



Human chorionic gonadotropin treatment for obesity: a rebuttal

Dear Sir:

We wish to thank Doctors Hirsch and Van Itallie for their letter (1) in the October 1973 issue of this journal, in which they re-analyzed portions of the data and also attempted to evaluate the experimental procedure and to interpret the results appearing in our original communication (2). It is reassuring that their re-analysis of our double-blind study, in which drug and placebo were assigned to patients on a random basis with injections planned daily 6 times/week for 6 weeks, duplicated the results of one portion of our analysis. Both their and our analyses indicated the mean percent of body weight lost by all patients in the human chorionic gonadotrophin (HCG) group was significantly greater ($P < 0.001$) than the mean percent body weight lost by all patients in the placebo group.

After this initial agreement, they consider several other points in their letter. We wish to comment on these as well as present some of our original data in a new manner. The thrust of their comments will be presented first with our response following.

1) There was a correlation between number of injections and weight lost.

Perhaps we should have pointed out in the original article that the number of injections was determined primarily by the length of time a patient stayed in treatment. We assumed readers would be aware that patients in studies, particularly obese patients, tend to drop out early when results do not come up to their expectations, when the treatment requires effort on their part, or when they believe they are experiencing undue side effects. The reasons for the patients not finishing treatment are: patient 5, moved; patient 6, work conflict; patients 10 and 24, discouraged and asked to quit; patient 20, didn't like diet; patient 21,

went on vacation; patients 32 and 38, reached weight goal; patients 22, 23, and 35, finished series but missed an injection; patients 2, 19, 25, 26, 33, 36, and 37, unknown.

Hirsch and Van Itallie state that: "A t test of the differences in the number of injections received, 33.85 for the experimental group vs. 29.05 for the placebo, shows a barely significant t value of 1.95 at 38 degrees of freedom." They also state: "One must ask why the placebo group received fewer injections on the average than those in the treated group." The patients were randomly assigned so that variability of patient characteristics (including predisposition to stay with dietary programs) was randomly distributed between the groups, which guarantees the validity of tests of significance. Therefore, the logical answer to the question posed would be that the placebo group received fewer injections because they stayed in treatment for a shorter average time, as they were less successful in losing weight. Considering the randomization and the study design, the variable most apt to be causative in the increased weight loss appears to be the HCG.

Hirsch and Van Itallie further state that "A correlation coefficient relating number of injections received to percent weight loss shows a correlation of 0.683. Such a correlation, occurring with a t value of 3.97 and 18 degrees of freedom, is a highly significant observation. In fact, one can say that within the placebo group, nearly one-half of the variance observed is the result of the number of injections."

The key difference between our points of view and analyses seems to be in regard to random assignment of treatment to patients. Differences between treatment groups on post hoc measures of performance, correlates with the response variable (e.g., number of injections with weight loss), or other behavior of the patient are confounded with treatment effects. With the random assignment of treatments to patients, the only valid conclusion to be reached, and the usual justification for running a controlled experiment, is that it is the treat-

ment that was or was not important. Usually a researcher does not say that post hoc correlates with the response variable (weight loss) may cause differences in the response variables (weight loss) among the treatment groups. Correlations usually have been used to indicate when two variables "co-vary," not as measures of cause and effect. Certainly a highly significant correlation is not a clear measure of strength of that correlation, as significance is so heavily dependent on sample size. Post hoc analyses, post hoc correlational statistics, systematic deletion of data, after the fact hypotheses, and preconceived notions just do not carry the same weight as a controlled randomized study for determining the existence of treatment effects.

2) When the patients receiving less than 36 injections are excluded from analysis, the significance of the increased mean percent weight loss of the HCG group compared with that of the placebo group "declined sharply."

They point out that their analysis indicates the significance level drops from less than 0.001 to a value close to 0.02. We agree.

