STATISTICS IN MEDICINE, VOL. 2, 417–427 (1983)

HOW MUCH OF THE PLACEBO 'EFFECT' IS REALLY STATISTICAL REGRESSION?

CLEMENT J. McDONALD* AND STEVEN A. MAZZUCA

Regenstrief Institute for Health Care and Department of Medicine, Indiana University School of Medicine, 1001 W. Tenth St., Indianapolis, Indiana 46202, U.S.A.

AND

GEORGE P. McCABE, JR.

Department of Statistics, Purdue University, West Lafayette, Indiana, U.S.A.

SUMMARY

Statistical regression to the mean predicts that patients selected for abnormalcy will, on the average, tend to improve. We argue that most improvements attributed to the placebo effect are actually instances of statistical regression. First, whereas older clinical trials susceptible to regression resulted in a marked improvement in placebo-treated patients, in a modern series of clinical trials whose design tended to protect against regression, we found no significant improvement (median change 0·3 per cent, $p > 0{\cdot}05$) in placebo-treated patients. Secondly, regression can yield sizeable improvements, even among biochemical tests. Among a series of 15 biochemical tests, theoretical estimates of the improvement due to regression by selection of patients as high abnormals (i.e. 3 standard deviations above the mean) ranged from 2·5 per cent for serum sodium to 26 per cent for serum lactate dehydrogenase (median 10 per cent); empirical estimates ranged from 3·8 per cent for serum chloride to 37·3 per cent for serum phosphorus (median 9·5 per cent). Thus, we urge caution in interpreting patient improvements as causal effects of our actions and should avoid the conceit of assuming that our personal presence has strong healing powers.

KEY WORDS Drug treatment Placebo Statistical regression Computerized medical record Clinical trial

Investigators have variously claimed that the placebo is powerful,[1] that it meters its curative effects in proportion to the severity of the illness[1] and that it influences both objective and subjective outcomes.[2–4] Some authors have advocated a legitimate place for placebo therapy in patient care.[2] Patients do tend to improve in association with placebo treatment. This association, however, does not by itself prove that the placebo treatment causes the improvement. This paper considers the degree to which statistical regression toward the mean could account for the improvements associated with placebo therapy.

We exclude from the scope of our discussion placebo therapy associated with intense conditioning[5,6] or body invasion, i.e. needle sticks or surgical incisions. In the first case, the improvements can be attributed to Pavlovian mechanisms and in the second, to neuroendocrine mechanisms.[7] Most medical prescribing is not associated with either of these circumstances.

At the outset, we emphasize that our question regarding the strength of the placebo effect does

* To whom reprint requests should be addressed.

0277–6715/83/040417–11$01.10

Received 27 September 1982
Revised 28 February 1983

not diminish the requirements for placebo-treated controls in clinical trials. Such controls are imperative to prevent bias and ensure proper assignment of cause.

STATISTICAL REGRESSION

Statistical regression describes a tendency of extreme measures to move closer to the mean when they are repeated. It is a well known phenomenon first described by Galton in 1885[8] and since reviewed and described in medical settings by many authors.[9-12] It explains such diverse phenomena as the observations that individuals who score high in tests tend to do less well in repeat tests and the observation that sons of tall fathers tend, on the average, to be shorter than their progenitors.[8] The amount of improvement due to regression can be large and statistically significant.[13] Given observations about paired random variables, X_1 and X_2, regression toward the mean occurs whenever the two variables are positively correlated and have identical distributions.[14]

For bivariate normal random variables with common mean μ, common variance, and correlation ρ (often referred to as the test/retest or reliability coefficient), regression toward the mean is usually expressed as follows:[15]

$$E(X_2 | X_1 = x_1) = \mu + \rho(x_1 - \mu) \tag{1}$$

The expected change between base line and repeat observations is thus,

$$\Delta = E(X_2 - X_1 | X_1 \geq x_1) = (\mu - x_1)(1 - \rho) \tag{2}$$

To convince oneself that Δ always represents a change toward the mean, consider the following. Under the assumption that X_1 and X_2 are positively correlated, the sign of Δ is determined by $(\mu - x_1)$. If the initial measure, x_1, is below the mean, $(\mu - x_1)$ is positive. Therefore, Δ represents an increase toward the mean. If the initial measure is above the mean, $\mu - x_1$ is negative. Consequently, Δ represents a decrease toward the mean. Only when $\rho = 1$, implying perfect reliability of the measure, does no change occur.

