

ORIGINAL RESEARCH

Obstetrics and Gynecology

# False negative point-of-care urine pregnancy tests in an urban academic emergency department: a retrospective cohort study

Sarah Kleinschmidt MD | Julianne N. Dugas MPH | Kerrie P Nelson PhD |  
James A Feldman MD, MPH

Department of Emergency Medicine, Boston Medical Center, Boston University School of Medicine, Boston, Massachusetts, USA

**Correspondence**

James A. Feldman, MD, MPH, Department of Emergency Medicine, Boston Medical Center, Boston University School of Medicine, 818 Harrison Ave, Boston, MA 02118, USA.  
Email: [jfeldman@bu.edu](mailto:jfeldman@bu.edu)

**Funding and support:** CTSA grant UL1TR001430 provided support for this study through the Clinical Data Warehouse and REDCap electronic data capture tools hosted at Boston University. JAF is supported in part by UL1TR001430.

## Abstract

**Study objective:** To determine the prevalence of false negative point-of-care (POC) urine pregnancy tests among emergency department (ED) patients and among those with abdominal pain or vaginal bleeding.

**Methods:** We identified all female patients, ages 14–50 years without prior hysterectomy who had a negative POC urine pregnancy test (beta subunit of human chorionic gonadotropin [ $\beta$ -hCG]) performed by trained clinical staff in the ED between September 1, 2017 and December 31, 2018, as well as a subgroup we defined a priori as “high risk” for early pregnancy complications based on a triage chief complaint (text) of abdominal pain or vaginal bleeding. We identified those with a positive urine  $\beta$ -hCG, serum  $\beta$ -hCG >5 mIU/mL, or a diagnosis of pregnancy within 3 months of the initial ED visit (index visit). We used structured chart review with American College of Obstetrics and Gynecology guidelines to determine pregnancy diagnosis and outcomes (ectopic, intrauterine, abnormal including spontaneous abortion, and unknown), the date of conception, and whether the pregnancy was present at the index visit.

**Results:** Of 10,924 visits with a negative urine pregnancy test result that were screened for a pregnancy outcome, 171 (1.6%, 95% confidence interval [CI] = 1.4, 1.8) had a pregnancy present at the index visit. Diagnoses were ectopic (n = 12, 7.0%), intrauterine (n = 71, 41.5%), abnormal (n = 77, 45.0%), and unknown (n = 11, 6.4%). Of the 2732 patients with high-risk complaints, 97 (3.6%, 95% CI = 2.9, 4.3) had a pregnancy present at the index visit (relative risk of a pregnancy diagnosis 3.9, 95% CI = 2.9, 5.3), including 10/12 ectopic (83%), 58/77 abnormal (75%), and 25/71 intrauterine pregnancies (35%). Serum  $\beta$ -hCG ranged from 2 mIU/mL to above assay (median = 119.5, interquartile range = 957.5).

**Conclusion:** Although false negative urine pregnancy tests were uncommon, multiple pregnancy diagnoses were missed, including ectopic pregnancies. False negatives were more common among patients with abdominal pain or vaginal bleeding. Concurrent serum  $\beta$ -hCG levels demonstrated a broad distribution.

Supervising Editor: Nicholas Caputo, MD, MSc.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2021 The Authors. *JACEP Open* published by Wiley Periodicals LLC on behalf of American College of Emergency Physicians

**KEYWORDS**

diagnostic errors, emergency medicine, false negative reactions, MESH, pregnancy tests

## 1 | INTRODUCTION

### 1.1 | Background

When a patient of childbearing potential presents to an emergency department (ED), determining whether they are pregnant is of critical importance to explain symptoms, rule out life-threatening complications, prevent iatrogenic harm, and allow a patient to make informed and timely decisions about pregnancy management. Because patient report and physical examination have limited sensitivity in early pregnancy, objective testing is routinely performed.<sup>1–3</sup> Testing methods vary by institution, but urine point-of-care (POC) detection of the beta subunit of human chorionic gonadotropin ( $\beta$ -hCG) is one of the most common because it is non-invasive, low cost, provides rapid results, and is routinely used at our center to screen for pregnancy.<sup>4</sup>

### 1.2 | Importance

Despite the widespread use of urine pregnancy tests, there is limited evidence of their real-world performance. One recent review concluded that “no relevant health technology assessments, systematic reviews, meta-analyses, randomized controlled trials, or evidence-based guidelines were identified” regarding the diagnostic accuracy or clinical utility of urine pregnancy testing in the ED.<sup>5</sup> POC urine pregnancy testing has a reportedly high sensitivity (97%) with diagnostic thresholds of  $\approx 20$  mIU/mL.<sup>6</sup> However, false negative results have been reported with several case reports of patients with ectopic pregnancy, ruptured ectopic pregnancy, or molar pregnancies who had negative urine  $\beta$ -hCG tests.<sup>7–9</sup>

One study reported a false omission rate (FOR) of 0.34%<sup>6</sup> with another study reporting in abstract a false negative rate of 10.8% (95% CI; 9.3 to 12.6%) corresponding to a FOR of 0.9%.<sup>10</sup> Of note, the former and lower of these estimates only examined false negative results discovered at the index visit via concurrent testing, adding a potential downward bias by not including pregnancies identified in follow-up. For example, if a patient who has a negative urine POC  $\beta$ -hCG result is then diagnosed 3 weeks later with an 8-week intrauterine pregnancy (IUP) by ultrasound, obstetric dating guidelines suggest that this index visit would have occurred a full 3 weeks after conception; however, such a delayed diagnosis would not have been identified as a false negative case by prior study methods.<sup>6,11</sup>

### 1.3 | Goals of this investigation

We sought to determine the prevalence of false negative urine pregnancy test results in a large urban ED to establish whether this test

should appropriately be used to rule out pregnancy, both among a general population and among patients with symptoms concerning for early pregnancy complications. A high rate of false negative urine tests could inform clinician decisions as well as operational changes to support exclusive or more frequent serum testing. When available, we examined the concurrent serum  $\beta$ -hCG level to evaluate the reason for false negative results. By identifying patients who had concurrent pregnancy testing (serum  $\beta$ -hCG or ultrasound) as well as those who had pregnancy diagnosed in follow up from the ED visit, we aimed to provide a more accurate estimate of the performance of POC urine  $\beta$ -hCG testing in routine clinical care.

