.(154, 159) A reanalysis of the multicenter trial(174) found that the frequency of total operative interventions was similar but more caesarian deliveries were performed in the lactate group (16.5 vs. 12.4%; RR: 1.33; 95% CI: 1.02–1.74).

What is the advantage of measuring fetal scalp lactate using a small POCT lactate analyzer over measuring pH?

A small volume, approximately 1 μL of blood, is required to measure fetal scalp lactate whereas 40–90 μL of sample is required to measure blood pH. POCT devices for measurement of lactate reduce TAT to a significant level.

For analysis of lactate by POCT strip analyzers, the volume of blood required is only 1 μL, while for pH, either measured on a POCT analyzer or on a blood gas analyzer, much more sample is needed (40–90 μL, depending on the device or cassette used). The use of POCT lactate measurements thus represents an attractive option to obtain quick TAT with a small volume of blood. The disadvantage, of course, is that lactate alone is measured, while a full blood gas including pH and lactate gives lot more data and confidence in data interpretation but needs a larger volume of blood and more training and competency for staff performing the test. The Clinical and Laboratory Standards Institute recommends a storage time of ≤15 minutes for whole blood lactate at room temperature.(175)

What should be done for the validation of small POCT lactate analyzers

For validation of fetal lactate devices, precision analysis and method comparison between POCT devices and laboratory-based blood gas analyzer or plasma lactate is recommended. The total allowable error tolerance should be much less than the biological variability of lactate in the normal range, as studies have shown that biological variability in patients with elevated lactate levels is lower.

Precision analysis may be performed using the CLSI: EP5A Complex Precision Protocol. In addition to the usual validation criteria, desirable precision when testing patients and evaluating them for lactic acid acidosis should be tighter and total allowable error should be half that of normal.(176) Comparison studies should not be done on 2 portable devices but rather between the POC device and a central lab blood gas analyzer or plasma lactate method. The lack of a reference method for lactate analysis is one problem in the validation of a POC device to perform fetal scalp lactate. One investigator used the average of the Roche and Vitros plasma lactate as an internal reference method, as the Roche and Vitros methods were originally calibrated against different comparator methods.(177) Another approach would be to use a blood gas analyzer as the reference method, since many blood gas analyzers have calibration schemes that return similar results across the range of lactate values normally encountered. When comparing blood gas analyzers to commercial plasma lactate assays, the blood gas analyzers exhibit a negative bias at higher lactate values.(177, 178)

What is the clinical utility and cost effectiveness of universal (vs selective) umbilical cord blood gas analysis?

Paired umbilical cord blood gas testing (Arterial and Venous) is recommended either immediately after delivery or from a clamped section of the umbilical cord within 2-3 minutes of delivery to provide appropriate care to newborns at birth,
quality management and training purposes. There is currently insufficient evidence to suggest that universal (every birth) collection of paired blood gas samples is more effective than collection in situations where the infant has low Apgar scores, or poor outcome is suspected.

Umbilical artery blood gas and acid-base analysis provides objective information about fetal metabolic condition, specifically the presence of hypoxic (respiratory) or metabolic acidosis, at the time of birth. Guidelines from the American College of Obstetricians and Gynecologists (ACOG)(168) and National Institute for Health and Care Excellence (NICE)(155) recommend selective testing of umbilical cord blood gas shortly after delivery. NICE guidelines recommend that if the baby is in “poor condition”, then the umbilical cord should be double clamped and paired (arterial and venous) samples should be drawn and sent for blood gas analysis. ACOG guidelines previously recommended that physicians should attempt to obtain paired umbilical cord blood gas samples in circumstances of cesarean delivery for fetal compromise, low 5-minute Apgar score, severe growth restriction, abnormal fetal heart rate tracing, maternal thyroid disease, intrapartum fever, or multifetal gestation (these guidelines have since been withdrawn).(155) In contrast, guidelines from the Society of Obstetricians and Gynecologists of Canada (SOGC)(179) recommend universal (all births) testing of paired umbilical cord blood gas samples drawn either immediately after delivery or from a clamped section of the umbilical cord within 2-3 minutes of delivery. Reasons given for universal sampling include that the information may help provide appropriate care to the newborn at birth, and having blood gas data from all births can assist in quality assurance and quality improvement activities.(179) Other advantages cited for universal umbilical blood gas analysis include staff becoming more proficient at collection, processing and testing of samples so accurate results will exist when really needed, and umbilical blood gas results will always be available for interpretation when there are adverse outcomes to delivery (medicolegal argument).(180)

Does arterial umbilical cord pH predict adverse neurological outcomes?