Two other aspects of equation (2) deserve attention. First, the amount of 'improvement' is proportional to $(1 - \rho)$. Thus the less reliable the measure, the greater the expected improvement. Secondly, statistical regression is proportional to $(\mu - x_1)$, the distance between the mean and the baseline measure, i.e. the more abnormal the initial measure, the larger is the expected improvement.

What we have described for a single variable has multivariate extensions. Here, multiple observations about a patient can be described as a point in multidimensional space. For a population of patients there is a mean point. The distance between any point and the mean point can be measured in terms of a distance function such as the Euclidean norm. In many cases, individuals whose distance from the mean is extreme on first observation will move closer to the mean upon repeat observation. For example, this will occur when each of many measurements are mutually independent and have the properties required for univariate regression. In this case equation (2) will apply to each measure individually and therefore the repeat measure will be closer to the mean point than the initial measure. Discussion of the general multivariate case is complex and beyond the scope of this paper.

The variables used to judge the success of therapy are especially susceptible to statistical regression. By definition, the pretreatment values of these measures will be extreme (abnormal). Because the pre- and post-measures are taken in the same individual they will tend to be positively, but imperfectly, correlated. Finally, under many circumstances (for example, when the treatment

has no effect and the disease process is not rapidly progressive), the statistical distribution of the before and after observations will be the same. (Note that the requirements for regression refer to the distribution of the larger population from which we draw patients for treatment, not the smaller population we treat.) Thus, the conditions for statistical regression to occur are often satisfied. We emphasize that this does not mean that all treated populations improve, nor that all improvement is due to regression. What it means is that regression may provide the illusion of efficacy when a drug has no effect.

The question posed in the introduction now becomes a question of size. Is the regression 'effect' large enough to account for the improvements observed in placebo treated patients? To answer this question, we obtain respective estimates of the size of the improvement expected from regression and that observed with placebo treatment.

THE ESTIMATED AMOUNT OF IMPROVEMENT OBSERVED WITH PLACEBO TREATMENT

Most information about the size of the placebo 'effect' comes from observations with placebo treated patients in clinical trials. In a widely quoted article, Beecher[1] reviewed the effect of placebo therapy in 15 different studies and reported that 35·2 per cent of the pooled population of patients from all these studies improved after placebo therapy. This is neither the average improvement per patient nor the average direction of change. The percentage reported is the number of patients who improved divided by the number treated, and contains no information about patients who worsened under placebo treatment. (Using the same technique, a clinic weighing patients with a scale accurate to the gram could demonstrate a 50 per cent weight loss each visit when there was no change in the average weight.) Thus, this measure is not a useful one for our purpose. The magnitude of improvement and the number of patients who worsened under placebo therapy were not reported in many of the papers of that era.

Since we were unable to find published estimates of the *average* amount of improvement in placebo treated patients, we obtained our own estimate from a random sample of 30 placebo controlled clinical trials reported in the 1979 *Abridged Index Medicus*. We obtained the mean percentage change in placebo treated patients by comparing the last reported baseline measure with the last measure obtained during treatment. The signs of these changes were adjusted to ensure that improvements were always represented by a positive, and deteriorations by a negative, value. When more than one variable was reported in the study, we used the variable with the median percentage change as our index variable. We included papers reporting both subjective and objective variables because the literature about placebos emphasizes the positive effect of placebo treatment in both classes of observations.

The index variable in 17 of our articles was a biological, physiological or anatomical measurement, i.e. an objective measure. In the remaining 13 articles it was a measure of behaviour, perception or pain, i.e. a subjective measurement. Many of these studies employed techniques that would tend to reduce the effect of statistical regression on the placebo treated patients. Half of the studies took two or more pretreatment measures and used either the average or the last of these measures as their baseline for measuring improvement. Most of the subjective measurements were averages of multipoint measurements (e.g. psychological scales). Such averages are more reliable than comparable single point measures.