## 2 | MATERIALS AND METHODS

### 2.1 | Study design and setting

We performed a retrospective observational cohort study at an urban academic medical center with  $\approx 140,000$  visits per year. The study was determined to be exempt by the Institutional Review Board of Boston University Medical Campus/Boston Medical Center. We followed the STROBE checklist for observational research studies.<sup>12</sup>

### 2.2 | Selection of participants

Our primary objective was to determine the proportion of false negative urine  $\beta$ -hCG results in the total population of patients who had a POC urine  $\beta$ -hCG performed (ie, FOR). The FOR is defined as the proportion of the individuals with a negative test result for which the true condition is positive.<sup>13</sup> Our secondary objectives were to identify the FOR among patients at high risk for early pregnancy complications who are recommended to have a pregnancy testing obtained in the ED clinical practice guidelines, to characterize the type of pregnancies with false negative results, and to describe the distribution of available concurrent serum results.

We included all female patients 14–50 years of age with visits from September 1, 2017 to December 31, 2018. We used the Clinical Data Warehouse (an organized and accessible research database that consolidates electronic medical records across our hospital system) to query the electronic medical record (Epic Systems software, Verona, WI) and identify all subjects who had a negative POC urine pregnancy test performed in our ED. For any subjects with multiple visits during the study period meeting inclusion criteria, only the first eligible visit was selected as an index visit, with follow-up data contributing to outcome determination, rather than being considered as a separate case for inclusion. We did include patients with multiple  $\beta$ -hCG results at the index visit if the first result was a

negative urine test. We excluded any patient with a documented history of hysterectomy.

Within this overall sample ( $n = 10,924$ ), we used the Data Warehouse to identify for chart review all subjects who met inclusion criteria and also had possible evidence of pregnancy at the index visit or within 90 days ( $n = 383$ ), including any positive urine  $\beta$ -hCG test, any non-zero serum  $\beta$ -hCG test, any pelvic ultrasound, or any visit with an ICD-10 diagnosis suggesting pregnancy (Table 1 and Figure 1).

We identified *a priori* a “high risk” subgroup that met the above inclusion criteria and also presented with a chief complaint suggestive of possible pregnancy. We based these clinical criteria on the National Quality Forum (NQF) indicator for patients requiring pregnancy testing in the ED,<sup>14,15</sup> including any triage chief complaint of abdominal pain, abdominal cramping, pelvic pain, pelvic cramping, or vaginal bleeding. We used independent review and a consensus process to evaluate various abbreviations and misspellings of the above (ie, “abd pain,” “adbominal pain” as deliberate misspelling), generally including these to maximize sensitivity. We excluded complaints that clearly indicated pain localized to the upper abdomen (ie, “left upper quadrant pain”), as well as complaints designated as post-operative or sickle cell-related. We included all other complaints, including patients with multiple complaints (“wrist pain, abdominal pain”), complaints with unclear location (“pain in belly”), and patients with additional symptoms such as diarrhea or nausea (Supporting Information Appendix 1).

### 2.3 | Methods of measurement

Urine POC testing was performed with the Sekisui OSOM Card Pregnancy Test (Instructions for Use [IFU] assay sensitivity of 20 mIU/mL), by ED staff who received annual training and competency verification per hospital policies. Although the IFU indicate that this assay should only be used within 8 weeks of reported last menstrual period (LMP), we included any case in which it was used and negative, reflecting real-world use in our ED. The central laboratory used the Abbott Architect i2000 to perform serum  $\beta$ -hCG testing, interpreted as “negative” for  $\beta$ -hCG levels  $\leq 5.00$  mIU/mL and “positive” for  $\beta$ -hCG levels  $\geq 25.00$  mIU/mL. Specimens with  $\beta$ -hCG levels  $> 5.00$  and  $< 25.00$  mIU/mL are therefore reported with concentrations only, noting that these levels “may be indicative of early pregnancy” and should be clinically correlated. Ultrasounds were performed by the radiology or antenatal testing units according to established protocols and interpreted by a radiologist or gynecologic documentation.

### 2.4 | Outcome measures

We followed recommended procedures to enhance the validity of chart review studies.<sup>16</sup> Study data were collected and managed using REDCap electronic data capture tools hosted at Boston University, and both this and the Clinical Data Warehouse were supported by CTSI (1UL1TR001430).<sup>17</sup>

#### The Bottom Line

“False negative urine pregnancy tests in the emergency department risk iatrogenic harm as well as delayed diagnosis of pregnancy complications. This retrospective study found a prevalence of 1.6%, including multiple ectopic pregnancies, and a 3.6% rate among females presenting with abdominal pain or vaginal bleeding.

We refined our final outcome measures through an iterative process, after initially developing outcome criteria through literature review and trial review of a small number of cases. All charts were then reviewed using these outcome criteria by 2 emergency medicine physicians and 2 research assistants who had been trained in the REDCap data collection form, search procedures within the electronic medical record, and specific coding definitions. We then performed an agreement analysis using a sample of 50 cases randomly selected from those patients found to have a pregnancy present on initial coding ( $n = 310$ ), with each case blindly and independently reviewed by each of these 4 abstractors for appropriate inclusion based on eligibility criteria, pregnancy outcome, and determination of whether the pregnancy was present at the index visit. We found near perfect chance-corrected agreement for inclusion ( $K = 1.00$ ,  $SE = 0.00$ ) and whether serum was performed at the index visit ( $K = 0.91$ ,  $SE = 0.06$ ).<sup>18</sup> Because we found only substantial agreement ( $K = 0.64$ ,  $SE = 0.05$ ) for the primary outcome (whether a pregnancy was present at the index visit) and the secondary outcome ( $K = 0.73$ ,  $SE = 0.03$ ) for (the specific type of pregnancy), we reviewed both criteria with an obstetric specialist, arriving at our final outcome criteria as listed in Tables 2 and 3. These changes

**TABLE 1** Study inclusion criteria

<sup>a</sup>Inclusion criteria ( $n = 10,924$ )