Low umbilical artery pH at birth is a risk factor for subsequent short- and long-term neurological outcomes but is a poor diagnostic test for predicting intermediate or long-term complications related to childbirth.

A systematic review and meta-analysis of studies found that low arterial umbilical cord pH (studies used cut-offs between 7.0 and 7.2) had a strong, consistent and temporal association with neonatal morbidity and mortality (hypoxic ischemic encephalopathy, seizures, intraventricular hemorrhage or periventricular leukomalacia) and possibly the long term outcome of cerebral palsy.(181) Despite increasing evidence that low arterial artery umbilical pH is associated with poor neonatal and long term outcomes, it is important to remember that few infants born with low umbilical arterial pH will develop cerebral palsy(182), and most infants that go on to have neurodevelopmental morbidities as children will be born with normal umbilical arterial pH.(183, 184) Thus, low umbilical artery pH at birth is a risk factor for subsequent short- and long-term neurological outcomes but is a poor diagnostic test for predicting intermediate or long term complications related to childhood.

Does implementation of universal (as opposed to selective) umbilical cord blood gas analysis improve neonatal outcomes in centers implementing this practice?

There is limited evidence that universal cord blood gas testing as opposed to selective testing improves the neonatal outcomes.

A single center study was conducted over a 3-year period after implementing universal cord blood gas testing found that outcomes as measured by percent of infants with umbilical artery pH <7.1 and number of nursery admissions, did not change over the 3-year period, nor did universal testing affect the distribution of Apgar scores at 5 minutes or percent of infants requiring resuscitation.(185) Another study found that universal umbilical cord blood gas
screening could identify infants with encephalopathy that were missed by other screening methods and predicted that universal screening would facilitate the timely initiation of therapeutic hypothermia to prevent secondary brain injury in the setting of hypoxic ischemia. (186) We could not find any published evidence that universal testing leads to more reliable blood gas values as measured by number/percent of samples meeting Westgate criteria for valid samples (based upon relationship between arterial and venous pH and pCO2 in paired samples) or percent of samples rejected by the laboratory or testing personnel.

Is universal umbilical cord blood gas analysis cost effective?

There is insufficient evidence to conclude that universal umbilical cord blood gas analysis is cost effective in most hospital practices with decreased nursery admissions.

One study performed at a single medical center (the same medical center observing improved neonatal outcomes associated with universal umbilical cord blood gas analysis) found that universal cord blood gas was cost effective due to cost savings associated with decreased nursery admissions resulting from universal cord blood gas testing. (187) This large study from Australia observed that the initial cost of universal umbilical cord blood gas analysis was high, but that there was significant savings from reduced special care nursery costs. Because findings are limited to one medical center, there is insufficient evidence to conclude that universal umbilical cord blood gas analysis is cost effective in most hospital practices.

Summary

POCT is gaining popularity in the field of reproductive medicine from predicting ovulation and diagnosing pregnancy, to managing premature rupture of membranes and fetal distress at birth. By providing quick results at the site of patient care, POCT can offer the opportunity for faster medical intervention. The AACC Academy formed an expert committee to examine the key clinical questions and published literature surrounding the use of POCT in fertility and reproduction. This guidance was a consensus of expert opinion based on current peer-reviewed literature. These recommendations updated the previous Laboratory Medicine Best Practice Guidelines (LMPG) for Evidence-Based POCT and provide recommendations for best practice in the utilization of POCT. A discussion of supporting literature, as well as challenges and limitations of POCT was provided for each recommendation.

