THE B FILES

Case studies of bias in real life epidemiologic studies

Bias File 5. How blind are the blind? The story of Vitamin C for common cold

Compiled by

Madhukar Pai, MD, PhD

Jay S Kaufman, PhD

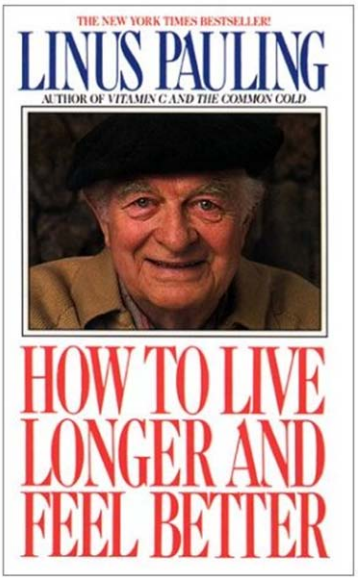
Department of Epidemiology, Biostatistics & Occupational Health

McGill University, Montreal, Canada

madhukar.pai@mcgill.ca & jay.kaufman@mcgill.ca

Bias File 5. How blind are the blind? The story of Vitamin C for common cold

The story



The causal effect of Vitamin C (ascorbic acid) on common cold is an old and enduring controversy. A Cochrane review in 2007 (Hemila 2007) reviewed the evidence from 30 randomized controlled trials, involving 11,350 participants. The review results suggested that regular ingestion of vitamin C had no effect on common cold incidence in the ordinary population. It reduced the duration and severity of common cold symptoms slightly, although the magnitude of the effect was so small its clinical usefulness was doubtful. Despite lack of strong evidence, Vit C continues to be widely sold and used for cold.

Vit C was particularly popular in the 1970s when Linus Pauling (a double Nobel laureate) advocated large dose vitamin C for colds and other conditions such as cancer. Scientists, however, were skeptical of his claims. Randomized trials on Vitamin C came under scrutiny and a randomized trial published in 1975 (Karlowski 1975) became a classic example of how lack of adequate blinding (also called masking) in a trial can result in serious bias (Chalmers TC, 1975; Weiss NS, 2006). What went wrong and why?

The study

From the original abstract (Karlowski et al 1975):

Three hundred eleven employees of the National Institutes of Health volunteered to take 1 gm of ascorbic acid or lactose placebo in capsules three times a day for nine months. At the onset of a cold, the volunteers were given an additional 3 gm daily of either a placebo or ascorbic acid. One hundred ninety volunteers completed the study. Dropouts were defined as those who missed at least one month of drug ingestion. They represented 44% of the placebo group and 34% of those taking ascorbic acid. Analysis of these data showed that ascorbic acid had at best only a minor influence on the duration and severity of colds, and that the effects demonstrated might be explained equally well by a break in the double blind.

The bias

According to the authors, this study was designed during the summer and rushed into operation to take advantage of the rise in upper respiratory infections expected to occur in the fall. There was no time to design, test, and have manufactured a placebo that would be indistinguishable from ascorbic acid.

Early in the trial, the investigators discovered that some of the volunteers had bitten into and tasted the contents of their capsules, and claimed to know whether they were taking ascorbic acid or the lactose placebo. The seriousness of this problem became apparent at the end of the trial, when the investigators asked each participant to fill out a questionnaire which asked them to guess which substance they had been taking. The results of the questionnaire survey is shown in the table.

TABLE V* **Results of Questionnaire on a Prophylactic Drug Ingested by Each Volunteer**

	Suspected Drug			
Actual Drug	Ascorbic Acid	Placebo	Unknown	Total
Ascorbic acid	40	12	49	101
Placebo	11	39	39	89
Total	51	51	88	190

NOTE: $x^2 = 28.6$, $p < 0.001$.
* From Karlowski et al. [12].

A significant number of the volunteers had correctly guessed their medication. Although there were no differences in the frequency of colds between the two treatment groups, there were distinct differences in duration and severity favoring ascorbic acid. The differences in duration of symptoms were totally eliminated, and the differences in severity largely eliminated when those who had guessed their therapy were excluded in a retrospective analysis. In other words, Vit C appeared to significantly reduce the duration of cold symptoms, but the effect disappeared when the analysis excluded participants who were 'unblinded.' The authors concluded that there was a clear association between knowledge of the medication taken and severity and duration of cold symptoms. They stated that the study data "strongly favor the possibility that the effects of ascorbic acid on symptoms are the result of the power of suggestion." (the famous "placebo effect").

The lesson

In this example, because a subject's suspicion of the group to which he or she had been assigned so strongly influenced the results, and because a subject's suspicion was much more often right than wrong, the validity of the trial was seriously compromised (Weiss NS 2006).

Blinding is a critical component of any randomized controlled trial, and every effort must be made to ensure blinding, especially when the outcomes are "soft" (e.g. pain, duration of symptoms, and such subjective measurements). Blinding is not critical for "hard" outcomes such as death. Schulz and Grimes (2002), in their nice overview on blinding, describe the potential benefits of blinding (shown in the box).

Panel 1: **Potential benefits accruing dependent on those individuals successfully blinded**

Individuals blinded	Potential benefits
Participants	Less likely to have biased psychological or physical responses to intervention More likely to comply with trial regimens Less likely to seek additional adjunct interventions Less likely to leave trial without providing outcome data, leading to lost to follow-up
Trial investigators	Less likely to transfer their inclinations or attitudes to participants Less likely to differentially administer co-interventions Less likely to differentially adjust dose Less likely to differentially withdraw participants Less likely to differentially encourage or discourage participants to continue trial
Assessors	Less likely to have biases affect their outcome assessments, especially with subjective outcomes of interest

Schulz KF, Grimes DA. Lancet. 2002 Feb 23;359(9307):696-700.

However, blinding is not always possible or feasible. For example, blinding is a big challenge in surgical intervention trials, although sham procedures have been tried, often with ethical concerns. Even if blinding is feasible, participants are known to deliberately taste or test their pills (in one study, participants actually had the pills subject to biochemical analyses!). Accidental unblinding (unmasking) can also happen. For example, the participant might accidentally learn about her treatment status if the drug were to cause a distinctive side effect (e.g. slowing of pulse rate, discoloration of urine). One typical example is from placebo-controlled trials of zinc lozenges for common cold, where zinc lozenges have a characteristic taste that is easy to differentiate from placebo lozenges. Another example is the Aspirin Myocardial Infarction Study (AMIS) trial, a double-blind placebo-controlled trial to test the effect of aspirin on the survival of people who had experienced a prior heart attack. In this trial, participants tasted the pills to identify aspirin.

Furthermore, the issue of whether all trials must check on whether blinding was successful is contentious. Some claim that all trials must do an assessment of blinding during or at the end of the trial, while others point out that this approach does not always work. For example, if a drug is truly efficacious, then the patient (and the treating doctor) can easily guess the treatment that was given with high degree of certainty, even without tasting or testing the pills. So, when a treatment is truly effective, unmasking of blindness is expected to happen. Also, it is not clear how exactly success of blinding can be measured accurately. Asking participants to guess is one method, but other approaches are also used. The typical approach is to ask the participants and clinic staff to guess to which group the participant was assigned. Ideally, in a trial with one half of the participants on treatment and the other half on placebo (or control drug), the guesses would be correct about 50% of the time in each group. If the guess rate is much higher than 50%, then this would suggest some degree of unblinding. If the guess rate is much lower than 50%, then it is possible that participants did know but were trying hard to not admit it.

A recent review found that methods used to assessing the success of blinding, analysis and reporting the results were inconsistent and questionable (Boutron I et al. 2005). Another review found very trials actually try to check whether blinding was successful (Fergusson D et al. 2004). Regardless of all these complexities, everyone agrees that randomized controlled trials must clearly report who was blinded (e.g. patient, physician, outcome assessor or all of them), how exactly blinding was implemented (e.g. using a placebo that looked and tasted the same as the drug), and whether any efforts were made during or at the end of the trial to check on whether blinding was successful. Merely reporting a trial as single or double blind is not sufficient and does not help readers to judge the quality of the trial.

Sources and suggested readings*

Hemilä H, Chalker E, Treacy B, Douglas B. Vitamin C for preventing and treating the common cold. *Cochrane Database of Systematic Reviews* 2007, Issue 3. Art. No.: CD000980.

Karlowski TR, Chalmers TC, Frenkel LD, Kapikian AZ, Lewis TL, Lynch JM. Ascorbic acid for the common cold. A prophylactic and therapeutic trial. JAMA. 1975 Mar 10;231(10):1038-42.

Chalmers TC. Effects of ascorbic acid on the common cold. An evaluation of the evidence. Am J Med. 1975 Apr;58(4):532-6.

Weiss NS. Clinical Epidemiology. 3rd Edition. Oxford Univ Press, 2006.

Schulz KF, Grimes DA. Blinding in randomised trials: hiding who got what. Lancet. 2002 Feb 23;359(9307):696-700.

Boutron I, Estellat C, Ravaud P. A review of blinding in randomized controlled trials found results inconsistent and questionable. J Clin Epidemiol. 2005 Dec;58(12):1220-6. Epub 2005 Sep 30.

Fergusson D, Glass KC, Waring D, Shapiro S. Turning a blind eye: the success of blinding reported in a random sample of randomised, placebo controlled trials. BMJ. 2004 Feb 21;328(7437):432. Epub 2004 Jan 22.

*From this readings list, the most relevant papers are enclosed.

Image credit: Wikipedia

Acknowledgement

We are grateful to Professor Stan Shapiro, McGill University, Seattle, for his input and suggestions.

