# Sensitivity and Specificity of Five Screening Tests for Diabetes in Ten Countries

*Kelly M. West, M.D., and John M. Kalbfleisch, M.D., Oklahoma City*

SUMMARY

The sensitivity and specificity of each of five screening tests were estimated in each of three to ten countries by testing subjects drawn from the general populations of adults over thirty-four years of age. This permitted comparisons among countries and among the different tests (fasting, postprandial, and postglucose urine tests, and fasting and postprandial blood glucose values).

Sensitivity and specificity of each test varied widely among populations. For example, the sensitivity of the two-hour urine glucose ranged from 17 per cent in Nicaragua to 100 per cent in East Pakistan. Apparently specificity and sensitivity of such tests are influenced by many factors including both the circumstances under which the tests are performed and the characteristics of the population tested. It is, therefore, not possible to predict prevalence rates reliably by extrapolating from the results of screening tests. However, we believe the data for specific populations on the sensitivity and specificity of various tests will provide a rough guide in predicting the cost-effectiveness of alternative approaches to case detection in those particular countries. For instance, these results suggest that roughly 56 per cent of the occult diabetics in Costa Rica in this age group would be detected by a two-hour urine glucose, but only about 41 per cent of those in whom this test was positive would prove to have diabetes.

Even modest changes of criteria in defining either "diabetes" or "abnormality" of the screening results produced marked changes in rates of sensitivity and specificity. With few exceptions, tests which were more sensitive were, comparably, less specific, and the reverse was also true. Rates of "diabetes" were markedly influenced by modest changes in diagnostic criteria. DIABETES *20*:289-96, May, 1971.

As blood glucose determinations have become more economical and more widely available, urine glucose tests have assumed a less important and somewhat different role in screening for diabetes. Yet there remain a variety of circumstances under which urine tests are useful as part of the process of case detection. Moreover, in less affluent societies, the availability of blood glucose determinations is often quite limited. Although a large body of experience exists concerning the factors which influence the concentration of urine glucose, the sensitivity and specificity of urine glucose determinations as screening procedures have rarely been measured systematically under controlled conditions.[1,2] Groups of negative screenees have seldom been systematically tested with glucose tolerance tests. Remein and Wilkerson pointed out that their results in a selected population (adult, outpatient volunteers) would not necessarily reflect precisely the sensitivity and specificity of such tests in a general population.[1] More information is needed concerning the effects of various factors on the sensitivity and specificity of both blood and urine screening procedures. We found, for example, that in marked contrast to our previous experiences elsewhere, in East Pakistan the urine test for glucose was a highly sensitive index of the presence of diabetes even in the older segment of the population (see below).

Several different screening tests were performed in the course of a recent series of surveys.[3-5] Because oral glucose was administered to *all* subjects it was possible to relate the results of the screening tests, both positive *and negative,* to the status of glucose tolerance. This also made it possible to compare the sensitivity and specificity of five different screening tests, and with each test to compare the results among three to ten countries.

## METHODS

Results of studies of the prevalence of diabetes and the relationship of prevalence to various epidemiologic factors have been reported previously for twelve age-matched populations in eleven countries.[3-5] Urine glucose determinations were performed in ten of these countries. The characteristics of these ten groups have been described, as well as the frequency distribution of two-hour blood glucose values for each population. With rare exceptions these samples from general populations

From the Department of Medicine, University of Oklahoma School of Medicine, Oklahoma City, Oklahoma 73104.

included only subjects over thirty-four years of age. The frequency distribution of ages was very similar in all countries.[3,4] Mean age in these ten groups only varied from forty-eight years in Pakistan to fifty-two years in Venezuela.

A total of 4,262 subjects had urine glucose determinations two hours after an oral glucose load of one gram per kilogram of body weight. A specimen of venous blood was also obtained from all subjects at two hours for a determination of glucose concentration. With rare exceptions tests were performed in the morning. In Uruguay, Venezuela, and Malaya, each of a total of 1,530 persons also had other tests as described below. In this subgroup of 1,530, the glucose load was administered in the fasting state in 418 instances. In 1,112 instances, it was given 2 to 10 hours after a meal. However, with few exceptions (less than 10 per cent) those who were not in the fasting state were tested between two and four hours after a meal. Of these 1,112 subjects, 591 also had a "postprandial" blood glucose determination immediately prior to the glucose load. Each of the 1,530 subjects had a urine glucose test immediately prior to, and two hours after, the glucose load. The duration of fasting was determined only by the prevailing administrative circumstances at the thirty-nine sites at which the 1,530 persons were tested.