Over the period of placebo treatment, the index variable improved in 16, remained the same in 1 and worsened in 13 of the selected reports (see Table I). The mean improvement was 9·9 per cent and the median—a more reliable measure of central tendency in skewed samples such as this—was

Table I. Percentage change observed during placebo treatment in 30 randomly selected placebo-controlled drug trials

Median variable	Source*	Length of study	Number of subjects	Baseline average†	Adjusted percentage change
Menstrual pain rating[16]	F2	6 mo	5	2·7	−11·1
Serum alkaline phosphatase[17]	F2	48 mo	10	4·4	−6·8‡
Work done (treadmill)[18]	T2	24 hr	11	10·2	−3·9‡
Plasma tocopherol[19]	TX, F2	7 day	14	0·250	216·0‡
Schizophrenia severity rating[20]	F3	3–7 day	6	2·3	−17·4
Papule grading[21]	F1	11 wk	20	3·1	35·5‡
Plasma growth hormone[22]	T1	5 hr	12	28·3	−3·2‡
Arterial pH[23]	T5	5 hr	7	7·35	0·3‡
Raskin depression scale[24]	T5	6 mo	17	8·8	25·0
Hamilton depression scale[25]	T2	3 mo	17	30·6	0·3
Drug craving scale[26]	T4	8 wk	51	−18·4	−32·6
Pain rating[27]	T3	4 mo	16	4·0	5·0
Systolic blood pressure (standing)[28]	F2	3 wk	22	150·0	−2·7‡
Diastolic blood pressure (standing)[29]	F4	10 day	27	120·0	−1.3‡
Bunney–Hamburg psychosis scale[30]	F2, F3	5 wk	13	7·5	10·7
Agitation rating[31]	F2	2 day	12	43·0	11·2
Acne cyst grading[32]	T2	12 wk	24	0·3	33.3‡
Systemic vascular resistance[33]	T1	1 hr	10	2·33	1·2‡
Hyperactivity rating by mother[34]	F1	6 wk	5	15·3	12·4
Number of spells of nocturnal enuresis/2wks[35]	T2	2 wk	22	10·6	−3·8
Healing peptic ulcer rating (endoscopy)[36]	T2	4 wk	24	4·0	24·0‡
Cardiothoracic ratio[37]	T1	6 wk	12	0·58	0·0‡
LDL cholesterol[38]	T4	2 wk	15	325·5	1·2‡
Arterial PO_2[39]	F1	6 hr	5	58·2	5·5‡
Pulmonary wedge pressure[40]	T1	6 hr	8	29·1	−9·6‡
Maximum expiratory flow[41]	T2	7 wk	21	0·43	37·2‡
Pain rating (post tooth extraction)[42]	F1	2 hr	17	3·4	−58·8
Opiate withdrawal rating[43]	F1	2 hr	5	12·6	−4·8
Foot infections[44]	Tx	6 wk	20	1·0	50·0‡
Rhinitis severity rating[45]	F2	8 wk	31	2·0	−15·0

* F indicates figure in article; T = table; Tx = taken from text

† Percentage change from placebo group baseline mean; positive values indicate change toward normal range; negative, away from normal range.

‡ Indicates objective measurement.

0·3 per cent. A number of factors could account for the difference between the size of the improvement in our, and in Beecher's sample of papers. First, our estimate took into account patients who worsened as well as those who improved, providing a valid estimate of the magnitude of the improvement.

Secondly, our sample of papers included only three reports about pain. Beecher's included nine. Placebos may have a greater effect on pain than on other conditions. Finally, the differences in the study design of the modern papers compared to the older papers would tend to mute the effect of statistical regression in the newer papers and thus reduce the total improvement observed compared with the older papers.

The inclusion of objective measurements in our report did not account for the lesser overall effect, since the average improvement observed in subjective parameters was actually less than that in the objective parameters.

THE SIZE OF THE REGRESSION EFFECT

The ideal way to judge the relative contribution of statistical regression to the improvement observed in placebo treatment would be to compute the expected amount of regression (using equation (2)) for the patient data reported in the cited studies and compare it with the observed improvement. Such an approach, however, would require information about the mean, μ, for the population from which the treated patients were selected and the test/retest reliability coefficient, ρ, of the measures reported. This information was not available in any of the cited papers. In fact, except for biochemical tests and blood pressure measures, we could find little information about the test/retest reliability of clinical measures in general.

The alternative was to obtain an order of magnitude estimate of the size of the regression effect by using data available in the medical literature. Harris[46] and Cotlove[47] reported detailed information about the mean (μ) S.D. (σ), and test/retest reliability (ρ) of 15 biochemical measures. Specimens for their study were obtained in a standardized fashion from normal volunteers on a weekly basis for 10 weeks. Because these were biochemical measurements obtained in a highly standardized fashion in a reference laboratory, these data will yield conservative estimates of the size of the regression 'effect' expected in ordinary practice.

