THE RELATIVE MERITS OF THE VARIOUS BIOLOGICAL TESTS FOR PREGNANCY

By H. P. Ferreira, M.D.
Chelsea Hospital for Women

Among the many purely magical methods of pregnancy diagnosis practised by the Ancient Egyptians, were tests based on the germination of seeds and rate of plant growth when treated with pregnancy urine (Bayon, 1939). That there is some scientific basis for these tests is shown by the general agreement that a factor contained in pregnancy exerts a beneficial effect on plant growth.

It is of interest to note that Orban (1947) (cited by Cowie, 1948, p. 48), has described a test depending on the rate of opening of gladioli blooms when placed in pregnancy urine.

Modern tests date from Aschheim and Zondek's (1927) observation of the effects of chorionic gonadotrophin on the ovaries of immature mice, and to their subsequent application of this observation as a basis for pregnancy diagnosis.

It is difficult to present a balanced, yet critical, account of the many tests evolved since that time. The use of a particular test in one's laboratory implies a certain bias in its favour. For this reason it is important to have a clear conception of the desirable features of the ideal pregnancy test, and with reference to this to see how much and in what ways tests fall short of the ideal. As defined in the Lancet Editorial (1948): 'Ideally a pregnancy test should be reliable, convenient, simple and speedy; and those involving the patient's participation, fall short of the ideal if they are irksome or require her to visit anyone but her own doctor.'

On the grounds of reliability we may exclude many of the more exotic tests. Thus the development of nupital colouration in the Japanese Bitterling as a positive response is too irregular for use in pregnancy diagnosis (Owen, 1936).

Tests such as the Konsuloff test based on the chromatophore hormone and its many modifications can similarly be condemned. The chromatophore hormone is present in the blood and urine of normal men and women, but is present in increased quantities during pregnancy.

Konsuloff (1934b) using hypophysectomized frogs showed that darkening of the skin caused by dispersal of the pigment granules in dermal melanophores occurred in 1-4 hours following injection of pregnancy urine. He claimed an accuracy equal to that of the Aschheim-Zondek test.

Brühl and Reikoff (1935) encountered many false positives from cases of cancer, tuberculosis and venereal diseases.

Modifications of this test include the use of Ringer perfused frogs' limbs (Raza and Spurrell, 1937), isolated pieces of frog's skin (Simola and Rivas, 1936) and scales from the Prussian carp (Binet, Verne and Luxembourg, 1934) as indicators of the presence of the hormone.

Tests based on the presence of oestrogens such as the Allen-Doisy test are also unreliable and tend to yield a high percentage of false positives.

Tests utilizing male rodents, depend on the growth-promoting effect of gonadotrophins on the testes, and particularly the seminal vesicles, of male rats and mice.

With experience many authors, Kraus (1932), Brouha and Hinglais (1931), found a high degree of accuracy which approached that of the classical Aschheim-Zondek. However, positive results could not be read with accuracy until the eighteenth day.

The remaining pregnancy tests to be considered depend on the effect of gonadotrophins on the ovaries of female rodents, female amphibia and on the testes of male amphibia.

The Aschheim-Zondek Test

The first of these chronologically is the Aschheim-Zondek test utilizing immature, 3-4 weeks old, female mice.

Aschheim and Zondek (1928a and b) examined 276 samples, with only four mistakes.

Five mice were used for each test, 1.2-2.4 ml. of early-morning urine (subdivided into six doses) was injected subcutaneously and the animals killed at 100 hours.

In the original test each mouse received
decreasing doses; thus the first received six doses of 0.5 ml., the second six of 0.4 ml., and so on. It is customary nowadays in most laboratories to give six doses of 0.4 ml. to all five mice, the test being read on the morning of the fifth day.

A positive test is denoted by hyperaemia of the ovaries and haemorrhagic follicles and luteinization and formation of corpora lutea atretica.