Acknowledgement

The authors wish to acknowledge the information manufacturers provided in their package inserts, some of which is cited in this guidance document. Neither the authors of this guidance document nor the AACC Academy are endorsing any manufacturer product. The mention of specific products is intended only as an example supporting key points made in the guidance, and the examples mentioned are not representative of all products that are available world-wide.

The committee developing this guidance would like to remember Dr. Harvey Kincaid, Jr. who contributed to this document, but unfortunately passed before the document was completed. We send our condolences to his family and are thankful for all of Harvey’s contributions.
References


41. Regidor P-A, Kaczmarczyk M, Schiweck E, Goeckenjan-Festag M, Alexander H. Identification and prediction of
the fertile window with a new web-based medical device using a vaginal biosensor for measuring the circadian and
42. Papaioannou S, Aslam M, Al Wattar BH, Milnes RC, Knowles TG. User’s acceptability of OvuSense: a novel vaginal
44. Maijala A, Kinnunen H, Koskimäki H, Jämsä T, Kangas M. Nocturnal finger skin temperature in menstrual cycle
changes in physiology and enable prediction of the fertile window: Observational study. J Med Internet Res.
46. Luo L, She X, Cao J, Zhang Y, Li Y, Song PXK. Detection and prediction of ovulation from body temperature
wearables capture the changes in temperature associated with the menstrual cycle. Biosci Rep. 2018;38(6).
49. Shilaih M, Clerck V, Falco L, Kübler F, Leeners B. PulserRate measurement during sleep using wearable sensors,
50. Braunstein GD, Rasor J, Danzer H, Adler D, Wade ME. Serum human chorionic gonadotropin levels throughout
53. Lenton EA, Neal LM, Sulaiman R. Plasma concentrations of human chorionic gonadotropin from the time of
54. Lohstroh PN, Overstreet JW, Stewart DR, Nakajima ST, Cragun JR, Boyers SP, et al. Secretion and excretion of
56. Chasen ST, Martinucci S, Perni SC, Kalish RB. First-trimester biochemistry and outcomes in twin pregnancy. The
57. Food and Drug Administration, Centers for Devices and Radiological Health. Guidance for industry and FDA
reviewers/staff: Guidance for over-the-counter (OTC) human chorionic gonadotropin (hCG) 510(k)s Rockville, MD: DHHS,
59. Assay of protein hormones related to human reproduction: Problems of specificity of assay methods and
chorionic gonadotropin (hCG) variants by monoclonal antibodies and mass spectrometry. Tumour Biol. 2014;35(2):1013-
22.
61. Gottlieb M, Wnek K, Moskoff J, Christian E, Bailitz J. Comparison of Result Times Between Urine and Whole
62. Nerenz RD, Gronowski AM, Grenache DG. Inappropriate off-label use of a qualitative, point-of-care hCG device.


159. Fetal Scalp Lactate Testing During Intrapartum Pregnancy with Abnormal Fetal Heart Rate: A Review of Clinical Effectiveness, Cost-Effectiveness, and Guidelines. Ottawa, Canada: Canadian Agency for Drugs and Technologies in Health (CADTH); March 2018.


**TABLE 1** Summary of the major methods developed to estimate the onset of LH surge for predicting ovulation in the menstrual cycle (summarized data from Godbert S, et. al.)(8)

<table>
<thead>
<tr>
<th>Number of studies</th>
<th>Method</th>
<th>Detection Rate within ±1 day of LH Surge (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>Fixed days</td>
<td>52.0-63.8 %</td>
</tr>
<tr>
<td>8</td>
<td>LH peak</td>
<td>73.2-90.6 %</td>
</tr>
<tr>
<td>4</td>
<td>Estimated LH surge</td>
<td>94.5-97.6 %</td>
</tr>
<tr>
<td>Authors</td>
<td>Study Design</td>
<td>Size(N)</td>
</tr>
<tr>
<td>-------------------------</td>
<td>-------------------------</td>
<td>---------</td>
</tr>
<tr>
<td><strong>Conception Rate Outcomes in Fertile Women</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Robinson JE et al 2007</td>
<td>Prospective RCT</td>
<td>653</td>
</tr>
<tr>
<td>Johnson S et al 2020</td>
<td>Open-label clinical trial</td>
<td>844</td>
</tr>
<tr>
<td>Yeh PT et al 2019</td>
<td>Meta-analysis of 4 RCTs</td>
<td>1,487</td>
</tr>
<tr>
<td>Bouchard TP et al 2018</td>
<td>Prospective effectiveness study</td>
<td>256</td>
</tr>
<tr>
<td><strong>Pregnancy Avoidance Outcomes in Fertile Women</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fehring RJ et al 2008</td>
<td>Retrospective study</td>
<td>204</td>
</tr>
<tr>
<td>Fehring RJ et al 2014</td>
<td>Comparison efficacy study</td>
<td>160</td>
</tr>
<tr>
<td>Fehring RJ et al 2017</td>
<td>Comparison efficacy study</td>
<td>816</td>
</tr>
</tbody>
</table>
### Use of Cycle Applications to compute and identify fertile window

