Screening Tests for Diabetes

A Study of Specificity and Sensitivity

Arnold B. Kurlander, M.D., M.P.H., Albert P. Iskrant, M.A.,† M. E. Kent, M.D., M.P.H.‡*

Washington, D.C.

In January 1952, a multiple screening project was initiated at the Out-patient Department of D. C. General Hospital in Washington, D. C. One of the purposes of this project was to study the results of some of the tests for diabetes in relation to the diagnosis established in the Medical Diagnostic Clinic and thus to investigate the efficiency of these tests as case-finding procedures for diabetes.

The D. C. General Hospital is a municipal hospital providing free medical care to the indigent. It functions as a teaching hospital for Georgetown and George Washington Medical Schools. The staff of the Medical Diagnostic Clinic consists of resident physicians and internes under the direct supervision of the Chief of the Outpatient Department who is an internist certified by the American Board of Internal Medicine.

METHODS OF STUDY

Persons included in this study were referred from the admitting office of the Out-patient Department of the hospital to the Medical Diagnostic Clinic. As many as possible of these persons received a multiple screening examination which included a venous specimen of blood tested by the Wilkerson-Heftmann method to show a blood sugar level above 130 mg. per 100 ml. and if positive again tested to show a blood sugar level above 180 mg. per 100 ml. A urine specimen was tested for sugar by a copper reduction method (Clinitest). In all, 964 persons attending the Out-patient Clinic received a screening examination. Of this number, 953 received the copper reduction test of urine and 928 the Wilkerson-Heftmann test of blood. (The Clinitron, an apparatus for the mechanical performance of the Wilkerson-Heftmann test, was out of order for a brief time.) From these screenees, persons who were going to the Diagnostic Clinic were selected at random to receive a complete diagnostic study requiring three-and-a-half hours, including special examinations for diabetes. This special group of screenees selected for diabetes examination received additional screening tests for diabetes, using the original screening samples of blood and urine: a quantitative blood sugar test by the Somogyi-Nelson method and a test of the urine by a bismuth reduction test (Galatest). This resulted in a Somogyi-Nelson blood test for 605 persons and a bismuth reduction test for 648 persons.

Of the persons referred to the Diagnostic Clinic, only 295 received the complete battery of tests and special diagnostic study for diabetes. This number is less than half of those who were requested to return for a glucose tolerance test and indicates the need for personnel for follow-up work in studies such as these. The diagnostic study for diabetes included a standard three-hour 100 gm. oral glucose tolerance test following three days of full customary meals. The Somogyi-Nelson quantitative method was used for the blood sugar tests on venous blood. When the original blood-sugar level was over 180 mg. per 100 ml., the glucose tolerance test was omitted and a single blood sugar determination after a full meal was substituted.

RESULTS OF TESTS FOR DIABETES

The criteria for positivity of the glucose tolerance tests were based more on the shape than on the height of the blood sugar curve. In general, a curve which returned to 110 mg. per 100 ml. at the end of three hours was considered normal. This held true even though the peak of the curve may have been above 180 mg. per 100

* Assistant Chief of Division of Chronic Disease and Tuberculosis, Public Health Service, Department of Health, Education and Welfare, Washington 25, D. C.

† Chief of Program Analysis in the same division.

‡ Medical Officer-in-Charge, D. C. General Hospital, Multiple Screening Project, Washington, D. C. Currently Medical Officer, Internal Medicine Outpatient Service, Walter Reed Army Medical Center, Washington 12, D. C.

ml. Conversely, any curve which was above 110 mg. per 100 ml. at the end of three hours was usually considered diagnostic of diabetes. Exceptions to these criteria were made in cases where other diseases were believed to account for abnormality or where special features of the curve seemed to demand a different interpretation. For cases screening over 180, the Somogyi-Nelson test was repeated two hours after eating and a venous blood sugar reading over 150 was considered diagnostic.

