PROCEEDINGS OF THE SOCIETY FOR EXPERIMENTAL BIOLOGY AND MEDICINE **184,** 201–205 (1987)

Verification of Early Pregnancy Tests in a Multicenter Trial (42468)

LOIS JOVANOVIC,* MUKUL SINGH,* BRIJ B. SAXENA,* JAMES L. MILLS,† DAN TULCHINSKY,‡ LEWIS B. HOLMES,‡ JOE LEIGH SIMPSON,§ BOYD E. METZGER,§ ANDREW LABARBERA,‖ JEROME AARONS,‖ MARGO I. VAN ALLEN,# AND THE NICHD-DIEP STUDY GROUP[1]

**Cornell University Medical College, New York, New York 10021; †National Institutes of Health, Bethesda, Maryland 20852; ‡Harvard University, Cambridge, Massachusetts 02215; §Northwestern University, Chicago, Illinois 60611; ‖University of Pittsburgh, Pittsburgh, Pennsylvania 15213; #University of Washington, Seattle, Washington 98105*

Abstract. Tests for the diagnosis of early pregnancy have been available since 1974. However, no studies have systematically verified the accuracy of routine clinical laboratories in measuring human chorionic gonadotropin (hCG) prior to the time that pregnancy is clinically evident. We have conducted such a study in association with the NICHD-funded Diabetes in Early Pregnancy Study (DIEP). The purpose of this study was to elucidate the etiology of malformations in pregnancies complicated by diabetes mellitus, which probably occurs within the first few weeks of pregnancy, and therefore uniformity of pregnancy testing was necessary among the five centers to find an association of a teratogen at the time of organogenesis. We confirmed that routine clinical laboratories, in fact, could measure accurately hCG at the time of the missed menses; however, detection was not necessarily possible prior to that time. We conclude that in order to assure accurate diagnosis of early pregnancy, tests should ordinarily be delayed until time of the missed menses. When the test is used at this time, it is a reliable tool for early pregnancy testing and thus can be used to resolve questions relating to early pregnancy pathophysiology.

Early diagnosis of pregnancy is important for numerous psychological, social, and medical reasons. Although tests for the early diagnosis of pregnancy have been available for over 5 years (1, 2), validation of those tests commonly used to detect pregnancy 2–5 days after the first missed menstrual period has not been reported, nor has its use been proven to be a reliable tool for multicenter clinical trials designed to answer questions related to teratogenesis. The Diabetes in Early Pregnancy Project, an NICHD-funded five university trial, was designed to elucidate the pathogenesis of malformations of infants of diabetic mothers (3). Because the first 8 weeks of pregnancy are the critical period of organogenesis (4), it was essential that events during this period be monitored in diabetic and nondiabetic (control) women. To accomplish this goal, subjects were recruited prior to conception and conception was documented by basal body temperature measurement and a human chorionic gonadotropin test within 2 days of missed menses. Early diagnosis of pregnancy was especially necessary in order to determine whether bleeding during this early interval was caused by early spontaneous abortion or merely from a delayed period. A validation procedure was sought to ensure that the diagnosis for pregnancy was made accurately as early as feasible in all five centers, and thus the spontaneous abortion rates could be compared. This paper describes the validation procedure and results.

Materials and Methods. Each of the five universities recruited nonpregnant women in their childbearing years. One-half of the subjects recruited were known to have insulin-dependent diabetes mellitus (IDDM); the other half were women with no history of diabetes mellitus, either personally or in first-degree relatives. All women were attempting

[1] NICHD: J. L. Mills, M. Withiam, E. Harley, G. Reed; Cornell University Medical College: L. Jovanovic, A. Duckles, J. Doyle; Harvard University: L. B. Holmes, C. Wasson, C. B. Bogal; Northwestern University: J. L. Simpson, B. E. Metzger, C. Ober; University of Pittsburgh: J. Aarons, R. Laporte, P. Meyer; University of Washington: R. H. Knopp, R. A. Brown, B. Plovie, M. L. Hanson.

