

Brief report

Masking of physicians in the Growth Failure in Children with Renal Diseases Clinical Trial

Russell M. Boyle¹, Vernon M. Chinchilli¹, and Diane A. Shasky²

¹Department of Biostatistics and ²Investigational Drug/Oncology Clinic, Medical College of Virginia, Virginia Commonwealth University, Box 32 MCV Station, Richmond, VA 23298, USA

Received July 30, 1992; received in revised form November 3, 1992; accepted November 4, 1992

Abstract. Masking – hiding identities of treatments from the patient, physician and/or statistician – is a critical element in clinical trials. Wherever possible, masking is implemented to eliminate observational bias or systematic error. In this paper, general concepts of masking in clinical trials are examined. Specific masking procedures used in the "Growth Failure in Children with Renal Diseases" (GFRD) Clinical Trial are described. A method to evaluate the "success" of this masking procedure for physicians is introduced. For each randomized patient at each clinical center, the clinic director was asked to predict which treatment (1,25-dihydroxyvitamin D₃ or dihydrotachysterol) was assigned. Results showed that 72% of responses initially indicated "absolutely no idea" of treatment. Additional analyses revealed that the number and percentage of "correct" guesses were essentially equal for the two treatment groups and that a patient's time on treatment did not affect the mask. We conclude that the mask of physicians in the GFRD Clinical Trial was well maintained.

Key words: Masking – Placebo – Clinical trials – Bias

Introduction

Of critical importance in the design and execution of a clinical trial is the masking of interventions. "Masking," defined as hiding the identities of treatments from the patient, physician and/or statistician, is implemented to minimize or eliminate observational bias. Potential effects of bias introduced in the *absence* of masking are noteworthy [1]. Patients who know that they are receiving a "new" drug may expect or realize positive results, while patients

who are aware that they are assigned to a "standard" group may feel that this treatment is inferior. With knowledge of treatments, physicians may more closely follow patients on a particular treatment arm and, again, expect better results. Also, it is possible that statisticians or others interpreting results may tend to favor the "new" group by "seeing" differences that are really not present. Therefore, masking treatments in clinical trials, where feasible, is recommended.

A clinical trial can be non-masked, single-, double- or triple-masked. In a non-masked trial, patients, physicians and statisticians know the identities of all interventions. In a single-masked trial, only the patient is unaware of the intervention, while a double-masked design prevents both the patient and investigator responsible for following the patient from knowing the treatments. A triple-masked clinical trial is one in which, in addition to the patient and investigator, those monitoring response variables are unaware of identities of the interventions [2]. Of the various methods of achieving masking in trials in which the intervention is an oral medication, most common is the masked placebo, virtually identical to the "treatment" drug but missing any active ingredient. Placebos *must* be closely matched to active drugs in form, size, shape, color and taste to be effective. Trials that compare multiple active drugs of different sizes and shapes can also be masked by assigning regimens of two or more medications (e. g., active "A" plus placebo "B" versus active "B" plus placebo "A").

The "Growth Failure in Children with Renal Diseases" (GFRD) Clinical Trial is a multicenter, triple-masked study described in detail elsewhere [3–5]. Briefly, children 1.5 through 10 years of age with a glomerular filtration rate between 20 and 75 ml/min per 1.73 m² were enrolled in a 6-month control period for baseline measurements. Subjects meeting compliance criteria with a glomerular filtration rate between 20 and 60 ml/min per 1.73 m² were then randomized by the data coordinating center to "A" or "B", either 1,25-dihydroxyvitamin D₃ (calcitriol) or dihydrotachysterol (DHT). Patients were followed in the treatment period for at least 6 months and as long as 5 years. Differences in treatment effect on linear growth, inverse

Correspondence to: R. M. Boyle

Table 1. Masking of physicians in the Growth Failure in Children with Renal Diseases Clinical Trial
94 unscored responses from 26 clinic directors

Response	No.	No. (%) correct
I am certain it was calcitriol	0	-
I am certain it was DHT	0	-
I think it was probably calcitriol	14	6 (43%)
I think it was probably DHT	12	5 (42%)
I have absolutely no idea which treatment was assigned	68	-
	94	
Of the 68 responses with "absolutely no idea":		
Guessed calcitriol	39	23 (59%)
Guessed DHT	22	15 (68%)
Still no idea	7	-
	68	

DHT, Dihydroxyacetone.

serum creatinine and frequency of hypercalcemia were the primary end-points.