In doing their re-analysis, 8 of 20 patients in the HCG group and 10 of 20 in the placebo group were deleted from consideration, leaving a total study sample of 22 rather than 40. Deleting observations can reduce the probability of significance irrespective of the relation between injections and treatment effects. The size of the standard error of the difference in means is basically a function of

$$1/\sqrt{\text{number of observations.}}$$

Decreasing the number of observations increases the standard error of the differences in means which reduces the *t* value. The number of degrees of freedom are reduced as well. Together these two can account quite easily for a drop in significance levels. Actually the 0.02 level found by Hirsch and Van Itallie is usually considered medically significant. Both Grinker, Hirsch and Smith (3) and Campbell, Hashim and Van Itallie (4) have considered $P < 0.05$ significant previously and in this letter, Hirsch and Van Itallie state, "shows a barely significant *t* value of 1.95 at 38 degrees of freedom." This yields a *P* value between 0.05 and 0.1. So by their own criteria, 0.02 is clearly significant.

When only the patients receiving 36 injections are considered, the mean loss in the HCG group is 21.75 lb versus 14.62 in the placebo group, a difference of 7.13 lb that is not too different from the 8.91 lb difference found when the weight loss of all starting patients is considered.

Too often, in studies of the treatment of obesity, dropouts are excluded from final analysis with the investigators frequently justifying this action by stating they are "uncooperative." Stunkard and McLaren-Hume (5) and Albrink (6) made appeals for an analysis of all starting patients. Of the most importance is the impact of a treatment modality on all starting patients rather than those finishing treatment. If a patient drops out because there are side effects from an active drug or lack of effect from a placebo, this is important. This was our reason for considering all starting patients in our analysis.

We have re-analyzed the total mean percent of body weight lost by each group on a weekly basis for those remaining in treatment and returning on a weekly basis, and we find the loss in the HCG group is greater than in the placebo group at each weekly interval with significances ranging from $P < 0.02$ to < 0.001 .

For interest, the mean weight loss per planned injection for both the HCG and the placebo group is shown in Fig. 1.

3) They question why the placebo group received fewer injections on the average than the treated group, whether this was a random occurrence and "Did the patients or someone know who was who?"

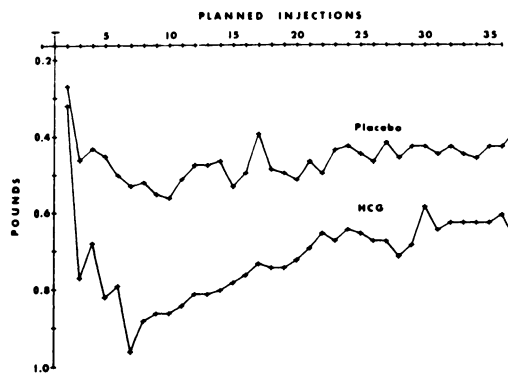


Fig. 1. Mean weight loss per planned injection.



In response, both the HCG and the placebo were prepared in identical appearing vials, randomized, labeled, and shipped from Colorado by Asher with instructions to Harper in California where the study was done. The code was not revealed to Harper until the study was over and the data had been received in Colorado. In addition, Harper saw the patients only at the initial and final visits. They were seen along with regular injectable patients in an office where the flow rate of patients was often 30/hr or more. Injections were given to all patients by a single girl. They were each seen daily for counseling, questioning, et cetera, by one of a number of other paramedical assistants. Few patients were seen continuously by one paramedical assistant for more than a few days at a time. Such a system would make biasing a study difficult, even if there were some way of knowing what was in each vial.

Another finding militating strongly against "cheating" was that HCG patients lost no more in the hip and thigh measurements per pound of weight lost than the placebo patients. There were some technical difficulties in making certain that measurements were done at the same levels before and after treatment; at the editor's suggestion, this material was deleted before the article was published.

Because the late Simeons and his followers have been about as vocal about selective loss of fat in the hips and thighs with HCG as they have about enhanced ability to follow a 500-kcal diet, if Harper or his assistants were aware of the coding and somehow cheated on reporting weight loss, hunger, and feeling of well-being, wouldn't it be likely they would have cheated on measurements also?

4) Because patients were told "the slightest deviation from any of the details will result in utter disaster," the placebo patients receiving fewer injections (and thus appearing to have more missed injections) might not unexpectedly show some difference from the treated group in their feeling and hunger response.

