Diabetologia 18, 179–185 (1980)

Diabetologia

Editorial Review

The Crux of the UGDP

Spurious Results and Biologically Inappropriate Data Analysis

C. Kilo, J. Ph. Miller, and J. R. Williamson

Washington University School of Medicine, St. Louis, Missouri, USA

In order to determine whether vascular complications associated with asymptomatic adult onset diabetes mellitus could be reduced or prevented by treatment with blood glucose lowering agents, a long term prospective clinical trial known as the University Group Diabetes Program (UGDP) was initiated in 1961. Over 1,000 non-ketotic, noninsulin-requiring maturity onset diabetics in 12 different medical centres were assigned one of five treatment schedules (approximately 200 subjects each) as follows: 1) Placebo (PLBO) – diet alone; 2) Tolbutamide (TOLB) – diet plus a fixed dose of 1.5 g of tolbutamide per day; 3) Insulin standard (ISTD) – diet plus a fixed dose of insulin (ranging from 10–16 units per day depending on body surface area); 4) Insulin variable (IVAR) – diet plus insulin given in variable dosage in an effort to normalize fasting blood glucose levels; and 5) Phenformin (PHEN) – diet plus a fixed dose of 100 mg of phenformin per day.

The conclusions drawn from the study [1, 2, 3] have evoked major controversies regarding: 1) the safety of oral hypoglycaemic agents and 2) the potential of oral agents and of insulin to prevent vascular complications associated with diabetes mellitus. These issues obviously have important implications regarding the pathogenesis of diabetic vascular disease as well as the care and management of patients. The study, however, has been criticized by statisticians and clinical investigators alike [4–15] and several significant problems have been noted in its design, execution and analysis. For these reasons it is important to ascertain whether interpretations and conclusions different from those reached by UGDP investigators are also compatible with the data. Our own review of published UGDP data and unpublished patient records recently released by the UGDP Coordinating Center indicates that UGDP data are indeed equally compatible with interpretations contrary to those of UGDP investigators [16–18].

We concur with Feinstein [15] that differences in philosophical approaches to data analysis and interpretation are at the heart of the UGDP controversy. The central issue concerns the selection and use of methods of data analysis consonant with sound statistical principles as well as with the biopharmacological nature of the problem(s) under study.

UGDP investigators contend [19] that: 1) "The main difficulty with the UGDP is not its design, execution or analysis, but rather that it reached an unpopular conclusion," 2) "No amount of criticism . . . can alter the findings of the UGDP," and 3) "The unfortunate aspect of the controversy is that it has served as a distraction from the real implications of the study concerning the absence of efficacy of the treatments tested." The position of UGDP investigators appears to be that the likelihood of their findings being flawed by spurious results or accounted for by factors other than those they have suggested is so remote it can be ignored.

It is our contention that: 1) statistical methods of analysis employed by UGDP investigators were inappropriate to the biopharmacological nature of the questions addressed and the conclusions drawn, 2) UGDP conclusions regarding inefficacy of the treatments tested must be tempered in light of effects of noncompliance (failure to take medications) and 3) in light of the anomalous sex ratio of cardiovascular death rates in the placebo group, UGDP conclusions based on comparison with that group are suspect and should not be extrapolated to the care and management of the general diabetic population. In this communication, we consider the impact of these problems and findings on the analysis and interpretation of UGDP data.

We agree with Schor [5] and O'Sullivan and D'Agostino [12] that data from large, expensive studies such as the UGDP, which cannot be easily repeated, should be thoroughly analyzed by a variety

0012-186X/80/0018/0179/$01.40

Table 1. UGDP cardiovascular death rates by sex

	Men		Women		Men/Women
	%	#	%	#	
PLBO	11.1[a]	(7/63)[b]	2.1	(3/143)	5.3[c]
TOLB	17.5	(11/63)	10.6	(15/141)	1.6
ISTD	8.8	(5/57)	5.2	(8/153)	1.7
IVAR	8.7	(4/46)	7.0	(11/158)	1.2
PHEN	20.0	(13/65)	11.5	(16/139)	1.7

[a] Cardiovascular death rate (%) during the tolbutamide phase of the study except for phenformin (PHEN). In that group figures apply to the phenformin phase

[b] Number of cardiovascular deaths followed by number of subjects at risk

[c] Ratio of cardiovascular death rate for men divided by that for women

of relevant approaches. If results from analyses based on different assumptions are in agreement, the conclusions drawn gain credibility. If results differ, then any conclusions drawn must be tempered in accordance with assumptions on which the analyses are based, with the biological nature of the question(s) addressed, and with the clinical circumstances of the study. Concordance of the results with findings from other relevant investigations must also be considered. In addition, observations gain credibility if they are homogeneous across relevant subpopulations and if they are consistent with biological principles.