*Must meet ALL criteria*

Date of index visit: January 9, 2017–December 12, 2018

Age at index visit: 14–50 years old

Female

Urine pregnancy test negative at index visit

No history of hysterectomy

<sup>b</sup>Criteria for chart review ( $n = 383$ )

Meets ANY criteria at index visit or within 90 days

Urine pregnancy test positive (other than initial negative test)

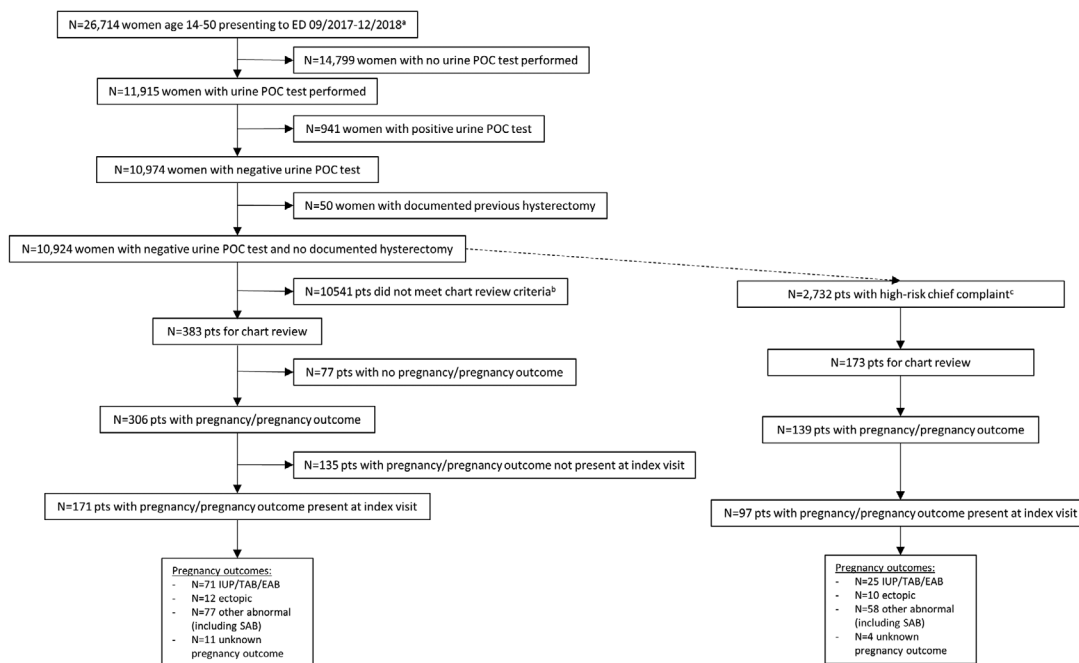
Serum  $\beta$ -hCG result other than  $< 1$  mIU/mL

Pelvic ultrasound performed

ICD-10 suggestive of pregnancy

<sup>c</sup>High-risk complaint

Any triage chief complaint (text) of abdominal pain, abdominal cramping, pelvic pain, pelvic cramping, or vaginal bleeding (see Supporting Information Appendix 1)



<sup>a</sup>Inclusion criteria (n=10,924): Meets ALL of the following: index visit 9/1/2017-12/31/2018; age 14-50 yrs. old at index visit; female; negative urine pregnancy test at index visit; and no history of hysterectomy

<sup>b</sup>Chart review criteria (n=383): Meets ANY of the following: positive urine pregnancy test (other than initial negative test); serum  $\beta$ -hCG result other than  $<1$  mIU/mL; pelvic ultrasound performed; or ICD-10 suggestive of pregnancy

<sup>c</sup>High-risk chief complaint (n=2,732): Any triage chief complaint (text) of abdominal pain, abdominal cramping, pelvic pain, pelvic cramping, or vaginal bleeding (see Appendix 1)

**FIGURE 1** Inclusion flow chart

**TABLE 2** Pregnancy dating criteria

Present at index visit (any of below):

Positive serum  $\beta$ -hCG result ( $>5$  mIU/mL), positive urine  $\beta$ -hCG, or ultrasound diagnostic for pregnancy at index visit

Return and pregnancy diagnosis within 7 days

Visit  $>14$  days s/p calculated start date of pregnancy (ACOG guidelines)

Not present at index visit (any of below):

Negative serum  $\beta$ -hCG at index visit ( $<1$  mIU/mL)

Visit  $\leq 14$  days s/p calculated start date of pregnancy (ACOG guidelines)

Indeterminate:

Any case not meeting either above criteria, including:

Irregular menses and no ultrasound results available

Not enough information documented (ie, missing LMP or ultrasound results)

ACOG guidelines summary:

Start date of pregnancy = first day of LMP if known and regular, unless one of the following are met:

- ultrasound with measured CRL before 8 weeks, 6 days has discrepancy of  $>5$  days

- ultrasound with CRL between 9–13 weeks, 6 days has discrepancy of  $>7$  days

- ultrasound with gestational sac but no CRL, only if LMP not recorded

CRL, crown rump length.

included simplifying the pregnancy dating criteria to strictly follow American College of Obstetrics and Gynecology (ACOG) criteria and simplifying outcome categories (eg, combining therapeutic abortions with IUPs) to be clinically relevant.<sup>11</sup> All charts were then reviewed for final outcome measures by 1 of the 2 emergency medicine MD reviewers, with final consensus reached with the second emergency medicine MD reviewer for any outcome assessments that differed from our initial determination. The study team was not blinded to the study hypothesis.

The primary outcome measure was defined as a pregnancy that was present at the index visit. To determine these outcomes, we reviewed any available clinical data, including within ED, inpatient, or outpatient services. Cases were categorized using predetermined explicit criteria either as not representing a pregnancy, a pregnancy that was present at index visit, a pregnancy that was not present at the index visit, or indeterminate due to insufficient data to establish a start date (Table 2). Pregnancies were determined to be present at the index visit if they were discovered at the index visit or if the index visit occurred at least 14 days after the start date of the pregnancy as calculated using ACOG criteria that incorporates LMP and ultrasound results. Supporting Information Appendix 2 provides detailed examples of the application of the ACOG guidelines to determine the timing of conception and date a pregnancy in relationship to the index visit.<sup>11</sup> (It should be noted that pregnancy start dates occur  $\approx 14$  days before ovulation and conception due to an historical reliance on LMP.)