0037-9727/87 $1.50

actively to become pregnant during the period of recruitment. Subjects were taught to record their basal body temperature (BBT) in order to document ovulation. A blood sample was obtained for estimating circulating human chorionic gonadotropin (hCG) in serum or plasma if the period was missed by more than 1 day or if the temperature remained elevated 16 days post-thermal nadir in BBT.

The following five methods were utilized in the assays:

I. The laboratory at the Brigham and Women's Hospital, Boston, Massachusetts, utilized a radioimmunoassay (RIA) of hCG by a modification of the double antibody technique described by Vaitukaitis *et al.* (5). The high titer antiserum was obtained from rabbits immunized with the β subunit of hCG (6). After an initial incubation of standards (2nd IRP-hCG) and samples with the first antibody at 37°C for 2 hr, I^{25}-labeled hCG was added and incubated overnight at 4°C for 15–17 hr. Thereafter the sheep anti-rabbit γ globulin was added as a second antibody. The antibody-bound hormone then was precipitated by 6% polyethylene glycol and separated by centrifugation.

II. The laboratory at the University of Pittsburgh utilized a serum radioimmunoassay kit, provided by GammaDab clinical assays, Travenol Laboratories, Deerfield, Illinois. This is also a double antibody procedure with a processed rabbit anti-β-hCG serum and a sheep anti-rabbit serum. The hCG standards and controls in this kit are calibrated according to the World Health Organization immunoassay standard (the First International Reference Preparation (IRP) 75/537).

III. The laboratory at the University of Washington, Seattle, Washington, used a serum β-hCG radioimmunoassay kit manufactured by American Biochemical, Portland, Oregon. This kit also employs the double antibody separation technique and the IRP-hCG standards.

IV. The laboratory at Northwestern University, Chicago, Illinois, performed the test on plasma instead of serum. A modification of the double antibody technique of Vaitukaitis *et al.* (4), was used; the reference standard used was the Second International Reference Preparation.

V. In the laboratory at Cornell University Medical College, New York, New York, the intact hCG molecule was measured by radioreceptorassay (RRA), utilizing bovine ovarian receptors as described by Saxena *et al.* (1). The Second International Reference was used.

VI. The central hCG laboratory at Duke University, used by the trial as the central laboratory for confirmation of the local laboratory results and for data analysis, also modified the technique of Vaitikaitis *et al.* (4). The Second International Reference was used.

To verify uniformity of assays among centers, seven test samples were prepared by adding known weighed amounts of purified hCG (1200 mIU/ml) and prepared by adding 1 ml of male serum. The samples for Northwestern were prepared exactly the same way except that the hCG was added to 1 ml of male plasma. The concentrations of hCG per 1-ml sample were as follows: 0, 150, 300, 600, 1200, 2400, and 4800 mIU/ml.

All seven samples were first assayed for hCG by RRA to verify that the sample had been measured properly. Each sample then was divided into two aliquots. One was labeled as a "research" sample, whereas the other was labeled with a fictitious name. This allowed each center to send the blood to their laboratory labeled with "research sample" which usually means "extra care" and with a patient name which may mean less than optimal handling. All samples were frozen immediately at −20°C and the 14 tubes were sent to each of the five universities. All specimens were received frozen at all centers within 24 hr of shipping.

Because there was relatively more difficulty in assaying lower levels of hCG, a second verification procedure later was undertaken to quantify the ability of all centers to measure less than 50 mIU/ml.