To check on this speculation, we calculated the percentage of missed visits in each group, and although the mean placebo treatment period is a little shorter than the HCG treatment period, there is no appreciable difference in the mean percentage of missed injections in either group while they remained in

treatment. When the difference was tested for significance, a *t* value of 0.385 was found, which at 38 degrees of freedom does not approach significance. The mean percent of injections missed while in treatment for each of the patients in the HCG group was 7.53 and in the placebo group, 10.75.

The majority of these missed injections were for legal misses, i.e., holidays (the study was done between August and February and holiday injections were not required). In addition, as mentioned in the previous article, some patients did not receive shots during the time of heavy menstrual flow and a few of the misses occurred because of trips and for other "excused" reasons. So, it is highly unlikely this factor affects either group appreciably and certainly not one group significantly more than the other. Unfortunately, our original data were presented so that it may have appeared that a patient receiving only a few injections had many skips rather than having discontinued treatment early.

5) Analysis of hunger rated as "none," "little," "some," "much," and general feeling as "poor," "fair," "good," or "excellent" requires a special type of statistical analysis and that lumping two categories together was an "arbitrary and totally unwarranted procedure from a statistical viewpoint."

Although statisticians may differ in their personal preference for handling categorical data, all to whom we have talked agree there is no reason two categories cannot be lumped together.

The hunger and feelings questions that were used to classify the patients into four categories each do not define continuous variables, but categorical, having the scaling properties of order (perhaps Hirsch and Van Itallie might better have characterized these variables as "nonmetric" rather than "nonparametric"). The field of statistics provides several alternatives for the analysis of data. Some are based on parametric and others on nonparametric statistical procedures. We chose a parametric analysis which required the definition of a single random variable for each patient called "percent of none or little hunger responses." The none or little responses were totaled for each subject and this percentage was based on the total injections received. This variable can be



considered a continuous variable measured on two independent samples with 20 independent observations each. Thus, as with the weight loss data analysis, we followed the usual normality assumptions and performed a *t* test for significant differences which yielded the results originally reported. A comparable procedure was used to analyze the feeling of well-being data.

Perhaps a real criticism of these data presented in Table 3 of the original article should be leveled at similar tests performed on the complements "some" and "much" hunger and feeling "poor" and "fair." As these responses are dependent on the "none" and "little" hunger and "good" and "excellent" feeling responses, it is not surprising that the HCG group means were significantly different from the placebo group means.

In regard to degree of hunger, the categories of little or none could have been kept separate at the time of each patient visit as we did, or they could have been lumped together. The

same is true of degrees of feeling. We have found it advantageous in gathering data to be as definitive as possible, because data can be "lumped" at the time of analysis if desired, but there is no way of subdividing these larger groups later. Generally, we attempted to assess whether or not the patients were hungry and whether or not they felt good.

6) Insufficient data are given in the article to permit re-analysis of hunger and feeling of well-being data as a function of receiving this mixture of hormones or the placebo.

We agree and have prepared a table showing for each patient the percentage of responses indicating little or no hunger or feeling good to excellent. The percentage of each patient's response indicating some to much hunger and feeling fair to poor can be obtained by subtracting these percentages from 100. To save space, we have not subdivided the little or none hunger categories or the feeling good to excellent categories. The means of the individual responses in each category vary slightly

TABLE 1
Percentage of patient responses (visit 2 to 37)

HCG			Placebo		
Patient No.	Little or no hunger	Feeling good to excellent	Patient No.	Little or no hunger	Feeling good to excellent
1	81.81	100.00	3	3.23	29.03
2	75.86	82.75	4	9.37	25.00
5	36.00	75.00	6	4.17	16.67
7	94.11	97.05	10	75.76	78.79
8	69.69	87.87	11	15.63	70.97
9	71.87	68.75	12	82.76	86.21
14	93.75	100.00	13		91.67
15	85.29	90.00	16		100.00
18	82.35	82.35	17	80.56	91.67
21	75.00	96.87	19	73.68	78.95
22	82.75	93.10	20	73.68	47.37
23	79.41	91.17	24	36.67	76.67
28	75.00	83.33	25	83.33	75.00
29	79.41	82.35	26	66.67	66.67
30	88.46	80.76	27	73.33	80.00
32	58.06	70.96	31	32.35	73.53
34	66.66	84.84	33	00.00	100.00
36	80.76	80.76	35	64.52	74.19
38	85.71	76.19	37	76.92	88.46
40	79.41	91.17	39	62.96	70.37
Mean percent of response	77.08* ± 2.90 ^a	85.81** ± 2.05		50.87 ± 8.81	71.06 ± 5.32