Secondary outcomes included determination of pregnancy as normal intrauterine, ectopic, or other abnormal, with definitions shown

**TABLE 3** Pregnancy outcome definitions

Evidence of pregnancy (any of the following):
Ultrasound showing fetal heart rate or yolk sac, whether intrauterine or ectopic
Serum hCG $\beta$ -hCG $\geq 5$ mIU/mL without other documented cause (eg, menopausal, malignancy)
Positive urine pregnancy test without other documented cause
Pathology confirmation of products of conception
Any patient without any of above criteria is determined to have no evidence of pregnancy.
<i>Specific pregnancy outcome criteria</i>
Normal intra-uterine pregnancy (any of the following)
Ultrasound showing intrauterine yolk sac or fetal heart rate <i>and</i> no evidence of abnormal pregnancy as defined below
Delivery of one or more live children
Therapeutic abortion with products of conception on uterine pathology
Abnormal pregnancy, including spontaneous abortion (any of the following unless also meeting the criteria for ectopic pregnancy)
Serial serum hCG showing fall or inappropriate rise, at or within 2 weeks of index visit
Ultrasound result other than IUP with concurrent or prior serum hCG $>2000$
Ultrasound result other than IUP after previous IUP
Documentation of likely spontaneous abortion, pregnancy of unknown location or other complication per gynecologic documentation, including at time of pregnancy or in future follow-up (ie, maternity record)
Ectopic pregnancy if meets <i>all</i> of the following
Ultrasound or CT results consistent with ectopic pregnancy
Serum $\beta$ -hCG $>25$ mIU/mL
Documented diagnosis of or treatment for likely or confirmed ectopic pregnancy per gynecologic documentation
Unknown pregnancy outcome
Any pregnancy not meeting above criteria for either normal intrauterine, abnormal or ectopic pregnancy, including due to lack of data or follow-up (eg, single visit without diagnostic ultrasound results)

in Table 3. Normal IUPs included any with a yolk sac or fetal heart-beat within the uterus, whether the patient later delivered, received a therapeutic abortion, or was lost to follow-up. Ectopic pregnancies were defined by imaging and serum results along with gynecologic documentation, in concordance with ACOG diagnostic guidelines.<sup>19</sup> The “other abnormal” category included spontaneous abortion but does not include the ectopic pregnancies as defined above. Pregnancy outcome was listed as unknown if there was not enough data to establish any above criteria.

## 2.5 | Primary data analysis

We report the prevalence of false negative urine pregnancy tests among the study population, and within the subgroup who presented

with concerning chief complaints, as well as the relative risk for a pregnancy diagnosis in this subgroup. We also report the range and distribution of values for concurrent serum  $\beta$ -hCG values at the index visit and 95% confidence intervals (CI) as appropriate. We used SAS v9.4 (SAS Institute Inc., Cary, NC) for all analyses.

## 2.6 | Sensitivity analysis

Four pre-planned sensitivity analyses were conducted as follows. Because serum  $\beta$ -hCG results  $\leq 5$  mIU/mL are required to rule out pregnancy, we considered values  $>5$  mIU/mL to represent evidence of pregnancy. However, due to a non-diagnostic zone of serum  $\beta$ -hCG results between 5 and 25 mIU/mL, we performed a sensitivity analysis (Analysis A) excluding from the pregnancy outcome measure any patient who only met outcome criteria through a serum result 5–25 mIU/mL ( $n = 19$ ).

Given the inherent uncertainty in dating of early pregnancy by both first trimester ultrasound and LMP, we performed sensitivity analyses to estimate the effects of this uncertainty on the outcome measure of whether pregnancy was present at the index visit. In sensitivity Analysis B, we excluded all pregnancies previously categorized as present but whose date of conception was  $\leq 7$  days after the index visit ( $n = 20$ ). Separately, in sensitivity Analysis C, we included all pregnancies previously categorized as not present but whose date of conception was  $\leq 7$  days before the index visit ( $n = 21$ ). To estimate the lowest possible number of pregnancies present at the index visit, we performed sensitivity Analysis D by excluding all previously included patients with either indeterminate serum  $\beta$ -hCG results (Analysis A) or with  $\leq 7$  days between index visit and date of conception (Analysis B) (total  $n = 39$ ).

## 2.7 | Sample size

Based on a reported prevalence of 10% false negative urine  $\beta$ -hCG testing on the most similar prior study<sup>10</sup> and with  $\approx 10,000$  unique negative urine  $\beta$ -hCG tests obtained in our ED annually, we estimated that our study would have adequate sample size to determine the overall proportion of false negative tests with narrow 95% CIs as well as among the relevant clinical subgroups (IUP, ectopic pregnancy, other abnormal pregnancy). Based on a prevalence of false negative tests in the ED of 10%, a minimum sample size of  $n = 385$  would be required to achieve a 3% margin of error for the 95% CI.

## 3 | RESULTS

We identified a total of 10,924 unique patients meeting inclusion criteria, of which 383 met criteria for chart review and 306 had evidence of pregnancy during the index visit or 90-day follow-up period (Figure 1). Of these, a total of 171 patients had pregnancies that were present at the index visit, reflecting a 1.6% (171/10,924) falsely negative omission rate of urine POC testing (95% CI = 1.4, 1.8). Demographics are shown

**TABLE 4** Demographics of patients found to have pregnancies present at index visit

	Pregnancy present at index visit (n = 171)
Age (years), mean (SD)	29.20 (6.87)
Race, n (%)	
American Indian or Alaska Native	0 (0.00)
Asian	5 (2.92)
Black or African American	73 (42.69)
Hispanic or Latino	58 (33.92)
Native Hawaiian or other Pacific Islander	0 (0.00)
White	22 (12.87)
Missing	13 (7.60)
English as primary language, n (%)	103 (60.23)
Serum $\beta$ -hCG performed at index complaint	124 (72.51)
Same-day diagnosis, n (%)	126 (73.68)
High-risk chief complaint, n (%)	97 (56.73)
Serum $\beta$ -hCG performed with high-risk complaint, n (%)	84 (86.6)

in Table 4, including a mean age of 29.2 years and a majority of women being non-White but English-speaking.