It was therefore possible in the group of 1,530 to compare the results of the glucose loading test (blood glucose concentration two hours after the load) with each of five screening procedures: (1) a fasting urine glucose test; (2) a postprandial urine glucose test; (3) a urine glucose obtained two hours after the glucose load; (4) a fasting blood glucose; and (5) a postprandial blood glucose value. Although each of the subgroups were drawn from the same universe of three, lumped, age-matched populations, there were differences in some of the characteristics of these subgroups. For example, the subgroup who had fasting urine tests contains a smaller number of Malayans than in certain other subgroups. In some instances these differences probably accounted for part of the differences among the tests in sensitivity and specificity. However, the effects of this imperfect matching can be largely corrected by calculating for each test the average of results for the three countries. In this way the results from each country are given equal weight with each of the tests. These averages are presented in table 1, which also gives the number tested and the results for each country.

Under certain circumstances, glucose tolerance may be improved by eating prior to the load (priming effect).[6] Under our experimental conditions, however, two-hour blood glucose levels were only slightly higher in those tested in the fasting state, and the small difference between the results in fasting and nonfasting subjects was not statistically significant.[3,4]

In East Pakistan and six countries of Central America the glucose load was also administered in the morning; in more than 95 per cent the load was given two to five hours after breakfast.

The Clinistix method was used to measure urine glucose. The accuracy of this semiquantitative method has been evaluated by Moran et al.,[7] by Ackerman et al.,[8] and by O'Sullivan, Kantor and Wilkerson.[9] These studies concerning the accuracy of the Clinistix method are reviewed in the *Discussion.* All observers were required to follow the manufacturer's instructions and to determine whether the specimen was negative or positive. They did not have the option of classifying the test as "trace" or "borderline". In some countries no attempt was made to record the degree of positivity. However, in most cases (see below) glycosuria was graded as "light", "medium", or "dark" according to the interpretation chart provided by the manufacturer. Blood glucose levels were measured by the AutoAnalyzer method. Subjects whose two-hour blood glucose values exceeded 149 mg./100 ml. were classified as diabetic. Under the *Results,* data are presented showing the effects of changing these diagnostic criteria. In the *Discussion,* evidence is reviewed indicating that the two-hour values reflect closely the status of tolerance as revealed by full glucose tolerance tests.

Sensitivity is defined as the percentage of the diabetics who are identified by the screening test. Specificity is defined as the percentage of normals correctly classified by the screening test. Thus, specificity figures provide no information concerning the specificity of *positive* screening tests. We have, therefore, expressed our results in a way which also shows the "specificity" of the positive tests by indicating in the tables the percentage of positives which were true positives.

## RESULTS

### *Five screening tests in Uruguay, Venezuela and Malaya*

Table 1 shows the sensitivity and specificity of five screening tests in each of three populations. These data permit comparisons among countries and among tests. Table 1 also shows results for each test when data for all three countries are lumped. In order to show the effect of giving equal weight to the results from each

TABLE 1

Sensitivity and specificity of five screening tests in three countries

| | A Total tested | B Diabetic* | C Nondiabetic | D Pos. tests | E True pos. | F False pos. | G Neg. tests | H True neg. | I False neg. | Sensitivity† (per cent) | Specificity† (per cent) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Fasting urine glucose | | | | | | | | | | | |
| Uruguay | 136 | 11 | 125 | 7 | 7 | 0 | 129 | 125 | 4 | 63.6 | 100.0 |
| Venezuela | 198 | 20 | 178 | 5 | 4 | 1 | 193 | 177 | 16 | 20.0 | 99.4 |
| Malaya | 84 | 3 | 81 | 1 | 1 | 0 | 83 | 81 | 2 | 33.3 | 100.0 |
| Totals | 418 | 34 | 384 | 13 | 12 | 1 | 405 | 383 | 22 | 35.3 | 99.7 |
| Average for 3 countries | | | | | | | | | | 38.9 | 99.8 |
| Urine glucose 2-4 hours after eating | | | | | | | | | | | |
| Uruguay | 348 | 22 | 326 | 17 | 9 | 8 | 331 | 318 | 13 | 40.9 | 97.5 |
| Venezuela | 282 | 15 | 267 | 5 | 3 | 2 | 277 | 265 | 12 | 20.0 | 99.3 |
| Malaya | 482 | 17 | 465 | 24 | 9 | 15 | 458 | 450 | 8 | 52.9 | 96.8 |
| Totals | 1,112 | 54 | 1,058 | 46 | 21 | 25 | 1,066 | 1,033 | 33 | 38.9 | 97.6 |
| Average for 3 countries | | | | | | | | | | 37.9 | 97.8 |
| Urine glucose 2 hours after oral glucose | | | | | | | | | | | |
| Uruguay | 484 | 33 | 451 | 62 | 25 | 37 | 422 | 414 | 8 | 75.8 | 91.8 |
| Venezuela | 480 | 35 | 445 | 25 | 18 | 7 | 455 | 438 | 17 | 51.4 | 98.4 |
| Malaya | 566 | 20 | 546 | 59 | 16 | 43 | 507 | 503 | 4 | 80.0 | 92.1 |
| Totals | 1,530 | 88 | 1,442 | 146 | 59 | 87 | 1,384 | 1,355 | 29 | 67.0 | 94.0 |
| Average for 3 countries | | | | | | | | | | 69.1 | 94.1 |