These tests are read macroscopically or mounted in glycerine and examined under a low-power lens. Despite a high degree of accuracy there are certain disadvantages. Zondek (1930e) found that 6-7 per cent. urines were so toxic that tests were virtually impossible. This defect has been overcome by the use of ether-treated urine or by various concentration methods, Cowie (p. 26). Certain strains of mice were found to be entirely unsuitable (Hummel, 1942). Zondek (1930e) described a rapid test by use of an alcohol precipitation concentration method. Positives could be reliably read at between 51-57 hours.

The principal defects of the Aschheim-Zondek test as now performed are:

1. Length of time required for completion of the test.
2. Inaccuracies resulting from expressing results on the examination of one or two animals where the others have died.

The Frank Rat Test

A closely-related test is the Frank rat test for pregnancy. The technique was originally described by Reiprich (1933 and 1934). Immature female rats are injected with 10-14 ml. subcutaneously in 2-3 doses over 6-9 hours.

With experience the test could be read at 24 hours with an accuracy similar to the Aschheim-Zondek. A positive result was indicated by hyperaemia of the ovaries which, in the rat (in contrast to the mouse) are rather pale and anaemic normally. Walker and Walker (1938) 'rediscovered' the test which has since been extensively investigated and found reliable. Salmon, Geist, Salmon and Frank (1942) used three rats for their six-hour test. Each rat was injected with a single dose of 2 ml. overnight urine and the animal killed with coal gas after six hours. With this technique they claimed practically 100 per cent. accuracy. Kupperman and Greenblatt (1946) have reported 90.5 per cent. accuracy in 752 tests read at the second hour.

Bunde (1947) achieved an accuracy of 84.5 per cent. using two rats per test and 90.5 per cent. with three animals.

Apart from a report by Farris (1944), who claimed positive results in non-pregnant women in mid-cycle, no false positive results have been recorded.

Elek (1953), using an immature 3-6 weeks' old Wistar No. 1 strain rat titrated for sensitivity once yearly, found 9 per cent. false negative and 5 per cent. false positive results.

The tests were read visually 9-12 hours following injection. The animals were chloroformed and the reading made with the circulation intact. Death of the animal or vessel damage caused rapid blanching of the ovary.

Economically the test is surprisingly expensive, the rats cost approximately 5s. each and unless a constant stream of tests are being performed many rats may be wasted on attaining maturity. However, Kupperman, Greenblatt and Noback (1943) found that adult rats could be used, providing they were in the dioestrus or metoestrous phase of their cycles as judged by vaginal smear examination.

The Friedman Test

Friedman (1929a) introduced the test which bears his name. He showed that a single injection of pregnancy urine was sufficient to produce ovulation in the young virgin rabbit.

The test as performed in most laboratories using the method today, is extremely simple. 10 ml. of early morning urine brought to a pH of 7.4 and filtered, is injected into the marginal ear vein of the animal. The dose is repeated on the following day, and test read on the third day. The ovaries are inspected by opening the abdomen under anaesthesia.

A positive test is denoted by haemorrhagic follicles and fresh corpora lutea; the clear follicles often seen in the rabbit ovary do not constitute a positive test (Fig. 1).

Rabbits may be used again three weeks after a positive test and sooner after a negative one. It is considered undesirable in this laboratory to use a rabbit more than three times, the animal being killed at the end of the third test.

Occasionally a rabbit proves insensitive to gonadotrophins and on this account it is desirable to use two animals for each test. As this is not economically feasible, Bishop (1934) advocated reading the test at 36 hours; if negative, the animal was injected with a known positive urine, and the test read after another 36 hours. If the test still remains negative, the animal is deemed refractory and the pregnancy test is repeated, using another animal. If, however, it is positive, the original pregnancy test is accepted as negative.

Many modifications designed (a) to increase the speed of the reaction and (b) to facilitate reading of the test have been devised, Cowie (pp. 34-36).

In the hands of most observers the Friedman test has attained a degree of accuracy comparable
with the Aschheim-Zondek reaction, together with a considerable saving of time.

Tests using amphibia must clearly be differentiated into those employing the female toad *Xenopus laevis*, and those employing male toads and frogs.