A majority of these patients (72.5%, n = 124) had a positive serum  $\beta$ -hCG drawn at the index visit. An additional 45 patients had pregnancies diagnosed during the follow-up interval that were determined to be present at the index visit. Specific pregnancy outcomes included 71 IUPs (41.5% of total), 12 ectopic pregnancies (7.0%), 77 other abnormal pregnancies (45.0%), with 11 pregnancies of unknown outcome (6.4%). The clinical course for each ectopic case is summarized in Table 5, showing that all but one of these cases had symptoms defined as “high risk” or had been referred for pregnancy testing (n = 1). (Note that despite this chief complaint, this one patient met the inclusion criteria of having a negative POC urine  $\beta$ -hCG.) All of these cases had serum  $\beta$ -hCG testing performed at the index visit, although in one case, this result was not known to the physicians at the index visit and the ectopic pregnancy was diagnosed in follow-up.

Of the 2732 patients defined as a “high-risk” subgroup by chief complaint, a total of 139 patients had pregnancies within the follow-up interval, of which 97 were present at the index visit. Within this “high-risk” subgroup, the false negative omission rate of urine pregnancy testing was 3.6% (97/2,732, 95% CI = 2.9, 4.3) and accounted for over half (56.7%) of the total false negative results. These included 10 patients with ectopic pregnancies (83% of total ectopics), 25 patients with IUPs (35%), 58 with other abnormal pregnancies (75%) and 4 with unknown pregnancy outcomes (36%). The relative risk of a pregnancy diagnosis was almost 4 times greater in the “high-risk” subgroup as compared to the risk of a pregnancy diagnosis to those without “high-risk” complaints (relative risk of a pregnancy diagnosis 3.9, 95% CI = 2.9, 5.3) (Table 6).

Concurrent serum  $\beta$ -hCG values ranged from 2 mIU/mL to above assay at >225,000 mIU/mL. As seen in Figure 2, a majority of these results fall in the lower end of the range ( $\leq$ 500 mIU/mL), including some below the stated threshold for our urine POC test at 20 mIU/mL. However, a large majority fall above this stated threshold, with a broad distribution and a second peak of values above 25,000 mIU/mL.

Multiple sensitivity analyses were performed to account for indeterminate serum  $\beta$ -hCG results and for inherent uncertainty in pregnancy dating, with results shown in Table 7. Resulting estimates of false negative prevalence ranged from 1.2% (with all borderline and reclassified cases excluded) to 1.8% (with all included) for our overall population. A similar analysis on those with concerning chief complaints yielded a range of false negative prevalence values from 2.9% to 3.9%.

#### 4 | LIMITATIONS

As a single center study, our results reflect a specific POC assay, specific processes, and contextual factors that may limit external validity. Different urine POC assays with claimed lower limit of detection (LLOD) of 20 U/L have different sensitivities and testing thresholds. For example, the OSOM hCG Combo Test used by Woo et al. OSOM did not reliably give a positive result until the concentration reached 100 U/L, and the performance of the Sekisui OSOM Card Pregnancy Test used at our site has not been reported.<sup>20,21</sup> Our findings also reflect the specific implementation and clinical practice at our center, rather than the diagnostic performance of the test itself. All tests were performed by trained clinical with regular quality assurance and monitoring but results may differ in absence of these systems.

Our results also include some patients who were >8 weeks from their LMP despite a specific manufacturer’s caution about the performance of the assay beyond 8 weeks of gestation (ie, “After 8 weeks of gestation, fragmented hCG represents the dominant molecule in urine. Therefore, using this test kit beyond 8 weeks gestation can produce false negative results during pregnancy”). This may have increased our false negative results, but also reflects real-world clinical practice in which many busy emergency medicine physicians or other clinical staff may not ask, document, trust or change testing decisions based on the stated LMP. Our results would not generalize to settings where the LMP is universally recorded and where testing with urine  $\beta$ -hCG is restricted to patients according to a specific assay’s IFU.

Our results also are limited by the retrospective chart review design, which may have underestimated pregnancy outcomes among patients who were lost to follow up or whose follow-up care was provided at another facility. However, we took multiple steps to try to minimize missing data resultant from lost to follow up in this study, including review of all subsequent available records, both inpatient, outpatient, and linked EMR systems to identify outcomes. Any pregnancy outcomes not captured by follow-up data would bias toward a conservative estimate of falsely negative results.

There is inherent uncertainty in determining the date of conception for naturally occurring pregnancies through either LMP or ultrasound. We, therefore, followed specific ACOG criteria, consulted with

**TABLE 5** Clinical course for ectopic pregnancies present at index visit (n = 12)

Age (years)	Chief complaint (text)	Serum $\beta$ -hCG at index visit (mIU/mL)	Ectopic diagnostic criteria	Management	Complications
28	Hand laceration	n/a	ultrasound with 2 mm yolk sac in adnexa	Methotrexate (MTX)	Unknown (lost to follow-up)
38	Vaginal bleeding	82	Uterine pathology with no products of conception	MTX (2 doses)	None
42	Abdominal cramping	216	Uterine pathology with no products of conception	MTX (2 doses)	Infertility
36	Abdominal pain	469	Uterine pathology with no products of conception	MTX	None
33	Pregnancy problem	1140	Uterine pathology with no products of conception	MTX (2 doses)	Recurrent ectopic pregnancy leading to salpingectomy
22	Abdominal pain	30	Uterine pathology with no products of conception	MTX	Delayed diagnosis. (Serum resulted after patient eloped from index visit.)
32	Vaginal bleed-pregnant	449	Inappropriate rise in hCG, "likely tubal abortion" per gynecology notes	Patient refused	Persistent hydrosalpinx with salpingitis
28	Abdominal pain	223	Uterine pathology with no products of conception	MTX	None
24	Abdominal pain-pregnant	628	Surgical pathology	Emergent surgery after rupture	Rupture, hemorrhage (blood loss 315 mL)
39	Abdominal pain	330	Inappropriate rise in hCG, no intrauterine pregnancy, patient refused uterine curettage	MTX	None
24	Vaginal bleed-pregnant	86	Uterine pathology with no products of conception	MTX	None
35	Vaginal bleeding	172	Uterine pathology with no products of conception	MTX	None

**TABLE 6** Relative risk of pregnancy at index visit

	High risk	Not high risk	Total
Pregnant at index	97	74	171
Not pregnant at index	2635	8188	10753
Total	2732	8192	10924

	Point estimate	95% CI	
Risk difference	0.26	0.019	0.034
Risk ratio	3.93	2.91	5.30

CI, confidence interval.