Difference between HCG and placebo significant at: * $P < 0.002$; ** $P < 0.02$.

^a SEM.

from the ones given in the original article, because in the re-analysis, the responses at the time of the initial visit (injection 1) were excluded and those on the day following injection 36 included (the patient thus reported on his response since the previous visit) (Table 1).

Our analysis of the difference between the mean percentage of responses of having little or no hunger for each patient in the HCG and the placebo group indicated a significantly greater value in the HCG category. The t value was 3.41, which with 36 degrees of freedom, gives a $P < 0.002$. The mean percentage of responses of feeling good to excellent for each patient was also significantly greater in the HCG group than in the placebo group. Here the t value was 2.59 which with 38 degrees of freedom yields a $P < 0.02$. It should be noted that the P value concerning the hunger rating given in the original article was less than 0.001 as compared with 0.002 in this re-analysis and that the P value of the feeling good to excellent rating changed from an original $P < 0.001$ to $P < 0.02$.

We erred in our original analysis in calculating degrees of freedom and are happy it is possible to correct this error. It should be noted, however, that the level of significance of the difference in means of the hunger ratings changed little, and although it changed more for the difference in the means of the feeling ratings, this difference still remains significant.

It is unfortunate that data on the hunger feelings of patients 13 and 16 are missing. If these patients had had high "little" or "none" hunger ratings, it might conceivably have changed the level of significance of the difference between the means found in the HCG and placebo groups. To test this, we took the highest percent of little or no hunger responses achieved by any patient in either group (94.11% by patient 7 in the HCG group) and arbitrarily assigned these values to patients 13 and 16. When this was done, a t value of 2.78 was found, which with 38 degrees of freedom, still gives a significant P value of less than 0.01.

We are happy to have this chance to respond to the letter by Hirsch and Van Itallie. It has allowed us to present further insights concerning the effectiveness of HCG and to correct two P values that were in error. Through the

assistance of Hirsch and Van Itallie's re-analysis, we think it is fair to conclude as we did in the original article that there were several significant differences between the HCG and placebo groups.

First, the HCG group lost significantly more mean weight when all patients are considered ($P < 0.001$, our initial analysis), lost significantly more weight per injection when all patients are considered ($P < 0.025$, our initial analysis), lost a significantly greater mean percentage of their starting weight when all patients are considered ($P < 0.001$, Hirsch and Van Itallie's and our initial analysis), and when those receiving 36 injections are considered (P close to 2 in 100, Hirsch and Van Itallie), and lost a significantly greater mean percentage of their starting weight when those returning at the end of each of the 6 weeks are considered ($P < 0.02$ to $P < 0.001$, our re-analysis). The figure shows graphically that the difference in weight loss started early and that a difference of approximately 0.2 lb per injection persisted throughout the study.

Second, the mean of the percentage of daily patient responses indicating little or no hunger and feeling good to excellent was significantly greater in the HCG group than in the placebo group ($P < 0.002$ and $P < 0.02$, respectively, our re-analysis).

Third, as the number of injections is primarily a function of the time the patients remained in the study, Hirsch and Van Itallie's analysis suggesting that HCG patients received more injections than placebo patients ("barely significant t value of 1.95 at 38 degrees of freedom") is consistent with our contention that patients find the diet more bearable when they receive HCG. We thus conclude, as we did previously, that HCG did have a significant effect on the parameters studied.

W. L. Asher, M.D.

Executive Director
American Society of
Bariatric Physicians
333 W. Hampden Ave.
Englewood, Colorado 80110

Harold W. Harper, M.D.