Ascorbic Acid for the Common Cold

A Prophylactic and Therapeutic Trial

Thomas R. Karlowski, MD; Thomas C. Chalmers, MD; Lawrence D. Frenkel, MD; Albert Z. Kapikian, MD; Thomas L. Lewis, MD; John M. Lynch, MD

• **Three hundred eleven employees of the National Institutes of Health volunteered to take 1 gm of ascorbic acid or lactose placebo in capsules three times a day for nine months. At the onset of a cold, the volunteers were given an additional 3 gm daily of either a placebo or ascorbic acid. One hundred ninety volunteers completed the study. Dropouts were defined as those who missed at least one month of drug ingestion. They represented 44% of the placebo group and 34% of those taking ascorbic acid. Analysis of these data showed that ascorbic acid had at best only a minor influence on the duration and severity of colds, and that the effects demonstrated might be explained equally well by a break in the double blind.**

(*JAMA* 231:1038-1042, 1975)

ASCORBIC acid has been repeatedly recommended as a prophylactic and therapeutic agent for the common cold, and a large number of studies of its efficacy in this use have been carried out.[1] Nevertheless, a careful review of the literature has failed to reveal a long-term prospective double-blind trial with high doses of ascorbic acid that was designed to measure as well as distinguish between the prophylactic and therapeutic effects of the agent. We undertook a study in an effort to fill this void.

See also p 1073.

From the Clinical Center (Drs. Karlowski, Chalmers, Frenkel, Lewis, and Lynch) and the National Institute of Allergy and Infectious Diseases (Dr. Kapikian), National Institutes of Health, Bethesda, Md. Dr. Chalmers is now with the Mount Sinai Medical Center and School of Medicine of the City University of New York, New York. Dr. Karlowski is now in private practice in Waterloo, Iowa. Dr. Frenkel is now in private practice in Washington, DC.

Reprint requests to Mount Sinai Medical Center, Fifth Ave and 100th St, New York, NY 10029 (Dr. Chalmers).

METHODS

A random sample of 2,500 National Institutes of Health employees received a questionnaire asking about the number of colds they had experienced during the previous 12 months and whether they were interested in joining a study of the effectiveness of ascorbic acid in preventing and ameliorating the common cold. Nearly 600 persons returned questionnaires, and 473 of these persons were considered by the Employee Health Service for entrance into the study. A preliminary history and physical examination were then carried out on these employees.

Volunteers were excluded from joining the study if they fulfilled any one of the following criteria: (1) history of diabetes, gout, renal stones, or respiratory symptoms of probable allergic origin, (2) women who were pregnant, suspected of being pregnant, or anticipated pregnancy in the following year, (3) those taking orally administered anticoagulants or medications that produce side effects involving the respiratory system, (4) those unwilling or unable to refrain from taking vitamin preparations containing ascorbic acid, and (5) those found to have an elevated blood uric acid level.

Of the initial 473 potential volunteers, 323 were admitted to the study and given a prescription for the study drug. Twelve persons dropped out of the study before taking an appreciable number of capsules and were eliminated from further consideration. Thus, 311 employees were classified as having been admitted to the study.

Randomization

An unrestricted randomization was used. The numbers 0, 1, 2, 3 were generated from pseudorandom sequence by the usual method on a CDC 3100 computer. The first 70 identification numbers were assigned to the volunteers who reported that they had experienced four or more colds during the previous year; the remaining 241 numbers were assigned to volunteers reporting fewer than four colds. The treatment groups are described in Table 1.

Treatment Groups

The maintenance dose of study drug was two capsules taken three times a day with meals. When a volunteer had a cold, the dose was doubled (supplemental drug was recommended: two additional capsules three times a day). Each test capsule contained 500 mg of ascorbic acid, 180 mg of spray-dried lactose, and 5 mg of magnesium stearate. Each placebo

capsule contained 645 mg of lactose and 5 mg of magnesium stearate. Validity of the randomization was checked by comparing the frequency of the following characteristics in the four groups and the two combinations: age (less than 35, 35 to 55, and more than 55 years), sex, race, number of cigarettes smoked daily (zero, one to ten, more than ten), history of nonrespiratory allergy, and previous regular intake of vitamins. The only significant ($P < .05$) discrepancy was in the distribution of a history of allergy between group 0 and groups 2 and 3. Similar comparisons were carried out for each combination of the six characteristics, for example, age and sex, age and race. No differences significant at the .05 level were found.

The study drug was dispensed by the pharmacy in bottles of 200. Volunteers were asked to return at monthly intervals with their bottles for refills of their prescription. At this time, they were interviewed for symptoms of side effects, and a check was made on the capsules remaining in their bottle to ascertain how carefully they were taking their study drug. Distribution of the capsules actually ingested by the volunteers was very similar for the ascorbic acid and placebo groups and indicated reasonable compliance with the prescribed study. Of the volunteers who completed the study, 99% took at least four of their six capsules per day, and 87% took at least five capsules per day.

If a cold developed, the volunteers were instructed to return to have their symptoms and clinical observations recorded and to receive supplemental study drug to be taken for the first five days of their colds. A routine throat culture was obtained and the volunteers were encouraged to have nasal washings and blood titers for viral isolation and identification performed. The volunteers were seen three times a week for the duration of their colds. Symptoms were recorded daily and observations were made and recorded for each return visit. Twenty different but interrelated symptoms were graded on a 0 to 3+ scale. The number of days home from work was also recorded. The end of the common cold was defined as that point in time when the individual being monitored failed to fulfill the criteria as described in the definition outlined in the next section.

Table 1.—Experimental Design

Treatment Group	Daily Dose	
	Maintenance	Supplemental
0	Placebo	Placebo
1	Placebo	Ascorbic acid, 3 gm
2	Ascorbic acid, 3 gm	Placebo
3	Ascorbic acid, 3 gm	Ascorbic acid, 3 gm

Table 2.—Results of Questionnaire on Drug

Actual Drug	Suspected Drug		Do Not Know	Total
	Ascorbic Acid	Placebo		
Prophylactic				
Ascorbic acid	40*	12*	49	**101**
Placebo	11*	39*	39	**89**
Total	**51**	**51**	**88**	**190**
Supplemental (therapeutic)				
Ascorbic acid	20†	11†	39	**70**
Placebo	10†	12†	40	**62**
Total	**30**	**23**	**79**	**132**

*χ^2=28.6; P<.001.
†χ^2=1.8; P<.25.

Table 3.—Colds in Treatment Groups

Group	No. of Colds	No. of Persons	Average No. of Colds per Person	
Completed study				
0	65	46	1.41	1.36
1	56	43	1.30	
2	52	44	1.18	1.27
3	76	57	1.33	
Total	**249**	**190**	1.31	
Did not complete study				
0 and 1	27	69	0.82*	
2 and 3	25	52	1.07*	

*Prorated to a nine-month basis.

Treatment Failures

In this study, an individual was defined as having a common cold if he complained of the acute onset of at least two symptoms indicative of conditions in either of the following categories: (1) sneezing, nasal congestion, rhinorrhea, and postnasal drip, or (2) laryngitis, pharyngitis, dysphagia, and bronchitis.

Selected for analysis were the number of colds per person, in total and according to various prerandomization characteristics, mean duration of colds and time at home, and summation of severity scores. Since most of the distributions were quite skewed, limiting the usefulness of a comparison of means, the Wilcoxon test for shift was employed. In the case of severity of symptoms, the scores were ranked according to magnitude, and the Wilcoxon two-sample test applied. The value of the Wilcoxon test was then converted to a standard normal deviate.

Estimates of the numbers required for the study were based on the estimated number of colds to be expected, on the basis of previous experience. It was calculated that if 100 volunteers were assigned to the maintenance ascorbic acid group and the same number to the placebo group, there would be a 95% chance that a 30% reduction in colds among the treated volunteers would be detected, provided that such reduction in fact exists. An extra 100 patients were admitted to allow for dropouts.

It was assumed that there was no need to try to detect a reduction of less than 30% because in the possible application of the study results to the general population, less than a 30% reduction would not be worth the trouble involved in taking two capsules three times a day.

Cessation of Study

It was decided to continue the study for a period of one year, provided that (1) the dropout rates from the group treated with ascorbic acid and the placebo group did not become significantly different (the level of significance was to be taken at .15); (2) the number of persons under study did not fall below 200; (3) at six months from the beginning of study, the number of colds in the ascorbic acid-treated group was not significantly greater than the number in the placebo group (level of significance, .05). In the event that one of these situations occurred, the study would be stopped.

It was stopped nine months after the last subjects had entered, when the number remaining dropped below 200 and it was apparent that more of the dropouts were in the placebo group ($P = .10$).

Maintenance of Double Blind

Early in the study, we discovered that some of the volunteers had tasted the contents of their capsules and professed to know whether they were taking the ascorbic acid or the placebo. The magnitude of the problem was not realized until completion of the study, when a questionnaire was submitted to each of the participants asking them to guess which substance they had been taking. The results of the questionnaire (Table 2) made it mandatory to perform the analyses both in toto as well as according to the participants' impression as to what they were taking. The data in the tables indicate that there was more tasting of the prophylactic capsules than of the therapeutic ones.

Table 4.—Distribution of Colds According to Knowledge of Capsule Contents

Treatment	No. of Colds Per Person 0-1	≥2	Total
Placebo			
Knew	19	20	39
Did not know	19	20	39
Guessed wrong	9	2	11
Subtotal	47	42	89
Ascorbic acid			
Knew	25	15	40
Did not know	36	13	49
Guessed wrong	4	8	12
Subtotal	65	36	101
Total	112	78	190

RESULTS

Average Cold Rates

Table 3 shows the number of colds per person among those who did and did not complete the study. Those receiving ascorbic acid prophylactically had 1.27 colds in nine months and those receiving a placebo, 1.36 ($P > .50$). Knowledge of the medication ingested did not appreciably change the numbers of colds per person (Table 4), except that those who guessed wrong had an interesting distribution of colds. The frequency of colds by month is indicated in Table 5. There was some increase in the winter months and a sharp drop-off in the spring, when the study was terminated. Not much ascorbic acid effect is apparent in the monthly rates. Also, no differences significant at the .05 level occurred in the average number of colds per person according to the previously determined characteristics: age, sex, race, number of cigarettes smoked daily, history of allergy, or previous vitamins.

