Special Article

A Summary of Criticisms of the Findings and Conclusions of the University Group Diabetes Program (UGDP)

Holbrooke S. Seltzer, M.D., Dallas

The purpose of the University Group Diabetes Program (UGDP) was to determine whether or not control of blood glucose levels helps to prevent or delay vascular disease in non-insulin-requiring diabetics. Patients were followed in twelve university-affiliated treatment centers ("clinics"), and all data were sent to a coordinating center for analysis. About 200 patients in each of five treatment groups were treated with diet and, respectively, placebo (PLBO), a standard dose of tolbutamide (TOLB), a standard dose of insulin (ISTD), a variable dose of insulin (IVAR), or a standard dose of phenformin (PHEN).[1]*

During eight and one-half years of observation of the first four treatment groups, there were eighty-nine total deaths and sixty-one cardiovascular deaths, whose respective distributions were as follows: PLBO, 21 and 10; TOLB, 30 and 26; ISTD, 20 and 13; IVAR, 18 and 12.[2] Subsequently, after eight and one-third years of observation, the PHEN group had sustained thirty-one total and twenty-six cardiovascular deaths.[3]

The principal conclusions of the UGDP study were that the combination of diet and either tolbutamide or phenformin therapy is no more effective than diet alone in prolonging life; and that diet and either tolbutamide or phenformin may be less effective than diet alone, or diet and insulin, with respect to cardiovascular mortality.

Many investigators, including the author, strongly question the validity of these conclusions because of numerous flaws in the design and execution of the UGDP study and in the analysis and interpretation of its findings. The most cogent criticisms are summarized below:

(A) *Defects in design*: (1) The only enrollment criteria were non-insulin-requiring diabetes diagnosed within the previous year and confirmed by a glucose tolerance test, an estimated five-year prognosis, and absence of a history of ketoacidosis. These criteria allowed 46 per cent of all admissions to have one or more stigmata of existing cardiovascular disease, such as hypertension, significant ECG abnormality, angina pectoris, or history of digitalis use.[1] (2) The delayed introduction of the phenformin group created "two pools" of noncomparable patients.[4,5] (3) Diverse patient populations were used (different treatment centers admitted patients from mass screening programs, charity clinics, or private sources).[4,5] (4) No data were recorded on familial longevity, smoking habits, previous myocardial infarction, or cardiac murmurs[4] (although a family history of diabetes and cardiovascular disease was elicited). (5) Some historical baseline risk factors were not clearly defined, e.g., the criteria for angina pectoris and for congestive heart failure and "use of digitalis."[4] (6) No evaluation was made of nonvascular "co-morbid states" that might affect prognosis, such as chronic hepatic, renal, or pulmonary disease.[4] (7) Some baseline tests, e.g., serum creatinine, were not performed in a central laboratory; (the frequency of elevated creatinine levels was disproportionately high in two clinics).[2,4] (8) No plans were made to control other variables, besides hyperglycemia, which might predispose to vascular disease. (9) Contrary to clinical practice, fixed daily dosages of TOLB (1.5 gm.) and PHEN (100 mg.) were used.[4]

(B) *Defects in execution*: (1) Sixty-nine patients, or one of every fifteen admissions, were nondiabetic by the stated UGDP criteria.[1,4,5] (2) In four clinics the oral glucose tolerance test protocol was not followed;

*The first four treatment groups (PLBO, TOLB, ISTD, IVAR) included 823 patients admitted to all twelve clinics between February, 1961, and February, 1966. The fifth group (PHEN) included 204 patients admitted to only six clinics between August, 1962, and February, 1966. Termination dates were October 7, 1969, for the TOLB findings[2] and January 6, 1971, for the PHEN data.[3]

Chief of Endocrinology, Veterans Administration Hospital, and Professor of Internal Medicine, the University of Texas Southwestern Medical School, Dallas, Texas.