Using Harris and Cotlove's data for each of the 15 variables and equation (3) [derived from equation (2) by dividing by X_1, multiplying by 100 per cent and substituting $(\mu + 3\sigma)$ for X_1], we determined the percentage change expected when we selected for repeat observation results that exceeded the usual upper normal limits (i.e. greater than $\mu + 3\sigma$).

$$\text{percentage improvement} = \frac{3\sigma(1-\sigma) \times 100}{\mu + 3\sigma} \tag{3}$$

We looked at the high rather than the low end of the normal range to ensure a conservative percentage estimate. We set our threshold at one standard deviation above the usual upper normal limit of 2σ to reflect usual treatment practices: we usually do not treat patients until they are well into the abnormal range. The first column of Table II shows the results of these computations. The sizes of the improvements range from 2·5 per cent for serum sodium to 26 per cent for serum lactate hydrogenase. The median of the 15 tests improved by 10 per cent.

Do changes of this size really occur in practice? To answer this question, we reviewed the computer-stored records for 12,000 patients who visited the Wishard Memorial Hospital General Medicine Clinic between August 1978 and August 1980. For each biochemical variable we selected patients who had two or more observations and examined only the first (T_1) and the last (T_2) measurements. For each biochemical variable, the T_1 and T_2 measures were positively correlated and their distributions were approximately the same. Thus, the conditions for occurrence of statistical regression are satisfied. Statistical regression is independent of time's arrow and can be observed whether we look forward or backward in time. Because we were concerned that changes from T_1 to T_2 might be too readily attributed to physician's interventions, we selected patients whose values were three standard deviations above the mean at T_1 and computed the percentage change backwards in time to T_2. The results are listed in Table III. Every one of the 15 tests improved (became less abnormal from T_2 to T_1). The median percentage change of the 15 measurements was 9·5, a rate comparable to our theoretic estimates.

For completeness, we performed the same analysis selecting patients who were abnormal at T_1 and computed the amount of change forward to T_2. Again, we saw improvement in all 15 parameters. In the case of each of the biochemical measurements, the change obtained in the

Table II. Percentage improvement expected from regression for 15 biochemical variables with the given mean (μ) standard deviation (σ) and test/retest correlation (ρ) when initial measures are 3 standard deviations above the mean (see text)

Test	Percentage improvement	μ*	σ*	ρ†
1. Sodium	2·5	139·4	1·9	0·34
2. Potassium	10·6	4·1	0·29	0·38
3. Chloride	4·2	104·6	2·2	0·29
4. Carbon Dioxide	10·4	27·3	1·9	0·40
5. Calcium	7·8	2·55	0·12	0·37
6. Magnesium	8·4	0·81	0·07	0·59
7. Inorganic Phosphorus	13·0	3·49	0·51	0·56
8. Total Protein	6·7	6·97	0·50	0·62
9. Albumin	8·8	4·2	0·35	0·56
10. Uric Acid	15·0	4·62	1·16	0·63
11. Urea Nitrogen	16·6	13·5	3·1	0·60
12. Glucose	10·0	94·5	9·7	0·57
13. Cholesterol	6·9	205	36	0·8
14. SGOT	19·0	14·5	4·7	0·47
15. LDH	26·0	328	72·0	0·51

* From Table 1 of Reference 47

† Computed from the equation, $\rho = \dfrac{1}{1 + (S_p/S_g)}$ where S_p/S_g is obtained from Table 5 of Reference 46.

Table III. Observed change in patients selected for 3-standard deviation abnormalcy

Test name	Number in total population	Average of total population at T_1	Average of total population at T_2	Average in selected population at T_2*	Percentage change from T_1 to T_2
Sodium	8373	139·3	139·0	153·7	6·2
Potassium	8946	4·13	4·05	6·6	27·1
Chloride	8365	102·3	101·9	110·0	3·8
Bicarbonate	8352	25·7	26·1	30·7	9·5
Calcium	7197	9·83	9·75	11·6	7·0
Magnesium	905	2·07	2·0	3·0	20·8
Phosphorus	1017	3·57	3·72	7·0	37·3
Total protein	7128	7·81	7·80	9·0	6·9
Albumin	7096	4·29	4·27	5·2	11·3
Uric acid	7417	6·21	6·23	9·7	13·7
Blood–urea–nitrogen	9340	15·6	15·8	36·0	21·2
Glucose	5968	154·7	155·4	224·5	8·7
Cholesterol	7243	223·0	219·6	367·3	13·0
SGOT	7267	43·5	38·5	117·5	25·6
Lactate dehydrogenase	6097	222·7	220·8	279·8	8·9

* Selected population consists of patients selected for values > 3 standard deviations above mean at T_2.

forward direction was comparable to that in the backward direction. Although it is possible to conceive of a causal mechanism that could explain the observed changes in some of these variables, the most reasonable explanation for the improvement seen in all 15 variables, when examined both forward and backward in time, is statistical regression.