9201 Sunset, Suite 912
Los Angeles, California 90069

References

1. HIRSCH, J., AND T. B. VAN ITALLIE. The treatment of obesity. *Am. J. Clin. Nutr.* 26: 1039, 1973.
2. ASHER, W. L., AND H. W. HARPER. Effect of human chorionic gonadotrophin on weight loss, hunger, and feeling of well-being. *Am. J. Clin. Nutr.* 26: 211, 1973.
3. GRINKER, J., J. HIRSCH AND D. SMITH. Taste sensitivity and susceptibility to external influence in obese and normal weight subjects. *J. Personal Soc. Psychol.* 22: 320, 1972.
4. CAMPBELL, R., S. HASHIM AND T. VAN ITALLIE. Studies of food-intake regulation in man. Responses to variations in nutritive density in lean and obese subjects. *New Engl. J. Med.* 285: 1402, 1971.
5. STUNKARD, A., AND M. McLAREN-HUME. The results of treatment of obesity. *Arch. Internal Med.* 103: 79, 1959.
6. ALBRINK, M. Chorionic gonadotropin and obesity? *Am. J. Clin. Nutr.* 22: 681, 1969.

Total and ionizable iron in common Indian cooked foods

Dear Sir:

A vast amount of information has been accumulated on the total iron content of various foods (1-3). As all of the total iron is not available for absorption, a knowledge of "available" or "ionizable" iron is of much greater value (4-7). These two measurements, however, do not allow for the possible effects of cooking in the release of iron from food conjugates and utensils. Therefore, the calculation of dietary iron content and intake from standard tables has consistently given variable results in comparison to actual analytical values (8-10). In view of this, it has been considered pertinent to determine the amount of total and ionizable iron in some of the common Indian foods, and to evaluate iron intake by actual chemical analysis.

The seventeen dishes selected were those frequently prepared in Central and North-Western India. These were prepared (11-13) with tap water, which was found to contain 0.4 mg iron per liter. The freshly cooked food was immediately analyzed for iron. All samples were thoroughly mixed; additional double-distilled water was added to the more viscous

samples to ensure homogeneity. A representative sample was taken for the estimation of total (14) and ionizable (4) iron. Aykroyd's (3) food composition table was used for calculating the iron content of the prepared food.

The results in Table 1 show that the analytical and calculated values of total iron were significantly different ($P < 0.05$). The difference between these values in respect to ionizable iron was even more significant ($P < 0.001$). In general, the values of ionizable and total iron as obtained by chemical analysis were found to be higher than those calculated from Aykroyd's tables, presumably because of the effects of cooking in the release of iron from food (15, 16) and iron utensils (17, 18). Also, iron contributed by tap water in some of the diets may be significant.

Furthermore, on examination of Table 1, it becomes apparent that total iron and ionizable iron contents are not parallel. Wheat roti seems to supply greater ionizable iron than maize roti, whereas bajra (millet) roti contains the highest amounts of total as well as ionizable iron among the cereal preparations. However, the three types of dals have total and ionizable iron in the same range.

The present data were used to calculate the iron intake of people belonging to lower and higher income groups consuming maize and wheat as the staple items in their diets, respectively. On this basis, the intake of total and ionizable iron in the higher income group was found to be 22.0 to 36.9 mg and 14.2 to 24.2 mg, respectively; slightly higher than that in lower income group, i.e., 21.6 to 34.8 mg and 14.1 to 22.5 mg, respectively. Although total iron intake in the two groups is apparently at par with the recently recommended dietary iron allowance for adult Indians (18, 19), nevertheless, the incidence of iron deficiency and, consequently anemia, is quite high in this country (20, 21). One possible explanation of this paradoxical observation may be that iron absorption depends not only on the amount of iron ingested but also on its form or chemical nature (22, 23). Thus, if only the ionizable fraction is considered to be available for absorption, the intake of ionizable iron falls short of the recommended allowance. This may partly explain the above paradox. Alternatively, it is difficult to calculate precisely the amount

