

SEQUENTIAL ANALYSIS FOR DIAGNOSING
DIABETIC PATIENTS

Damodar Y. Golhar
Management Department
Western Michigan University
Kalamazoo, MI 49008

Stephen M. Pollock
Industrial and Operations Engineering Department
The University of Michigan
Ann Arbor, MI 48109

Technical Report 85-36
November 1985
Revised June 1986

SEQUENTIAL ANALYSIS FOR DIAGNOSING DIABETIC PATIENTS

Damodar Y. Golhar
Management Department
Western Michigan University
Kalamazoo, MI 49008

Stephen M. Pollock
Industrial and Operations Engineering
The University of Michigan
Ann Arbor, MI 48109

May 1986

Abstract

We use a truncated sequential probability ratio test (SPRT) to diagnose diabetic patients. It is hypothesized that, for a given oral glucose tolerance test (OGTT), the differences between successive observations are more diagnostic than the observations alone. Using such differences in a SPRT, OGTT test data for 950 subjects is analyzed and thresholds for classifying diabetes and normal subjects are established. This sequential approach provides the same diagnosticity as that obtained by using the National Diabetic Data Group criteria, but it is more efficient. In particular, the procedure significantly reduces the sampling cost and average patient waiting time.

SEQUENTIAL ANALYSIS FOR DIAGNOSING DIABETIC PATIENTS

Introduction

Different groups in the diabetes community have established diagnostic criteria for defining diabetes and other manifestations of glucose intolerance. The oral glucose tolerance test (OGTT) is often recommended for diagnosing diabetes, but interpretation of fasting plasma glucose levels and glucose levels during an OGTT differ widely. Hence OGTT responses labelled "abnormal" by one set of criteria may be classified "normal" by another, resulting in widely varying estimates of the prevalence of diabetes. The National Diabetes Data Group (1) notes:

"Substantial differences exist in the diagnostic criteria used in practice by diabetologists and there is no consensus as to the dividing line between normal and diabetic glucose levels. This lack of agreement is primarily caused by the fact that there is no clear division between diabetics and nondiabetics in their fasting plasma glucose concentration or their response to an oral glucose load."

Disagreement among experts is not helped by excessive reliance on absolute levels of test responses, which are subject to error due to biological influences such as diet, smoking, alcohol consumption by the patient prior to the OGTT test and laboratory errors. However, differences between successive test responses might be more diagnostic than the absolute test results, in that they could possibly cancel out common errors. Thus statistical analysis of successive differences might provide meaningful criteria for diagnosing diabetic patients accurately with fewer samples. In addition, the use of a sequential analysis allows for the opportunity for rapid decision making. This paper addresses such a procedure, and demonstrates its use with an example of a diagnostic differentiation between diabetic and normal patients. The sequential test diagnoses these patients with a substantial decrease in the number of tests, and length of test time.

Classification of diabetes includes three clinical categories: diabetes mellitus (hereinafter referred to as "diabetes"), impaired glucose tolerance (IGT) and gestational diabetes. Diabetes mellitus is characterized by either fasting hyperglycemia or levels of plasma glucose (PG) during an oral glucose tolerance test above defined limits. For impaired glucose tolerance patients, PG levels during an OGTT lie above normal but below the levels of diabetes. The gestational diabetes class is restricted to pregnant women in whom the onset or recognition of diabetes or IGT occurs during pregnancy. Thus, diabetic women who later become pregnant are not included in this class.

For our analysis a dataset was obtained for non-pregnant adults who have a family history of diabetes. The sample was collected by the Diabetes Research and Training Center (DRTC) at the University of Michigan, Ann Arbor, over a period of more than 20 years, as part of an ongoing study of diabetes. A long term follow-up, involving multiple OGTT's over a period of up to twenty years, found that of 1115 non-pregnant adults in the dataset, 741 were normal, 209 had diabetes, and 165 had IGT.

The methodology presented here only deals with the dichotomous differentiation between "diabetes" and "normal" diagnosis. It does not include the third category of "IGT". (Further work is needed to incorporate this category of patients into a sequential testing paradigm -- for sequential tests between three hypotheses see Borgan (1979), Wetherill (1975), and the references listed therein.) However, not including IGT does not detract from our argument about the diagnostic superiority of the sequential test procedure, based upon OGTT differences, discussed here.