| | Total tested | Diabetic* | Nondiabetic | Pos. tests | True pos. | False pos. | Neg. tests | True neg. | False neg. | Sensitivity (per cent) | Specificity (per cent) | "Specificity" of pos. test‡ (per cent) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Fasting blood glucose (values > 119 mg./100 ml. considered positive) | | | | | | | | | | | | |
| Uruguay | 134 | 11 | 123 | 10 | 7 | 3 | 124 | 120 | 4 | 63.6 | 97.6 | 70.0 |
| Venezuela | 120 | 13 | 107 | 5 | 3 | 2 | 115 | 105 | 10 | 23.1 | 98.1 | 60.0 |
| Malaya | 25 | 1 | 24 | 1 | 1 | 0 | 24 | 4 | 0 | 100.0 | 100.0 | 100.0 |
| Totals | 279 | 25 | 254 | 16 | 11 | 5 | 263 | 249 | 12 | 44.0 | 98.0 | 68.8 |
| Average for 3 countries | | | | | | | | | | 62.2 | 98.6 | 76.7 |
| Blood glucose 2-4 hours after eating (values> 129 mg./100 ml. considered positive) | | | | | | | | | | | | |
| Uruguay | 341 | 22 | 319 | 13 | 10 | 3 | 328 | 316 | 12 | 45.5 | 99.1 | 76.9 |
| Venezuela | 150 | 12 | 138 | 11 | 7 | 4 | 139 | 134 | 5 | 58.3 | 97.1 | 63.6 |
| Malaya | 100 | 2 | 98 | 2 | 1 | 1 | 98 | 97 | 1 | 50.0 | 99.0 | 50.0 |
| Totals | 591 | 36 | 555 | 26 | 18 | 8 | 565 | 547 | 18 | 50.0 | 98.6 | 69.2 |
| Average for 3 countries | | | | | | | | | | 51.3 | 98.4 | 63.5 |

* Two-hour blood glucose above 149 mg./100 ml. Of eighty-eight diabetics tested in these three countries, fifty-five (63 per cent) had no previous history of the disorder. The numbers of new and previously known diabetics by country are given in table 3.

† Sensitivity is columns E/B, specificity is H/C, and "specificity" of positive tests is D/E.

‡Percentage of positive tests that were true positives.

country, irrespective of the number tested, the average sensitivity and specificity in the three countries is given for each test.

Table 1 shows that in the 1,530 subjects, the urine glucose test after oral glucose was sensitive (67.0 per cent) and specific (94.4 per cent). Yet, the utility and role of the test is not fully evident until the "specificity" of the positive urine test is given. Table 1 shows that only 40.4 per cent were true positives. In certain circumstances, the disadvantage of this propensity to produce false positives would be outweighed by the other considerations and advantages. For example, even though a majority of the two-hour urine glucose values were false, the prevalence of diabetes in those with positive

tests (40.4 per cent) was more than nineteen times higher than in those with negative two-hour urine tests of whom only 2.1 per cent had diabetes. Since only 9.5 per cent of these subjects (146 of 1,530) had positive tests and 67.0 per cent of all diabetics had positive tests, it would apparently be possible to identify a substantial majority of the diabetics in this group, even if blood glucose tests were performed only in the small subgroup with positive urine tests. One method of doing this would be to administer a carbohydrate load and draw blood only on those with glycosuria.