TABLE 1
Summary of the tests

Type of Investigation	Number of Persons
Multiple screening tests	964
Urine test by copper reduction method (Clinitest)	953
Urine test by bismuth reduction method (Galatest)	648
Blood test by Wilkerson-Heftmann method	928
Blood test by Somogyi-Nelson method	605
Complete diagnostic study for diabetes	295

The population screened is described as to race, sex and age in Table 2. It will be noted that the population, in addition to being indigent, is predominantly non-white; more than half the persons tested were non-white females. The median ages were in the fifties for white males, in the forties for white females and non-white males, and in the thirties for non-white females. Obviously, then, the screened population is much older than the general population of the United States but, because of their age, a good group to screen for diabetes.

TABLE 2
Race, sex and age distribution of 964 persons receiving multiple screening tests

Age Group	White		Non-white	
	Male	Female	Male	Female
Total	66	70	274	554
15-19	1	3	17	38
20-29	2	7	35	118
30-39	3	13	49	150
40-49	18	15	66	125
50-59	26	19	50	62
60-69	8	11	40	45
70-79	4	2	16	14
80-89	4		1	2

COMPARISON OF TEST METHODS

In the following tables only cases were used that received the tests being compared. Obviously then, fewer cases are included in tables where the four tests are being compared than where two tests are being compared. Where only one test is

TABLE 3
Results of diabetes screening by method and hours after eating for the 478 persons tested by all four methods

Method	Hours since eating	Number tested	Positive Number	Positive Per cent**
Blood sugar test	Total	478	41	8.6
Wilkerson-Heftmann method				
positive—over 130 mg. per ml.	Fasting*	236	17	7.2
	2 or more	131	12	9.2
	Less than 2	104	10	9.6
	Not stated	7	2	—
Blood sugar test	Total	478	41	8.6
Somogyi-Nelson method				
positive—over 130 mg. per ml.	Fasting*	236	17	7.2
	2 or more	131	12	9.2
	Less than 2	104	10	9.6
	Not stated	7	2	—
Test for sugar in urine	Total	478	89	18.6
Copper reduction method (Clinitest)				
positive—trace or more	Fasting*	236	33	14.0
	2 or more	131	32	24.4
	Less than 2	104	23	22.1
	Not stated	7	1	—
Test for sugar in urine	Total	478	55	11.5
Bismuth reduction method (Galatest)				
positive—trace or more	Fasting*	236	16	6.8
	2 or more	131	22	16.8
	Less than 2	104	16	15.4
	Not stated	7	1	14.3

* The individuals had not eaten since the night before. All tests were made between 8:30 a.m. and 10:30 a.m.
**Percentages have been calculated in this and subsequent tables, even when numbers are small, for ease in rough comparisons. The calculation to one decimal place in no way signifies that small differences are meaningful.

being analysed the numbers are correspondingly larger. For example, where only the Somogyi-Nelson method is involved, as in Table 9, 23 previously undiagnosed diabetics are included; in Table 6 only 21 are included, as two of the cases did not receive all four tests. Unequal numbers were used to make maximum use of the small number of cases of diabetes.

Table 3 shows the number tested for diabetes by each of the four methods for those persons receiving all four of the tests, the number of hours since last eating and the number screening positive. (Readings of 130 mg. per 100 ml. for the Wilkerson-Heftmann and trace or more for the urine determination were considered positive.) It is noted that the percentage screening positive is quite high. "Glycosuria" was found in 18.6 per cent by the copper reduction test of the urine and in 11.5 per cent by the bismuth reduction test. "Hyperglycemia" was found in 8.6 per cent by the Wilkerson-Heftmann and the Somogyi-Nelson methods. As may be expected, the percentage screening positive is lower for persons fasting. (Fasting is defined as not having eaten since the night before. All screening was performed between 8:30 a.m. and 10:30 a.m.)