For each patient the dataset contains the results of a large number of OGTT, repeated over a period of months or even years. For the analysis reported on here, one randomly selected OGTT was chosen for each patient, to represent a "typical" test result. Strictly speaking, then, the diagnostic results we

present are applicable only to randomly selected persons from a population of those with a family history of diabetes.

The Test (OGTT):

A three-hour glucose tolerance test was performed on each individual, blood samples being drawn with the subject in the sitting position immediately before a glucose dose of 75 gm. was administered, and at half hour intervals for the duration of the three-hour test. All subjects were instructed to follow a weight maintaining diet containing at least 250 gms of carbohydrates per day for three days prior to the test. All subjects were ambulatory and free of known disease.

Diagnosis Using NDDG Criteria

For each test, an individual's response to a glucose load can be classified as being either normal, non-diagnostic, impaired or diabetic according to the criteria described by the National Diabetes Data Group (1) as given in table 1. (Table 1 contains only OGTT criteria for venous plasma. The description of other classical symptoms of diabetes is omitted.) Using these criteria, the 950 patients were diagnosed as shown in table 2. The number of patients diagnosed at 0-hr, .5-hr, ..., 2-hr is also given. The fraction of patients correctly diagnosed as being "normal" was $1 - \beta = .98$; the missed diagnosis rate was $\alpha = .043$. All "normal" diagnoses required two hours to be diagnosed; only 37% of "diabetic" diagnoses could be made by using only a fasting level reading. Note that out of 950 patients, only 2.5% of the subjects were classified as "non-diagnostic". The normal clinical procedure would be to repeat the OGTT test on these patients. Thus, it is reasonable to assume that the ultimate diagnosticity, including re-testing, is no different than that indicated by the given α and β .

A. Diabetes Mellitus in Nonpregnant Adults

Any of the following are considered diagnostic of diabetes:

- i) Elevated fasting glucose concentration on more than one occasion of venous plasma ≥ 140 mg/dl. If the fasting glucose concentration meets this criterion, the OGTT is not required. Virtually all persons with FPG > 140 mg/dl will exhibit an OGTT that meets or exceeds the criteria in (ii) below.
- ii) Fasting plasma glucose < 140 mg/dl, but sustained elevated glucose concentration during the OGTT on more than one occasion. Both the 2-hour sample and some other sample taken between administration of the 75 gms glucose dose and 2 hours later must have venous plasma ≥ 200 mg/dl.

B. Impaired Glucose Tolerance (IGT) in Nonpregnant Adults

Three criteria must be met: the fasting glucose concentration must be below the value that is diagnostic for diabetes; the glucose concentration two hours after a 75 gms oral glucose challenge must be between normal and diabetic values; and a value between .5-hr, 1-hr or 1.5-hr OGTT value later must be unequivocally elevated.

- i) Fasting value: venous plasma < 140 mg/dl.
- ii) .5-hr, 1-hr or 1.5-hr OGTT value: venous plasma ≥ 200 mg/dl. and
- iii) 2-hr OGTT value: venous plasma of between 140 and 200 mg/dl.

C. Normal Glucose levels in Nonpregnant Adults

- i) Fasting venous plasma < 115 mg/dl.
- ii) 2-hr OGTT venous plasma value < 140 mg/dl. and
- iii) OGTT values between .5-hr, 1-hr or 1.5-hr later of venous plasma < 200 mg/dl.

Glucose values above these concentrations but below the criteria for diabetes (A. above) or IGT (B. above) should be considered non-diagnostic for these conditions.

TABLE 1

National Diabetes Data Group Diagnostic Criteria (1)

	ACTUAL	DIAGNOSIS			NO. OF PATIENTS DIAGNOSED AT					TOTAL
		NORMAL	DIABETES	NON DIAGNOSTIC	N=1 (0-hr)	N=2 (.5-hr)	N=3 (1-hr)	N=4 (1.5-hr)	N=5 (2-hr)	
	NORMAL	726 (97.98%)		15 (2.02%)						741
	DIABETES		200 (95.7%)	9* (4.3%)	78		1			209

* Includes one patient diagnosed IGT

TABLE 2

Diagnosis of 950 Patients Using Criteria of Table 1.

Diagnosis Using Sequential Analysis of Test Differences

The sequential probability ratio test (SPRT) was developed by Wald (2) for hypothesis testing. For a desired diagnosticity (i.e., for given levels of type I and type II errors) Wald developed a "standardized" cumulative test score and upper and lower thresholds. If the score crosses either lower or upper threshold then the null or alternate hypothesis, respectively, holds good. If the score is between these two thresholds, then another observation is taken. The advantage of such a sequential test is that it provides the same α and β as that of non-sequential tests, but produces a smaller average sample size. The use of a truncated SPRT in quality control ("CUSUM" chart) has shown that, while retaining the same power as other established quality control procedures, minimizes the expected total cost. (Further discussion on the truncated SPRT is in Wald (2).)