DISCUSSION

The numerical size of the improvement we observed in placebo-treated patients and that of the estimate that would occur in biochemical variables due to statistical regression were remarkably similar. Because these estimates are based upon different clinical variables and different patient populations, we cannot determine the true proportion of the improvement due to statistical regression. Observations used to judge the success of therapy, however, are ripe for statistical regression toward the mean. Moreover, improvements of as much as 26 per cent could be expected, even from biochemical variables measured under highly standardized conditions. Together, these two facts suggest that regression accounts for an important share of the improvement observed with placebo treatment.

Many of the unusual characteristics of the placebo 'effect' could be explained by assuming that statistical regression is responsible for the observed improvements. By definition, statistical regression is a random phenomenon and therefore could explain the observation that the placebo effect is 'not uniform, constant or predictable' in individual patients.[48] Regression is proportional to the degree of the baseline abnormalcy and therefore could explain the observation that the placebo effect is 'most effective when stress (anxiety or pain, for example) is greatest.'[1] Statistical regression is proportional to measurement unreliability [$(1-\rho)$ of equation (2)] and single point human observations tend to be less reliable than comparable objective measurements. This could explain the traditional wisdom that subjective measurements are more susceptible to placebo influence than objective measurements.

Finally, conclusive proof of a causal role of placebo treatment requires a controlled trial comparing placebo treated with non-treated patients. One of the early proponents of the importance of the placebo effect noted the lack of studies making direct comparisons between these two groups.[49] We found one modern trial that did compare placebo treated and untreated control groups.[50] In this study of blood pressure control, both control groups improved by the same amount. One could argue that the improvements were the same because the investigator contact required in obtaining follow-up had a placebo effect equal to that of the placebo treatment itself. However, the results are more easily explained by statistical regression operating equally in the two control groups.

Some have argued that placebo treatment causes negative as well as positive effects. Clearly, side effects such as nausea, vomiting and drowsiness occur in association with placebo treatment. A few studies have reported that placebo treated patients had unusual rates of adverse effects, of a kind unique to the experimental treatment. These observations may be examples of the causal effects of placebo therapy but they can also be explained by ascertainment bias in such studies.

Regardless of the relative importance of statistical regression to the placebo effect, regression is important in its own right. It can produce improvements in biochemical variables that are large enough to be important. Similar or larger regression induced improvements are likely to occur in clinical observations that tend to be less reliable than analytical biochemical measurements. In fact, we can even expect regression to yield improvement in imagining findings, since such findings also vary spontaneously. For example, in a study of serial barium swallows, oesophageal varices disappeared on one or more occasions in 25 per cent of patients who had biopsy proven cirrhosis and manometry proven hypertension.[51]

It is also important because it extracts a price of increased sample size requirements from the clinical researcher and an increased risk of judging a clinical therapy effective when it is not from the clinician. This price can be minimized by taking steps to reduce intrapatient measurement variability. There are three approaches. The first, and most obvious, is to select the most reliable measures from among those that are practical. The second is to use the average of a number of

different measures of similar reliability rather than a single measure, since such an average is more reliable than any of its components. The last is to observe the chosen measures more than once, preferably at different points in time before beginning therapy. By using the average of enough pretreatment measures, one can reduce the size of regression to any predetermined level.[9,10] When the baseline variation in the observation is due only to random noise, and not to drifts in the baseline, circadian rhythms or other cycles, the regression effect can be eliminated entirely by two pretreatment observations of the outcome variables. In this case, the first observation is used to select patients for treatment, and the second to measure the change due to treatment.[12]

Proper use of these techniques will reward the clinical researcher with smaller sample size requirements and the clinician with better judgements.

ACKNOWLEDGEMENTS

This work was supported in part by grants from the U.S. Department of Health and Human Services, National Center for Health Services Research (PHS R18 HS2485 and PHS R18 HS04080) and the U.S. National Institute of Arthritis, Diabetes, Digestive and Kidney Diseases (1 P60 AM 20542).