*Urine tests in East Pakistan*

In East Pakistan, 513 subjects had urine glucose determinations before, and two hours after, oral glucose. The tests before oral glucose were performed on specimens obtained two to five hours after breakfast. Both postprandial and postglucose urine tests identified 100 per cent of the diabetics (9 of 9). Four of thirteen (33 per cent) postprandial positives were false, while 18 of 27 (67 per cent) postglucose specimens were false positives.

*Effects of changing criteria for interpreting tests and for diagnosis*

The blood glucose screening levels shown in table 1 for the fasting (120 mg. per 100 ml.) and the postprandial (130 mg./100 ml.) tests were arrived at arbitrarily. Sensitivity or specificity for each of these tests may be changed at will by lowering or raising the arbitrary values we used. But improvements in sensitivity were achieved at the price of roughly comparable deterioration in specificity, and the reverse was also true. For example, when the postprandial screening level was raised progressively from 120 mg./100 ml. to 130, 140 and 150, respectively, the "specificity" of positive tests (per cent of positives which are true positives) was improved progressively from 52 to 69 per cent, 83 and 100 per cent, respectively. On the other hand, these improvements in the "specificity" of positive tests were attended by a declining sensitivity which fell from 59 per cent to 50, 41, and 38 per cent, respectively.

Most of the subjects with false positive two-hour urine tests had minimal glycosuria. When data from East Pakistan, Malaya, Venezuela and Uruguay were lumped, 64 per cent of the false negative tests (69 of 108) were only slightly positive ("light" glycosuria by the Clinistix interpretation chart). Only 13 per cent of the subjects with "light" glycosuria proved to be diabetic in these four countries, while 50 per cent of those with "medium" glycosuria were diabetic, and 71 per cent with "dark" tests were diabetic.

Diagnostic criteria also had a substantial influence on sensitivity and specificity. For example, in Guatemala only 12 of 18 subjects (67 per cent) with "abnormal" tolerance (two-hour glucose more than 149 mg./100 ml.) had glycosuria after oral glucose, but 100 per cent (12 of 12) of those with two-hour blood glucose values over 189 mg./100 ml. had glycosuria after oral glucose.

We arbitrarily divided the subjects into "diabetic" and "nondiabetic" groups on the basis of their two-hour blood glucose levels after the oral glucose load. Those with values above 149 mg./100 ml. were classified as diabetic and those below 150 were regarded as normal. Two-hour values as low as 120 mg./100 ml. are sometimes considered abnormal. We believe that this level

TABLE 2

Effect on sensitivity and specificity of changing diagnostic criteria*

| Screening tests | Assuming 2-hour blood glucose values greater than 149 mg./100 ml. are diagnostic | | | Assuming 2-hour blood glucose values greater than 119 mg./100 ml. are diagnostic | | |
|---|---|---|---|---|---|---|
| | Sensitivity (per cent) | Specificity (per cent) | "Specificity of positive tests† | Sensitivity (per cent) | Specificity (per cent) | "Specificity" of positive tests† |
| Fasting urine glucose | 35.3 | 99.7 | 92.3 | 18.8 | 99.7 | 92.3 |
| Urine glucose 2-4 hours after eating‡ | 38.9 | 97.6 | 45.7 | 20.2 | 97.6 | 45.7 |
| Urine glucose 2 hours after oral glucose | 67.0 | 94.0 | 40.4 | 41.7 | 94.4 | 47.9 |
| Fasting blood glucose (over 119 mg./100 ml.) | 44.0 | 98.0 | 68.8 | 23.4 | 97.8 | 68.8 |
| Blood glucose 2-4 hours after eating‡ (over 129 mg./100 ml.) | 50.0 | 98.6 | 69.2 | **27.7** | 96.6 | 69.2 |

* These are data on the same 1,530 subjects for whom data are presented in table 1.

† Percentage of positive tests that were true positives.

‡ In a few instances (less than 10 per cent) these specimens were drawn 4 to 10 hours after eating.

is inappropriately low in middle age and older adults.[10] Nevertheless, since this level has been widely used, we also examined the sensitivity and specificity of the five screening tests under the assumption that all individuals with two-hour levels of 120 mg./100 ml. or more were "diabetic." Table 2 shows that the apparent sensitivity of all five tests declined substantially under these assumptions. Specificity of each of the tests was about the same with the two different diagnostic criteria.