**Female Toad—Hogben Test**

*Xenopus laevis* is unique in being the only female toad which responds by oviposition to injections of gonadotrophins and at the same time spontaneous ovulation never occurs in captivity. Upon this observation made by Hogben (1930) was based the pregnancy tests described independently by Bellerby (1934a) and Shapiro and Zwarenstein (1934a).

The test procedure consists of injecting a urine extract into the dorsal lymph sac of the toad. Various methods of extraction have been advocated, but the method chosen by most laboratories is that of Scott (1940). By this technique, the chorionic gonadotrophin contained in 100 ml. of early-morning urine, is concentrated to a volume of 5 ml.

2.5 ml. of extract is injected into the test animal which is then placed in a glass jar containing a few inches of water. If the toad extrudes its ova within 18 hours the test is recorded as positive.

If negative, a second toad is injected and the final result read 48 hours after the commencement of the test (Fig. 2).

Concentration of the urine is essential to the accurate performance of this test for the following reasons:

1. The *Xenopus* toad is not particularly sensitive to gonadotrophins requiring a dose of 70-100 i.u. to produce a positive response; the use of untreated urine yields a high percentage of false negatives (Geoghegan and McGrath, 1944).

2. Toads are particularly sensitive to urinary toxins and these are eliminated by the extraction process.

3. Menopausal urines treated by precipitation techniques often give false positives, which does not occur when Scott’s extraction method is used (Landgrebe, 1939).

The Hogben test is used in the large pregnancy diagnosis laboratories at Edinburgh and Watford and at these centres excellent results are obtained. Thus Hobson (1932) records an accuracy of 99.8 per cent. in 37,020 and Schwabacher, 99.48 per cent. in 1,512 tests.

On the other hand, smaller laboratories do not achieve such excellent results, possibly due to failure to keep the animals in prime condition and maintain a regular check on sensitivity. Land-
Landgrebe (1939) has shown that pollution of tanks as a result of overcrowding leads to diminished food intake with resulting atrophy of the ovaries. He recommends tanks, 24 by 15 by 12 in., filled to a depth of 6 in. and fed by a trickle of running water at 22°C. as being adequate to house 20 toads.

The toads are fed once per week with raw chopped liver (cubes of 0.5 cm. sides) and the water changed on the day following the feed.

Hobson (1952) uses toads weighing 60-100 g., which have been shown to respond to a priming dose of 70 i.u. gonadotrophin. Furthermore, twice yearly he reprimes all negative toads; those not responding are discarded.

Landgrebe (1948) has shown that toads used every 10 days are more sensitive than those used at 21-day intervals. While the above standards are easy to maintain in a laboratory dealing with large numbers of pregnancy tests, small units find them tiresome to achieve, and out of all proportion to the value of the tests performed.

The Male Toad Pregnancy Test

In 1947 Galli Mainini reported a new pregnancy test employing the South American indigenous species of male toad, B. arenarum Hensel.

This test is based on the observation by Houssay and Lascano Gonzales (1929) that implantation of anterior pituitary in the testis of the male toad resulted in the release of spermatozoa. These pass through the renal tubules and are voided in the urine (Fig. 3).

Using male toads of 100 gm. or over Galli Mainini (1947) showed that 10 ml. untreated urine was sufficient to induce this response. Urine was collected by insertion of a glass cannula into the cloaca of the animal and the specimen examined with the 2/3 objective. The presence of spermatozoa, easily detected at this magnification, indicates a positive response. In a series of 3,156 tests Galli Mainini (1948) achieved an accuracy of 98-100 per cent.

Other workers, A. Pou de Santiago (1947) and Haines (1948) obtained results of comparable accuracy. Since 1947 the test has been carried out all over the world using nearly every species of male toad and frog with almost universal success (Bhaduri, 1951). Broadly speaking, all species of toads have been found suitable, responding only to injections of gonadotrophins, while frogs are much more variable in their behaviour. Of the frogs, Rana esculenta and Rana pipiens are probably the best. The former has been used extensively in Europe and the latter in the United States of America. In 1949, at this laboratory, we investigated the possibility of using the native species of toad, Bufo bufo (Fig. 4). This species, when mature, weigh between 20-40 g. An injection of 4 ml. of untreated urine proved fatal to 40 per cent. of toads, while smaller doses resulted in many false negatives.