One very useful analytical technique for these assessments and for elucidating biopharmacological interactions in clinical studies is subgroup analysis based on biomedically relevant criteria. Indeed, UGDP investigators themselves employed subgroup analysis techniques to examine effects of age, sex, and clinic differences. Obviously, it is desirable to anticipate subgroups of interest and to formulate hypotheses to be tested prior to initiation of the study, since findings derived from retrospective analyses after examination of the data are more likely to be biased (although it is impossible to assess the extent or even the presence of such bias in any specific analysis). For this reason potentially important findings derived from retrospectively formulated analyses can and should be evaluated in subsequent prospective studies.

Spurious Data: The Anomalous Sex Ratio of Cardiovascular Deaths in Placebo Subjects

Perhaps the first indication of anomaly in UGDP data was the finding that the two oral agents tested (tolbutamide and phenformin), which lower blood glucose by different mechanisms, were both associated with increased cardiovascular mortality compared to the placebo group, even though no evidence of tolbutamide toxicity had been reported in several previous studies. Furthermore, cardiovascular deaths in the IVAR group (whose fasting blood glucose values were maintained close to normal) were comparable to those in PLBO subjects. Thus, the effects of three different pharmacological agents (a sulfonylurea, a biguanide, and insulin) on cardiovascular deaths were contrary to expectation. Since the credibility of each of these observations hinges on the validity of the PLBO cardiovascular death rate, it is obviously important to examine critically that data in particular for any evidence of anomaly.

As first noted by O'Sullivan and D'Agostino [12], the sex ratio of cardiovascular death rates in placebo subjects (in data published by UGDP investigators [1]) and confirmed by the Biometric Committee [9] was 5.3 males to 1 female. This ratio is highly atypical for the general diabetic population in whom cardiovascular deaths in women tend to equal or exceed those in men [20–23]; it also differs markedly from that for each of the other treatment groups in the UGDP (Table 1). It is somewhat surprising therefore that neither UGDP investigators nor the Biometric Committee who reviewed the UGDP data commented on these discrepancies. Note that these important findings do not hinge on subgroup analysis or correction for noncompliance and therefore are not subject to any potential bias associated with those procedures (see below). It is of interest that the sex ratio of cardiovascular death rates in placebo subjects in the Bedford study, 0.7 males to 1 female in Table A 8.3 of the Biometric Committee report [9], is consistent with that from other surveys.

These findings document that cardiovascular deaths in UGDP PLBO subjects were atypical for diabetics in general as well as for all of the other UGDP treatment groups. For this reason UGDP conclusions based on comparison with the PLBO group are suspect and should not be extrapolated to the care and management of the general diabetic population. In addition, these findings account substantially for the anomalous low overall cardiovascular mortality rate in the placebo group, since about 70% of UGDP subjects were women. This, in turn, gave the false impression of increased cardiovascular death rates in each of the other treatment groups which was interpreted as evidence of tolbutamide and phenformin toxicity and lack of insulin efficacy in prevention of cardiovascular deaths.

The alternative to this interpretation (if one chooses not to question the validity of UGDP data) is that tolbutamide and phenformin are more harmful

to women than to men and that insulin is harmful to women but not to men even though cardiovascular death rates in women are actually lower than those in men in all three treatment groups. This alternative explanation, however, does not resolve the serious discrepancy in the sex ratio of UGDP PLBO cardiovascular deaths versus that for the other treatment groups and for diabetics in general.

