On the Effectiveness of Restaurant Inspection Frequencies

O. BENJAMIN KAPLAN, PHD, MPH

Introduction

Sanitarians inspect restaurants to detect unsafe or unclean conditions and prevent food poisoning. Although it is generally believed that about four inspections per restaurant per year are necessary, Zaki, et al., believe that one or two suffice if a restaurant is found "safe" on two consecutive semi-annual inspections.[1]

"Beliefs" are not too reliable as a guide unless supported by valid statistics. This paper will quantify the theoretical effectiveness of various inspection frequencies used to detect unsafe food conditions (or preparation practices). The quantification may help administrators choose appropriate frequencies.

Theoretical Considerations

A potentially dangerous event or condition, such as food being held for long periods at bacterial incubation temperatures, may not exist all of the time in a given restaurant. An unsafe condition may exist during, say, 10 per cent of a day, week, or year. Let p = probability of one unsafe condition (or of one cluster of unsafe conditions) per unit time; let $P = 100p$ = per cent of time such condition exists; and let $q = 1 - p$.

A hypothetical sanitarian who always detects any and all unsafe conditions inspects restaurants during randomly chosen days and hours. The number of inspections r when this sanitarian does detect an unsafe condition in a given restaurant depends on p and on the frequency of inspections, n, to which the restaurant is subject. By chance alone, the sanitarian's four yearly inspections might occur during the only four days of the year when the unsafe condition does (or does not) exist.

Address reprint requests to Dr. O. Benjamin Kaplan, Environmental Health Planning Coordinator, and Land Use Specialist, Department of Environmental Health Services, 1111 E. Mill St., San Bernardino, CA 92415. This paper, submitted to the Journal November 10, 1977, was revised and accepted for publication December 7, 1977.

The probability of an unsafe condition (p), the frequency of inspections (n) and the number of inspections that detect an unsafe condition (r) can be related by the binomial expansion formula, $nCr\ p^r q^{n-r}$, where $nCr = n!/[r!(n-r)!]$. The right hand column in Table 1 is derived from the binomial expansion.

Discussion and Conclusions

Table 1 shows that two inspections (per restaurant per year) will fail to detect an unsafe condition in 25 per cent of the restaurants which are "unsafe" 50 per cent of the year, and four inspections will fail to detect an unsafe condition in 65.6 per cent of the restaurants which are "unsafe" 10 per cent of the year. It can be calculated that even 30 inspections will fail to detect an unsafe condition in 4 per cent of the restaurants in which such condition exists 10 per cent of the year. For any given restaurant and year, an unsafe condition must exist over 50 per cent of the year to be "reliably" (error $\leq 6\%$) detected in at least one of four inspections.

It can be seen in Table 1 that two, four, or even eight inspections per year may not be adequate to categorize a restaurant's relative "unsafety". For instance, should an administrator want to estimate the relative "unsafety" of a restaurant on the basis of recorded detections (of unsafe conditions) per eight inspections, and should the subject restaurant be truly unsafe 50 per cent of the year, the expected ratio of detections to total inspections should be 4/8 (i.e., 50 per cent). But ratios higher and lower than 4/8 will be obtained 72.7 per cent of the time.* In a similar vein, it can be shown that it is extremely risky to decide, on the basis of a few prior inspections, whether a restaurant is so "safe" that it needs to be inspected only once yearly.

Other difficulties arise when deciding whether a restaurant is unsafe P per cent of the year on the basis of past inspection records:

*Per Table 1, 100 minus 27.3 equals 72.7.

TABLE 1—Probability π of r detections (of unsafe conditions) when performing n inspections per year on a restaurant having an unsafe condition P per cent of the year. P = 100p, π = nCr $p^r q^{n-r}$ and 100 π = percentage of restaurants in which r detections are obtained, given n and P. (r times out of n the unsafe condition is detected.) For easy reference, numbers referred to in the text are italicized.

P	n	r	100π
100%	1	0	0
		1	100
50%	1	0	50
		1	50
	2	0	25
		1	50
		2	25
	4	0	6.25
		1	25
		2	37.5
		3	25
		4	6.25
	8	0	0.39
		1	3.1
		2	10.9
		3	21.8
		4	*27.3*
		5	21.8
		6	10.9
		7	3.12
		8	0.39
10%	*4*	0	*65.6*
		1	29.9
		2	4.9
		3	0.4
		4	0.01
	8	0	43
		1	37.6
		2	14.9
		3	3.3
		≥4	≤0.5

(a.) P may vary from year to year;

(b.) Inspections are usually not performed at statistically representative times. Unsafe conditions may be (and in this author's experience often are) flagrantly common during holidays, weekends, and during evenings. Yet the inspection activity usually occurs during normal workdays, 8:00 am to 5:00 pm;

(c.) Not all sanitarians detect all unsafe conditions. Some sanitarians are more able or conscientious than others. Often, sanitarians are not effectively "calibrated" against each other;

(d.) FDA-like restaurant sanitation scales[2] penalize unsafe conditions with semi-arbitrary demerit points. These points have *not* been obtained from food-poisoning-risk regression equations, or from other statistical analyses of food-poisoning variables. By inspecting and evaluating restaurants according to FDA-like scales, one may segregate restaurants into sanitation level categories, but not into food-poisoning-risk categories.