Patients can also be classified by using the sequential procedure developed by Wald (2). After obtaining data at any stage of the test it is possible to make one of the following three decisions: 1) to diagnose "diabetic"; 2) to diagnose "normal"; or 3) to continue the test by making an additional observation. (Again, for the immediate purpose of this paper, the intermediate diagnosis of IGT is ignored.) Such a test procedure is carried out sequentially until either a diabetic or normal diagnosis is made, or until a pre-specified time limit (or number of tests) is reached. The sequential procedure is described in more detail in the appendix.

For our dataset the first three readings of the OGTT were considered: $x_1 = 0^{\text{th}}$ hour reading (FPG), $x_2 = .5$ hour reading, $x_3 = 1$ hour reading. Out of the 950 patients, 32 patients had either the second or third observation missing, and so were not included in this analysis. Hence only 918 patients were diagnosed by our method. Of them 713 were known to be normal patients and 205 were known diabetics. Since we argue that the differences between successive

readings are more diagnostic than the readings themselves, new variables y_2 and y_3 were defined, where $y_2 = x_2 - x_1$ and $y_3 = x_3 - x_2$.

In order to obtain functional forms for the distributions of variables x_1 , y_2 , y_3 , histograms and scatter plots were obtained for both diabetes and normal patients. Using chi-squared goodness of fit test, variable x_1 was found to be normally distributed, with different means and variances for both diabetic and normal groups. Variables y_2 and y_3 were found to be bivariate normal.

In order to use the sequential procedure, it is necessary to establish desired values of α and β . Values of $\alpha = .04$ and $\beta = .02$, were selected to be close to those obtained in the non-sequential test (.043 and .02, respectively). Classification is then performed by computing likelihood ratios at each stage (FPG, .5-hr, 1-hr) and comparing these with thresholds, as discussed in the appendix.

For the above mentioned values of α and β , diagnosis regions for realizations of x_1 , x_2 and x_3 can also be obtained, by using the sequential procedure. Thus a patient with a family history of diabetes could be diagnosed without actually computing likelihood ratios. In particular, these regions (which in some cases are simple thresholds) are:

- a) If $62 < x_1 < 126$, take an additional observation x_2 ; if $x_1 \geq 126$, the patient is diagnosed as "diabetic"; if $x_1 \leq 62$, the patient is diagnosed as "normal".
- b) For $62 < x_1 < 126$, figure 1(a) shows the threshold values of x_2 : if the observed value of x_2 lies in the continuation region, an additional sample x_3 should be taken.
- c) If a third sample x_3 is taken, figure 1(b) shows the threshold values of x_3 as a function of x_2 for selected values of x_1 (70, 90 and 110). For any $70 < x_2 < 230$, if the observed value of x_3 is above the curve