Volunteers taking placebo had colds of a mean duration of 7.14 days, while those taking 3 gm of ascorbic acid (groups 2 and 3) had colds of a mean duration of 6.59 days and those taking 6 gm had colds of a mean duration of 5.92 days. Thus, each 3-gm increment of ascorbic acid would appear to shorten the mean duration of a cold by approximately half a day. However, these differences were eliminated by taking into account the correct guesses of medication ingested (Table 6).

Analyzing the severity of the 20 recorded symptoms of a cold in each treatment group and the association between knowledge of capsule content and severity presented a complicated statistical problem (Table 7). Each symptom on each day of the cold had been graded 0 to 3, depending on whether it was absent, mild, moderate, or severe. The total score for each symptom was obtained by adding the digits for all colds, and this sum was divided by the number of colds to give the average score per volunteer. For

Table 5.—Average No. of Colds per Person by Month, According to Treatment

Month	Ascorbic Acid: Completed Study	Ascorbic Acid: Did Not Complete Study	Ascorbic Acid: Combined	Placebo: Completed Study	Placebo: Did Not Complete Study	Placebo: Combined
1. September	0.139	0.140	0.139	0.135	0.118	0.127
2. October	0.099	0.152	0.116	0.191	0.044	0.127
3. November	0.119	0.065	0.106	0.146	0.109	0.133
4. December	0.139	0.080	0.127	0.146	0.057	0.121
5. January	0.168	0.130	0.161	0.191	0.172	0.186
6. February	0.208	0.153	0.202	0.202	0.095	0.182
7. March	0.198	0.200	0.198	0.169	0.056	0.150
8. April	0.129	0	0.121	0.124	0.111	0.122
9. May	0.069	0	0.065	0.056	0	0.054

Table 6.—Mean Duration of Colds: Completed Study

Group	No. of Colds	Duration, Days	"Blinded" Subjects: No. of Colds	"Blinded" Subjects: Duration, Days	"Unblinded" Subjects: No. of Colds	"Unblinded" Subjects: Duration, Days
0	65	7.1	30	6.3	16	8.6
1	56	6.5	18	6.7	15	4.7
2	52	6.7	14	6.4	8	7.0
3	76	5.9	30	6.5	13	4.8

each of 20 symptoms, the distributions of clinical scores among the two groups of subjects were compared by ranking the scores according to magnitude and applying the Wilcoxon two-sample test. The value of the Wilcoxon test was then converted to a standard normal deviate (Z statistic). When the effects of therapeutic ascorbic acid were examined (columns 2 and 3 of Table 7), no trends in the shifts were encountered, so these were combined for comparisons of the prophylactic ascorbic acid and placebo (column 4). Here there was a distinct tendency for the ascorbic acid volunteers to have less severe symptoms. (A positive Z statistic indicates that those scores of the groups designated first in the column tended to be higher than the scores of the groups designated second, and vice versa for a negative statistic.) In only two of the 20 symptoms did the shift favor the placebo, and four of the 18 symptoms whose shift favored ascorbic acid were significant, two at the .05 and two at the .01 levels.

The effects of knowledge on the clinical score (columns 5 and 6) were assessed by comparing again the respective Z statistics. This was done by finding the difference between the Z statistics and then dividing by $\sqrt{2}$. Positive values of the resulting statistics (column 7) are in keeping with the tendency of subjects who know they are getting placebo to rate the severity of their symptoms higher than those subjects on placebo without knowing it, and the opposite tendency in the subjects receiving ascorbic acid with and without knowledge. Fifteen of the 20 and all nine of the symptoms complained of by more than 50% of the subjects are positive Z figures in column 7, strongly indicating an association between severity of symptoms and correct guessing of the medication received. The difference in symptoms between the placebo and ascorbic acid groups was lessened when the symptoms of those who did not know what they were taking were analyzed separately (column 6 of Table 7). The placebo is favored in an additional four symptoms, and now none of the differences favoring ascorbic acid is significant.

Table 7.—Comparison of Distributions of Clinical Scores of Symptoms, Based on Average Score per Volunteer

		Z Statistic					
				Groups 0 and 1 vs Groups 2 and 3			
Symptom	Subjects With Zero Scores, %	Group 0 vs Group 1	Group 2 vs Group 3	Total	Knew	Did Not Know	Difference*
Sneezing	22	−1.10	0.14	−1.43	−0.02	−1.51	1.05
Stopped-up nose	15	−1.59	0.43	1.05	1.12	0.79	0.23
Runny nose, watery	16	−1.41	−0.24	0.09	0.83	−0.51	0.95
Runny nose, thick	77	0.57	0.83	−2.06†	−1.82	−0.21	−1.14
Postnasal drip	19	0.58	−1.31	2.21†	2.89‡	0.55	1.65
Sore throat	20	0.11	−0.43	1.31	1.31	−0.67	1.40
Pain on swallowing	51	0.29	−0.38	2.63‡	2.83‡	0.90	1.36
Hoarseness	30	−0.73	−0.64	2.79‡	2.35†	1.46	0.63
Cough, dry	39	0.36	−0.60	1.99†	1.86	0.74	0.79
Cough, productive	58	−1.00	−0.29	1.09	0.96	1.02	−0.04
Chest pain	73	−0.40	−0.24	1.14	0.16	1.30	−0.81
Headache	41	0.37	−1.14	0.79	1.12	0.79	0.23
Eyes tearing	66	1.55	0.01	0.14	0.41	−0.04	0.32
Earache	81	0.16	0.42	1.88	1.44	1.10	0.24
Loss of appetite	72	1.70	−0.03	1.52	0.66	1.54	−0.62
Feverishness	55	2.13†	0.24	0.18	0.02	−0.18	0.14
Chills	67	0.59	−0.76	0.22	1.25	0.13	0.79
Night sweats	80	1.54	0.64	1.20	1.21	0.75	0.33
General aches and pains	59	0.13	−0.39	1.55	0.71	0.73	−0.01
Feel below par	9	0.34	0.39	1.59	2.38†	0.58	1.27
Stay at home	54	−0.46	0.04	0.14	0.07	0.03	0.03

*Values in previous two columns divided by $\sqrt{2}$.
†$P<.05$.
‡$P<.01$.

Viral Studies

Only 20% of the volunteers had complete virus isolation studies. A virus was isolated or a serologic response demonstrated in 15 of the 39 volunteers (33%, an infection rate consistent with previous studies in civilian adults). Although the virus infection rate in the prophylactic placebo groups combined was higher than that in the ascorbic acid groups combined, the difference was not statistically significant.

Side Effects

No important side effects could be determined in either the placebo or ascorbic acid groups. A battery of laboratory tests that included measurements of the albumin-globulin ratio, alkaline phosphatase, total bilirubin, calcium, cholesterol, glucose, lactic acid dehydrogenase, phosphorus, serum glutamic oxaloacetic transaminase, urea nitrogen, and uric acid in 20 randomly selected volunteers failed to reveal any difference between the ascorbic acid and placebo groups.

COMMENT

This study was designed during the summer and rushed into operation to take advantage of the rise in upper respiratory infections expected to occur in the fall. There was no time to design, test, and have manufactured a placebo that would be indistinguishable from ascorbic acid. It did not occur to the investigators that a substantial number of the volunteers, presumably fully informed about the purpose of the study and the importance of the double blind, would not be able to resist indefinitely the temptation to learn which medication they were taking. In retrospect, this phenomenon is understandable in any study that continues for as long as nine months. The increasing placebo-ascorbic acid disproportion in dropout rates made the investigators suspicious that the study might not be completely blind, and this was confirmed by the data from the routine

end-of-study questionnaire, an essential ingredient of all clinical trials.

The Power of Suggestion

Depending on one's point of view, it is either an unfortunate or fortunate aspect of the study. It would have been gratifying to have performed a flawless clinical trial; on the other hand, it has turned out to be a unique opportunity to gain some insight into the importance of perfect blinding in trials with subjective endpoints. An association between severity and duration of symptoms and knowledge of the medication taken seems to have been clearly established.

Most pertinent then are the following questions: Did the participants who had less severe and shorter colds than formerly guess correctly that they were receiving ascorbic acid because they expected it to be effective, while those who had more severe colds assumed that they must be taking the placebo? Or did those who knew they were taking ascorbic acid or placebo because they had tasted their capsules have less or more severe colds as a result of suggestion? In an attempt to determine which might be the appropriate explanation, those who guessed correctly and confessed to tasting were compared with those who did not admit to tasting but did guess correctly, possibly by chance. Unfortunately, the numbers were too small to draw definite conclusions, and, in addition, the possibility remains that a number may have tasted and not confessed to having done so.

In any event, the effects of ascorbic acid on the number of colds seem to be nil (an average of 0.11 colds per person per year), and the effects on severity, although statistically significant if tests are allowable when the blinding has been broken, are clinically insignificant. In view of the absence of any information on possible toxicity if the medication is taken in such high doses over a period of years, it does not seem worthwhile to take two capsules or tablets three times a day for the rest of one's life to achieve such a small and equivocal benefit. Furthermore, recent studies in animals show that ascorbic acid mobilizes calcium from bone,[2] and this could be a disastrous long-term side effect, albeit difficult to prove, in people with a tendency to osteoporosis, and other side effects of long-term administration are possible.[3] No study of these has ever been conducted.

Small Effect in Other Studies

This study is in conformity with the rest of the better clinical trials of prophylaxis.[4] A review of nine reasonably well-controlled trials in 3,940 volunteers has shown an average difference in number of colds per person per year of 0.09 ± 0.06 (1 SE), and in duration of 0.11 ± 0.24 days, both favoring ascorbic acid. The only other study that included a questionnaire at the end did not reveal any breaking of the blind.[5] A study among Navajo Indian children[6] did not include a questionnaire, but school children might not be expected to break the blind purposely. However, observations of the colds were made by the children's teachers, and the latter might have had an irresistible curiosity about the nature of the pills ingested. In this study, there was little effect on number of colds, and the statistically significant severity effect was not found in the older boys.