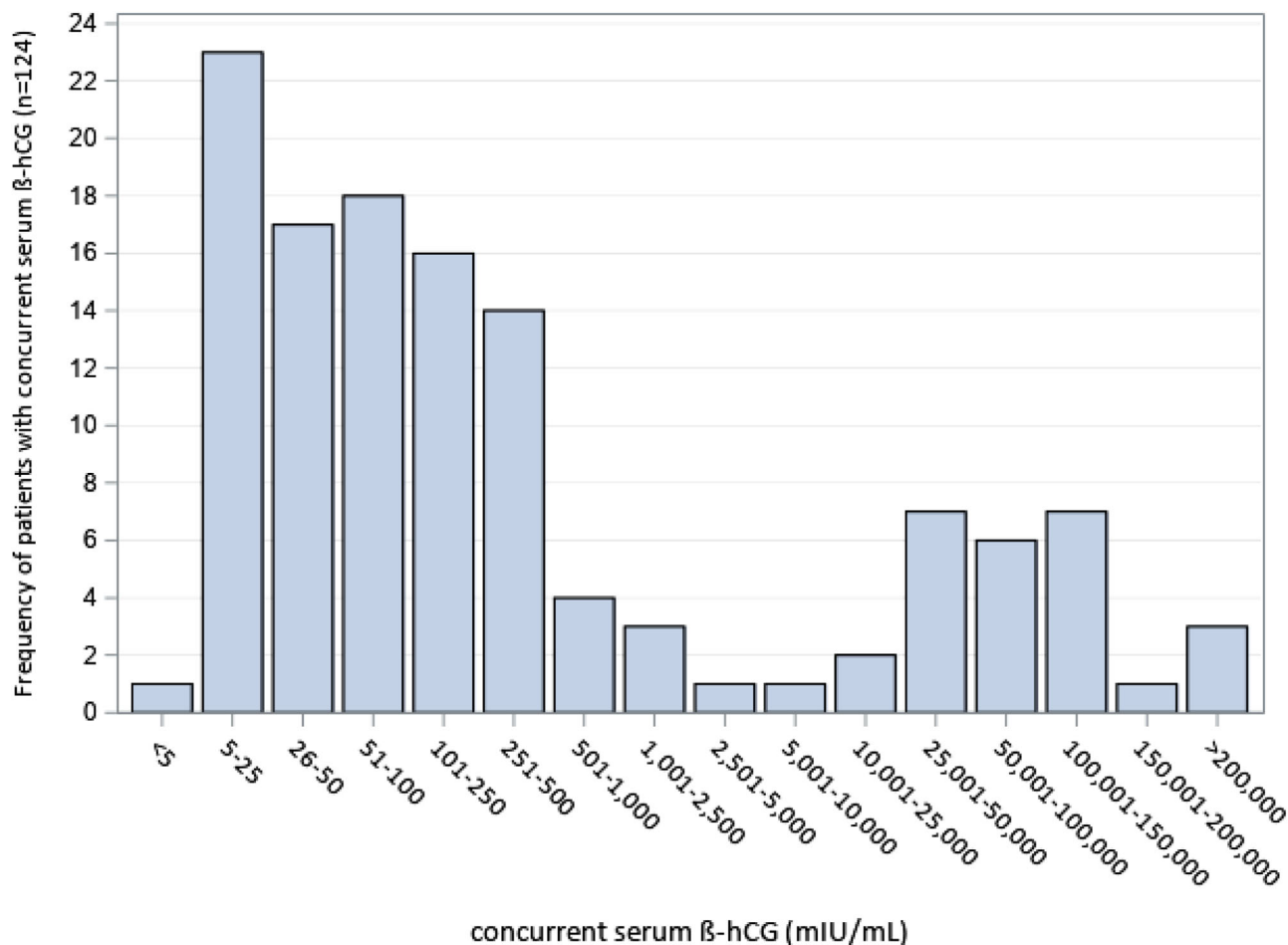
an obstetric specialist (GM) to review coding criteria, and performed a sensitivity analysis with any case having  $\leq 7$  days between estimated date of conception and the index visit. The majority of our missed pregnancy cases were confirmed by testing at the index visit and our sensitivity analysis showed a narrow range of false negative results. Ideally, future studies could prospectively obtain serum  $\beta$ -hCG tests for all

patients with negative urine testing, although this may not be logistically feasible in an ED setting.

Our sample size, although based on the most similar prior ED study, precluded an analysis of patient level factors associated with false negative results, as well as any analysis of other factors affecting decision-making.

## 5 | DISCUSSION

We found that POC urine pregnancy testing yielded a false negative omission rate of 1.6%. Because we included patients who had concurrent pregnancy testing (serum  $\beta$ -hCG or ultrasound) as well as those who did not have testing at the index visit, our study provides a more accurate and complete understanding of the limitations associated with using a POC urine  $\beta$ -hCG approach to screen patients of childbearing potential for pregnancy. Our estimate is higher than the prior estimates of Griffey who reported a 0.34% false negative rate but only considered those with concurrent serum  $\beta$ -hCG testing,<sup>6</sup> as well as Woo,<sup>10</sup> who found a false negative rate of 10.8% with a FOR of 0.9% (including follow-up diagnoses within 3 months of the index visit).