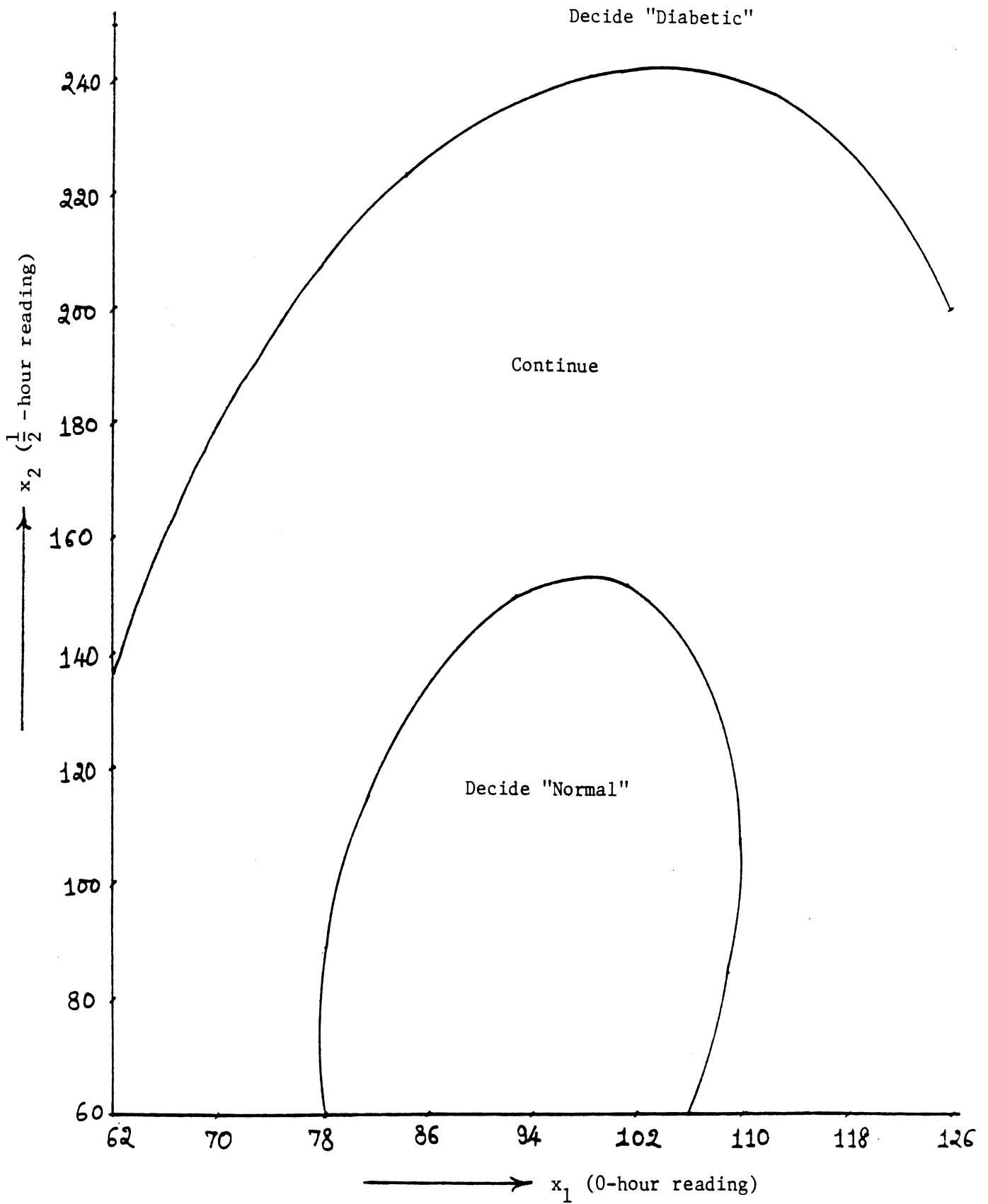


FIGURE 1(a)

Threshold Decision Values of First Two OGTT Readings

Decide "Diabetic" if (x_2, x_3) is above x_1 curve.