It was therefore decided to use Scott's (1940) concentration technique which was simplified to suit our requirements (Law, 1949). 2.5 ml. of
the extract was injected into the dorsal lymph sac of the test animal, and the test was read in the usual way at the end of three hours. In a series of 732 routine pregnancy tests there were 6 false negatives and 0 false positives (Haines and Ferreira, 1954).

In addition to routine pregnancy diagnosis we have found the test invaluable in the follow-up and diagnosis of hydatidiform mole and chorion epithelioma. A titre of 1:200 of unconcentrated urine in distilled water is regarded as strong presumptive evidence of molar pregnancy. The simplicity and cheapness of performance of the test has allowed us to repeat tests at weekly or fortnightly intervals on patients following evacuation of hydatidiform moles—in one case these tests continued over a period of eight months before the test became negative.

The speed with which results may be obtained makes the test useful as an aid to the diagnosis of ectopic gestations; 13 out of 15 proven cases yielded positive tests within three hours.

Toads may readily be obtained in this country in early spring, less easily in early and late summer and are almost unobtainable during the winter months, October-February, when they are hibernating. It is therefore desirable to establish a stock during the spring and summer to last through the winter months.

Details of animal husbandry are described by Law (1949). Toads are fed once a week; small motile grubs and larvae, such as mealworms and gentles are suitable. These are placed in the cage in a shallow dish, preferably at night.

Toads in good condition may be used repeatedly; and at intervals as short as a week, though possible, the interval is extended to two or three weeks.

Conclusions

All the reliable modern pregnancy tests in current use are based on the excretion of chorionic gonadotrophin, which appears in detectable amounts in the urine at a very early stage of pregnancy. One of our own positive tests was obtained two days after the first missed period. Frazer and Wohlzogen (1950) have shown that the threshold dose of a standard gonadotrophin required to produce a positive response in the various test animals, is approximately as follows:

- Friedman rabbit 15 i.u., immature mouse 10 i.u., Xenopus laevis 70 i.u., and Bufo bufo 7 i.u. The classical Aschheim-Zondek reaction is a very sensitive and accurate method. Its main defect is the slowness of the reaction which is accentuated in doubtful tests which have to be repeated. Furthermore, considerable experience is required in the reading of the test. Nevertheless, it still retains its value as a yardstick to which the sensitivity of the more recently devised tests may be related.

The Frank rat test is rapid and, in the hands of experienced workers, accurate. Three rats per test ensures consistently accurate results, Bunde (1947), Elek (1953). Unless rats are being used for other purposes in a laboratory, the test is too expensive for routine use.

Similar criticism on economic grounds can be levelled at the Friedman test. Providing the recommended precautions are strictly adhered to, the test is sufficiently accurate for routine pregnancy diagnosis. It can be recommended for use in units where the number of tests per annum are few. Use of each animal for three tests considerably reduces the cost.

The Hogben test is best suited for use in pregnancy diagnosis laboratories, or at most, units where a large number of tests are performed. Unless the strict routine in the care and use of these toads is observed, there is a serious loss of accuracy. Even under presumably ideal conditions Schwabacher reports in 1949: 'it was found that 2.86 per cent. of the positive tests failed to be diagnosed by the first test toad, a year later the figure rose to 13.89 per cent. This means that our colony appears to be less sensitive now than 12 months ago.'

The male toad pregnancy test appears at the moment to be the test of choice. The choice of species depends entirely on local supplies. While some species of male frogs have proved suitable, others, e.g., Rana temporaria and the male Xenopus laevis, are unreliable, either by responding to certain non-specific stimuli or by spontaneous seasonal spermatiation. If the difference between the behaviour of toads and frogs is understood, a great deal of confusion and conflicting reports can be avoided.