Since this single anomalous result accounts for, and renders untenable, virtually all of the controversial UGDP conclusions, the question is often asked, "How could such a devastating anomalous result occur in a study as carefully planned and executed as the UGDP?" In our opinion, the anomalous sex ratio of cardiovascular death rates in the placebo group is either a spuriously low cardiovascular mortality per se, or it is the consequence of randomization failure in assigning subjects to treatment groups. If evidence of the latter were demonstrable (by differences in baseline variables), the problem could be corrected by statistical procedures, e. g., by a multiple logistic regression model similar to that utilized by UGDP investigators. The Biometric Committee did, in fact, document a highly significant randomization failure in the assigment of subjects by sex to treatment groups within clinics [9]; however, it was their opinion that this randomization failure did not account for the higher cardiovascular mortality rate in tolbutamide-treated subjects (compared to placebo subjects) reported by UGDP investigators. Since the randomization failure they observed and the anomalous cardiovascular death rate in PLBO subjects are both concerned with sex ratio differences in treatment groups, we still have reservations regarding the Biometric Committee's assessment that the randomization failure they discovered was of no consequence. If they are correct, however, then the anomalous sex ratio of cardiovascular death rates in the PLBO group is indeed a spurious result which is not correctable by any statistical procedure. The determination that a result is spurious ultimately rests on demonstrating that it is biologically incoherent, i. e., inconsonant with other data (as discussed above) and/or established biological principles.

In the context of these considerations, it is useful to recall the logic of statistical significance tests. The observed difference between two groups is compared to the probability of observing that large a difference (or larger) between two random samples drawn from the same population. Only when this latter possibility is quite low are the two groups considered to differ significantly from each other. It is customary to indicate the probability (P) that such a finding might be due to chance or random selection, i. e., $P < 0.01$. Unfortunately, even when an apparent significant difference is observed, it is impossible to know whether the difference is indeed real, or whether it is due to sampling "extremes" from the same population. A spurious result (difference) may be due to either group being too high or too low. Although in general, spurious results can neither be foreseen nor prevented, their occurrence can be minimized. As pointed out by O'Sullivan and D'Agostino [12], the fact that the UGDP decision to discontinue tolbutamide was based on inspection of paired (PLBO and TOLB) data, selecting for extreme divergence without considering the basis for the divergence (i. e., possible anomalous behavior of the PLBO group), maximized the likelihood of obtaining a spurious result due to random fluctuation of cardiovascular death rates (in either the PLBO or the TOLB group or both).

Additional evidence supporting the interpretation that the placebo cardiovascular death rate in the TOLB phase of the study was spuriously low due to random fluctuations is provided by the findings that [1] the PLBO cardiovascular death rate more than doubled during the subsequent insulin phase of the study, while that for both insulin treated groups was identical for both time periods and [2] the overall cardiovascular death rate for the entire study period was identical in all three groups.

In view of these and other considerations [18], it is our interpretation that the cardiovascular death rate in the placebo group, more specifically in placebo women, was spuriously low due to random fluctuation [18].

The Problem of Bias

A major concern in the design and analysis of clinical trials is the elimination of bias. 1) Bias in selection of subjects who participate in a study may preclude extrapolation of any important new findings to the target population, e. g., noninsulin-requiring diabetics in general. In view of the fact that a high proportion of diabetics selected for participation in the UGDP were from clinic populations in whom compliance is a serious problem (as documented by UGDP investigators themselves), it has been questioned whether the conclusions drawn from the study are applicable to diabetics in general. 2) Subjects assigned to different treatment groups should be alike in every respect (regarding risk factors relevant to the outcome of the study), otherwise any medication effect (or lack thereof) observed may not be ascribable to the medication itself. It has been well documented that the tolbutamide treatment group was allotted substantially more than its share of sub-

jects with risk factors associated with increased cardiovascular mortality. Although UGDP investigators used a multivariate logistic model in an effort to compensate for these disparities, details of their methodology have not been released, and the validity of the model is not known. Feinstein [4] has criticized the batch randomization procedure employed for assigning subjects to treatment groups as well as the fact that no allowance was made in the multivariate logistic model for interaction of multiple risk factors in the same subject. 3) Failure of subjects to take assigned medications (noncompliance) and/or changes in their biomedical characteristics during the course of the study, as well as administrative decisions to modify treatment regimens (in light of effects observed), may bias analyses against demonstration of medication effects and must be carefully considered in interpreting the data.

Treatment Failure Due to Noncompliance versus Primary Pharmacological Inefficacy

Obviously, the importance of distinguishing between medication failure due to primary pharmacological inefficacy versus that due to noncompliance is that if it is the latter, increased efforts should be devoted to educate and motivate subjects to comply with the treatment regimen; if it is the former, different pharmacological agents or approaches to therapy must be sought.