Special thanks to Ed Quigley (Horsham, Pennsylvania) and Kerry Lee (Duke University Medical Center) for their thoughtful criticisms of the manuscript and to Shelly Glazener for her superb secretarial support.

Thanks also to Herman Rubin and Myra Samuels (Purdue University) for helpful discussions about regression toward the mean.

APPENDIX

Here we present a brief proof that the expected value in the results of an observation in the upper 'tail' of a distribution will tend to decrease with repeat determinations. The proof is given for bounded positive random variables. The case for extreme values in the lower tail of the distribution can be obtained by a very similar argument with reversal of the scale. The more general case of unbounded random variables requires more technical details than is feasible to present here.

Let X_1 and X_2 be jointly distributed random variables which represent a baseline observation and a repeat observation, respectively. An extreme value of X_1 is defined to be an observation in the set $\{X_1 \geqslant a\}$ where a is some constant. The average value of X_2 for extreme values of X_1 is denoted $E(X_2|X_1 \geqslant a)$. In an analogous fashion, $E(X_1|X_1 \geqslant a)$ denotes the average value of the extreme X_1s.

Given that X_1 and X_2 have identical marginal distributions, inequality (4) will hold.

$$E(X_2|X_1 \geqslant a) \leqslant E(X_1|X_1 \geqslant a) \tag{4}$$

First we write:

$$E(X_1|X_1 \geqslant a) = k\int_0^\infty P(X_1 \geqslant \max\,(a, c))\,dc \tag{5}$$

and

$$E(X_2|X_1 \geqslant a) = k\int_0^\infty P(X_1 \geqslant a,\ X_2 \geqslant c)\,dc \tag{6}$$

We will prove that inequality (1) holds by showing the integrand in (2) is always greater than or equal to that in (3). To see that this is so, consider the cases in which $a \geqslant c$ and the case in which $a > c$. For the case that $a < c$,

$$P(X_1 \geqslant \max\,(a, c)) = P(X_1 \geqslant a)$$

Obviously, this value is greater than $P(X_1 \geq a, X_2 \geq c)$. For $a < c$,

$$P(X_1 \geq \max(a, c)) = P(X_1 \geq c)$$
$$= P(X_2 \geq c)$$

The latter equality follows from the assumption of identical marginal distributions for X_1 and X_2. The last expression is obviously greater than or equal to $P(X_1 \geq a, X_2 \geq c)$. This concludes the proof. Note that the inequality in (1) is strict wherever any of the statements 'greater than or equal to' above can be replaced by 'greater than' for a non-trivial set of cs.