**FIGURE 2** Serum  $\beta$ -hCG results (mIU/mL) for patients with concurrent serum testing at index visit

**TABLE 7** Sensitivity analyses for false negative rate susceptibility to indeterminate  $\beta$ -hCG results and dating uncertainty

Sensitivity analysis	Among all patients				Among high-risk patients			
	$\Delta n$	Total pregnancies	FN rate	95% CI	$\Delta n$	Total pregnancies	FN rate	95% CI
A. Excluding indeterminate $\beta$ -hCG (5-25 mIU/mL) from pregnancy outcome	-19	152	1.4	1.2, 1.6	-16	81	3.0	2.4, 3.7
B. Excluding $\leq 7$ days after index from pregnancy present at index visit	-20	151	1.4	1.2, 1.6	-3	94	3.4	2.8, 4.2
C. Including $\leq 7$ days before index in pregnancy present at index visit	+21	192	1.8	1.5, 2.0	+5	102	3.7	3.1, 4.5
D. Excluding indeterminate hCG and $\leq 7$ days after index (both A and B)	-39	132	1.2	1.0, 1.4	-19	78	2.9	2.3, 3.6

CI, confidence interval; FN, false negative

The majority of concurrent serum results were above the stated limits of detection for urine testing and in the lower ranges of concentrations (Figure 2), confirming that POC assays may be susceptible to false negative results at low  $\beta$ -hCG concentrations due to dilute urine samples.<sup>21</sup> Our results also show a second peak of concurrent high serum  $\beta$ -hCG concentrations (>25,000 mIU/mL), consistent

with the previously-documented "hook effect" from excess  $\beta$ -hCG fragments.<sup>6,9,22</sup> However, we also found a broad range of serum values, consistent with prior study of the FDA's Manufacturer and User Facility Device Experience (MAUDE) database, which showed that most false negative results could not be attributed to the lower limit of detection or the hook effect.<sup>23</sup> Although some false negative results



may reflect process errors, such as faulty documentation or deviations from assay procedures, previous studies have shown equivalent false negative rates between POC and laboratory urine testing.<sup>4</sup>

Clinical decisions around pregnancy testing should incorporate the indication for the test, the implications of a false negative result, the pre-test suspicion for pregnancy, as well as test characteristics. For instance, urine testing meets current guidelines for most radiography and clinical trials, and some studies suggest that sensitivity may be further strengthened by history and clinical suspicion.<sup>24,25</sup> However, our results suggest that a urine POC testing strategy will miss a small but measurable number of pregnancy diagnoses, and failure to diagnose an early healthy pregnancy still risks harm via teratogenic medications, radiographic exposure, missed opportunities for prenatal care, and impacts to patient autonomy and trust of healthcare (as illustrated in Supporting Information Appendix 2).

Missed ectopic pregnancies have both higher mortality and medicolegal risks, as well as a documented higher risk of false negative urine results due to lower  $\beta$ -hCG values.<sup>26,27</sup> We found a total of 12 ectopic pregnancies that presented with negative urine testing, consistent with prior studies and case reports.<sup>6,7,8</sup> In addition, we identified a priori a “high-risk” subgroup of patients with symptoms suggestive of ectopic or other early pregnancy complications, including abdominal pain, pelvic pain and vaginal bleeding, who indeed had a higher rate of false negative urine tests (3.6% compared to 1.6% for the overall sample).

Given the anecdotal risks of false negative urine testing, many emergency physicians already use serum testing to exclude pregnancy. For instance, we found that the majority of false-negative urine tests were recognized through serum testing at the index visit, implying that many clinical staff at our site employed a combined or serial testing strategy in which  $\beta$ -hCG serum testing is performed for high-risk patients despite a negative urine test. However, this practice is not universal or standardized, with even the most recent National Quality Forum Consensus Standards for Emergency Care (ED-018-08) only requiring that “women, ages 14–50 years old who present to the ED with a chief complaint of abdominal pain, have a pregnancy test (*urine or serum*) ordered in the ED”<sup>15</sup> (italics added). Although almost all patient encounters meet this standard,<sup>9</sup> our current results suggest that urine testing alone is not sufficient to rule out ectopic pregnancy, and that further decision tools or guidelines could benefit from the greater sensitivity of serum testing if high risk populations can be identified.

One reported benefit to urine POC testing is decreased length of stay and resource utilization in the age of overcrowding.<sup>28</sup> One study of urine pregnancy tests did confirm that a POC test had a much shorter turnaround time, taking 7.6 versus 67.4 minutes for laboratory testing, although this does not include the highly-variable time to obtain a urine sample.<sup>29</sup> Although a POC approaches using both serum and capillary whole blood have been explored and validated, all of these studies have also reported similar negative predictive values for both whole blood and urine (97.9% and 97.6%, respectively).<sup>29,30</sup> Even the i-Stat POC assay designed for use with whole blood has demonstrated false negatives due to a hook effect at high  $\beta$ -hCG levels.<sup>31</sup> While further trials are ongoing, it may be impossible for an antibody-driven, POC test to reach the sensitivity of quantitative serum laboratory testing.

## 6 | CONCLUSIONS

We examined a large retrospective cohort of female patients of childbearing age who had negative POC urine pregnancy testing in our ED. We found that an important number of these results were falsely negative, including patients with ectopic, intrauterine, and abnormal pregnancies. The relative risk of false negative results was even greater among patients with abdominal pain or vaginal bleeding, a subgroup that included almost all of our ectopic cases. These false negative tests can contribute to diagnostic errors and pose a risk to patient safety and clinicians should use caution when relying on urine POC testing to rule out pregnancy. Serum  $\beta$ -hCG testing, either as the initial screening test or obtained following a negative urine  $\beta$ -hCG test, is the gold standard test for ruling out pregnancy in an ED patient of childbearing potential.

## ACKNOWLEDGMENTS

We thank Ms. Linda Rosen and the Clinical Data Warehouse at Boston Medical Center for the development of the cohort through the data warehouse, and Dr. Glenn Markenson, Professor of Obstetrics & Gynecology, Boston University School of Medicine, Boston MA for his role in developing the coding structure and reviewing specific cases for determining the dating of conception, as well as research assistants Samantha Roberts, MPH and Tyler Pina who contributed countless hours to chart review. Ms. Kathleen Shea assisted in the manuscript preparation.

## AUTHOR CONTRIBUTIONS

All of the listed authors have participated actively in the entire study project, including study design, data acquisition, analysis, and manuscript preparation. This research represents part of the residency research requirement for SK. SK and JAF developed the design and conduct of the study. JND and KPN participated in the data analysis, interpretation and manuscript preparation. SK and JAF drafted the original manuscript. All authors participated in and approved the final submission. JAF assumes responsibility for the paper as a whole.

## CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

## REFERENCES

1. Bastian LA, Piscitelli JT. Is this patient pregnant?: can you reliably rule in or rule out early pregnancy by clinical examination? *JAMA*. 1997;278(7):586-591.
2. Saul T, Lewiss RE, Rivera MDR. Accuracy of emergency physician performed bedside ultrasound in determining gestational age in first trimester pregnancy. *Crit Ultrasound J*. 2012;4(1):22.
3. Goyal M, Hersh A, Luan X, et al. Frequency of pregnancy testing among adolescent emergency department visits. *Acad Emerg Med*. 2013;20(8):816-821.
4. Lazarenko GC, Dobson C, Enokson R, Brant R. Accuracy and speed of urine pregnancy tests done in the emergency department: a prospective study. *CJEM*. 2001;3(4):292-295.
5. Pejic W, Argáez C. *Point-of-Care Urine Pregnancy Screening in the Emergency Department: Diagnostic Accuracy, Clinical Utility, and Guidelines*. Ottawa, Ca.: CADTH; 2017:1-7.

6. Griffey RT, Trent CJ, Bavolek RA, Keeperman JB, Sampson C, Poirier RF. Hook-like effect" causes false-negative point-of-care urine pregnancy testing in emergency patients. *J Emerg Med.* 2013;44(1):155-160.
7. Sheele JM, Bernstein R, Counselman FL. A ruptured ectopic pregnancy presenting with a negative urine pregnancy test. *Case Rep Emerg Med.* 2016;2016:7154713.
8. Daniilidis A, Pantelis A, Makris V, Balaouras D, Vrachnis N. A unique case of ruptured ectopic pregnancy in a patient with negative pregnancy test—A case report and brief review of the literature. *Hippokratia.* 2014;18(3):282-284.
9. Hunter CL, Ladde J. Molar pregnancy with false negative  $\beta$ -HCG urine in the emergency department. *West J Emerg Med.* 2011;12(2):213-215.
10. Woo K-M, Director T, Sweeney C, Cristales D, Deguia J, Baumlin K. False negative point-of-care urine pregnancy results in the emergency department: quantifying a needle in the haystack in the clinical setting. *Ann Emerg Med.* 2015;66(4):S150.
11. ACOG. Committee opinion no 700: methods for estimating the due date. *Obstet Gynecol.* 2017;129(5):e150.
12. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *PLoS Med.* 2007;4(10):e296.
13. Nagy GK. False negative rate. *Acta Cytologica.* 1997;41(3):778-780.
14. Schuur JD, Tibbetts SA, Pines JM. Pregnancy testing in women of reproductive age in US emergency departments, 2002 to 2006: assessment of a national quality measure. *Ann Emerg Med.* 2010;55(5):449-457.
15. National Quality Forum (NQF). *National Voluntary Consensus Standards for Emergency Care: A Consensus Report.* Washington, DC: NQF; 2009.
16. Kaji AH, Schriger D, Green S. Looking through the retrospectroscope: reducing bias in emergency medicine chart review studies. *Ann Emerg Med.* 2014;64(3):292-298.
17. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform.* 2009;42(2):377-381.
18. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33(1):159-174.
19. ACOG. ACOG practice bulletin no. 191: tubal ectopic pregnancy. *Obstet Gynecol.* 2018;131(2):e65.
20. Nerenz RD, Gronowski AM. Confirmation of hCG point-of-care qualitative device limitations in the emergency department. *J Emerg Med.* 2015;48(5):609-610.
21. Kamer SM, Foley KF, Schmidt RL, Greene DN. Analytical sensitivity of four commonly used hCG point of care devices. *Clin Biochem.* 2015;48(6):448-452.
22. Winder AD, Mora AS, Berry E, Lurain JR. The "hook effect" causing a negative pregnancy test in a patient with an advanced molar pregnancy. *Gynecol Oncol Rep.* 2017;21:34-36.
23. Nerenz RD, Gronowski AM. Qualitative point-of-care human chorionic gonadotropin testing: can we defuse this ticking time bomb?. *Clin Chem.* 2015;61(3):483-486.
24. Abushouk AI, Sanei Taheri M, Pooransari P, Mirbaha S, Rouhipour A, Baratloo A. Pregnancy screening before diagnostic radiography in emergency department; an educational review. *Emerg (Tehran).* 2017;5(1):e60.
25. Minnerop MH, Garra G, Chohan JK, Troxell RM, Singer AJ. Patient history and physician suspicion accurately exclude pregnancy. *Am J Emerg Med.* 2011;29(2):212-215.
26. Stovall TG, Kellerman AL, Ling FW, Buster JE. Emergency department diagnosis of ectopic pregnancy. *Ann Emerg Med.* 1990;19(10):1098-1103.
27. Robertson JJ, Long B, Koefman A. Emergency medicine myths: ectopic pregnancy evaluation, risk factors, and presentation. *J Emerg Med.* 2017;53(6):819-828.
28. Rooney KD, Schilling UM. Point-of-care testing in the overcrowded emergency department—can it make a difference?. *Critl Care.* 2014;18(6):692.
29. Fromm C, Likourezos A, Haines L, Khan ANGA, Williams J, Berezow J. Substituting whole blood for urine in a bedside pregnancy test. *J Emerg Med.* 2012;43(3):478-482.
30. Gottlieb M, Wnek K, Moskoff J, Christian E, Bailitz J. Comparison of result times between urine and whole blood point-of-care pregnancy testing. *West J Emerg Med.* 2016;17(4):449-453.
31. Wilgen U, Pretorius CJ, Gous RS, Martin C, Hale VJ, Ungerer JPJ. Hook effect in Abbott i-STAT  $\beta$ -human chorionic gonadotropin ( $\beta$ -hCG) point of care assay. *Clin Biochem.* 2014;47(13-14):1320-1322.

#### AUTHOR BIOGRAPHY



Sarah Kleinschmidt, MD, is a Resident of Emergency Medicine at Boston University School of Medicine, Boston, Massachusetts.

#### SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**How to cite this article:** Kleinschmidt S, Dugas JN, Nelson KP, Feldman JA. False negative point-of-care urine pregnancy tests in an urban academic emergency department: a retrospective cohort study. *JACEP Open.* 2021;2:e12427. <https://doi.org/10.1002/emp2.12427>