The detection and prevention powers of even highly frequent inspections are bound to be limited unless the above-mentioned problems are solved. With so much variability and margin for error, it is not surprising that a change in inspection frequency may appear to be inconsequential.

Although the average inspection activity appears to be somewhat ritualistic, it may be useful and important in a variety of other contexts: deterring the unsanitary behavior of restaurant personnel (who fear "getting caught"); collecting information about community problems (sewage disposal, vectors); disseminating health information, etc. Hence, the inspection activity merits improvement, not abandonment.

REFERENCES

1. Zaki, M., et al. A progressive approach to the problem of food infections. Am. J. Public Health 67:44–49, 1977.
2. U.S. Food and Drug Administration. Sanitation Program Formulator (SPIF) Vol. II. U.S.G.P.O., July 1976.

Use of an Edit Feedback System in Data Collection Quality Control

BETH GOLDMAN, MPH, AND OSCAR JONES, BS

Growing concern over individual privacy has affected the research world in the form of more and more restrictions on collecting data on human subjects. Within the federal government this concern has manifested itself in legislation such as the Privacy Act of 1974 and limits on new forms requiring personal identifiers. Fortunately, in some research involving statistical analysis, collecting data without personal identifiers is appropriate. However, since the lack of personal identifiers precludes correcting errors or collecting missing data after forms are completed, data of unacceptably low quality may result. An edit/feedback system may overcome this problem. Such a system does not lead to the immediate correction of errors, but effectively improves the quali-

From the Venereal Disease Control Division, Bureau of State Services, Center for Disease Control, Atlanta, GA. Address reprint requests to Technical Information Services, Bureau of State Services, Center for Disease Control, Atlanta, GA 30333. This paper, submitted to the Journal September 16, 1977, was revised and accepted for publication January 3, 1978.