UGDP investigators chose to ignore medication changes by subjects in their analysis of the data. This conservative approach avoids the risk of unbalancing the randomization achieved (in assigning subjects to treatments) that might occur if noncompliers are excluded from analysis. Furthermore, they argued that the conservative strategy of including subjects who changed medication strengthened the credibility of their observation that cardiovascular mortality was higher in tolbutamide and phenformin treated subjects than in PLBO subjects since inclusion of such subjects would tend to obscure medication effects and bias the data against demonstrating any effects among compliers.

By the same logic, however, if no medication effect is discerned, as was the case for insulin (and the the failure of oral agents to maintain the lower blood glucose levels achieved initially), the conclusion that the medication (insulin) lacks pharmacological efficacy loses credibility if based on analyses including noncompliers! Thus, while it would be valid to conclude that insulin was no better than diet alone in reducing cardiovascular deaths for the group as a whole (including noncompliers) when used under conditions governing the UGDP, it would be inappropriate to imply that insulin lacks the potential (pharmacological efficacy) to reduce cardiovascular deaths if administered more physiologically under circumstances more conductive to patient compliance. Indeed, the overwhelming majority of insulin-treated UGDP subjects received only a single daily injection of intermediate (duration of action) insulin, whereas recent studies have shown that much better control of blood glucose levels (with normalization of other cardiovascular risk factors) can be achieved with two or more daily injections of mixed insulin (and with less risk of hypoglycemic reactions; [25, 26]).

In order to determine whether a beneficial effect of insulin was obscured by inclusion of subjects who changed medication, it is necessary to reanalyze the data after correcting for noncompliance. Thus, subgroup analysis is mandatory in order to differentiate medication failure due to noncompliance (for whatever reason) from that due to pharmacological inefficacy. In our opinion, it is not enough merely to exclude subjects known to have changed medication; the analysis should be restricted to those known (or at least likely) to have taken the medication. Subjects who have been unavailable for follow-up for considerable time periods, i. e., one year or longer, are highly suspect in this regard; presumably they would have had to obtain their medication from sources other than UGDP investigators. In view of the fact that a high proportion of the IVAR group who died of cardiovascular causes was assigned very low doses of insulin, as discussed below, it is most unlikely that they would have pursued medication from sources outside the UGDP.

There is no general agreement on how to correct for noncompliance. The simplest, and possibly the most effective, approach is to exclude from analysis all subjects who fail to comply for significant periods of time. We therefore reanalyzed the data after excluding all subjects who had changed medication and/or missed four consecutive quarterly visits or more by the end of the insulin phase of the study. The efficacy of these criteria is attested to by the fact that excluded subjects missed an average of 25 quarters (over 6 years), i. e., one-half the duration of the study; those retained missed an average of only 2.5 quarters of follow-up during the entire study [18]. The magnitude of the medication change/dropout problem is attested to by the fact that by the end of the insulin phase of the study, 42% of all insulin-treated and placebo subjects had been off their originally assigned medication and/or were unavailable for follow-up for one half of the duration of the study!

Since so many subjects had dropped out and/or changed medication, it was essential to check the distribution of baseline cardiovascular risk factors, by treatment groups, among those retained for data analysis. Inspection of the data revealed that disproportionate numbers of IVAR (relative to PLBO and ISTD) subjects who subsequently died of cardiovascular causes were at increased risk at the time they entered the study in terms of age > 70, diastolic blood pressure > 110 mmHg and fasting blood glucose levels > 150 mg/dl. This situation appeared to be attributable to the fact that: 1) more IVAR than PLBO subjects possessed these risk factors at baseline, and 2) more PLBO and ISTD subjects than IVAR subjects who dropped out and/or changed medication possessed these risk factors [18]. When all subjects with baseline risk factors in excess of these values were excluded, so that the remaining subjects were more comparable (at baseline) in terms of age, severity of diabetes and diastolic blood pressure, we found that cardiovascular mortality was four times higher in PLBO (16.9%) and ISTD (14.7%) than in IVAR subjects (4.4%) [16]. If subjects with baseline FBS values up to 200 mg/dl were included, a two fold difference in mortality between IVAR and the other two groups was still demonstrable. Thus, evidence of insulin efficacy in this subgroup is independent of the flawed PLBO group since the difference in mortality is present between ISTD and IVAR subjects as well as between PLBO and IVAR subjects.