When the dividing line between normal and abnormal two-hour blood glucose values was set at 150 mg./100 ml., the prevalence of "diabetes" among these 1,530 subjects was 5.8 per cent. On the other hand, when the dividing line was lowered to 120 mg./100 ml., the apparent prevalence rate became 11.0 per cent. Thus, even when testing methods are the same, relatively modest differences in diagnostic criteria may account for marked differences in apparent rates of prevalence.

*Comparisons among countries*

We had hoped that these studies would allow us to establish, at least for certain age groups, some general relationships between the results with the various tests, as a basis for the development of formulas for estimating prevalence rates from the results of screening tests. One might hypothesize, for example, from certain of the data in table 1 (results for all three countries lumped) that postprandial urine tests and fasting blood glucose determinations would identify in this age group about one third of those who have diabetes, and that urine glucose determinations after a carbohydrate load would identify about two thirds of the diabetics. The establishment of such relationships would also make it possible to estimate prospectively the cost-effectiveness of various approaches to diabetes screening. Unfortunately, our data, when analyzed in detail, showed that the sensitivity and specificity of these tests varied markedly among populations even when populations were matched for prevalence of diabetes and for age. Table 3 shows, for example, that the sensitivity of a urine glucose test two hours after oral glucose in ten populations ranged from 11 per cent in Nicaragua to 100 per cent in East Pakistan. The rate of falsely positive tests in those with glycosuria ranged from 14 per cent in Honduras to 83 per cent in Malaya. Table 1 also shows substantial differences among countries in the sensitivity and specificity of the screening tests.

Table 3 shows that in some countries a substantial majority of the positive two-hour urine tests occurred in persons with normal glucose tolerance. However, this table also shows that this was in some cases mainly the result of the low number of true positives rather than a high frequency of false positives among the nondiabetics.

## DISCUSSION

We recognize that a single two-hour blood glucose value is probably not so definitive as a full glucose tolerance test in classifying individuals as diabetic or nondiabetic. Doubtless a small percentage of our subjects would have been classified differently based on results with one or more full glucose tolerance tests. But interindividual changes in results of duplicate tests are not

**TABLE 3**

Marked variation in ten populations of sensitivity and specificity of urine glucose two hours after oral glucose

| | | No history of diabetes | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Known diabetics | | | Diabetics discovered | | | Nondiabetics | | | | | | |
| | Total | Urine test | | | Urine test | | | Urine test | | | Sensitivity* | Specificity* | "Specificity" of positive tests† |
| Country | no. | No. | Pos. | Neg. | No. | Pos. | Neg. | No. | Pos. | Neg. | (per cent) | (per cent) | (per cent) |
| Uruguay | 484 | 17 | 15 | 2 | 467 | 16 | 10 | 6 | 451 | 37 | 414 | 63 (76) | 91.4 (91.8) | 21 (40) |
| Venezuela | 480 | 7 | 7 | 0 | 473 | 28 | 11 | 17 | 445 | 7 | 438 | 39 (51) | 98.4 (98.4) | 61 (72) |
| Malaya | 566 | 9 | 7 | 2 | 557 | 11 | 9 | 2 | 546 | 43 | 503 | 82 (80) | 91.8 (92.1) | 17 (27) |
| East Pakistan | 513 | 0 | 0 | 0 | 513 | 9 | 9 | 0 | 504 | 29 | 475 | 100(100) | 94.2 (94.2) | 24 (24) |
| Costa Rica | 470 | 4 | 3 | 1 | 466 | 23 | 12 | 11 | 443 | 22 | 421 | 52 (56) | 94.8 (95.0) | 36 (41) |
| El Salvador | 265 | 2 | 2 | 0 | 263 | 7 | 2 | 5 | 256 | 7 | 249 | 29 (44) | 97.3 (97.3) | 22 (36) |
| Guatemala | 413 | 6 | 5 | 1 | 407 | 12 | 7 | 5 | 395 | 7 | 392 | 58 (67) | 98.0 (98.2) | 50 (52) |
| Honduras | 343 | 0 | 0 | 0 | 343 | 16 | 6 | 10 | 327 | 1 | 326 | 38 (38) | 99.7 (99.7) | 86 (86) |
| Nicaragua | 383 | 5 | 2 | 3 | 378 | 18 | 2 | 16 | 360 | 8 | 378 | 11 (17) | 96.9 (97.7) | 20 (33) |
| Panama | 345 | 0 | 0 | 0 | 345 | 10 | 4 | 6 | 335 | 14 | 321 | 40 (40) | 95.8 (95.8) | 22 (22) |
| Totals | 4,262 | 50 | 41 | 9 | 4,212 | 150 | 72 | 78 | 4,062 | 175 | 3,887 | 48 (57) | 95.5 (95.7) | 29 (39) |

* Numbers in parentheses are results when calculations are based on the entire population sample including known diabetics.