In the second verification procedure, male serum recovered from 500 ml of blood was used to prepare unknowns in the manner described above except that the purified hCG was weighed and added at concentrations of 1, 10, 20, 30, 40, and 50 mIU/ml. These samples were sent to each center in duplicate following the same procedures outlined above, on a separate occasion. The rationale for the "second" verification was based on observations that the hCG level at the time of missed period is ap-

TABLE I. FIRST VERIFICATION (CONCENTRATIONS OF hCG ARE EXPRESSED IN mIU/ml)

Tube number		RRA confirmation of weighed amounts	University of Pittsburgh[a]	Brigham and Women's Hospital	University of Washington	Cornell	Northwestern
1	Research	150	230	80	85	180	76
	Pseudo PT		245	80	91	228	109
2	Research	300	320	168	181	252	163
	Pseudo PT		320	qns	160	238	256
3	Research	600	285	400	408	600	527
	Pseudo PT		420	370	438	516	599
4	Research	1200	500	690	750	1330	1224
	Pseudo PT		451	730	808	1044	782
5	Research	2400	490	1200	1240	2196	3170
	Pseudo PT		500	910	1380	2400	1323
6	Research	4800	500	2050	2160	4536	3155
	Pseudo PT		449	2450	2720	2652	2278
7	Research	0	34	10	7	22.8	0.7
	Pseudo PT		13	10	5	22.8	0.1
	Cutoff for positive	10	25	30	30	30	30
Decision on tube No. 7		Not pregnant	Pregnant	Not pregnant	Not pregnant	Not pregnant	Not pregnant
(Research, Pseudo PT)		Not pregnant	Not pregnant	Not pregnant	Not pregnant	Not pregnant	Not pregnant

Note. qns, quality not sufficient; mIU/ml, milli international units per milliter; Research, sample labeled as a special research sample; Pseudo PT, sample coded with a fictitious patient name. Pittsburgh used the first International Standard and therefore to compare Pittsburgh's data with the other centers, divide by 2.

TABLE II. SECOND VERIFICATION (CONCENTRATIONS OF hCG ARE EXPRESSED IN mIU/ml)

Tube number		Central lab weighed amounts	University of Pittsburgh	Brigham and Women's Hospital	University of Washington	Cornell	Northwestern
1	Research	100	200	83	100	102	100
	Pseudo PT		220	87	100	108	95
2	Research	77	150	69	75	90	90
	Pseudo PT		150	70	80	96	133
3	Research	0	5	10	5	5	1
	Pseudo PT	0	5	10	5	5	2
4	Research	61	95	45	58	80	124
	Pseudo PT		105	51	52	86	72
5	Research	124	218	105	120	148	180
	Pseudo PT		210	103	120	136	176
6	Research	30	60	26	26	30	32
	Pseudo PT		60	24	28	27	62
	Cutoff for positive	30	25	30	30	30	10
Decision on tube No. 7		Equivocal (repeat needed)	Pregnant	Equivocal (repeat needed)	Equivocal (repeat needed)	Equivocal (repeat needed)	Equivocal (repeat needed)

proximately 200 mIU/ml (8). Since the DIEP was designed to assess the incidence of spontaneous abortions in women with diabetes as compared with that of normal controls, false positive or negative pregnancy tests coupled with subsequent vaginal bleedings would negate the study.

Results. In the first verification procedure there was general agreement (Table I), with only one center showing inconsistencies. For example, the zero tubes were assayed, respectively, as 7, 0.7, 34, 22.8, and less than 10 mIU/ml (the 34 mIU/ml tube would have been called "positive" (see Table I).

In the second verification, which included the central laboratory, all laboratories identified correctly the serum without hCG (Table II). However, the unknown with a quantity of hCG in the ranges of LH cross-reactivity at the time of expected period was called positive by two of the laboratories. The other four laboratories considered the result equivocal.

All six laboratories had adequate intra- and interassay agreement ("research" vs pseudo patient). All laboratories were able to detect hCG in levels compatible with those observed at the time of the first missed menstrual period (less than 200 mIU/ml). Even the large interassay variation in one center failed to affect the ability of that laboratory to detect low levels of hCG.