In this country the supply of the indigenous species, B. bufo is inadequate except during the spring. With experience we are finding it possible to maintain our colony throughout the year. The animals are easy to keep and require little attention. A small stock of about two dozen toads would be quite adequate to supply the requirements of the average laboratory unit. The time taken in injecting the animal and reading the test amounts to a minute or two. The use of a hormone concentration method is essential for accuracy and adds but little to the time taken.

BIBLIOGRAPHY

ASCHHEIM, S., and ZONDEK, B. (1928a), Ibid., 7, 8.
ASCHHEIM, S., and ZONDEK, B. (1928b), Ibid., 7, 1404, 1453.
Continued on page 369.
been taken from an area free of new growth. If both methods are employed the overall cancer detection will be increased. Errors in judgment are mainly responsible for 'false positive' reports, and it is, therefore, imperative to confirm 'positive smears' by histological methods before any therapeutic measures are taken. 'Positive smears' which are unsupported by histology should be repeated; if follow up smears continue to show cells believed to be malignant, more extensive repeat biopsies must be carried out, but on no account should a patient be submitted to major surgery or radiotherapy on the strength of one positive smear report.

A survey of the 3 years ending December 31, 1952, routine cytological screening carried out at the Hammersmith Hospital gave the following results: The total number of patients suffering from malignant disease of the uterus seen in the Gynaecological Department was 64; 30 of these were squamous and 34 columnar-celled malignancies.

Of the 30 squamous cancers 26 were diagnosed correctly by routine smear procedure, 3 or 10 per cent., were missed and for one an unsuitable smear had been submitted. This compares well with American figures of 'false negatives.'

Of the 34 adenocarcinomata 31 were diagnosed correctly, two were missed and for one an unsuitable film had been submitted. This gives an incidence of error of only 5.8 per cent., which is much lower than the corresponding American figures.

More important than these absolute figures is the fact that out of these 64 carcinomata 8 were entirely unsuspected clinically and would not have been diagnosed and treated at this early stage had cytological investigation not been carried out.

The incidence of false positive reports has been regrettably high, but as the only criterion of comparison is the histological report, which is regarded as the ultima ratio although it is sometimes based on a single biopsy only, it is possible that cases now listed as cytological errors, may, in the course of time, prove to have been correctly diagnosed cytologically. One case included in the 64 carcinomata had been listed as a 'false positive' for 18 months until the patient was re-admitted with secondary vault metastases. (Martin and Kenny 1950 and Martin 1953).

In the light of our present knowledge, however, and accepting the histological report as the final judgement, we have had 9 'false positive' reports, all of which concerned erroneous diagnoses of columnar-celled malignancies; most of these came from patients with ulcerating polypi, one from a patient with a sloughing endometrium and one from a woman diagnosed as an atypical glandular hyperplasia.

To sum up: of our 62 carcinomata investigated cytologically, omitting the two cases with unsuitable smears, 57 or 91.9 per cent. were diagnosed correctly, 5 or 8.7 per cent. were 'false negatives,' and 8 were detected by the smear. Cytology is a valuable aid to cancer diagnosis and an easily applicable method of cancer detection in the unsuspected case.

BIBLIOGRAPHY

AYRE, J. ERNEST, (1951), Cancer Cytology of the Uterus, New York, Grune and Stratton.


VINCENT MEMORIAL HOSPITAL (1950), The Cytological Diagnosis of Cancer, Philadelphia, W. B. Saunders Co.


Continuation of Bibliography—H. P. Ferreira, M.D., from page 359.


BISHOP, P. M. F. (1934), Brit. med. J., 1, 1186.


ELEK, S. D. (1953), Personal Communication.


GECHIEN, F. and MCGREATH, J. (1944), Irish J. med. Sci., 6"m Ser., No. 228, 509.


LANCEWIT, II, 477.


SCHWARZACHER, (1951), Med. Presse, 226, 82.

SCOTT, J. L. (1940), Brit. J. exp. Path., 21, 536.


SIMOLA, P. E., and RIVAS, L. (1936), Surgeon, 9, 24.


ZONDEK, B. (1930e), Klin. Wschr., 9, 964.