† Percentage of positive tests that were true positives.

infrequent when "full" oral glucose tolerance tests are performed employing three or more blood glucose determinations.[11-13] Moreover, even the most elaborate test leaves something to be desired in dividing general populations into "diabetics" and "nondiabetics."[10,14] Jackson and his associates have presented and reviewed evidence to support the sensitivity and specificity of the two-hour value.[2]

In order to further evaluate the validity of the two-hour value as a diagnostic instrument we analyzed data of Unger who performed glucose tolerance tests on a population of food handlers.[11] Each subject had a fasting glucose determination on venous blood, and the blood glucose was also determined on specimens drawn one and two hours after 100 gm. of oral glucose. We found that the two-hour value was above 149 mg./100 ml. in thirty of 152 (19.7 per cent) subjects. Thirty of 152 had values over 398 when the result was expressed as the sum of the three values. If it is assumed that two-hour values above 149 mg./100 ml. and summed values above 398 are diagnostic (in each case, the thirty highest of 152 values), the two diagnostic methods yield very similar results. Only two of the thirty (7 per cent) with "diabetic" two-hour values would have been reclassified as "nondiabetic" by the results of the full glucose tolerance test, and only two of the 122 (1.6 per cent) with "nondiabetic" two-hour values would have been classified as "diabetic" by the full glucose tolerance test.

Previous studies[7-9] have shown that the Clinistix method is highly reliable in identifying the presence of glucose at concentrations above 0.1 per cent. Specimens containing between 0.05 per cent and 0.1 per cent were usually found to be negative, but positive results were not infrequent at this concentration. Specimens containing less than 0.05 per cent glucose were always negative in the study of Ackerman and his associates.[8] Thus, the test is highly specific for glucose. In our studies, many (one to three per country) different physicians, medical students and technicians performed the urine glucose tests. It is quite possible that there were some differences among observers in "reading" the results. However, in view of the simplicity of the test, we believe it is unlikely that our results were affected substantially by inter-individual differences in the performance or interpretation of the tests.

Although 4,262 persons were tested, some of the percentages in the tables are based on small numbers. The high sensitivity of the fasting blood glucose in Malaya, for example, has limited meaning because only twenty-five persons were tested. For this reason, the numbers tested and the numbers of positives and negatives are given in each instance. In several cases the sensitivity figures should be considered only crude estimates because of the small numbers on which they are based.

Our data permit only limited conclusions concerning the reasons for the differences among populations in the sensitivity and specificity of the screening procedures. We were, however, able to account for some of these differences. For instance, the unusual sensitivity of the urine tests in East Pakistan was explained by the frequency distribution of blood glucose values. There were very few people with mild abnormalities of glucose tolerance.[3] The people who had diabetes had gross hyperglycemia.[3] It is not surprising, then, that the urine tests were a highly sensitive index of the presence of diabetes. On the other hand, some of the differences among countries (such as between Nicaragua and Guatemala) could not be attributed to differences in the frequency distribution of blood glucose values because the two groups were very similar in this respect.[4]

One of the factors which affects both sensitivity and specificity of screening procedures is the amount of previous screening. Seventy-five per cent (150 of 200) of the diabetics in the total group of 4,262 had no previous history of diabetes, but there were substantial differences among countries in this respect. In Uruguay, seventeen of thirty-three diabetics had a history of diabetes, while none of the nine diabetics in East Pakistan had a history of the disorder. Tests for diabetes are commonly performed in Uruguay and rarely done in East Pakistan. In order to "correct" for this variable in comparing results in the various countries, we have shown in the tables sensitivity and specificity rates based on calculations which included all subjects in these samples of general populations. Thus, differences among countries in results are not attributable to differences in the amount of previous screening. In table 3, results for each country are also shown when calculations for sensitivity and specificity excluded the small subgroup with known diabetes. Because of certain shortcomings in our coding methods we were not able to calculate the sensitivity and specificity for all tests in Uruguay, Venezuela and Malaya with the known diabetics excluded, but table 3 shows the number of new and previously known diabetics for each country. In these three countries only 33 of 1,530 subjects (2.2 per cent) had known diabetes and only thirty-three of eighty-eight (37 per cent) of those with diabetes were known diabetics.