The two challenge experiments[7,8] revealed no influence on number of colds, but in one there was slight but significant effect on the severity score. No information was available on breaking of the blind.

Some corroborative data suggest a possible mechanism for an effect on severity of colds, if such exists. A drop in white blood cell ascorbic acid with onset of a cold has been demonstrated,[9] and an antihistamine-like action of ascorbic acid has been shown in some volunteers[10] and patients.[11]

Caveats

Several caveats are necessary with regard to interpretation of the present study. Viral culture data were not obtained on enough of the colds to be sure that a significant effect on a particular viral infection was not being missed. The volunteers were all healthy and it is possible that serious late effects of colds might be prevented by ascorbic acid, especially in the elderly or infirm, who might be consuming a subnormal amount in their food.

Finally, there are two claims made by ascorbic acid advocates that have not been tested in this trial. Many people are convinced of their ability to abort a cold by taking ascorbic acid at the first sign and repeating the dose every few hours for a day or two. The therapeutic increment in this study was begun after delays of 1 to 24 hours. Obviously, an increased awareness could lead to many false impressions that a cold is starting, and there would result a false impression that ascorbic acid had prevented it. Now, a competent trial by the British General Practitioner Clinical Trials Group has shown no suggestion of an early ascorbic acid effect.[12] We know of no attempt to study whether or not herpes labialis is aborted by the early ingestion of large doses of ascorbic acid.

Dennis A. George, MD, George M. Shaffer, MD, Nancy Mullen, RN, and Helen Auth, RT, assisted with this study.

References

1. Pauling L: *Vitamin C and the Common Cold.* San Francisco, WH Freeman & Co, Publishers, 1970.
2. Thornton PA, Omdahl JL: Further evidence of skeletal response to exogenous ascorbic acid. *Proc Soc Exp Biol Med* 132:618-621, 1969.
3. Spero LM, Anderson TW: Ascorbic acid and common colds. *Br Med J* 2:354, 1973.
4. Chalmers TC: Effects of ascorbic acid on the common cold: An evaluation of the evidence. *Am J Med,* to be published.
5. Anderson TW, Reid DBW, Beaton GH: Vitamin C and the common cold: A double-blind trial. *Can Med Assoc J* 107:503-508, 1972.
6. Coulehan JL, Reisinger KS, Rogers KD, et al: Vitamin C prophylaxis in a boarding school. *N Engl J Med* 290:6-10, 1974.
7. Walker GH, Bynoe ML, Tyrell DAJ: Trial of ascorbic acid in prevention of colds. *Br Med J* 1:603-606, 1967.
8. Schwartz AR, Togo Y, Hornick RB, et al: Evaluation of the efficacy of ascorbic acid in prophylaxis of induced rhinovirus 44 infection in man. *J Infect Dis* 128:500-505, 1973.
9. Hume R, Weyers E: Changes in leucocyte ascorbic acid during the common cold. *Scott Med J* 18:3-7, 1973.
10. Zuskin E, Lewis AJ, Bouhuys A: Inhibition of histamine-induced airway constriction by ascorbic acid. *J Allergy Clin Immunol* 51:218-226, 1973.
11. Valic F, Zuskin E: Pharmacological prevention of acute ventilatory capacity reduction in flax dust exposure. *Br J Ind Med* 30:381-384, 1973.
12. Ineffectiveness of Vitamin C in treating coryza, General Practitioner Research Group. *Practitioner* 200:442-445, 1968.

for a 24-month interval. A total of 65,751 male and female recruits of all races were screened and those with positive screening tests for hemoglobin S were confirmed to be hemoglobin S-positive by hemoglobin electrophoresis. The incidence of the sickle cell trait among 7,986 black recruits remained at 8% with no significant difference between men and women. A total of 57,665 nonblack recruits (45,500 men and 12,165 women) were screened. Twenty-seven nonblack recruits (25 men and two women have been confirmed to have sickle cell trait. In cases where blood could be tested from one or both of their parents (11 cases so far), the mother had AS hemoglobin in six cases, so there is no reason to suspect this male preponderance is more than statistical variation at present. This gives a refined estimate of sickle cell trait in this nonblack population of .046%.

As pointed out in the previous communication, some articles have implied that sickle cell trait in the nonblack population indicated "obvious Negro admixture." One Rh genotype (cDe) has been thought to be the "African blood-group marker" or an indication of Negro ancestry.[2] Of the 11 nonblack families tested, none have had this genotype. An attempt to determine ancestor history has been fruitful in about half of the cases, with definite Mediterranean origin in about one fourth. Another one fourth are of European and English ancestry; the remainder of the nonblack recruits with sickle cell trait do not know their ancestry.

The opinions expressed in this communication are those of the author and do not represent official Navy Department policy.

CAPT C. J. McGREW, JR., MC, USN (RET)
Spring Hill, Fla

1. Castro O, et al: Splenic infarction in sickle cell trait: Are whites more susceptible? *N Engl J Med* 291:630-631, 1974.
2. Gelpi AP: Migrant populations and the diffusion of the sickle-cell gene. *Ann Intern Med* 79:258-264, 1973.

Television Film: "Primates"

To the Editor.—I have just read the editorial by Dr. Kingman, Executive Director of the National Society for Medical Research, on the television film, "Primates" (231:392, 1975).

Dr. Kingman states, "Dr. Geoffrey Bourne . . . was scheduled to appear prior to the showing of the film in an attempt to place it in perspective, but withdrew." This statement is totally incorrect. When I previewed the film, I was very worried about certain segments and made a three-minute videotape introduction that I asked the Public Broadcasting Service (PBS) to show prior to the showing of the film. This they refused to do. However, the Atlanta station, and later on, I believe, the St. Louis station, did show the introduction. What the PBS network offered instead was a discussion with Wiseman (the producer) to be presented *after* the film. I refused to appear on this discussion because the network obviously wanted to use it to stir up more controversy, and I thought that no good purpose could be served by this. Second, I saw no reason to dignify the producer of the film by debating with him in public a film in which he had so grossly misrepresented us.

Dr. Kingman also says, "In spite of warnings that Wiseman should not be permitted to film within the Yerkes Laboratories, Bourne permitted the film to be made." First, when I was approached about letting Wiseman make a film of the Center, I had never heard of him or any of his films. Second, after arrangements had been made for him to film at the Center, I was given secondhand information that a Hollywood documentary producer thought that Wiseman would not do a good job and that his technique might not be suitable for our Center. No other warnings or comments were received by me or any of my staff. Before Wiseman started filming, he was asked by one of our staff if his technique would not produce an antivivisection type of film. He replied that it could, but that was not his intention. When asked how he would avoid it, he replied that he would see that there was suitable narration to explain what was going on.

GEOFFREY H. BOURNE, DSC, DPHIL, *Director*
Yerkes Primate Research Center
Emory University
Atlanta

Liver Scans to Detect Metastases

To the Editor.—We read with great interest the article entitled "Liver Scanning in Patients With Suspected Abdominal Tumor" by Fee and others (230:1675, 1974). However, we would like to comment on their result. The techniques to evaluate hepatic metastases at laparotomy are several; namely, simple inspection and palpation, needle biopsy, and marginal wedge biopsy. Fee et al mentioned that at exploratory laparotomy, the liver was examined (we presume by inspection and palpation) in every case and biopsies performed in questionable cases (not in all cases). In our experience, patients with a positive liver scan had a perfectly normal liver by inspection and palpation, but biopsy (needle or marginal wedge type) indicated an abnormal liver. So, if the scan was positive and the surgeon inferred a normal liver on the basis of only inspection and palpation, it would be wrong to term the scan as false-positive. With needle or wedge biopsy, one can reduce the incidence of false-positive liver scans. Similarly, even if one gets a negative finding by needle biopsy, it would be mistaken to conclude a false-positive liver scan. We have observed cases where the scan was abnormal, needle biopsy was normal, but liver abnormality was finally established at autopsy. It means that the needle missed the trouble area. We would therefore caution against assuming a false-positive liver scan without having definite proof of a normal liver by all means. Biopsy in some selected cases might be the major reason for the high incidence of false-positive liver scans in the study of Fee et al.

T. K. CHAUDHURI, MD
L. G. MAY, MD
G. DE LOS SANTOS, MD
VA Center
Hampton, Va

CORRECTIONS

Cannell, Not Connell.—In the SPECIAL COMMUNICATION, "Controlled Study of the Cytotoxic Food Test," published in the Feb 17 issue (231:728-730, 1975), the surname of coauthor Barry Cannell, MD, was misspelled Connell in the byline, affiliation footnote, and in the Table of Contents.

Wrong Affiliation.—Lawrence D. Frenkel, MD, a coauthor of the ORIGINAL CONTRIBUTION, "Ascorbic Acid for the Common Cold: A Prophylactic and Therapeutic Trial," published in the March 10 issue (231:1038-1042, 1975), is **not** in private practice, as indicated in the affiliation footnote on page 1038. Dr. Frenkel is a Fellow in Immunology, Allergy, and Infectious Diseases, and an Instructor in Pediatrics at the Georgetown University School of Medicine, Washington, DC.

Effects of Ascorbic Acid on the Common Cold

An Evaluation of the Evidence

THOMAS C. CHALMERS, M.D.
New York, New York

Of 14 clinical trials of ascorbic acid in the prevention and treatment of the common cold, the data from 8 were considered well enough gathered to be creditable and to warrant combining for an over-all assessment of efficacy. Differences in mean prorated numbers of colds per year and durations of illness were 0.09 ± 0.06 (± 1 standard error) and 0.11 ± 0.24, respectively, favoring ascorbic acid over the placebo. These are minor and insignificant differences, but in most studies the severity of symptoms was significantly worse in the patients who received the placebo. In one study lasting 9 months, a large number of the volunteers tasted their capsules and correctly guessed what group they were in. All differences in severity and duration were eliminated by analyzing only the data from those who did not know which drug they were taking. Since there are no data on the long-term toxicity of ascorbic acid when given in doses of 1 g or more per day, it is concluded that the minor benefits of questionable validity are not worth the potential risk, no matter how small that might be.