It may be argued that 1) since the subjects remaining in this subgroup constitute such a small fraction of those recruited into the study and 2) since they were selected by "a posteriori" criteria, the data are likely to be biased and it is inappropriate to draw any conclusions from them. We would emphasize that although the numbers of subjects remaining in the subgroup are much smaller than for the whole group (N = ~ 70 for each treatment group compared to ~ 200 originally), they still comprise over 30% of recruited subjects, more importantly, however, they are the only relevant subjects to whom the question of pharmacological efficacy of insulin should be addressed. The selection criteria used are medically relevant and they were applied uniformly to data obtained at baseline from all subjects. A. Bradford Hill has noted [27], ". . . Large numbers in themselves are worse than useless if the groups are not comparable (or, we would add, appropriate to the question addressed), since they encourage confidence in an erroneous opinion." While there is a greater likelihood of selection bias in this subgroup than in the total study group, the presence of bias is by no means a certainty. On the other hand, as discussed above, the data on which UGDP investigators based their conclusion that insulin was inefficacious was clearly biased against demonstrating any medication effect.

The fact that insulin efficacy in reducing cardiovascular deaths was demonstrable only in IVAR subjects with relatively mild diabetes is not particularly surprising in light of the fact that insulin was administered in a single daily injection (almost without exception) regardless of the severity of diabetes. Indeed, single daily injections of insulin would be expected to come much closer to normalizing consequences of insulin deficiency in mild diabetics than in those with more severe diabetes who need, and benefit from, more physiological insulin injections as discussed above.

Thus, while further studies are needed to confirm and extend evidence of insulin efficacy in this subgroup, the UGDP conclusion that insulin was not efficacious must be tempered in accordance with this finding and with other considerations discussed above and below.

Regardless of any question regarding the interpretation and significance of the findings discussed above, the credibility of the UGDP conclusion that insulin was inefficacious in reducing cardiovascular deaths is greatly weakened by the fact that almost one-half (45%) of the 29 IVAR subjects who died of cardiovascular causes were virtually untreated with insulin [18].

1) Five of the 29 received an average of only 10 units of insulin or less per day for an average of only 2.2 quarters and were off insulin on their last visit prior to death (one had not been seen for 2 years and another for 9 years prior to death).

2) Another 5 had not returned for follow-up visits for an average of 3.5 years (range 1–7 years) prior to death (3 were on 10 units of insulin or less per day and 4 of the 5 had been seen less than 4 follow-up visits each).

3) Three more subjects died shortly after entering the study with only 3 follow-up visits each (2 of these received only 15 units of insulin per day).

Thus, these 13 subjects received very small amounts of insulin (in subjects who completed 35 quarters of follow-up, daily insulin dosage ranged from 5–40 units in the 1st quarter and 5–240 units in the 35th quarter [3]) and/or died shortly after entering the programme well before any substantial beneficial effect of insulin on vascular complications would be anticipated. These clinical circumstances preclude attributing cardiovascular death to lack of insulin efficacy.

The Question of Tolbutamide and Phenformin Toxicity and Efficacy

Tolbutamide and phenformin were both discontinued before termination of the study because cardiovascular deaths appeared to be excessive relative to those in placebo subjects and there was no evidence that either medication was efficacious in reducing non-fatal events. UGDP investigators also reported they were unable to identify any marker by which tolbutamide-treated subjects likely to die of cardiovascular causes might be identified [28]. On the contrary we have verified that fully 50% of all tolbutamide-treated subjects who died of cardiovascular causes had fasting blood glucose values of 200 mg/dl or greater in the year prior to death. In contrast, the frequency of cardiovascular deaths in tolbutamide-treated subjects with fasting blood glucose values below 200 mg/dl was no greater than that for the other treatment groups. No similar association between fasting blood glucose levels and cardiovascular deaths was evident in the other treatment groups [18].