REFERENCES

1. Beecher, H. 'The powerful placebo', *Journal of the American Medical Association*, **159**, 1602–1606 (1955).
2. Vogel, A., Goodwin, J. and Goodwin, J. 'The therapeutics of placebo', *American Family Physician*, **22**, 105–109 (1980).
3. Gallimore, R. and Turner, J. 'Contemporary studies of placebo phenomena'. *Psychopharmacology in the practice of medicine*, Edited by M. E. Jarvick. Appelton Century-Crofts, New York, 47–57, (1976).
4. Benson, H. and McCallie, D. 'Angina pectoris and the placebo effect', *New England Journal of Medicine*, **300**, 1424–129 (1979).
5. Lown, B. 'Verbal conditioning of angina pectoris during exercise testing', *American Journal of Cardiology*, **40**, 630–634 (1977).
6. Wolf, S. 'Effects of suggestion and conditioning on the action of chemical agents in human subjects—the pharmacology of placebos', *Journal of Clinical Investigation*, **29**, 100–109, (1950).
7. Wintrobe, M., Lee, G., Boggs, D., Bithell, T., Athens, J. and Foerster, J. *Clinical Hematology*, 7th Edition, Lea & Febiger, Philadelphia, 1974, p. 252.
8. Galton, F. 'Regression towards mediocrity in hereditary stature', *Journal of the Anthropological Institute of Great Britain and Ireland*, **15**, 246–263 (1885–86).
9. David, C. E. 'The effect of regression to the mean in epidemiologic and clinical studies', *American Journal of Epidemiology*, **104**, 493–498 (1976).
10. Gardner, M. J. and Heady, J. A. 'Some effects of within person variability epidemiologic studies', *Journal of Chronic Diseases*, **26**, 781–795 (1973).
11. Shepard, D. S. 'Reliability of blood pressure measurements: implications for designing and evaluating programs to control hypertension', *Journal of Chronic Diseases*, **34**, 191–209 (1981).
12. Ederer, F. 'Serum cholesterol changes: effects of diet and regression toward the mean', *Journal of Chronic Diseases*, **25**, 277–289 (1972).
13. Silverman, G. 'Placebo effect and changes in response set with retesting: a further source of bias', *Neuropharmacology*, **18**, 1019–1021 (1979).
14. Tversky, A. and Kahneman, D. 'Judgement under uncertainty: heuristics and biases', *Science*, **185**, 1124–1131 (1974).
15. Anderson, T. W. *Introduction to Multivariate Statistical Analysis*, Wiley, New York, 1958, p. 29.
16. Chan, W., Dawood, M. and Fuchs, F. 'Relief of dysmenorrhea with the prostaglandin synthetase inhibitor ibuprofen: effect on prostaglandin levels in menstrual fluid', *American Journal of Obstetric Gynecology*, **135**, 102–108 (1979).
17. Kershenobich, D., Uribe, M., Suarez, G., Mata, J., Perez-Tamayo, R. and Rojkind, M. 'Treatment of cirrhosis with colchicine: a double-blind randomized trial', *Gastroenterology*, **77**, 532–536 (1979).
18. Thadani, U. and Parker, J. 'Propranolol in angina pectoris: duration of improved exercise tolerance and circulatory effects after acute oral administration', *American Journal of Cardiology*, **44**, 118–125 (1979).
19. Bell, E., Brown, E., Milner, R., Sinclair, J. and Zipursky, A. 'Vitamin E absorption in small premature infants', *Pediatrics*, **63**, 830–832 (1979).
20. Shopsin, B., Klein, H., Aaronsom, M. and Collora, M. 'Clozapine, chlorpromazine, and placebo in newly hospitalized, acutely schizophrenic patients', *Archives of General Psychiatry*, **36**, 657–664 (1979).
21. Cook, C., Center, R. and Michaels, S. 'An acne grading method using photographic standards', *Archives of Dermatology*, **115**, 571–575 (1979).
22. Liuzzi, A., Chiodini, P., Oppizzi, G. *et al.* 'Lisuride hydrogen maleate: evidence for a long lasting dopaminergic activity in humans', *Journal of Clinical Endocrinology and Metabolism*, **46**, 196–202 (1978).