<table>
<thead>
<tr>
<th>Study</th>
<th>Study Type</th>
<th>Sample Size</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stanford JB et al 2020</td>
<td>Prospective cohort study</td>
<td>8363</td>
<td>Analysis of the tracked self-reporting data to assess the influence of using cycle apps on per cycle probability of conception. After adjusting for potential confounders, use of cycle apps was associated with increased fecundability of 12-20% per cycle of attempt. (27)</td>
</tr>
<tr>
<td>Soumpasis I et al 2020</td>
<td>Big data analysis</td>
<td>32,595</td>
<td>Compare users’ perceived cycle characteristics with actual cycle characteristics using anonymized cloud data collected from women trying to conceive. There exist wide variations in cycle length and the ovulation day. The info was analyzed to provide population ranges and probability table for better timing of the fertile window. (15)</td>
</tr>
</tbody>
</table>

### Timing Insemination in Intrauterine Insemination (IUI) for Infertility Treatment

<table>
<thead>
<tr>
<th>Study</th>
<th>Study Type</th>
<th>Sample Size</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cantineau AE et al 2014</td>
<td>Meta-analysis of 14 RCTs</td>
<td>2279</td>
<td>A systematic review to evaluate the effectiveness of different synchronization methods for IUI in subfertile couples. More research is needed to determine if any difference exists in safety and effectiveness among methods used for synchronizing ovulation and insemination. (30)</td>
</tr>
<tr>
<td>El Hachem, H et al 2017</td>
<td>Prospective cohort study</td>
<td>232</td>
<td>Compare the cumulative live birth rates (LBR) between those using u-LH strips to detect LH surge verse those using ultrasound monitoring (US) to time sperm insemination. Urinary LH strip vs. ultrasonography groups Live birth rates -12.4% vs. 9.2% Conception rates- 19.9% vs. 13.3% (31)</td>
</tr>
</tbody>
</table>

### Confirming Sufficient Ovarian Stimulation Before Oocyte Retrieval in IVF

<table>
<thead>
<tr>
<th>Study</th>
<th>Study Type</th>
<th>Sample Size</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cozzolino M et al 2020</td>
<td>Prospective cohort study</td>
<td>356</td>
<td>Urine LH testing was performed at home 12 h after the GnRHa trigger to replace serum LH for confirming LH surge. • OPK use resulted in an overall sensitivity of 99.7% (355/356) and specificity of 66.7% (2/3). • There is a significant cost saving when u-LH kits ($2,010) was used to replace serum LH ($17,680). (33)</td>
</tr>
</tbody>
</table>
**TABLE 3.** Features of POCT and Laboratory testing hCG. Each row distinguishes a different feature of the test. The feature with an advantage is colored green in each row.