We would emphasise that this apparent relationship between high blood glucose levels and excess cardiovascular mortality in tolbutamide treated subjects is particularly subject to selection bias (more so than any of the other findings we have reported), since the 200 mg/dl fasting blood glucose level which provided optimal discrimination (between blood glucose levels associated with increased cardiovascular mortality versus those not) was determined by inspection of the data. On the other hand, it would be foolish indeed to ignore these observations on grounds that they were adduced from "a posteriori" criteria. In view of recently proposed mechanisms of cardiovascular toxicity attributed to tolbutamide [18] and the important implications of our findings regarding use of tolbutamide in poorly controlled diabetics, there is clearly a need for further studies of potential interaction between blood glucose levels and the putative cardiovascular toxic effects of tolbutamide.

UGDP investigators reported that lowered fasting blood glucose levels achieved initially with tolbutamide and phenformin were not maintained during follow-up visits. Their analysis, however, did not take into consideration the biological heterogeneity of the population under study or the different mechanisms by which these agents lower blood glucose levels [29]. Sulfonylureas (i. e., tolbutamide) act by facilitating release of endogenous insulin from the pancreas; with chronic use, insulin-mediated extrapancreatic effects may also come into play. Phenformin acts strictly by non-insulin dependent extrapancreatic mechanisms and does not lower blood glucose levels in nondiabetics. We have examined the effects of these agents on fasting blood glucose values, reorienting the analysis in light of these biopharmacological considerations. First, we excluded subjects who dropped out and/or changed medication for a year or longer. Second, in view of the fact that 1) tolbutamide-induced release of endogenous insulin reserves should be inversely related to fasting blood glucose levels and 2) lowering of blood glucose by phenformin should be proportional to fasting blood glucose levels, we subdivided the remaining subjects according to their baseline fasting blood glucose values (i. e., < 110, 110–129, 130–149, 150–199, ≥ 200 mg/dl) [18]. We then determined mean fasting blood glucose values for each subject for the last three quarters of treatment prior to discontinuation of tolbutamide or phenformin. We found that mean fasting blood glucose values for tolbutamide-treated subjects averaged 20–25 mg/dl less ($t = 2.88$, $P < 0.005$) [18] than for placebo-treated subjects, except for those whose baseline fasting blood glucose values exceeded 200 mg/dl.

Phenformin was even more efficacious. Mean fasting blood glucose values of phenformin-treated subjects whose baseline fasting blood glucose values were over 110 mg/dl were 65 mg/dl lower ($t = 5.444$, $P < 0.001$) than those of placebo subjects. In contrast, mean fasting blood glucose values reported by UGDP investigators for these two groups (including subjects who changed medication) never differed by more than 8 mg/dl (Table F-1, [2]). Thus, both oral agents were still achieving substantial reduction of blood glucose levels after several years of use. It should be noted that this data is still undoubtedly biased against demonstrating medication effects, since it includes both primary and secondary treatment failures. Nevertheless, these findings demonstrate that the approach to data analysis is critical in differentiating between apparent lack of medication efficacy due to noncompliance vs true pharmacological inefficacy. In view of these findings, reanalysis of data (not released by the UGDP Coordinating Center) on other non-fatal cardiovascular events might also be informative.

Conclusions

1) The statistical methods utilized by UGDP investigators were inappropriate to the biopharmacological nature of the questions addressed and the conclusions drawn from the study.

2) The anomalous sex ratio of cardiovascular deaths in PLBO subjects constitutes a spurious result which renders suspect all UGDP conclusions based on comparison with that group.

Again, one must ask, "How and why did these problems happen and how can they be avoided in the future?" Our findings indicate the answer to the second part of the question; only UGDP investigators themselves can answer the first.