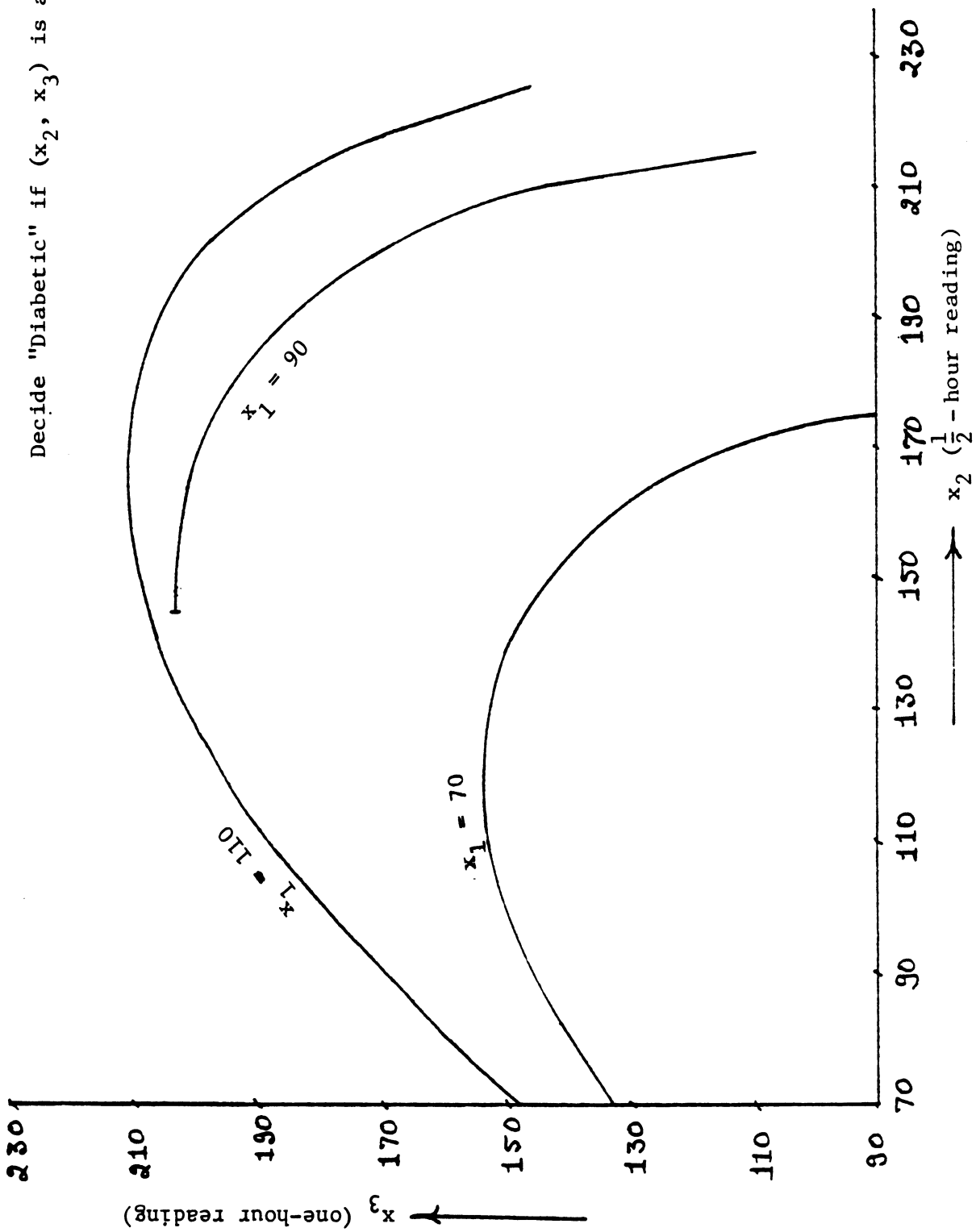


FIGURE 1(b)

Decision Thresholds For First, Second and Third OGTT Readings, Given Continuation.

corresponding to a given value of x_1 then the patient is diagnosed as "diabetic", otherwise the patient is diagnosed as "normal".

Table 3 gives the results for the 918 patients diagnosed by our method.

Cross validation was carried out by randomly selecting two-thirds of the patients in order to estimate the parameters. The remaining one-third of the patients were diagnosed using these parameters. The results obtained with cross validation were consistent with those shown in table 3. A "jack-knife" technique (i.e., using the data from all patients--except one--for estimation, and then diagnosing the left out patient and repeating this procedure for each patient) was also used on the dataset. The results are given in table 4 and as can be seen, are essentially the same as those given in table 3.

A comparison of tables 2 and 3 shows that the sequential approach is as powerful as non-sequential diagnosis, yet it is much more efficient in terms of diagnosing patients early. In particular, the sequential approach diagnoses patients using at most three observations, whereas the non-sequential diagnosis needs five observations for 91.7% of the patients. Note that the "non-diagnostic" classification of patients using the NDDG criteria is statistically the same as "mis-classification" using the sequential approach. However, it should be noted that a non-diagnosed patient will most likely be re-tested. Hence this error using the NDDG criteria might be less costly than "mis-classification" by a sequential approach, for which a mis-classified patient may or may not be subject to follow-up tests.

The expected number of samples for the sequential test is 2.41 as compared to 4.67 for the non-sequential test. Thus, the sequential procedure, on average, requires 2.26 fewer samples to be analyzed, a 48% reduction in sampling cost/OGTT test.

The laboratory cost of analyzing a single sample varies between \$2.50 and \$8.00. Taking the average of these two values, the cost of analyzing a sample

	DIAGNOSIS		NO. OF PATIENTS DIAGNOSED AT			TOTAL
	NORMAL	DIABETES	N=1 (0-hr)	N=2 (.5-hr)	N=3 (1-hr)	
ACTUAL	706 (99.00%)	7 (1.00%)	2	310	401	713
STATE		9 (4.39%)	100	25	80	205

TABLE 3

Diagnosis of 918 Patients Using Sequential Method.