In Table 4, the results of the tests by the Wilkerson-Heftmann method showing the blood sugar both over 130 mg. per 100 ml. and 180 mg. per 100 ml. are compared with the results of testing by copper reduction and bismuth reduction methods. While the figures are small, it appears that a person who screens positive to the Wilkerson-Heftmann test at 180 mg. per 100 ml. is likely to screen positive to urine tests. Of those screening positive at 130 mg. per 100 ml. but not at 180 mg. per 100 ml. to the Wilkerson-Heftmann test, the proportion screening positive to the urine tests is rather low, being lowest when the subjects were fasting. Conversely, many persons with tests negative to the Wilkerson-Heftmann procedure at 130 mg. per 100 ml., had positive urine tests. Within the limits of these data, then, and without

TABLE 4

Results of blood sugar tests by the Wilkerson-Heftmann method compared with tests for sugar in the urine by the copper reduction and bismuth reduction methods for the 636 persons tested by the three methods

Blood sugar—Wilkerson-Heftmann method			Sugar in urine—Copper reduction method (Clinitest)				Sugar in urine—Bismuth reduction method (Galatest)			
Results (findings)	Hours since eating	No. tested	Positive No.	Positive Per cent	Negative No.	Negative Per cent	Positive No.	Positive Per cent	Negative No.	Negative Per cent
Blood sugar over 180 mg. per 100 ml.	Fasting	7	7	100.0	0	0.0	6	85.7	1	14.3
	2 or more	5	5	100.0	0	0.0	5	100.0	0	0.0
	Less than 2	6	6	100.0	0	0.0	6	100.0	0	0.0
	Total	18	18	100.0	0	0.0	17	94.4	1	5.6
Blood sugar over 130 mg. but not over 180 mg.	Fasting	10	2	20.0	8	80.0	2	20.0	8	80.0
	2 or more	7	4	57.1	3	42.9	3	42.9	4	57.1
	Less than 2	4	1	25.0	3	75.0	1	25.0	3	75.0
	Not stated	2	0	0.0	2	100.0	0	0.0	2	100.0
	Total	23	7	30.4	16	69.6	6	26.1	17	73.9
Blood sugar under 130 mg.	Fasting	285	31	10.9	254	89.1	14	4.9	271	95.1
	2 or more	169	30	17.8	139	82.2	17	10.1	152	89.9
	Less than 2	124	20	16.1	104	83.9	12	9.7	112	90.3
	Not stated	17	2	11.8	15	88.2	4	23.5	13	76.5
	Total	595	83	13.9	512	86.1	47	7.9	548	92.1

TABLE 5

Comparison of results of blood sugar tests by Somogyi-Nelson and Wilkerson-Heftmann methods for the 572 persons tested by both methods

Blood sugar tests by Somogyi-Nelson method mg. per 100 ml.	Blood sugar tests by Wilkerson-Heftmann method mg. per 100 ml.: Less than 130	130-179	180 or more	Total No.	Total Per cent
Less than 130	509	7	0	516	90.2
130-179	8	22	1	31	5.4
180 or more	0	3	22	25	4.4
Total Number	517	32	23	572	
Total Per cent	90.4	5.6	4.0		100.0

consideration of the diagnosis of diabetes, it would seem that the urine test is not a good indication of the blood sugar result in cases with a positive blood sugar test over 130 and less than 180 mg. per 100 ml. of blood.

In Table 5 are shown the results of the blood sugar tests by the Somogyi-Nelson method and the Wilkerson-Heftmann method for the 572 persons tested by both methods. The agreement between the results of these two methods was almost perfect as far as percentages are concerned, although there was a difference in individual instances. The results of the two methods were in disagreement in 19 of the 572 cases in which tests were made by both methods. In 12 of the 19 instances of disagreement the difference was more than 10 mg. per 100 ml.

SPECIFICITY AND SENSITIVITY OF TESTS

Sensitivity may be defined as the ability of the test to classify as positive those who have the condition being screened for and is calculated as the percentage screening positive of those diagnosed as having diabetes. For example, the Somogyi-Nelson method at 130 mg. per 100 ml. screened positive 15 persons of the total of 23 persons diagnosed as having diabetes and so the sensitivity rating is 15/23 or 65.2 per cent (see Table 9). *Specificity* may be defined as the ability of the test to classify as negative those who do not have the condition being screened for and is calculated as the percentage screening negative of those determined not to have diabetes. For example, the Somogyi-Nelson method at 130 mg. per 100 ml. screened negative 245 persons of the total of 250 persons confirmed as not having diabetes and so the specificity rating is 245/250 or 98.0 per cent (see Table 9). Another measurement commonly used to evaluate screening is the proportion of the positives who are diagnosed as having the condition. Obviously this latter measurement of the test is a function of the prevalence of the condition (unlike specificity and sensitivity indices, which are independent of the prevalence) but is a good measurement of the efficiency of the test for areas of similar prevalence.