serum, instead of whole blood, determinations were made on 280 patients, and only three, instead of four, glucose values were obtained on twenty-three patients.[1] (3) Many baseline data were missing. Fourteen measurements of pretreatment health status were to have been obtained on all 1,027 patients. For any single measurement, however, from nine to 123 patients had no values recorded. This resulted in a total of 522 missing values (3.6 per cent of all intended measurements), which in turn modified the relative frequencies of baseline risk factors among the five treatment groups.[1,3,4] (4) The clinical prognosis of many patients was clearly less than five years. Of these adult-onset diabetics, 46 per cent were fifty-five to seventy-nine years of age on admission, and the same percentage had known cardiovascular disease.[1] Actually, sixty of the eighty-nine deaths in the first four treatment groups occurred within five years after enrollment. Moreover, three of four patients who died of rheumatic heart disease did so within two years after admission; all four were receiving digitalis when enrolled, and three of them were more than fifty-eight years of age.[2] (5) Patients admitted to the twelve clinics differed greatly with respect to baseline health status (charity patients with a considerable incidence of cardiovascular and other diseases were used in some clinics, whereas others used private patients without other diseases). The 1,027 patients, therefore, represented a heterogeneous study population instead of the homogeneous one required for a prospective therapeutic trial like that of the UGDP.[4,5] (6) Among the five treatment groups, as well as among clinics, baseline risk factors were also unevenly distributed. This was due to simple randomization of patients without subsequent "stratification" to correct for chance preponderance of antecedent risk factors in one or more treatment groups (which actually occurred). (7) The use of fixed amounts of TOLB, PHEN, and ISTD resulted in an unknown number of patients in each group being overtreated or undertreated in terms of blood glucose control at various times during follow-up. (8) Smoking, hypertension, obesity, hypercholesterolemia, and hypertriglyceridemia were not controlled during the study; and hyperglycemia itself was inadequately controlled in all except half of the IVAR patients.[2] (9) Autopsy verification of cardiovascular mortality was obtained in 50 per cent of TOLB deaths but in only 20 per cent of PLBO deaths.[2]

(C) *Defects in over-all statistical analysis:* (1) No attempt was made to assess the "effect of diabetes" per se, i.e., the influence of longstanding diabetes (albeit recently diagnosed) on vascular disease and mortality. (2) The criteria defining some baseline risk factors differed in the published TOLB and PHEN reports[2,3] from those originally used. For example, when resting electrocardiograms were originally read by the complete "Minnesota Code,"[1,6] 19 per cent of all patients had "one or more ECG abnormalities;"[7] when ultimately interpreted by a "modified Minnesota Code,"[8,9] only 4 per cent of 999 adult-onset diabetics had a "significant ECG abnormality."[2,3] (3) The mode of admitting phenformin patients (the group was started eighteen months after the first four groups, the patients were enrolled in only six of the twelve clinics, and there were three PHEN admissions for each admission to the other four groups in those six clinics) created "two pools" of patients that should not have been combined for statistical analysis.[4,5] (4) Deaths in twelve clinics with diverse patient populations were pooled and analyzed, a questionable maneuver.[4,5] Actually, the three clinics that enrolled the sickest patients in the first four treatment groups sustained the most fatalities (twenty-three, eighteen, and twelve deaths), whereas the three clinics that admitted the healthiest patients sustained the least (one, two and three deaths).[2] As for the twenty-six cardiovascular deaths in TOLB patients, two clinics contributed thirteen of them and another contributed four.[2] (5) Evaluation of a "data-dictated endpoint" is statistically improper, since it maximizes the possibility of finding a significant difference by picking an extreme variation for analysis.[10] After only two years of follow-up, for example, cardiovascular mortality in IVAR patients was twice as high as in other treatment groups—the same self-limited phenomenon that subsequently occurred in the TOLB group.[2,10] (6) The numbers of total and cardiovascular deaths in each treatment group were too small to permit meaningful statistical comparisons between groups. (7) There were no significant differences among total fatalities per treatment group; this is pertinent because in a toxicity study (which the UGDP turned out to be) it is statistically untenable to analyze deaths from specific cause (e.g., cardiovascular) separately from total deaths.[10] In this regard, the seven cancer deaths in the PLBO group compared to only two in the TOLB category also represented a significant difference, but in the opposite direction.[4]