	DIAGNOSIS		NO. OF PATIENTS DIAGNOSED AT			TOTAL
	NORMAL	DIABETES	N=1 (0-hr)	N=2 (.5-hr)	N=3 (1-hr)	
ACTUAL	705 (98.88%)	8 (1.12%)	2	310	401	713
STATE	9 (4.39%)	196 (95.61%)	100	25	80	205

TABLE 4

Diagnosis of 918 Patients Using "Jack-knife" Technique For Sequential Method.

is assumed to be about \$5.25. A conservative assumption is that 2000 OGTT tests are performed in a typical hospital each year. The resulting expected cost savings is thus over \$23,000/year per hospital. However, we should emphasize that the real savings are in time and inconvenience to the patient, achieved at no sacrifice to the quality of the test.

The average patient waiting time for the sequential test is 0.71 hours, compared to 1.84 hours for the non-sequential test. Thus the sequential approach, with equivalent diagnosticity, reduces a patient's average waiting time by at least one hour. With fewer samples to take, the nursing staff would, of course, also save time.

Conclusion

For classifying patients between diabetic and normal groups only, our results show that a sequential test of glucose level differences is more efficient than the NDDG procedure suggested in (1). For the same level of diagnosticity, patients are classified with fewer observations compared to a non-sequential test. The savings in sampling cost and patient waiting time, along with staff time, are significant. A similar analysis including the third diagnostic category IGT is currently under study.

References

1. National Diabetes Data Group, "Classification and Diagnosis of Diabetes Mellitus and Other Categories of Glucose Intolerance", *Diabetes*, Vol. 28, No. 12, pp. 1039-1057, 1979.
2. Wald A., "Sequential Analysis", Dover Publications Inc., New York, 1947.
3. Armitage, P. "Restricted Sequential Procedures", *Biometrika*, Vol. 44, pp. 9-26, 1957.
4. Wetherill, G. B., "Sequential Methods in Statistics", Monographs on Applied Probability and Statistics, Chapman and Hall, 1975.
5. Borgan, O., "Comparison of Two Sequential Tests for Two-Sided Alternatives", *J. R. Statist. Soc. B*, Vol. 41, No. 1, pp. 101-106, 1979.

Appendix

The sequential probability ratio test (SPRT) by Wald (2) provides a rule for making one of the following three decisions at every decision stage: 1) diagnose "diabetic"; 2) diagnose "normal"; or 3) take an additional observation. The test terminates when either decision 1) or 2) is made. The number 'N' of observations required, before the test is terminated by such a procedure, is a random variable, since the value of N depends on the outcome of the observations. The objective is to minimize the expected number of observations and still achieve a pre-specified level of significance α and power $1 - \beta$, where

$$1 - \alpha = \text{Pr.}\{\text{Decide that the patient has diabetes given} \\ \text{the patient has diabetes}\}$$

$$1 - \beta = \text{Pr.}\{\text{Decide that the patient is normal given} \\ \text{the patient is normal}\}$$

The test is based on the following arguments. Let x_1, \dots, x_n be the first n test results for a patient. If $f_n(x_1, \dots, x_n|D)$ is the p.d.f. for these results given the patient has diabetes and $f_n(x_1, \dots, x_n|\bar{D})$ is the p.d.f. for these results given the patient is normal, then the likelihood function is:

$$L_n(\underline{x}) = L_n(x_1, \dots, x_n) = f_n(x_1, \dots, x_n|\bar{D})/f_n(x_1, \dots, x_n|D)$$

The procedure consists of taking observations x_1, x_2, \dots sequentially. At the n^{th} step, if

$$B < L_n(\underline{x}) < A$$

take another observation; if $L_n(\underline{x}) \leq B$ then decide that the patient has diabetes; and if $L_n(\underline{x}) \geq A$ then decide that the patient is normal. The values of A and B can be shown to be, approximately,

$$A \approx (1 - \beta)/\alpha \text{ and } B \approx \beta/(1 - \alpha)$$

Although Wald has shown that the sequential test procedure will eventually terminate, it is occasionally desirable to set a definite upper limit, say n_0 , for the number of observations. This can be achieved by truncating the

sequential process at $n = n_0$. Such truncation also requires a new rule for diagnosis at the n_0^{th} trial, if the sequential process does not lead to a final decision for $n \leq n_0$. A simple and reasonable rule for truncation at the n_0^{th} trial is the following: decide "diabetes" when $B < L_n(\underline{x}) \leq 1$, and decide "normal" when $1 < L_n(\underline{x}) < A$. Such a truncated procedure can be shown to produce a controllable increase in the desired levels of α and β , with the same expected number of samples needed to reach a decision [3].