The group included in this portion of the paper is limited to those persons for whom diagnostic observations were made for diabetes by a physician who did not know the results of any screening test.

It is important to point out here that the cases of diabetes (23) used in determining sensitivity ratings were those discovered as a result of a complete diagnostic survey including special examinations for diabetes. They do not include previously known diabetics, nor are they limited to diabetics with symptoms or those discovered through a screening examination. It might be expected that few, if any, unknown cases of diabetes would be present among those screening negative to the usual tests. This was not so among persons screened in the Out-patient Department at D. C. General Hospital. Whether these results were influenced by the age of the population tested, or existing conditions other than diabetes, or the nutritional status of those screened, is not known. The results were obviously influenced by the large proportion of persons screened who had not eaten recently. Studies such as this must be expanded to other population groups before these points can be settled.

Table 6 shows the results obtained through urine tests by the copper reduction (Clinitest) and bismuth reduction (Galatest) methods. It is noted that less than half of the previously undiagnosed (new) cases of diabetes showed positive reactions to these screening tests and that many nondiabetics also showed false positive reactions.

When only specimens showing a reaction graded 1-plus or more are considered positive, then the specificity of both tests is increased greatly. Of those persons diagnosed as not having diabetes only 2.9 per cent showed a false positive reaction to the copper reduction method and 1.7 to the bismuth reduction method considering reactions graded 1-plus or more. No nondiabetic screened positive to either test with a reaction graded 3-plus or more. It would appear then that, under conditions as they existed in this hospital, urine tests reported as showing a trace of sugar could be ignored as far as the follow-up of diabetes is concerned and that tests graded 3-plus or more are very likely to mean diabetes. Conversely, it would appear that many cases of early diabetes will be missed if dependence is placed on screening by tests of the urine, even if it is decided to follow up the slightest positive reaction.

Table 7 shows the results by the Somogyi-Nelson and the Wilkerson-Heftmann methods. A vast increase in efficiency over the urine tests is seen. At 130 mg. per 100 ml. the specificity ratings were 98.0 and 96.8. This means that 98 per cent of nondiabetics will screen below 130 mgs. per 100 ml. by the Somogyi-Nelson method and 96.8 per cent will screen negative by the Wilkerson-Heftmann method at the 130 level. These percentages seem satisfactory but the sensitivity levels may be considered low—65.2 and 56.5; that is 34.8 per cent of the newly diagnosed diabetics had Somogyi-Nelson readings below 130 and 43.5 per cent had negative screening results by the Wilkerson-Heftmann method at the 130 level.

The results of all four methods are summarized in

TABLE 6

Results of tests for sugar in the urine by copper reduction and bismuth reduction methods according to the presence or absence of diabetes in 269 cases in which both tests were used*

Diagnosis and result of tests	Copper reduction method (Clinitest)				Bismuth reduction method (Galatest)			
	Number		Per cent		Number		Per cent	
Not diabetes	239		100.0		239		100.0	
Negative	205		85.8		223		93.3	
Trace or more	34		14.2		16		6.7	
1+ or more		7		2.9		4**		1.7**
2+ or more		1		0.4		1**		0.4**
3+ or more		0		0.0		0**		0.0**
4+ or more		0		0.0		0**		0.0**
Newly discovered diabetes	21		100.0		21		100.0	
Negative	14		66.7		13		61.9	
Trace or more	7		33.3		8		38.1	
1+ or more		7		33.3		6		28.6
2+ or more		6		28.6		5		23.8
3+ or more		5		23.8		5		23.8
4+ or more		4		19.0		5		23.8
Previously known diabetes	9		100.0		9		100.0	
Negative	3		33.3		3		33.3	
Positive	6		66.7		6		66.7	

* Includes only cases with known diagnosis and known results.
**For 2 nondiabetics the report of the bismuth reduction test was "positive." They are included as a "trace or more" but not in the "1+ or more."