(D) *Defects in comparison of TOLB versus PLBO deaths:* (1) Twelve of fourteen baseline risk factors occurred more frequently in the TOLB group than in the PLBO patients, including four characteristics subsequently associated with the highest mortality in the

first four treatment groups.[4] (2) The widely different causes of death in the two groups[2] reflected their nonhomogeneous beginnings. In PLBO patients there were no fatalities either from myocardial infarction (MI) or arteriosclerotic heart disease (ASHD), whereas there were seven cancer deaths. Conversely, in TOLB patients there were ten MI and four ASHD deaths, but only two cancer deaths. Actually, the complete lack of MI and ASHD deaths in 205 adult-onset diabetics treated with PLBO and followed for four to eight years is both clinically and statistically incredible.[5] (3) Another surprising negative correlation was that none of the fourteen TOLB patients who died of MI or ASHD had a significant ECG abnormality at baseline, nor did the six similar fatalities in the ISTD and IVAR groups.[2] (4) Even among early dropouts, fatalities were assigned to their original treatment groups. For example, two cardiovascular deaths attributed to the TOLB group occurred after five and one-half and six and one-half years of follow-up, in patients who had taken the drug for only three months and twenty-one months, respectively.[2] (5) Since there were almost twice as many nonfatal myocardial infarctions* in PLBO as in TOLB patients, the sums of fatal and nonfatal MIs amounted to twenty-eight for PLBO and twenty-six for TOLB, i.e., no difference.[2] (6) Cardiovascular mortality in PLBO patients (48 per cent of total PLBO mortality, compared with 65 to 87 per cent of total mortality in the other four groups, respectively) was exceptionally low for any group of adult-onset diabetics. (7) The chi-square analysis and two other statistical technics used to show that cardiovascular deaths in TOLB patients were significantly greater than in PLBO patients were all based on the same noncomparable data obtained from originally heterogeneous treatment groups.[4,10] (8) The published cumulative death rates for TOLB compared to PLBO patients[2] differed sharply from the patterns of mortality that were actually observed year by year. Thus, cumulative cardiovascular mortality rates predicted by "life table methods" showed progressively increasing fatalities in TOLB patients through the eighth year of follow-up, compared to no PLBO mortality at all during the eighth year.[2] Actually, although deaths in TOLB patients did rise steadily for three years, they then leveled off and equaled those in the PLBO patients during the last fifteen months of follow-up.

(E) *Defects in interpretation:* (1) The nonhomogeneous baseline health status of patients in different treatment groups (especially TOLB compared to PLBO) prevented distinction of true drug effects from those of uncontrolled variables, like duration of diabetes and underlying cardiovascular disease, smoking, hypertension, obesity, and hyperlipidemia. (2) Patients treated with diet only (the PLBO group) did not truly represent "dietary control," since the majority of patients in all treatment groups were initially obese and remained so throughout the study.[12] (3) Similarly, the failure of fixed doses of TOLB, PHEN, and ISTD to maintain blood sugar control does not mean they are ineffective agents when properly used in conjunction with strict dietary control. (4) The ancillary UGDP conclusion that control of hyperglycemia does not prevent vascular disease[12] is untenable because fasting hyperglycemia was not controlled, nor were other variables predisposing to vascular disease. (5) Cardiovascular mortality was identical in patients treated with TOLB and PHEN, two compounds with entirely different structures and modes of action. This fact, together with the great excess of TOLB mortality in the six "nonphenformin clinics" compared to its non-excess in the six "phenformin clinics,"[13] suggests that over-all TOLB mortality was largely due to enrollment of too many patients with advanced cardiovascular disease in at least three of the first six clinics.* (6) Discontinuation of TOLB therapy after only four to eight years of follow-up prevented clear-cut interpretation of the thirty-six month spurt of cardiovascular deaths in TOLB patients that was followed by fifteen months of nonexcessive mortality before the drug was stopped on October 7, 1969.† (7) The high cardiovascular mortality attributed to TOLB administration does not necessarily apply to other sulfonylurea agents. (8) The UGDP findings in only 1,027 patients are not applicable to the four million adult-onset diabetics in this country.