Widespread sales of the book "Vitamin C and the Common Cold" by Professor Linus Pauling [1] have undoubtedly resulted in even greater sales of ascorbic acid to the self-prescribing public. There has also resulted a continuing controversy over the efficacy of the drug or vitamin in the medical literature, and the addition of several more clinical trials to the many that had been carried out before publication of Dr. Pauling's book. My purpose is to review the evidence for and against the efficacy of ascorbic acid in preventing colds, shortening their duration and alleviating their symptoms. The data suggest that ascorbic acid does have some effect on the severity of cold symptoms, but the effects are quantitatively so small, and the possibility of suggestion as the primary mechanism so large, that it hardly seems worthwhile for anyone to take all those pills for such a long time. This is especially true in view of the fact that there are as yet no data on long-term toxicity.

Fourteen clinical trials carried out by 11 investigators [2–13] between 1942 and 1974 have been reviewed. Five of the studies [2–5] (Table I) have been classified as poorly controlled because

From the Mount Sinai Medical Center, New York, New York 10029. Requests for reprints should be addressed to Dr. Thomas C. Chalmers, The Mount Sinai Medical Center, Fifth Avenue and 100th Street, New York, New York 10029. Manuscript accepted July 26, 1974.

TABLE I Ascorbic Acid and the Common Cold—Poorly Controlled Studies (neither randomized nor double blind)

Source	Ascorbic Acid			Placebo			Ascorbic Acid-Placebo Difference	
	No. of Subjects	Colds/ Person/ Year	Mean Duration (days)	No. of Subjects	Colds/ Person/ Year	Mean Duration (days)	Colds/Person/ Year	Duration (days)
Charleston, Clegg 1972 [5]	47	3.24	3.5	43	6.45	4.2	3.21	0.7
Barnes 1961 [4]	22	6.00	2.0	16	8.25	6.5	2.25	4.5
Dahlberg et al. 1944 [3]	1,259	0.40	Not reported	1,266	0.51	Not reported	0.05	Not reported
Glazebrook, Thomson 1942 [2]	335	0.43	2.5	1,100	0.52	4.9	0.09	2.5
Glazebrook, Thomson 1942 [2]	60	1.13	3.2	90	1.28	4.0	0.15	0.8
						Mean ± 1 standard error:	1.15 ± 0.66	2.12 ± 0.89

they did not employ the technics of randomization and double blinding. The data are summarized as the differences in the number of colds per year among those taking ascorbic acid and those taking a placebo, and the difference in the mean duration of colds, when available. The five studies by four groups of investigators add up to a mean difference of 1.2 ± 0.7 (±1 standard error) colds per year and a mean difference in duration of colds in days of 2.1 ± 0.9. Only the difference in mean duration is significantly different from zero ($P < 0.05$).

In the case of the randomized or double blind studies (Table II), the differences are smaller, amounting to 1/10 of a cold per year and an average difference in duration of 1/10 of a day per cold. These differences, although favoring ascorbic acid, are far from statistically significant. Two of the eight favored placebo in number of colds, and in three the duration

TABLE II Ascorbic Acid and the Common Cold—Reasonably Well Controlled Studies

Source	Ascorbic Acid			Placebo			Ascorbic Acid-Placebo Difference	
	No. of Subjects	Colds/ Person/ Year	Mean Duration (days)	No. of Subjects	Colds/ Person/ Year	Mean Duration (days)	Colds/Person Year	Duration (days)
Anderson et al. 1972 [7]	407	5.51	3.96	411	5.92	4.18	0.41	0.22
Anderson et al. 1974 [8]	583	6.03	3.28	578	6.00	3.18	−0.03	−0.10
Coulehan et al. 1974 [11]*	321	0.40	4.71	320	0.46	5.92	0.06	1.21
Wilson et al. 1973 [10]†	158	2.31	2.65	144	2.18	2.79	−0.13	0.14
Karlowski et al. 1974 [12]‡	101	1.69	6.80	89	1.81	6.30	0.12	0.50
Franz et al. 1956 [6]	44	1.27	Not reported	45	1.33	Not reported	0.06	Not reported
Cowan et al. 1942 [13]§	233	1.90	1.10	194	2.20	1.60	0.30	0.50
Cowan et al. 1942 [13]§	227	2.40	1.70	120	2.40	1.00	0	−0.70
						Mean ± 1 standard error:	0.09 ± 0.06	0.11 ± 0.24
Ritzel 1961 [9]¶	139	6.37¶	1.35	140	11.54¶	1.95	5.17¶	0.60

* Double-blind study with subjects assigned to ascorbic acid or placebo group alternately.

† Summary of several trials.

‡ Double-blind broken.

§ Blinding of subjects only, and subjects assigned to ascorbic acid or placebo group alternately.

¶ Highly inaccurate figure because the study lasted less than 2 weeks and the number of colds per person had to be multiplied by 26. Also the assurances on blinding and randomizing are taken from a review by Pauling [25] because they were not included in Ritzel's paper.

TABLE III Ascorbic Acid and the Common Cold—Challenge Experiments in Volunteers

Data	Walker et al. 1967 [14]		Schwartz et al. 1973 [15]	
	Ascorbic Acid	Placebo	Ascorbic Acid	Placebo
No. inoculated	47	44	11	10
No. colds	18	18	11	10
Mean duration of colds (days)	8	7	6	6
Mean severity score	16.5	16.5	5.2*	7.4*

* Fourth day symptoms.

was less on the placebo. In extracting the data from the nine well controlled [7–13] trials, the results in all the patients were combined in each study. Subgroup differences were thus averaged out. In the study by Wilson et al. [10] only the "whole colds" have been tabulated.

It is noteworthy that in all nine of the randomized or double blind studies, patients treated with ascorbic acid prophylactically tended to have less severe symptoms than those who received placebo. In Anderson's study [8] the subjects took increased doses of ascorbic acid with the onset of a cold, but in the study by Karlowski [12], the therapeutic increment was controlled by a placebo, and the therapeutic dose had less of an effect on the severity of symptoms than the prophylactic dose. In the studies by Wilson et al. [10] and by Coulehan et al. [11], the effects on symptoms seemed to be more striking in girls than in boys.

Two groups of investigators have tested the efficacy of ascorbic acid in preventing colds induced by experimental inoculation of viruses in normal volunteers (Table III). Walker et al. [14] gave 3 g of ascorbic acid daily for only 3 days before inoculation of a number of different viruses; they found absolutely no effect on either the number of colds, their duration or their severity. Schwartz et al. [15] gave 3 g daily for 2 weeks before inoculation of a rhinovirus and likewise found that ascorbic acid had no effect on the incidence of colds. However, in those who received the drug, symptoms were slightly less severe.

TABLE V* Results of Questionnaire on a Prophylactic Drug Ingested by Each Volunteer

Actual Drug	Suspected Drug			
	Ascorbic Acid	Placebo	Unknown	Total
Ascorbic acid	40	12	49	101
Placebo	11	39	39	89
Total	51	51	88	190

NOTE: $x^2 = 28.6$, $p < 0.001$.
* From Karlowski et al. [12].

The trial carried out by Karlowski et al. [12] among employees of the National Institutes of Health revealed data most pertinent to the discrepancy between the effects of ascorbic acid on the incidence of colds and on their severity. The quantitative data from that study are summarized in Table IV. There was a minute effect on incidence and a larger one on mean durations of colds. Analyses of the severity of symptoms revealed that volunteers who received ascorbic acid tended to have milder symptoms in 18 of 20 instances, the differences being statistically significant in 5.

However, a questionnaire at the end of the study revealed that a significant number of the volunteers had correctly guessed their medication (Table V). Many had tasted the contents of their capsules. When the severity scores were reanalyzed according to those who knew and those who apparently did not know what they were taking, the differences in symptoms between those taking the placebo and those taking ascorbic acid were lessened. In fact, the group receiving the placebo who thought they were receiving ascorbic acid had fewer colds than the group receiving ascorbic acid who thought they were receiving the placebo ($p = 0.05$). There were no differences in the durations of colds among those who did not know what they were taking; those who did

TABLE IV* Comparison of Ascorbic Acid and Placebo in Prevention and Treatment of Colds

Group		No. Persons Completing Study	Total No. of Colds	Mean† No. of Colds per Person	Mean† Duration of Colds (days)
Prophylactic	Therapeutic				
Placebo	Placebo	46	65	1.41 ± 0.19	7.14 ± 0.46
Placebo	Ascorbic acid	43	56	1.30 ± 0.18	6.46 ± 0.39
Ascorbic acid	Placebo	44	52	1.18 ± 0.16	6.71 ± 0.53
Ascorbic acid	Ascorbic acid	57	76	1.33 ± 0.15	5.92 ± 0.40

* From Karlowski et al. [12].
† ± 1 standard error.

know demonstrated an appreciable ascorbic acid "effect." Similarly the differences in severity were largely eliminated when knowledge of the pill taken was included in the analyses.

The data in this study strongly favor the possibility that the effects of ascorbic acid on symptoms are the result of the power of suggestion. However, no such effect was demonstrated in the only other study in which a questionnaire was employed to determine whether or not the blind had been broken [7].

If the minor effects of ascorbic acid are real, there are three studies of physiologic changes which might explain them. Hume and Weyers [16] found that the ascorbic acid level in leukocytes drops sharply on the first day of a cold. Zuskin et al. [17] found that ascorbic acid reduces the airway constriction induced by the inhalation of histamine in adults. Valik and Zuskin [18] also found that it diminished the airway constrictor effects of certain textile dusts.