TABLE 7

Results of blood sugar tests by Somogyi-Nelson and Wilkerson-Heftmann methods according to the presence or absence of diabetes for the 281 persons receiving both tests*

Diagnosis and result of test	Somogyi-Nelson method		Wilkerson-Heftmann method	
	Number	Per cent	Number	Per cent
Not diabetes	250	100.0	250	100.0
Negative	245	98.0	242	96.8
Positive (over 130 mg. per 100 ml.)	5	2.0	8	3.2
Newly discovered diabetes	23	100.0	23	100.0
Negative	8	34.8	10	43.5
Positive (over 130)	15	65.2	13	56.5
Previously known diabetes	8	100.0	8	100.0
Negative	1	12.5	1	12.5
Positive (over 130)	7	87.5	7	87.5

* Includes only cases with known diagnosis and known results.

Table 8. It is noted that even when all four tests were done on each screenee, over one-quarter of the previously undiagnosed diabetics screened negative to all tests (or 71.4 screened positive to at least one of the four tests) with the conditions under which screening was carried out at D. C. General Hospital at random times after eating. Fourteen of the 21 newly diagnosed diabetes cases were screened fasting and only 4 were tested within two hours of eating.

In cases ultimately proved to have newly discovered diabetes the blood sugar screening level (by the Somogyi-Nelson method) was as low as 84; in 8 of the 23 cases, the blood sugar was below 130 mg. per 100 ml. Table 9 shows the sensitivity and specificity ratings testing with the Somogyi-Nelson method at various levels. It will be noted that no case of undiagnosed diabetes screened below 80 mg. per 100 ml. and no person without diabetes screened above 190 mg. per 100 ml. This is a wide range, since only 23 cases of undiagnosed diabetes are included.

TABLE 8

Sensitivity and specificity levels of diabetes screening methods for the 21 newly diagnosed cases of diabetes and the 226 nondiabetics all of whom received the four tests (at random times after eating)

Method	Sensitivity Number Positive	Sensitivity Per cent	Specificity Number Negative	Specificity Per cent
Test for sugar in urine				
by copper reduction method	7*	33.3	192	85.0
by bismuth reduction method	8*	38.1	210	92.9
Blood sugar test				
Wilkerson-Heftmann method over 130 mg. per 100 ml.	12	57.1	219	96.9
Somogyi-Nelson method over 130 mg. per 100 ml.	14	66.7	221	97.8
Positive result on one of four tests	15	71.4		
Negative results on all four tests			183	81.0

* Trace or more

TABLE 9

Sensitivity and specificity ratings of the Somogyi-Nelson test method at various levels at random* times after eating for the 273 persons receiving this test and also a complete diagnostic study

Blood sugar level — If this level were considered positive (mg. per 100 ml.)	Sensitivity — This per cent of newly diagnosed diabetics (23) would screen positive	Specificity — This per cent of nondiabetics (250) would screen negative
80	100.0	40.8
90	95.7	66.8
100	87.0	82.4
110	73.9	91.2
120	69.6	97.2
130	65.2	98.0
140	56.5	98.8
150	52.2	99.2
160	39.1	99.6
170	30.4	99.6
180	26.1	99.6
190	21.7	100.0

* Without regard to lapse of time since eating.

It must again be emphasized that these results are among persons screened at random times after eating in an out-patient clinic population. When the population is classified according to time after eating, one finds that there were only four diabetics who received postprandial blood tests and three of them had positive tests. Obviously then, these results cannot be applied to nonclinic population groups where screening may be much closer to the time of eating and where the renal threshold for glucose may be different. The results of this study, however, do indicate the need for similar studies among other groups and under more controlled conditions to determine screening levels which will give optimum balance between sensitivity and specificity. The Division of Chronic Disease and Tuberculosis has already initiated such a study at Boston City Hospital (Boston, Mass.), under the direction of Dr. Hugh L. C. Wilkerson.

