21. Principal Requirements for Controlled Clinical Trials of Caries Preventive Agents and Procedures (1982). Technical Report of the Commission on Oral Health, Research and Epidemiology, *Int Dent J* 32:292-310.
22. SCHEFFE, H. (1959): **The Analysis of Variance.** New York: John Wiley & Sons.
23. SEARLE, S.R. (1971): **Linear Models.** New York: John Wiley & Sons.
24. SEARLE, S.R.; SPEED, F.M.; and HENDERSON, H.V. (1981): Some Computations and Model Equivalences in Analyses of Variance of Unequal-Subclass-Numbers Data, *Amer Statistician* 35:16-33.
25. YATES, F. (1934): The Analysis of Multiple Classifications with Unequal Numbers in the Different Classes, *J Amer Statist Assoc* 29:52-66.

Stratification Methods in Caries Clinical Trials: Discussion of Dr. Kingman's Presentation

H. SMITH, Jr.

Mt. Sinai School of Medicine, New York, New York 10025

J Dent Res 63(Spec Iss):777, May, 1984

Dr. Kingman has written an excellent paper on the use of stratification methods in clinical trials. The emphasis in the paper is on post-stratification methods that can be used with the two-way classification model, with or without the interaction term. His analysis is done using the least-squares techniques: method of fitting constants and the method of weighted squares of means.

The paper needs clarification on several assumptions made for the analyses. I quote: ". . .it makes sense to test for main effects in the presence of interaction if one can assume that the treatments are consistently ordered. . . one must also assume that the confounding due to unequal cell frequencies is not too severe." How severe? In the clinical trial time frame – say, 36 months – how do withdrawals from the trial affect this statement?

The same sort of caution on unequal cell sizes takes place in other places in the paper. It would be helpful if Dr. Kingman gave us some idea of how unequal or how unbalanced the data could be without affecting the analysis. Post-stratification will, in general, produce unequal cell sizes and may be a very important consideration.

As to the example, I can only say that whenever an author produces a real set of data he's asking for trouble. Not to be outdone, I have some real reservations about the trial. For example, the fluoride arms were weekly and daily rinses – what about the placebo? Weekly, daily, no rinse? Why were only 268 people randomly allocated to the placebo arm, while 278 people were randomly allocated to the other two arms? Yet, more people finished the trial on the placebo arm than on the two treated arms. The completion rates are 76% on a base of 268 for Placebo, 72% on a base of 278 for Weekly, and 70% on a base of 278 for Daily. Were the placebo patients handled differently? This is a strange result, in my experience.

Table 2 represents a set of data which looks as if the standard deviation is proportional to the mean, thus requiring a log transformation to stabilize the variance. In particular, the significant interaction alluded to in the paper can be attributed to the placebo group, with an initial MGSI of 3. Otherwise everything looks parallel.

Finally, this is a 36-month trial, and there is no way to indicate when the ultimate trends shown in Table 2 took place. For example, what did the Table look like after 12 and 24 months? These aspects are not shown.

All in all, I found the paper excellent, and I congratulate Dr. Kingman on his fine contribution to the conference.