None of the clinical trials has revealed any significant toxicity of ascorbic acid when given in doses as high as 3 to 6 g/day. It is known, however, that renal stones may complicate chronic acidification of the urine that may result from such a regimen. It has also been suggested that mobilization of calcium from bone may result from chronic ingestion of large doses of ascorbic acid [19], and this could be a very serious long-term side effect, especially in women because of their greater tendency to have osteoporosis. Late toxicity of oral hypoglycemic agents in patients with middle-age diabetes [20] and the late appearance of carcinoma of the vagina in the female offspring of women given the "harmless" stilbestrol as a means of preventing threatened abortion [21] are examples of totally unanticipated long-term toxicity. The latter is a particularly poignant one because the drug was totally ineffective in preventing abortion [22]. So the absence of any apparent short-term toxicity of ascorbic acid does not mean that there cannot be serious long-term effects, and in the case of pregnant women risks to the fetus have not been ruled out. In addition, a number of theoretic toxic effects [23] must be kept in mind. Considering the lack of efficacy in preventing colds and the very small effects on symptoms, which could be due to suggestion, it hardly seems worth the risk to encourage people to take large doses of ascorbic acid over long periods of time [24].

The conclusion drawn from this analysis of the published controlled trials is the opposite of that drawn from the analysis carried out by Pauling [25], of four trials, published before 1971. However, Pauling averaged "p" values from the different studies rather than differences in the number of colds, and he also omitted the second study by Cowan which was entirely negative.

The best way to conclude this review of the evidence for and against the efficacy of ascorbic acid in preventing the common cold and amelioration of its symptoms is to state that I, who have colds as often and as severe as those of any man, do not consider the very minor potential benefit that might result from taking ascorbic acid three times a day for life worth either the effort or the risk, no matter how slight the latter might be.

ACKNOWLEDGMENT

I wish to thank Stephanie Lee, Anne-Marie Kunzler and Biruta Schroeder for their help in the gathering of data for this paper.

REFERENCES

1. Pauling L: Vitamin C and the Common Cold, San Francisco, W. H. Freeman & Co., 1970.
2. Glazebrook AJ, Thomson S: The administration of vitamin C in a large institution and its effect on general health and resistance to infection. J Hygiene 42: 1, 1942.
3. Dahlberg G, Engel A, Rydin H: The value of ascorbic acid as a prophylactic against "common colds." Acta Med Scand 119: 540, 1944.
4. Barnes FE: Vitamin supplements and the incidence of colds in high school basketball players. NC Med J 22: 22, 1961.
5. Charleston SS, Clegg KM: Ascorbic acid and the common cold. Lancet 1: 1401, 1972.
6. Franz WL, Heyl HL, Sands W: Blood ascorbic acid level in bioflavonoid and ascorbic acid therapy of common cold. JAMA 162: 1224, 1956.
7. Anderson TW, Reid, DB, Beaton GH: Vitamin C and the common cold: a double-blind trial. Can Med Assoc J 107: 503, 1972.
8. Anderson TW, Suranyi G, Beaton GH: The effect on winter illness of large doses of vitamin C. Can Med Assoc J 111: 31, 1974.
9. Ritzel G: Kritische Beurteilung des Vitamins C als Prophylacticum und Therapeuticum des Erkaltungskrankenheiten. Helv Med Acta 28: 63, 1961.
10. Wilson AWM, Loh HS, Foster FG: The beneficial effect of vitamin C on the common cold. Europ J Clin Pharmacol 6: 26, 1973.
11. Coulehan JL, Reisinger KS, Rogers KD, Bradley DW: Vitamin C prophylaxis in a boarding school. N Engl J Med 290: 6, 1974.
12. Karlowski TR, Chalmers TC, Frenkel LD, et al.: A prophylactic and therapeutic trial of ascorbic acid for the common cold. JAMA (in press).
13. Cowan DW, Diehl HD, Baker AB: Vitamins for the prevention of colds. JAMA 120: 1268, 1942.
14. Walker GH, Bynoe ML, Tyrell DAJ: Trial of ascorbic acid in prevention of colds. Br Med J 1: 603, 1967.
15. Schwartz AR, Hornick RB, Tominaga S, Gleckman RA:

Evaluation of the efficacy of ascorbic acid in prophylaxis of induced rhinovirus 44 infection in man. J Infect Dis 128: 500, 1973.
16. Hume R, Weyers E: Changes in leucocyte ascorbic acid during the common cold. Scott Med J 18: 3, 1973.
17. Zuskin E, Lewis AJ, Bouhuys A: Inhibition of histamine-induced airway constriction by ascorbic acid. J Allergy Clin Immunol 51: 218, 1973.
18. Valic F, Zuskin E: Pharmacological prevention of acute ventilatory capacity reduction in flax dust exposure. Br J Indust Med 30: 381, 1973.
19. Thornton PA, Omdahl JL: Further evidence of skeletal response to exogenous ascorbic acid. Proc Soc Exp Biol Med 132: 618, 1969.
20. University Group Diabetes Program. A study of the effects of hypoglycemic agents on vascular complications in patients with adult-onset diabetes. Diabetes 19 (suppl 2): 1, 1970.
21. Herbst AL, et al.: Clear-cell adenocarcinoma of the genital tract. N Engl J Med 287: 1259, 1972.
22. Chalmers TC, et al.: The impact of clinical trials on the practice of medicine. Mt Sinai J Med NY 41: 753, 1974.
23. Brown R: Possible problems of large intakes of ascorbic acid. JAMA 224: 1529. 1973.
24. Anderson TW: Letter to the editor. Br Med J 2: 354, 1973.
25. Pauling L: The significance of the evidence about ascorbic acid and the common cold. Proc Natl Acad Sci USA 68: 2678, 1971.

Epidemiology series

Blinding in randomised trials: hiding who got what

Kenneth F Schulz, David A Grimes

Blinding embodies a rich history spanning over two centuries. Most researchers worldwide understand blinding terminology, but confusion lurks beyond a general comprehension. Terms such as single blind, double blind, and triple blind mean different things to different people. Moreover, many medical researchers confuse blinding with allocation concealment. Such confusion indicates misunderstandings of both. The term blinding refers to keeping trial participants, investigators (usually health-care providers), or assessors (those collecting outcome data) unaware of the assigned intervention, so that they will not be influenced by that knowledge. Blinding usually reduces differential assessment of outcomes (information bias), but can also improve compliance and retention of trial participants while reducing biased supplemental care or treatment (sometimes called co-intervention). Many investigators and readers naively consider a randomised trial as high quality simply because it is double blind, as if double-blinding is the sine qua non of a randomised controlled trial. Although double blinding (blinding investigators, participants, and outcome assessors) indicates a strong design, trials that are not double blinded should not automatically be deemed inferior. Rather than solely relying on terminology like double blinding, researchers should explicitly state who was blinded, and how. We recommend placing greater credence in results when investigators at least blind outcome assessments, except with objective outcomes, such as death, which leave little room for bias. If investigators properly report their blinding efforts, readers can judge them. Unfortunately, many articles do not contain proper reporting. If an article claims blinding without any accompanying clarification, readers should remain sceptical about its effect on bias reduction.

The rich history of blinding in clinical trials spans a couple of centuries.[1] Most researchers worldwide appreciate its meaning. Unfortunately, beyond that general appreciation lurks confusion. Terms such as single-blind, double-blind, and triple-blind mean different things to different people.[2] Moreover, many medical researchers confuse the term blinding with allocation concealment. The fact that such confusion arises suggests that both terms are misunderstood. Clear theoretical and practical differences separate the two. Blinding prevents ascertainment bias and protects the sequence after allocation.[3,4] By contrast, researchers use methods of allocation concealment primarily to prevent selection bias and to protect an assignment sequence before and until allocation. Furthermore, in some trials, blinding cannot be successfully implemented, whereas allocation concealment can always be successfully implemented.[4,5]

Blinding represents an important, distinct aspect of randomised controlled trials.[3] The term blinding refers to keeping trial participants, investigators (usually health-care providers), or assessors (those collecting outcome data) unaware of an assigned intervention, so that they are not influenced by that knowledge. Blinding prevents bias at several stages of a trial, although its relevance varies according to circumstance. Although initial forays into blinding might have used a blindfold,[1] the processes have now become much more elaborate. In this article, we focus on the attributes and benefits of blinding.

Potential effects of blinding

If participants are not blinded, knowledge of group assignment can affect responses to the intervention received.[3] Participants who know that they have been assigned to a group who will receive a new treatment might harbour favourable expectations or increased apprehension. Those assigned to standard treatment, however, might feel deprived or relieved. Despite evidence to suggest that new treatments are as likely to be worse as they are to be better than standard treatments,[6] participants probably assume that new treatments will be better than standard treatments—new means improved. In any case, knowledge of the intervention received, and perceptions of that treatment, can affect the psychological or physical responses of the participants. Knowledge of treatment allocation can also affect compliance and retention of trial participants (panel 1).

Blinding investigators—those who contribute to a broadly defined trial team including, but not limited to, trial designers, participant enrollers, randomisation implementors, health-care providers, intervention counsellors, and routine data collectors—is also important.[3] Investigators especially pertinent to blinding include health-care providers (such as an attending physician or nurse) and intervention counsellors—eg, someone who delivers a behavioural prevention message—who might interact with the participants throughout the trial. If investigators are not blinded, their attitudes for or against an intervention can be directly transferred to participants.[7] Their inclinations could also be manifested in differential use of ancillary interventions of supplemental care or treatment (co-interventions), differential decisions to withdraw participants from a trial, or differential adjustments to the medication dose (panel 1). Investigators might also encourage or discourage continuation in a trial on the basis of knowledge of the intervention group assignment.