SUMMARY

In a multiple screening project in the Out-patient Department of the D. C. General Hospital, the efficiency of four tests as diabetes case-finding procedures was investigated. Venous blood samples were tested for sugar by the Wilkerson-Heftmann and Somogyi-Nelson methods, and urine samples were tested by a copper reduction method (Clinitest) and by a bismuth reduction method (Galatest). Diagnoses were established by independent investigation in the Medical Diagnostic Clinic. The sensitivity and specificity of these four tests were compared. Sensitivity—the test's ability to classify as positive those who have the condition being screened for—was calculated as the percentage screening positive of those diagnosed as having diabetes. Specificity—the test's ability to classify as negative those who do not have the condition being screened for—was calculated as the percentage screening negative of those determined not to have diabetes.

Among this outpatient clinic population (screened at random times after eating) blood tests gave superior results to urine tests. The specificity ratings of blood sugar tests by the Somogyi-Nelson and Wilkerson-Heftmann methods were satisfactory—98.0 and 96.8 per cent respectively; these figures would indicate that almost all

the nondiabetics were properly classified by the screening test. However, the sensitivity ratings were low—65.2 and 56.5 per cent respectively, which would indicate that a relatively low percentage of the diabetics was detected by the screening test at the 130 mg. per 100 ml. level. In other words, 34.8 and 43.5 per cent respectively of those diagnosed as having diabetes had screening levels below 130 mg. per 100 ml. and theoretically would have been missed if that level had been arbitrarily set as the detection level.

The results of this study indicate a need for similar investigations among other population groups and under conditions subject to better control to determine the most efficient screening levels.

ACKNOWLEDGMENT

The authors gratefully acknowledge the assistance received from Dr. Joseph F. Fazekas, Chief of Staff of D. C. General Hospital; from Dr. Elizabeth H. Hill, Chief of the Out-patient Department; and from the Staff of that Department.

The Common Cause

This is the century of the biological sciences. The physical sciences have already achieved a high degree of consistency and conceptual maturity. In the life sciences, on the other hand, analytical understanding is still no more than rudimentary, unifying concepts are still scarce, and many fundamental principles remain to be discovered. The major job still lies ahead. At the same time, the urgency of the task grows, as mankind looks expectantly to new advances in agriculture, public health, and medicine, whose rational development depends on biological knowledge and understanding.

Biology has grown in volume and diversity to the point where it would be far beyond the capacity of any one individual to acquire competence in more than a limited sector of the field. Biologists, in the sense of miniature incarnations of universal biological knowledge, no longer exist. Biological science has become a group enterprise with many servants in varied stations. The single-celled organism has evolved into a multi-cellular one, and its health, survival and growth depend on the harmonious cooperation of its many specialized members. Anyone contributing to this collective task, constructively, competently, and conscientiously, thus becomes a biologist. Consequently, it takes all kinds of biologists to make the biological world, none of them able to carry on without the others. And biology needs their full diversity.

It needs the observer, the gatherer of facts, the experimenter, the statistician, the theorist, the classifier, the technical expert, the interpreter, the critic, the teacher, the writer. It needs the student of evolutionary history as much as it does the experimental physiologist; the precise recorder of morphological data as much as the analytical biophysicist and biochemist; the investigator of molecular interactions as much as the student of supramolecular organization, of the order of events in space and time. It needs the help of all hands at all stations, from the research man who conceives a new idea, to the assistants who prepare solutions or tend cultures or animals; from the mechanic who builds a new instrument, to the artist or photographer who prepares indelible records of microscopic specimens or physiological tracings; and last, not least, from the man who willingly gives of his time and effort in order to help obtain and distribute some of the most basic tools of science—fellowships, research grants, materials and jobs—to the one who willingly accepts them to good advantage. They all work for a common cause and should feel above the unjustified and undignified popularity contests that center on such monomaniac questions as who is "more important," the "fundamental" or the "applied" scientist; the explorer or the instructor; the technical expert or the philosopher. They are all needed—in their proper stations. And they should be rated not by *what* they are doing but by *how* they are doing it.

From "The Challenge of Biology" by Paul Weiss, Ph.D., M.D., in *Science,* July 1953.