Discussion. The Diabetes in Early Pregnancy (DIEP) Project depended on the ability of each collaborating center to diagnose pregnancy at the time of the first missed menstrual period. Accurate diagnosis of pregnancy based on serum hCG levels was also essential for the subsequent management of study participants. False positive pregnancy diagnosis would lead to a menses being misinterpreted as a spontaneous abortion. Since the DIEP was designed to assess the incidence of spontaneous abortions in women with diabetes as compared with that of normal controls, as well as to document teratogens early in pregnancy, false positive or negative pregnancy tests coupled with subsequent vaginal bleeding would negate the study. This study not only shows that it is possible for a multicenter study to use selected local laboratories for the diagnosis of early pregnancy, but also has implications for the clinician in general. Routine clinical laboratories contracted by each of the five univer-

sities to perform the serum hCG assay had no trouble detecting hCG normally seen at 3 weeks post conception (i.e., greater than 200 mIU/ml) (7). Two laboratories reached an endpoint in the ability to measure absolute amounts of hCG at higher levels (greater than 1200 mIU/ml); however, this plateau did not interfere with the sensitivity of the assay in the lower ranges. That not all assays measure hCG accurately when the levels are elevated could be important when assessing hydatidiform mole and multiple gestation (8).

Of relevance is that hCG levels before the first missed menstrual period are similar to those levels of LH in the initial phase of the menstrual cycle. A low level of hCG (<30 mIU/ml) in tube 6 was not consistently measured accurately. The "correct" interpretation of tube 6 would have been to consider the results as equivocal and to draw another blood sample 2–4 days later when the hCG level would be closer to 100 mIU/ml. If the result was still equivocal (<30 mIU/ml) then the test would be measuring LH.

The clinical lessons learned from our verification procedures can thus be summarized as follows:

(i) Routine clinical laboratories are indeed capable of diagnosing pregnancy at the time of missed period (equivalent to 14 days post conception).

(ii) However, routine clinical laboratories cannot diagnose pregnancy consistently before 14 days post conception. If doubt exists, tests should be repeated in 2 days.

(iii) Routine clinical laboratory hCG tests cannot be used to monitor hydatiform male or twin gestation.

(iv) The study points out the importance of introducing a verification procedure before clinical studies are initiated which have as their study design a need to document pregnancy accurately as early as possible.

1. Saxena BB, Hasan SH, Haour F, Schmidt-Gollsitzer IM. Radioreceptor-assay of human chorionic gonadotropin: Detection of early pregnancy. Science **184:** 793–795, 1974.
2. Kosasa TS, Levesque LA, Goldstein DR, Taymor ML. Clinical use of a solid phase radioimmunoassay specific for human chorionic gonadotropin. Amer J Obstet Gynecol **118:**784–786, 1974.
3. Mills JL, Fishl AR, Knopp RHJ, Ober CL, Jovanovic LG, Polk BF. Malformations in infants of diabetic mothers: problems in study design. Preventive Med **12:**274–286, 1983.
4. Mills JL, Baker L, Goldman AS. Malformations in infants of diabetic mothers occur before the seventh gestational week: Implications for treatment. Diabetes *28:*92–97, 1979.
5. Vaitukaitis JL, Braunstein GD, Ross GT. A radioimmunoassay which specifically measures human chorionic gonadotropin in the presence of luteinizing hormone. Amer J Obstet Gynecol **113:**751–755, 1972.
6. Vaitukaitis JL, Robbins JB, Neischling E, *et al.* A method for producing specific antisera with small doses of immunogenic hCG J Clin Endocrinol **33:** 988–993, 1971.
7. Saxena BB, Landesman R. Diagnosis and management of pregnancy by radio-receptorassay of human chorionic gonadotropin. Amer J Obstet Gynecol **131:** 97–107, 1978.
8. Jovanovic L, Landesman R, Saxena BB. Screening for twins. Science **33:**198, 1977.

Received September 9, 1986. P.S.E.B.M. 1987, Vol. 184.
Accepted October 31, 1986.