There are many other possibilities which might account for the differences among groups in the sensitivity and specificity of the screening procedures including

differences in hydration and in renal thresholds for glucose. In each of the populations we tested, the concentrations of the loading doses were the same (25 per cent). Water intake was not specifically restricted or measured prior to or during the tests, but in every country only a very small minority of the subjects consumed water during the test period. In contrast to the experimental conditions in Uruguay, Venezuela, Malaya and East Pakistan, urine specimens were not collected before the glucose load was administered to the Central Americans. This would tend to dilute the glucose excreted after administration of glucose and make the urine test less sensitive and more specific. This would not however, account for differences among the Central American countries because the same procedures were used in all countries with respect to urine collections. The frequency distribution of the duration of the postprandial period was similar in each of the ten populations. However, there were some small differences among countries in this respect that could explain a small part of some of the differences in the results relating to the postprandial urine tests. Although experimental conditions were similar in all of these countries, doubtless the differences mentioned above and other differences, if corrected, would tend to reduce the wide variations of the results. On the other hand, even greater differences might be expected among mass screening programs and in clinical practice with respect to factors such as hydration of the patients, details of procedures for timing and performing the tests, etc. These results indicate the extent to which the sensitivity and specificity of a relatively standardized procedure, such as the two-hour urine glucose, may be influenced by the characteristics of the subjects tested and other factors. A recent report of the Committee on Statistics of the American Diabetes Association has discussed concepts relating to determinations of sensitivity and specificity.[15]

These data also illustrate dramatically the extent to which differences in methods of testing and diagnostic criteria may affect reported rates of prevalence of diabetes. For example, table 1 shows that estimates of prevalence in a population might vary as much as fivefold depending on the methods and diagnostic criteria employed, even if all positive screenees had follow-up glucose loading tests.

While the results based on our lumped data cannot be applied to a specific country, we believe that our results in specific countries would be useful in predicting roughly the cost-effectiveness of various approaches to case finding in those particular countries.

Undoubtedly sensitivity and specificity are to some degree affected by prevalence. For example, if prevalence rates were quite low, specificity rates would tend to be high because false negatives would be infrequent. The effect of prevalence on sensitivity would depend on the frequency distribution of the hyperglycemic values. If a substantial portion of the diabetics had mild impairment of tolerance, false negatives would be frequent and sensitivity low. Probably this is usually the case. However, we did not observe a significant relationship between prevalence and either sensitivity or specificity. This does not, of course, discredit the theoretical relationships between prevalence and either specificity or sensitivity. The failure to demonstrate these relationships was probably attributable to the relatively small number of populations studied and the many other factors which influence sensitivity and specificity.

## ACKNOWLEDGMENT

These studies would not have been possible without the cooperation and participation of hundreds of colleagues in the ten countries where the surveys were performed. This list of persons is too long to permit full and detailed acknowledgments. The work was sponsored by the Interdepartmental Committee on Nutrition for National Defense and successor agencies including the Office of International Research of the National Institutes of Health and the National Clearinghouse for Nutrition and Health of the U. S. Public Health Service. The Institute of Nutrition of Central America and Panama cosponsored the work in Central America.

## REFERENCES

[1] Remein, Q., and Wilkerson, H. C.: Efficiency of screening tests for diabetes. J. Chronic Dis. *13*:6, 1961.

[2] Jackson, W. P. U.: Effectiveness, reproducibility, and weight-relation of screening-tests for diabetes. Lancet *2*:1101-05, 1968.

[3] West, K. M., and Kalbfleisch, J. M.: Glucose tolerance, nutrition, and diabetes in Uruguay, Venezuela, Malaya, and East Pakistan. Diabetes *15*:9-18, 1966.

[4] West, K. M., and Kalbfleisch, J. M.: Diabetes in Central America. Diabetes *19*:656-63, 1970.

[5] West, K. M., and Kalbfleisch, J. M.: Influence of nutritional factors on prevalence of diabetes. Diabetes *20*:99-108, 1971.

[6] Hayner, N. S., et al.: Carbohydrate tolerance and diabetes in a total community, Tecumseh, Michigan. Diabetes *14*:413-23, 1965.

[7] Moran, J. J., Lewis, P. L., Reinhold, J. G., and Lukens, F. D. W.: Enzymatic tests for glucosuria. Diabetes 6:358-62, 1957.