*As defined by the modified ECG criteria used in the published UGDP reports.[2,3] The latter were considered prognostically equivalent to a proved myocardial infarction.[8,9,11]

*The reasons for the equally excessive cardiovascular mortality in PHEN patients remain unclear for lack of published data enabling similar correlation of baseline risk factors and deaths per clinic.

†Phenformin was stopped fifteen months later (January 6, 1971). During that time two more cardiovascular deaths occurred—one each in ISTD and IVAR.[2,3] Two additional cardiovascular deaths occurred—one each in PLBO and PHEN—before the phenformin data were submitted for publication.[3] No deaths in the former TOLB group have been reported since October, 1969.

REFERENCES

[1] A study of the effects of hypoglycemic agents on vascular complications in patients with adult-onset diabetes: I. Design, methods and baseline results. University Group Diabetes Program. Diabetes *19* (Supp. 2):747-83, 1970.

[2] A study of the effects of hypoglycemic agents on vascular complications in patients with adult-onset diabetes: II. Mortality results, University Group Diabetes Program. Diabetes *19* (Supp. 2):787-830, 1970.

[3] Knatterud, G. L., Meinert, C. L., Klimt, C. R., Osborne, R. K., and Martin, D. B.: Effects of hypoglycemic agents on vascular complications in patients with adult-onset diabetes. IV. A preliminary report on phenformin results, University Group Diabetes Program. JAMA *217*:777-84, 1971.

[4] Feinstein, A. R.: Clinical biostatistics—VIII. An analytic appraisal of the University Group Diabetes Program (UGDP) study. Clin. Pharmacol. Ther. *12*:167-91, 1971.

[5] Schor, S.: The University Group Diabetes Program. A statistician looks at the mortality results. JAMA *217*:1671-75, 1971.

[6] Blackburn, H., Keys, A., Simonson, E., Rautaharju, P., and Punsar, S.: The electrocardiogram in population studies. Circulation *21*:1160-75, 1960.

[7] Klimt, C. R., Meinert, C. L., Miller, M., and Knowles, H. C.: University Group Diabetes Program (UGDP): A study of the relationships of therapy to vascular and other complications of diabetes. *In* Tolbutamide . . . After Ten Years, Butterfield, W. J. H., and Van Westering, W., editors, New York, Excerpta Medica Foundation, 1967, pp. 261-69.

[8] Rose, G. A., and Blackburn, H.: Cardiovascular survey methods, World Health Organization Monograph Series No. 56, Geneva, World Health Organization, 1968, pp. 137-54.

[9] Cornfield, J.: The University Group Diabetes Program: A further statistical analysis of the mortality findings. JAMA *217*:1676-87, 1971.

[10] O'Sullivan, J. B., Mahan, C. M., and D'Agostino, R. B.: Critique of the UGDP mortality analyses. Second International Diabetes Symposium, Curacao, November 1971. Academic Press, New York. In press.

[11] Meinert, C. L.: Personal communication.

[12] Goldner, M. G., Knatterud, G. L., and Prout, T. E.: Effects of hypoglycemic agents on vascular complications in patients with adult-onset diabetes. III. Clinical implications of UGDP results, University Group Diabetes Program. JAMA *218*:1400-10, 1971.

[13] Seltzer, H. S.: Tolbutamide mortality in the six "phenformin clinics." JAMA *218*:594, 1971.