Perhaps most importantly, blinding helps to reduce differential assessment of outcomes (often called information or ascertainment bias) (panel 1). For example, if outcome assessors who know of the treatment allocation believe a new intervention is better than an old

Lancet 2002; **359:** 696–700

Family Health International, PO Box 13950, Research Triangle Park, NC 27709, USA (K F Schulz PhD, D A Grimes MD)

Correspondence to: Dr Kenneth F Schulz
(e-mail: KSchulz@fhi.org)

Panel 1: **Potential benefits accruing dependent on those individuals successfully blinded**

Individuals blinded	Potential benefits
Participants	Less likely to have biased psychological or physical responses to intervention More likely to comply with trial regimens Less likely to seek additional adjunct interventions Less likely to leave trial without providing outcome data, leading to lost to follow-up
Trial investigators	Less likely to transfer their inclinations or attitudes to participants Less likely to differentially administer co-interventions Less likely to differentially adjust dose Less likely to differentially withdraw participants Less likely to differentially encourage or discourage participants to continue trial
Assessors	Less likely to have biases affect their outcome assessments, especially with subjective outcomes of interest

one they could register more generous responses to that intervention. Indeed, in a placebo-controlled trial in patients with multiple sclerosis[8] the unblinded, but not the blinded, neurologists' assessments showed an apparent benefit of the intervention.

Subjective outcomes—eg, pain scores—present great opportunities for bias.[3] Furthermore, some outcomes judged objective can be fraught with subjectivity, for example, salpingitis. In general, though, blinding becomes less important to reduce observer bias as the outcomes become less subjective, since objective (hard) outcomes leave little opportunity for bias. Knowledge of the intervention would not greatly affect measurement of a hard outcome, such as death.

Lexicon of blinding

Non-blinded (open or open label) denotes trials in which everyone involved knows who has received which interventions throughout the trial. Blinding (masking) indicates that knowledge of the intervention assignments is hidden from participants, trial investigators, or assessors.

The terminology single blind usually means that one of the three categories of individuals (normally participant rather than investigator) remains unaware of intervention assignments throughout the trial.[9] A single-blind trial might also, confusingly, mean that the participant and investigator both know the intervention, but that the assessor remains unaware of it.

In a double-blind trial, participants, investigators, and assessors usually all remain unaware of the intervention assignments throughout the trial.[3] In view of the fact that three groups are kept ignorant, the terminology double blind is sometimes misleading. In medical research, however, an investigator frequently also assesses, so in this instance the terminology accurately refers to two categories.

Triple blind usually means a double-blind trial that also maintains a blind data analysis.[10] Some investigators, however, denote trials as triple-blind if investigators and assessors are distinct people and both, as well as participants, remain unaware of assignments. Investigators rarely use quadruple blind, but those that do use the term to denote blinding of participants, investigators, assessors, and data analysts.[11] Thus, quintuple blind must mean that the allocation schedule has been lost and nobody knows anything. Contrary to Mae West's claim that "too much of a good thing can be wonderful", such is not always the case in blinding.

Confused terminology of single, double, and triple blinding permeates the literature,[3] with physicians, textbooks, and journal articles all offering different interpretations and definitions.[2] Not only do investigators not define double-blind trials consistently, in particular, but they make matters worse by frequently failing to report their definitions clearly in their articles. Building on the original blindfolding efforts,[1] and the once common double blindfold terminology,[12] we further obfuscate by offering additional definitions of single and double blinding (figure 1). More seriously, when we use double-blind or its derivatives in this article, we mean that steps have been taken to blind participants, investigators, and assessors to group assignments. In reporting randomised controlled trials, we urge researchers to explicitly state what steps they took to keep whom blinded.

Sparse reporting on blinding, however, is common. Many investigators neglect to report whether or not their trial was blinded. For example, reports of 51% of 506 trials in cystic fibrosis,[13] 33% of 196 trials in rheumatoid arthritis,[14] and 38% of 68 trials in dermatology[15] did not state whether blinding was used. When researchers have reported their study as double-blind, they frequently have not provided much further clarification.[14,16–18] For example, of 31 double-blind trials in obstetrics and gynecology, only 14 (45%) reports indicated the similarity of the treatment and control regimens (for example, appearance, taste, administration) and only 5 (16%) provided statements to indicate that blinding was successful.[18]

Masking or blinding

Some people prefer the term masking to blinding to describe the same procedure. Masking might be more appropriate in trials that involve participants who have impaired vision, and could be less confusing in trials in which blindness is an outcome.[3] Blinding, however, conveys a strong bias prevention message. Apparently, blinding terminology emerged when Benjamin Franklin and colleagues[19] actually blindfolded participants to shield them from knowledge in their assessments of the

Figure 1: **The authors: double blinded versus single blinded**

Figure 2: **The authors blinded and masked**

therapeutic claims made for Mesmerism. The imagery of blindfolding, a total covering of the eyes, conveys stronger bias prevention than masking, where eye holes could permit viewing (figure 2). Blinding also suggests a more secure procedure to some. The International Conference on Harmonization (ICH) guidance,[20] for example, primarily uses blinding terminology. (The ICH is an intensive tripartite collaboration between regulatory authorities in Europe, Japan, and the USA to develop common guidelines for the design, implementation, and reporting of clinical trials). We prefer blinding because it has a long history, maintains worldwide recognition, creates strong imagery, and permeates the ICH guidelines.[3]

Placebos and blinding

Interventions (treatments) sometimes have no effect on the outcomes being studied.[3] When an ineffective intervention is administered to participants in the context of a well-designed randomised controlled trial, however, beneficial effects on participants' attitudes sometimes occur, which in turn affect outcomes.[10] Researchers refer to this phenomena as the placebo effect.

A placebo refers to a pharmacologically inactive agent that investigators administer to participants in the control group of a trial.[3] The use of a placebo control group balances the placebo effect in the treatment group, allowing for independent assessment of the treatment effect. Although placebos can have a psychological effect, they are administered to participants in a trial because they are otherwise inactive. An active placebo is a placebo with properties that mimic the symptoms or side-effects—eg, dry mouth, sweating—that might otherwise reveal the identity of the (pharmacologically) active test treatment. Most researchers agree that placebos should be administered, whenever possible, to controls when assessing the effects of a proposed new treatment for a condition for which no effective treatment already exists.[9,10] Indeed, blinding frequently necessitates the use of placebos.

However, a proven effective standard treatment, if such exists, is usually given to the control group for comparison against a new treatment.[3] Thus, investigators might compare two active treatment groups without a placebo group. Even then, however, investigators frequently attempt to achieve blinding by use of the double-dummy method, in essence two placebos.[11,21] For example, for comparison of two agents, one in a blue capsule and the other in a red capsule, the investigators would prepare blue placebo capsules and red placebo capsules. Then both treatment groups would receive a blue and a red capsule, one active and one inactive.

Does blinding prevent bias?

Some investigators, readers, and editors overstate the importance of blinding in prevention of bias. Indeed, some consider a randomised trial as high quality if it is double blind—ie, as if double blinding is the sine qua non of a randomised controlled trial.[3] Unfortunately, scientific life is not that simple. A randomised trial can be methodologically sound and not be double blind or, conversely, double blind and not methodologically sound. Lasagna[12] captured that notion long ago: "Let us examine the placebo somewhat more critically, however, since it and 'double blind' have reached the status of fetishes in our thinking and literature. The Automatic Aura of Respectability, Infallibility, and Scientific Savoir-faire which they possess for many can be easily shown to be undeserved in certain circumstances."[12] Although double blinding suggests a strong design, it is not the primary indicator of overall trial quality. Moreover, many trials cannot be double blinded. Such trials must, therefore, be judged on overall merit rather than an inapplicable standard based on double blinding.

We do not, however, suggest that blinding is unimportant.[3] Intuitively, blinding should reduce bias, and available evidence supports that impression. Methodological investigations tend to show that double blinding prevents bias but is less important, on average, in prevention of bias than is adequate allocation concealment.[4,22,23]

What to look for in descriptions of blinding

In general, if researchers describe a trial as double-blind, readers can assume that they have avoided bias. Empirical evidence lends support to this recommendation. As suggested in the CONSORT guidelines,[24,25] however, investigators should not use only the single-blind, double-blind, or triple-blind terminology, but should also explicitly state who was blinded, and how. Moreover, if the researchers contend that the trial investigators, participants, and assessors were blinded—ie, double blind—then they should provide information about the mechanism (capsules, tablets, film, &c), similarity of treatment characteristics (appearance, taste, administration), and allocation schedule control—eg, location of the schedule during the trial, when the code was broken for the analysis, and circumstances under which the code could be broken for individual instances. Such additional information can lend support to or undermine claims of double-blinding (panel 2).[26–29]

If researchers properly report their blinding efforts, readers can judge those efforts. Unfortunately, many articles will not contain proper reporting. If a researcher claims to have done a blinded study, but does not provide accompanying clarification, readers should remain sceptical about its effect on bias reduction. For example, one trial[30] of prophylactic antibiotics claimed to be blinded, but the methods section of the report revealed that little or no blinding occurred.