<table>
<thead>
<tr>
<th><strong>POCT hCG</strong></th>
<th><strong>Laboratory hCG</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Fast turnaround of test results</td>
<td>Longer analytical time</td>
</tr>
<tr>
<td>Unprocessed samples (no centrifugation)</td>
<td>Sample processing requires centrifugation</td>
</tr>
<tr>
<td>Native urine and whole blood samples</td>
<td>Serum or plasma (clotting and centrifugation adds steps and time)</td>
</tr>
<tr>
<td>Portable</td>
<td>Fixed instrumentation (requires dedicated space)</td>
</tr>
<tr>
<td>Manual, visually interpreted tests</td>
<td>Automated methods</td>
</tr>
<tr>
<td>Manual result reporting for visual tests (devices and readers have interfaces available)</td>
<td>Electronic result reporting and instrument interfaces</td>
</tr>
<tr>
<td>Minimal maintenance</td>
<td>Ongoing maintenance required</td>
</tr>
<tr>
<td>Qualitative results predominantly (few blood quantitative hCG methods)</td>
<td>Quantitative results (can trend over time)</td>
</tr>
<tr>
<td>No dilution (limited range for quantitative tests)</td>
<td>Dilution above analytical measurement range (wide range of reportable results)</td>
</tr>
<tr>
<td>POCT does not match laboratory results</td>
<td>Results harmonized by same method</td>
</tr>
<tr>
<td>CLIA waived (urine) – blood is moderate complexity</td>
<td>CLIA moderate complexity</td>
</tr>
<tr>
<td>Ease-of-use (minimal training)</td>
<td>Technical degree with laboratory experience required</td>
</tr>
<tr>
<td>Room temperature storage (limited refrigeration)</td>
<td>Refrigerated reagents and controls</td>
</tr>
<tr>
<td>Unit-use tests (implement in low volume areas)</td>
<td>High throughput factory environment</td>
</tr>
<tr>
<td>Analyze one test at a time</td>
<td>Random access analyzers (multiple tests simultaneously)</td>
</tr>
<tr>
<td>Higher cost per test (individual packaging)</td>
<td>Bulk reagents (less cost per test)</td>
</tr>
</tbody>
</table>
TABLE 4: Summary of studies that assess the diagnostic performance of PROM tests