Drug procurement and storage, dosage calculation and logistics of masking the two treatment regimens were the responsibilities of a central core pharmacy. Patient medication packages were prepared and mailed to the clinical centers monthly. Because calcitriol and identical-appearing placebos were in capsule form and DHT and placebos were in tablets, patients were prescribed either active calcitriol and placebo DHT or active DHT and placebo calcitriol. Thus, patients took at least one capsule and one tablet daily to help ensure masking.

Preserving the triple-masked design was the responsibility of the core pharmacy. A procedure to break the mask (for an individual patient) was explained in the manual of operations, but the need never arose. Throughout the entire clinical trial, including the period of data analysis, only the core pharmacy knew the identities of the treatment assignments.

Methods

There were 94 patients randomized to treatment in the GFRD Clinical Trial. The data coordinating center sent a brief questionnaire to 26 clinic directors identifying each patient, the randomization date and the length of time on treatment. All 94 questionnaires were returned. In the following manner, we asked the physician to deduce the treatment assigned to his or her patient:

- () I am certain it was calcitriol.
 () I think it was probably calcitriol.
 () I have absolutely no idea which treatment was assigned.
 () I think it was probably DHT.
 () I am certain it was DHT.

If you answered "I have absolutely no idea which treatment was assigned," which treatment would you guess?

- () Calcitriol
 () DHT

This format was modeled directly from the Aspirin Myocardial Infarction Study and discussed by Krol [6]. Forcing a respondent with "absolutely no idea" to make a guess was designed to eliminate truly uncertain responses.

For purposes of analysis, scores were assigned as follows. For a patient actually randomized to calcitriol, a score of +4 was given for a correct "certain calcitriol" response and +2 for a correct "probably calcitriol", while -2 was assigned for an incorrect "probably DHT" and -4 for an incorrect "certain DHT." Scores were similarly assigned for DHT patients. When respondents with "absolutely no idea" were forced to guess, +1 was scored for a correct guess and -1 for incorrect. A score of 0 was given for "absolutely no idea" responses with no guess.

Because multiple observations from each of the 26 physicians were correlated, the 94 patient scores were not independent. To adjust for this, a data set comprised of the median scores for each physician was analyzed in a standard manner since these scores are independent. A one-sample Wilcoxon signed rank test ascertained if the mean response was significantly different from zero. Also, ordinal logistic regression was used to determine whether actual treatment group or length of time on treatment affected the mask. Data were analyzed using SAS version 6.06 [7].

Results

Table 1 presents a summary of responses before scoring. Interestingly, 68 or 72% of responses initially indicated "absolutely no idea" of treatment group; of these, 61 guessed a treatment while 7 did not guess. The number and percentage correct were essentially equal for the "probably calcitriol" and "probably DHT" responses, while a slightly higher proportion of the "guessed DHT" responses were correct compared with the "guessed calcitriol." A combination of 53 responses reported that the assigned treatment was "probably calcitriol" ($n = 14$) or guessed calcitriol ($n = 39$), while 34 thought the treatment was "probably DHT" ($n = 12$) or guessed DHT ($n = 22$). Median scores of the 26 physicians (representing 26 clinical centers) were distributed somewhat evenly around 0 ("no guess"), with 14 positive (correct) scores, 8 negative (incorrect) scores and 4 0s. The Wilcoxon test determined that these median scores were not significantly different from 0 ($P = 0.35$). When all 94 scores were used in ordinal logistic regression, neither treatment group ($P = 0.12$) nor length of time on treatment ($P = 0.90$) was significant.

Discussion

In this study, "ideal" results would indicate that masking of treatments was preserved. Thus, for calcitriol patients, approximately 50% of respondents would guess calcitriol and 50% DHT. For patients randomized to DHT, a similar return would be expected. Using our assigned scores, we would expect mean scores close to 0 for both treatment arms, implying that bias was not introduced.