Ideally, researchers should also relate if blinding was successful. Investigators can theoretically assess the success of blinding by directly asking participants, health-care providers, or outcome assessors which intervention they think was administered (panel 3). In principle, if blinding was successful, these individuals should not be able to do better than chance when guessing the intervention, for example. In practice, however, blinding might be totally successful, but participants, health-care providers, and outcome assessors might nevertheless guess the intervention because of ancillary information. Disproportionate levels of adverse side-effects might

Panel 2: **Descriptions of blinding**

"No patient, research nurse, investigator, or any other medical or nursing staff in the ICU was aware of the treatment assignments for the duration of the study. All statistical analysis was also done with masking maintained. Randomisation authorities were instructed to report any suspected breach of the masking procedures. No report was filed . . . The drug or placebo (vehicle without active drug) was prepared for syringe pump infusion or for volumetric pump infusion in indistinguishable syringes or bags."[26]

". . . in a double-blind, placebo-controlled manner . . . Neither the patients nor doctors could distinguish the placebo from sibutramine capsules. The taste of the capsules was identical provided they were swallowed whole as instructed. . . Results of biochemical analyses were completed before the randomisation code was broken at the end of the completed trial."[27]

"The study was double-blinded—that is, neither the women nor the study staff, including the biostatisticians at Family Health International, knew which group was using the nonoxynol 9 film. The nonoxynol 9 film contained . . . The placebo film contained . . . The films were identical in appearance, packaging, and labeling."[28]

"The doxycycline and placebo were in capsule form and identical in appearance . . . The randomization code was kept in the USA." (Note: the trial was conducted in Kenya) "Thus, all administration and assessments were done blinded to treatment assignment, and the investigators and patients were also blinded to the ongoing results of the study. The code was broken only after data collection had been completed."[29]

ICU=intensive care unit.

provide strong hints as to the intervention. Irrespective of painstaking efforts to do double-blinded trials, some interventions have side-effects that are so recognisable that their occurrence will unavoidably reveal the intervention received to both the participants and the health-care providers.[11,24] Even more fundamental than hints from adverse effects are the hints from clinical outcomes. Researchers usually welcome large clinical effects (except perhaps in equivalence trials). If they arise, health-care providers and participants would likely deduce—not always accurately of course—that a participant with a positive outcome received the active (new) intervention rather than control (standard). If indeed the active (new) intervention materialises as helpful (highly desirable) then their deductions would be correct more often than chance guesses.[24,31] Irrespective of their suspicions, end-of-trial tests of blindness might actually be tests of hunches for adverse effects or efficacy.[32,33]

Furthermore, individuals might be reluctant to expose any unblinding efforts by providing accurate responses to the queries—in other words, if they have deciphered group assignments, they might provide responses contrary to their deciphering findings to disguise their actions. That difficulty, along with interpretation difficulties stemming from adverse side-effects and successful clinical outcomes, leads us to question the usefulness of tests of blinding in some circumstances. Investigators should carefully consider the usefulness of assessing the success of their blinding efforts, but if they proceed, should provide the results of any assessments. At the very least, they should report any failure of the blinding procedure, such as non-identical placebo or active preparations. Published reports rarely contain assessments of blinding, but, if provided, readers should sceptically assess the information presented.

Panel 3: **Assessment of the success of blinding**

"We asked 126 staff members their opinions of which film was the placebo. Eighteen percent thought film A (the placebo) was the placebo, 13 percent thought film B (nonoxynol 9) was the placebo, and 69 percent had no opinion about which film was the placebo. Of the 68 peer educators (the staff members most likely to reflect the opinion of the participants), 16 percent thought film A was the placebo, 13 percent thought film B was the placebo, and 71 percent had no opinion."[28]

Double blinding proves difficult or impossible in many trials. For instance, in general, surgical trials cannot be double blinded. Specifically, a trial that compares degrees of pain associated with sampling blood from the ear or thumb cannot be double-blinded.[34] If researchers do not describe their trial as double-blind or the equivalent, it could still be scientifically strong. Apart from assessment of the other methodological aspects of the trial, readers would have to assess how much bias might have ensued due to absence of blinding. Readers should identify if anybody was blinded in the trial and what benefits might have accrued (panel 1). Indeed, blinding of outcome assessors is often possible and advisable, even in open trials.[11] For example, lesions can be photographed before and after treatment and assessed by someone not involved in the study.[11] We recommend placing greater credence in results when someone unaware of treatment assignments judges outcome measures.

Even that recommendation, however, is not absolute. As noted earlier, some hard outcomes, such as death, leave little room for ascertainment bias. In other words, blinding the assessor to hard outcomes might have little effect.

Conclusion

Blinding embodies a rich history spanning over two centuries. Most researchers worldwide understand blinding terminology, but confusion lurks beyond a general comprehension. Investigators should clearly explicate those blinded and not blinded in their trial, rather than only labeling their trial as single-blind, double-blind, or triple-blind. Readers should expect such clarity when reading and judging a trial report.

We thank Willard Cates and David L Sackett for their helpful comments on an earlier version of this report. Much of this material stems from our 15 years of teaching the Berlex Foundation Faculty Development Course.

References

1 Kaptchuk TJ. Intentional ignorance: a history of blind assessment and placebo controls in medicine. *Bull Hist Med* 1998; **72:** 389–433.

2 Devereaux PJ, Manns BJ, Ghali WA, et al. Physician interpretations and textbook definitions of blinding terminology in randomized controlled trials. *JAMA* 2001; **285:** 2000–03.

3 Schulz KF, Chalmers I, Altman DG. The landscape and lexicon of blinding in randomized trials. *Ann Intern Med* 2002; **136:** 254–59.

4 Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995; **273:** 408–12.

5 Schulz KF, Chalmers I, Grimes DA, Altman DG. Assessing the quality of randomization from reports of controlled trials published in obstetrics and gynecology journals. *JAMA* 1994; **272:** 125–28.

6 Chalmers I. What is the prior probability of a proposed new treatment being superior to established treatments? *BMJ* 1997; **314:** 74–75.

7 Wolf S. Effects of suggestion and conditioning on action of chemical agents in human subjects: pharmacology of placebos. *J Clin Invest* 1950; **29:** 100–09.

8 Noseworthy JH, Ebers GC, Vandervoort MK, Farquhar RE, Yetisir E, Roberts R. The impact of blinding on the results of a randomized, placebo-controlled multiple sclerosis clinical trial. *Neurology* 1994; **44:** 16–20.

9 Meinert CL. Clinical trials: design, conduct, and analysis. New York: Oxford University Press, 1986.

10 Pocock SJ. Clinical trials: a practical approach. Chichester: Wiley, 1983.

11 Day SJ, Altman DG. Statistics notes: blinding in clinical trials and other studies. *BMJ* 2000; **321:** 504.

12 Lasagna L. The controlled trial: theory and practice. *J Chronic Dis* 1955; **1:** 353–67.

13 Cheng K, Smyth RL, Motley J, O'Hea U, Ashby D. Randomized controlled trials in cystic fibrosis (1966–1997) categorized by time, design, and intervention. *Pediatr Pulmonol* 2000; **29:** 1–7.

14 Gøtzsche PC. Methodology and overt and hidden bias in reports of 196 double-blind trials of nonsteroidal antiinflammatory drugs in rheumatoid arthritis. *Control Clin Trials* 1989; **10:** 31–56.

15 Adetugbo K, Williams H. How well are randomized controlled trials reported in the dermatology literature? *Arch Dermatol* 2000; **136:** 381–85.

16 Mosteller F, Gilbert JP, McPeek B. Reporting standards and research strategies for controlled trials: agenda for the editor. *Controlled Clin Trials* 1980; **1:** 37–58.

17 DerSimonian R, Charette LJ, McPeek B, Mosteller F. Reporting on methods in clinical trials. *N Engl J Med* 1982; **306:** 1332–37.

18 Schulz KF, Grimes DA, Altman DG, Hayes RJ. Blinding and exclusions after allocation in randomised controlled trials: survey of published parallel group trials in obstetrics and gynaecology. *BMJ* 1996; **312:** 742–44.

19 Franklin B, Bailly JS, Lavoisier A. Rapport des commissaires chargés par le roi, de l'examen du magnetisme animal. Nice: Chez Gabriel Floteron, 1785.

20 Department of Health and Human Services, FDA. International conference on harmonisation: guidance on statistical principles for clinical trials. *Fed Regist* 1998; **63:** 49583–98.

21 Altman DG. Practical statistics for medical research. London: Chapman and Hall, 1991.

22 Khan KS, Daya S, Collins JA, Walter SD. Empirical evidence of bias in infertility research: overestimation of treatment effect in crossover trials using pregnancy as the outcome measure. *Fertil Steril* 1996; **65:** 939–45.

23 Moher D, Pham B, Jones A, et al. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *Lancet* 1998; **352:** 609–13.

24 Altman DG, Schulz KF, Moher D, et al. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med* 2001; **134:** 663–94.

25 Moher D, Schulz KF, Altman D. The CONSORT statement: revised recommendations for improving the quality of reports or parallel-group trials. *Lancet* 2001; **357:** 1191–94.

26 Bellomo R, Chapman M, Finfer S, Hickling K, Myburgh J. Low-dose dopamine in patients with early renal dysfunction: a placebo-controlled randomised trial. Australian and New Zealand Intensive Care Society (ANZICS) Clinical Trials Group. *Lancet* 2000; **356:** 2139–43.

27 James WP, Astrup A, Finer N, et al. Effect of sibutramine on weight maintenance after weight loss: a randomised trial. STORM Study Group. *Lancet* 2000; **356:** 2119–25.

28 Roddy RE, Zekeng L, Ryan KA, Tamoufe U, Weir SS, Wong EL. A controlled trial of nonoxynol 9 film to reduce male-to-female transmission of sexually transmitted diseases. *N Engl J Med* 1998; **339:** 504–10.

29 Sinei SK, Schulz KF, Lamptey PR, et al. Preventing IUCD-related pelvic infection: the efficacy of prophylactic doxycycline at insertion. *Br J Obstet Gynaecol* 1990; **97:** 412–19.

30 Baker KR, Drutz HP, Barnes MD. Effectiveness of antibiotic prophylaxis in preventing bacteriuria after multichannel urodynamic investigations: a blind, randomized study in 124 female patients. *Am J Obstet Gynecol* 1991; **165:** 679–81.

31 Quitkin FM, Rabkin JG, Gerald J, Davis JM, Klein DF. Validity of clinical trials of antidepressants. *Am J Psychiatry* 2000; **157:** 327–37.

32 Sackett DL, Gent M, Taylor DW. Tests for the blindness of randomized trials may not. *Clin Res* 1986; **34:** 711A.

33 The Canadian Cooperative Study Group. A randomized trial of aspirin and sulfinpyrazone in threatened stroke. *N Engl J Med* 1978; **299:** 53–59.

34 Carley SD, Libetta C, Flavin B, Butler J, Tong N, Sammy I. An open prospective randomised trial to reduce the pain of blood glucose testing: ear versus thumb. *BMJ* 2000; **321:** 20.