<table>
<thead>
<tr>
<th>Confirmed Clinical Diagnostic Criteria of PROM</th>
<th>N</th>
<th>Tests</th>
<th>Sens (%)</th>
<th>Spec (%)</th>
<th>PPV (%)</th>
<th>NPV (%)</th>
<th>Accu (%)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>48-hour follow up medical records review by an experienced physician who was blinded to the immunoassay test results</td>
<td>323</td>
<td>-Amnisure -ROM PLUS -SCA (all 3 present: ferning, pH and pooling)</td>
<td>93.4</td>
<td>91.7</td>
<td>95.0</td>
<td>91.9</td>
<td>96.0</td>
<td>(138)</td>
</tr>
<tr>
<td></td>
<td>322</td>
<td></td>
<td>95.0</td>
<td>97.0</td>
<td>94.8</td>
<td>87.4</td>
<td>97.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>323</td>
<td></td>
<td>95.0</td>
<td>97.0</td>
<td>94.8</td>
<td>87.4</td>
<td>97.1</td>
<td></td>
</tr>
<tr>
<td>Visualization of vaginal amniotic fluid and/or persistence of oligohydramnios</td>
<td>75</td>
<td>-Actim -Amnisure -pH test</td>
<td>100</td>
<td>90.2</td>
<td>96.7</td>
<td>97.8</td>
<td>100</td>
<td>(144)</td>
</tr>
<tr>
<td></td>
<td>82</td>
<td></td>
<td>90.2</td>
<td>100</td>
<td>95.2</td>
<td>100</td>
<td>86.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>63</td>
<td></td>
<td>85.7</td>
<td>95.2</td>
<td>97.3</td>
<td>76.9</td>
<td>88.9</td>
<td></td>
</tr>
<tr>
<td>Two of the following criteria: delivery within 48 hours to 7 days, chorioamnionitis, membrane rupture before delivery, adverse perinatal outcomes associated with prolonged PROM</td>
<td>99</td>
<td>-Amnisure -AmnioQuick Duo</td>
<td>93.2</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>94.9</td>
<td>(188)</td>
</tr>
<tr>
<td></td>
<td>(all)</td>
<td></td>
<td>97.3</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>98.0</td>
<td></td>
</tr>
<tr>
<td>Review of medical records after delivery by a physician blinded to the immunoassay test results</td>
<td>220</td>
<td>-Amnisure -AmnioQuick Duo -ferning -pH test</td>
<td>95.5</td>
<td>93.6</td>
<td>89.1</td>
<td>89.7</td>
<td>95.1</td>
<td>(143)</td>
</tr>
<tr>
<td></td>
<td>(all)</td>
<td></td>
<td>93.6</td>
<td>86.4</td>
<td>87.3</td>
<td>93.1</td>
<td>92.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>72.7</td>
<td>80.9</td>
<td>79.2</td>
<td>74.8</td>
<td>90.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>76.4</td>
<td>83.6</td>
<td>82.4</td>
<td>78.0</td>
<td>76.8</td>
<td></td>
</tr>
<tr>
<td>Review of medical records after delivery. ROM was considered when the membranes are absent during vaginal examination or a positive pad chart was obtained.</td>
<td>211</td>
<td>-Amnisure -SCA (2 of 3 present: ferning, pH or pooling)</td>
<td>95.7</td>
<td>78.1</td>
<td>100</td>
<td>100</td>
<td>75</td>
<td>(140)</td>
</tr>
<tr>
<td></td>
<td>(all)</td>
<td></td>
<td>78.1</td>
<td>100</td>
<td>100</td>
<td>75</td>
<td>36.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>75</td>
<td>36.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>77.5</td>
<td>100</td>
<td>100</td>
<td>25.5</td>
<td>36.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>62.6</td>
<td>100</td>
<td>100</td>
<td>25.5</td>
<td>36.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>86.6</td>
<td>100</td>
<td>100</td>
<td>49.0</td>
<td>36.3</td>
<td></td>
</tr>
<tr>
<td>Review of medical records after delivery by a physician blinded to the immunoassay test results</td>
<td>285</td>
<td>-ROM PLUS -SCA (visualization of fluid or 2 of 3 present: ferning, pH or pooling)</td>
<td>99</td>
<td>91</td>
<td>95</td>
<td>99</td>
<td>96.5</td>
<td>(142)</td>
</tr>
<tr>
<td></td>
<td>(all)</td>
<td></td>
<td>91</td>
<td>98</td>
<td>99</td>
<td>77</td>
<td>96.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>98</td>
<td>99</td>
<td>99</td>
<td>77</td>
<td>89.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>85</td>
<td>80</td>
<td>99</td>
<td>88</td>
<td>86.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>72</td>
<td>80</td>
<td>99</td>
<td>88</td>
<td>86.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>83</td>
<td>80</td>
<td>99</td>
<td>88</td>
<td>89.5</td>
<td></td>
</tr>
<tr>
<td>Medical record review after delivery by physicians blinded to the immunoassay test results</td>
<td>167</td>
<td>-Actim -Amnisure -SCA (visualization of fluid or 2 of 3 present: ferning, pH or pooling)</td>
<td>89.8</td>
<td>94.3</td>
<td>97.5</td>
<td>97.5</td>
<td>89.5</td>
<td>(139)</td>
</tr>
<tr>
<td></td>
<td>(all)</td>
<td></td>
<td>84.3</td>
<td>97.5</td>
<td>97.5</td>
<td>89.5</td>
<td>89.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>88.6</td>
<td>97.5</td>
<td>97.5</td>
<td>89.5</td>
<td>89.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>95.1</td>
<td>97.6</td>
<td>93.9</td>
<td>88.2</td>
<td>(93</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>91</td>
<td>95</td>
<td>95</td>
<td>81</td>
<td>91</td>
<td></td>
</tr>
</tbody>
</table>
### TABLE 5: Analytical detection limit of a few commercial PROM kits

<table>
<thead>
<tr>
<th>Kit</th>
<th>Target Protein</th>
<th>Detection Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actim</td>
<td>IGFBP-1</td>
<td>25 ng/mL</td>
</tr>
<tr>
<td>Amnisure</td>
<td>PAMG-1</td>
<td>5 ng/mL</td>
</tr>
<tr>
<td>ROM plus</td>
<td>PP12 and AFP</td>
<td>5 ng/mL for PP12; 150 ng/mL for AFP</td>
</tr>
</tbody>
</table>