23. Danahy, D., Tobis, J., Aronow, W., Chetty, K. and Glauser, F. 'Effects of isosorbide dinitrate on pulmonary hypertension in chronic obstructive pulmonary disease', *Clinical Pharmacology and Therapeutics*, **25**, 541–548 (1979).
24. Prusoff, A., Williams, D., Weissman, M. and Astrachan, B. 'Treatment of secondary depression in schizophrenia: a double-blind, placebo-controlled trial of amitriptyline added to perphenazine', *Archives of General Psychiatry*, **36**, 569–575 (1979).
25. Klaiber, E., Broverman, D., Vogel, W. and Kobayashi, Y. 'Estrogen therapy for severe persistent depressions in women', *Archives of General Psychiatry*, **36**, 550–554 (1979).
26. Report of the National Research Council Committee on Clinical Evaluation of Narcotic Antagonists. 'Clinical evaluation of naltrexone treatment of opiate-dependent individuals', *Archives of General Psychiatry*, **35**, 335–340 (1978).
27. Morrison, J. and Jennings, J. 'Primary dysmennorrhea treated with indomethacin', *Southern Medical Journal*, **72**, 425–428 (1979).
28. Miller, S. and Vertes, V. 'Ticrynafen and hydrochlorothiazide: a double-blind study of antihypertensive properties with an open crossover', *Journal of the American Medical Association*, **241**, 2174–2176 (1979).
29. Olivari, M., Bartorelli, C., Polese, A., Fiorentini, C., Moruzzi, P. and Guazzi, M. 'Treatment of hypertension with nifedipine, a calcium antagonistic agent', *Circulation*, **59**, 1056–1062 (1979).
30. Alexander, P., Van Kammen, D. and Bunney, W. 'Antipsychotic effects of lithium in schizophrenia', *American Journal of Psychiatry*, **136**, 283–287 (1979).
31. Carman, J. and Wyatt, R. 'Use of calcitonin in psychotic agitation or mania', *Archives of General Psychiatry*, **36**, 72–75 (1979).
32. Weimar, V., Puhl, S., Smith, W. and tenBroeke, J. 'Zinc sulfate in acne vulgaris', *Archives of Dermatology*, **114**, 1776–1778 (1978).
33. Strumza, P., Rigaud, M., Mechmeche, R., Rocha, P., Baudet, M., Bardet, J. and Bourdarias, J. P. 'Prolonged hemodynamic effects (12 hours) or orally administered sustained-release nitroglycerin', *American Journal of Cardiology*, **43**, 272–277 (1979).
34. Harley, J., Matthews, C. and Eichman, P. 'Synthetic food colors and hyperactivity in children: a double-blind challenge experiment', *Pediatrics*, **62**, 975–983 (1978).
35. Birkasova, M., Birkas, O., Flynn, M. and Cort, J. 'Desmopressin in the management of nocturnal enuresis in children: a double-blind study', *Pediatrics*, **62**, 970–974 (1978).
36. Lam, S., Lam, K., Lai, C., Yeung, C., Yam, L. and Wong, W. 'Treatment of duodenal ulcer with antacid and sulpiride: a double-blind controlled study', *Gastroenterology*, **76**, 315–322 (1979).
37. Aronow, W., Lurie, M., Turbow, M., Whittaker, K., Van Camp, S. and Hughes, D. 'Effect of prazosin vs placebo on chronic left ventricular heart failure', *Circulation*, **59**, 344–350 (1979).
38. Kuo, P. T., Hayase, K., Kostis, J. B. and Moreyra, A. E. 'Use of combined diet and colestipol in long-term (7–7½ years) treatment of patients with type II hyperlipoproteinemia', *Circulation*, **59**, 199–211 (1979).
39. Marks, K. H., Berman, W., Friedman, Z., Whitman, V., Lee, C. and Maisels, M. J. 'Furosemide in hyaline membrane disease', *Pediatrics*, **62**, 785–788 (1978).
40. Orlando, J. R., Danahy, D. T., Lurie, M. and Aronow, W. S. 'Effect of trimazosin on hemodynamics in chronic heart failure', *Clinical Pharmacology and Therapeutics*, **24**, 531–536 (1978).
41. Little, J. W., Hall, W. J., Douglas, R. G., Mudholkar, G. S., Speers, D. M. and Patel, K. 'Airway hyperreactivity and peripheral airway dysfunction in influenza A infection', *American Review of Respiratory Diseases*, **118**, 295–303 (1978).
42. Levine, J. D., Gordon, N. C. and Fields, H. L. 'The mechanism of placebo analgesia', *Lancet*, **1**, 654–657 (1978).
43. Gold, M. S., Redmond, D. E. and Kleber, H. D. 'Clonidine blocks acute opiate-withdrawal symptoms', *Lancet*, **2**, 599–601 (1978).
44. Ongley, R. C. 'Efficacy of topical miconazole treatment of tinea pedis', *Canadian Medical Association Journal*, **119**, 353–354 (1978).
45. Horan, J. D. and Johnson, J. D. 'Flunisolide nasal spray in the treatment of perennial rhinitis', *Canadian Medical Association Journal*, **119**, 334–338 (1978).
46. Harris, E., Kanofsky, P., Shakarji, G. and Cotlove, E. 'Biological and analytic components of variation in long-term studies of serum constituents in normal subjects. II. Estimating biological components of variation', *Clinical Chemistry*, **16**, 1022–1027 (1970).
47. Cotlove, E., Harris, E. and Williams, G. 'Biological and analytic components of variation in long-term studies of serum constituents in normal subjects. III. Physiological and medical implications', *Clinical Chemistry*, **16**, 1028–1032 (1970).

48. Shapiro, A. K. 'Factors contributing to the placebo effect and their implications for psychotherapy', *American Journal of Psychotherapy*, **18**, Supplement 1, 73–88 (1964).
49. Liberman, R. 'An analysis of the placebo phenomenon', *Journal of Chronic Diseases*, **15**, 761–783 (1962).
50. Medical Research Council Working Party, 'Randomized controlled trial of treatment for mild hypertension. Design of a pilot trial', *British Medical Journal*, **1**, 1437–1440 (1977).
51. Palmer, E. 'On the natural history of esophageal varices which are secondary to portal cirrhosis', *Annals of Internal Medicine*, **47**, 18–26 (1957).