Previous studies designed to assess masking produced mixed results. Participants and clinic staff were asked to identify treatments in the Lipid Research Clinics Coronary Primary Prevention Trial [8]. Roughly half of participants (56.0%) and physicians (55.2%) made the "correct" guess of cholestyramine. A similar study in the β -Blocker Heart Attack Trial showed that masking was not as successful, but still essentially free of serious bias [9]. Here, 63.6% of patients on propranolol and 60.0% of physicians of propranolol patients correctly guessed the treatment. In contrast, a clinical trial on ascorbic acid clearly failed to main-

tain the mask [10]. Of the participants in this study who guessed their treatment assignment, 78.4% of those on ascorbic acid and 76.4% of those on placebo correctly identified their treatment. The mask was assessed only after a relatively large number of subjects dropped out of the trial, and it was determined that more of these dropouts had been randomized to placebo than to ascorbic acid. Trial organizers discovered that some subjects had actually tasted the contents of the capsules and were able to determine their assigned group, underscoring the importance of taking time to ensure truly masked placebos.

In the GFRD Clinical Trial, our conclusion is that the mask for physicians was preserved. Most striking is the finding that 68 (72%) of the 94 responses and 16 (61%) of the 26 physicians initially expressed "absolutely no idea" which treatment was assigned to these patients. Also, there were no respondents who were "certain" that the treatment was either calcitriol or DHT. This suggests that physicians responsible for caring for these patients did not observe signs or side effects which would have revealed treatment identities. (These results could also imply that physicians did not want to be "wrong" in indicating "certain" or "probable" knowledge of treatment, but we have no reason to believe this was the case). For the 26 "probably calcitriol" and "probably DHT" responses, there were only 11 correct responses. As previously mentioned, 53 responses were "probably calcitriol" or guessed calcitriol, while only 34 indicated "probably DHT" or guessed DHT. There was no difference between these two groups of responses as to purported knowledge of each patient's actual treatment. Also, patients' time on treatment did not affect the mask. These results support the contention that both treatment arms were equally and adequately masked for physicians.

Masking is a tedious and exacting task. Maintaining a triple-mask in a multicenter clinical trial requires planning, commitment and a bit of serendipity. Assessing the mask for physicians and/or patients is strongly recommended so that potential bias from these sources is identified. In the

GFRD Clinical Trial, the data coordinating center, core pharmacy and all participating physicians collaborated to preserve the mask throughout the treatment period as well as the analytical phase. The results of this study suggest that this effort was successful.

Acknowledgements. This research was supported by grants (R01 DK 31370 and R01 DK 32431) from the National Institutes of Health, Bethesda, MD.

References

1. Pocock SJ (1983) *Clinical trials: a practical approach*. Wiley, New York, pp 90-92
2. Friedman LM, Furberg CD, DeMets DL (1981) *Fundamentals of clinical trials*. Wright, Boston, pp 59-63
3. Chan JCM, McEnery PT, Brouhard BH, Chinchilli VM, Greifer I (1990) History and organization of the Growth Failure in Children with Renal Diseases Study. *J Pediatr* 116: S3-S7
4. Chan JCM, Greifer I, Boineau FG, Mendoza SA, McEnery PT, Strife CF, Abitbol CA, Stapleton FB, Roy S, Strauss J (1990) Rationale of the Growth Failure in Children with Renal Diseases Study. *J Pediatr* 116: S11-S15
5. Chan JCM, McEnery PT, Chinchilli VM, Boyle RM, Massie MD, Jennings SJ, Kohaut EC, Dresner IG, Tejani A, Arbus GS (1990) Protocol of the Growth Failure in Children with Renal Diseases Study. *J Pediatr* 116: S17-S21
6. Krol WF (1983) Closing down the study. *J Controlled Clin Trials* 4: 505-512
7. SAS Institute (1990) *SAS procedures guide, version 6, 3rd edn*. SAS Institute, Cary, N. C.
8. Lipid Research Clinics Program (1984) The Lipid Research Clinics Coronary Primary Prevention Trial results: reduction in incidence of coronary heart disease. *Jama* 251: 351-364
9. Byington RP, Curb JD, Mattson ME (1985) Assessment of double-blindness at the conclusion of the β -Blocker Heart Attack Trial. *JAMA* 253: 1733-1736
10. Karlowski TR, Chalmers TC, Frenkel LD, Kapikian AZ, Lewis TL, Lynch JM (1975) Ascorbic acid for the common cold. *JAMA* 231: 1038-1042