References

1. University Group Diabetes Program (1970) A study of the effects of hypoglycemic agents on vascular complications in patients with adult-onset diabetes: Sections I and II. Diabetes 19 (Suppl 2):747–830
2. University Group Diabetes Program (1974) A study of the effects of hypoglycemic agents on vascular complications in patients with adult-onset diabetes: Section V. Evaluation of phenformin therapy. Diabetes 24 (Suppl 1):65–184
3. Knatterud G, Klimt CR, Levin ME, Jakobsen ME, Goldner MG (1978) Effects of hypoglycemic agents on vascular complications in patients with adult-onset diabetes: Section VII. Mortality and selected nonfatal events with insulin treatment. JAMA 240:37–42
4. Feinstein AR (1971) Commentary. Clinical biostatics – VIII. An analytic appraisal of the University Group Diabetes Program (UGDP) Study. Clin Pharmacol Ther 12:167–191
5. Schor SS (1971) The University Group Diabetes Program. A statistician looks at the mortality results. JAMA 217:1671–1675
6. Seltzer HS (1972) A summary of criticisms of the findings and conclusions of the University Group Diabetes Program (UGDP). Diabetes 21:976–979
7. O'Sullivan JB (1973) Discussion. In: Camerini-Davalos RA, Cole HS (eds) Vascular and neurological changes in early diabetes. Academic Press, New York, p 509–512
8. Schor SS (1973) Discussion. In: Camerini-Davalos RA, Cole HS (eds) Vascular and neurological changes in early diabetes. Academic Press, New York, p 512–516
9. Committee for the Assessment of Biometric Aspects of Controlled Trials of Hypoglycemic Agents (1975) Report. JAMA 231:583–608
10. Moss JM (1975) Letter to the Editor. JAMA 232:806–808
11. Bradley RF, Dolger H, Forsham PH, Seltzer H (1975) Settling the UGDP controversy? JAMA 232:813–817
12. O'Sullivan JB, D'Agostino RB (1975) Decisive factors in the tolbutamide controversy. JAMA 232:825–829
13. Feinstein AR (1976) Clinical biostatistics – XXXV. The persistent clinical failures and fallacies of the UGDP study. Clin Pharmacol Ther 19:78–93
14. Feinstein AR (1976) Clinical biostatistics – XXXVI. The persistent biometric problems of the UGDP study. Clin Pharmacol Ther 19:472–485
15. Feinstein AR (1979) How good is the statistical evidence against oral hypoglycemic agents? Adv Intern Med 24:71–94
16. Kilo C, Williamson JR, Choi SC, Miller JP (1979) Letter to the Editor: Insulin treatment and diabetic vascular complications. JAMA 241:26–27
17. Kilo C, Williamson JR, Choi SC, Miller JP (in press) Refuting the UGDP conclusion that insulin treatment does not prevent vascular complications in diabetes. In: Proceedings of the IV International Symposium on Early Diabetes
18. Kilo C, Miller JP, Williamson JR (in press) The Achilles heel of the University Group Diabetes Program. JAMA
19. Prout TE, Knatterud GL, Meinert CG (1979) Letter to the Editor. Science 204:362–363
20. Warren S, LeCompte PM, Legg MA (1966) The pathology of diabetes mellitus. Philadelphia, Lea and Febiger, p 188–199
21. Bradley RF (1971) Cardiovascular disease. In: Marble A, White P, Bradley R, Krall L (eds) Joslin's diabetes mellitus. Lea and Febiger, Philadelphia, p 417–477
22. Report of the Workgroup on Mortality (1975) Part 1. Scope and impact of diabetes. In: Report of the National Commission on Diabetes to the Congress of the United States, Vol III. DHEW Publication No (NIH) 76–102. U.S. Government Printing Office, Washington, p 177–247
23. U.S. Department of Health, Education, and Welfare (1977) Diabetes mortality. In: Diabetes data. DHEW Publication No (NIH) 78–1468. U.S. Government Printing Office, Washington, p 59–83
24. Cornfield J (1971) The University Group Diabetes Program. A further statistical analysis of the mortality findings. JAMA 217:1676–1686
25. Peterson CM, Jones RL, Koenig RJ, Melbin ET, Lehrman ML (1977) Reversible hematologic sequelae of diabetes mellitus. Ann Intern Med 86:425–429
26. Peterson CM, Koenig RJ, Jones RL, Sander CD, Cerami A (1977) Correlation of serum triglyceride levels and hemoglobin A_{1c} concentrations in diabetes mellitus. Diabetes 26:507–509
27. Hill AB (1971) In: Principles of medical statistics, Ninth edition. The Lancet Limited, London, p 266
28. Prout TE (1971) Prospective study of the vascular complications in adult-onset diabetes mellitus: The University Group Diabetes Program. Ann Intern Med 75:607–621
29. Shen S-W, Bressler R (1977) Clinical pharmacology of oral antidiabetic agents. Part 1. N Engl J Med 296:493–497

Prof. J. R. Williamson
Department of Pathology
Washington University School of Medicine
660 S. Euclid Avenue
St. Louis, MO 63110
USA