[8] Ackerman, R. F., Williams, E. F., Jr., Packer, H., Hawkes, J. H., and Akler, J.: Comparison of Benedict's solution, Clinitest, Tes-Tape, and Clinistix. Diabetes 7:398-402, 1958.

[9] O'Sullivan, J. B., Kantor, Nathan, and Wilkerson, H. C.:

Comparative value of tests for urinary glucose. Diabetes *11*:53-55, 1962.

[10] West, K. M.: Laboratory diagnosis of diabetes: A reappraisal. Arch. Intern. Med. *117*:187-91, 1966.

[11] Unger, R. H.: The standard two-hour oral glucose tolerance test in the diagnosis of diabetes mellitus in subjects without fasting hyperglycemia. Ann. Intern. Med. *47*:1138-53, 1957.

[12] McDonald, G. W., Fisher, G. F., and Burnham, C.: Reproducibility of the oral glucose tolerance test. Diabetes *14*: 473-80, 1965.

[13] O'Sullivan, J. B., and Hurwitz, D.: Spontaneous remissions in early diabetes mellitus. Arch. Intern. Med. *117*:769-74, 1966.

[14] Keen, H., and Jarrett, R. J.: Modern concepts in the diagnosis of diabetes: *In* The New Management of Stable Adult Diabetes, Charles Weller, Ed., Charles C Thomas, 1969, p. 16.

[15] Klimt, C. R., et al.: Standardization of the oral glucose tolerance test. Diabetes *18*:299-307, 1969.

## *Pancreatic Alpha Cell Function in Man*

(*Continued from page* 288)

In each of the three groups a forty-minute infusion of L-arginine at constant rate caused a rise in plasma glucagon levels. The mean maximum rise in twenty-eight normal subjects was 239 $\mu\mu$g. per milliliter, whereas in both diabetic groups it was similar and significantly greater, being 458 and 452 $\mu\mu$g. per milliliter respectively. If the arginine infusion was given to normal subjects in whom a glucose infusion was already in progress, the absolute levels of plasma glucagon achieved in response to arginine were, at all times, lower than when arginine was given alone. However, the rate and degree of rise from the fasting level were similar with or without glucose. This indicated that, in normal subjects, the plasma glucagon response to arginine is little affected, if at all, by hyperglycemia. This finding enhanced the unusual nature of the observation that the plasma glucagon response to arginine was greater than normal in diabetic subjects, all of whom were hyperglycemic at the time of the arginine infusion.

Further indication that alpha cell function may be disturbed in diabetes mellitus came from observations on patients with severe ketoacidosis or Kimmelstiel-Wilson disease. It was argued that the striking reduction in insulin requirements seen in some patients with Kimmelstiel-Wilson disease could be the consequence of diminished glucagon secretion. This idea was not supported by the finding of a normal plasma glucagon response to arginine in two such patients. In eight patients with severe diabetic ketoacidosis initial plasma glucagon levels averaged 587 $\mu\mu$g. per milliliter. With treatment and restoration of normal diabetic control the plasma glucagon levels declined towards normal.

The salient features of this paper lie in the description of normal pancreatic alpha cell function in man when fasting and in response to stimulation with glucose and arginine. The qualitative nature of the changes described does not differ from that reported earlier in which less specific immunoassays were employed. The importance of the absolute values of the normal subjects is that they form a yardstick against which another facet of abnormal pancreatic islet cell function in the diabetic can be assessed.

The discovery that there is hyperglucagonemia in both juvenile-type and adult-onset diabetics is amplified further by Müller, G. R. Faloona, Aguilar-Parada, and Unger (*New Eng. J. Med. 283*:109, 1970). The authors take up and reiterate an old suggestion that excessive glucagon secretion may be a pathogenetic factor in diabetes and adduce in support of this viewpoint the relative hyperglucagonemia of the diabetics both in the fasting state and in response to arginine.

In a discussion of the possible mechanisms by which an increased glucagon secretion results, the suggestion is made that glucagon release is related, in a reciprocal manner, to the intracellular concentration of glucose in the alpha cell. If glucose transport across the alpha cell membrane is dependent upon insulin, the findings in the diabetic subjects could be explained. The availability of a radioimmunoassay for pancreatic glucagon will permit further analysis of this particular hypothesis and elucidation of the physiology of the alpha cell which has been for so long the poor cousin in the islet family.

From *Nutrition Reviews,* Vol. 29, No. 1
January 1971, pp. 8-10