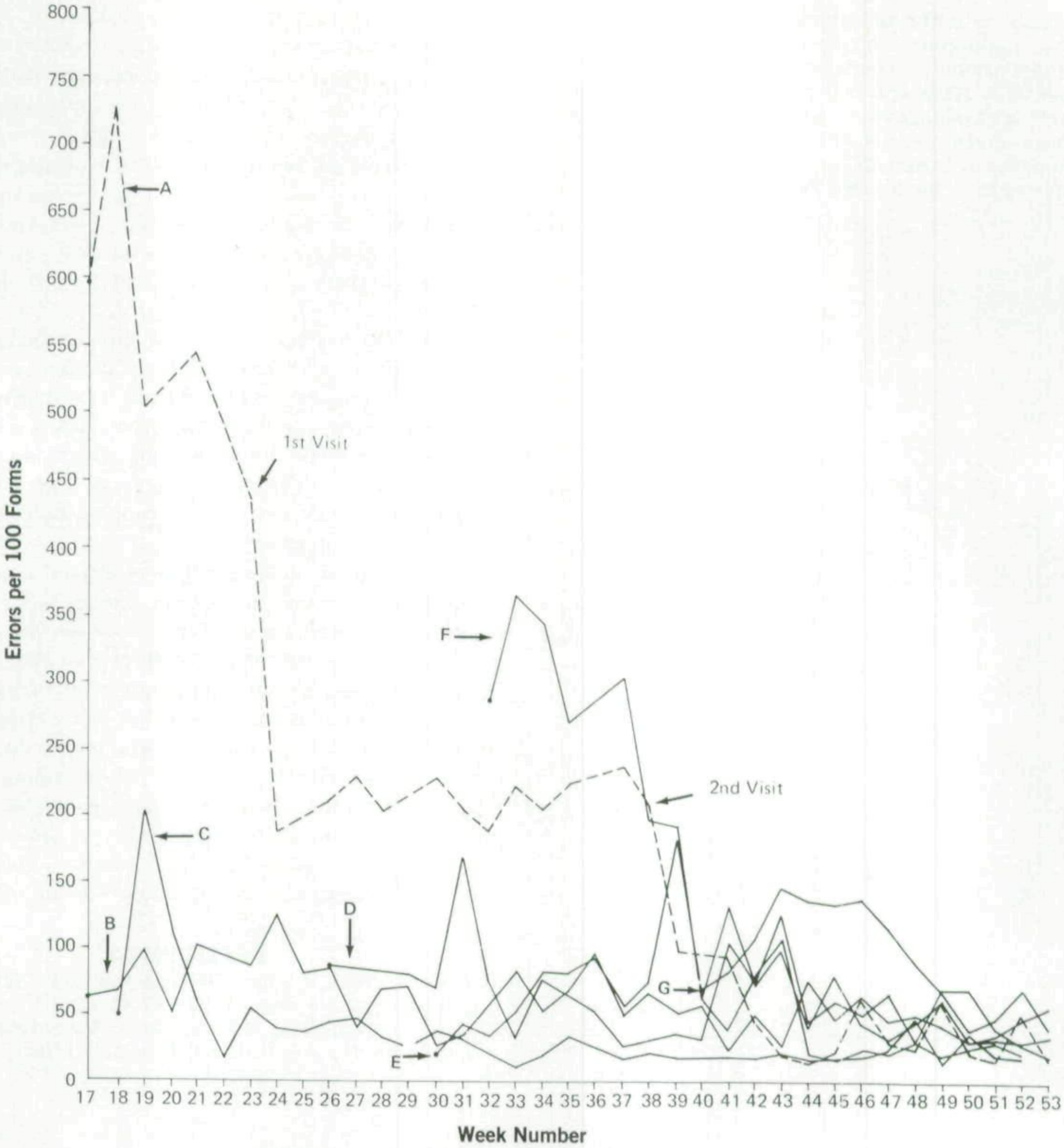

FIGURE 1—Error rates in clinics A–G, STD Study, 1976.

ty of data collection over time. This paper describes such an edit/feedback system and presents a case history of how this sytem was used to improve performance at one data-collection site.

The Sexually Transmitted Disease (STD) monitoring system was established in seven clinics across the United States to obtain data on the incidence and epidemiology of STD. Clinic personnel at each site record data from each patient visit on a one-page form with one carbon copy; the carbon copy, containing no personal identifiers in accordance with the Privacy Act, is sent to the Center for Disease Control for keypunching and analysis.

After keypunching, all data are "reviewed" by the edit routine. This edit procedure checks for missing or incomplete data, some inconsistencies within the data form and for "wild" values in numerical entries. A ten-page computer printout, generated on each batch of data forms for each clinic, shows the frequency of missing or erroneous data for each item on the form and also presents summary statistics on overall error rates per 100 entries and per 100 forms.

Approximately two weeks after a batch of forms reaches CDC, a copy of the edit report is sent to the clinic. If the error rate for that batch is unusually high, or if the clinic is making errors of an unusual nature, a letter is also sent with concrete suggestions for improving data quality. Thus, the edit reports serve to 1) allow the CDC to keep an ongoing record of each clinic's performance, 2) give the clinics rapid feedback on their performance, and 3) provide a diagnostic tool indicating the weakest areas of data collection.

Figure 1 shows the error rates for each of the seven clinics over a span of 36 weeks. (Note that all clinics did not enter the study at the same time.) Although all clinics varied initially, all sites eventually obtained acceptable error rates of approximately 20–30 errors per 100 forms, or about 0.5 errors per 100 entries.

Of particular note is the dramatic improvement of one clinic. Although Clinic A reduced its error rates from more than seven errors per form to five in the first six weeks, this error rate was still significantly higher than that of the other clinics. One author (B.G.) visited the clinic, observed the personnel completing forms, and discussed problems and questions with each staff member involved in form completion. From the meeting came the recommendations that this clinic hire additional clerical staff to monitor form completion and that the CDC provide each clinician with his/her own edit report. After the visit (see arrow 1) the clinic's error rate dropped to about two errors per form. During week 38, B.G. made a second visit, this time with individual clinician edit reports (see arrow 2). The error rate fell dramatically, and this clinic now consistently makes fewer errors than any other.

Without the edit/feedback system, these errors might not have become obvious until data analysis—too late to prevent considerable loss. Although other factors, such as increased staffing, experience with the form, and establishing good rapport between the Division and clinic personnel, obviously contributed to Clinic A's improvement, we feel that the edit/feedback system provided the concrete facts which gave direction to the behavior changes which ultimately led to their success.

Discussion

Data editing and correction systems are a vital part of any data collection system because, as Naus[1] states, ". . . errors are a pervasive part of life." Such mechanisms are best if built into the system at the onset and can thus be tailor-made to fit each particular set of data and circumstances. Editing, for example, can be done by hand or by machine; hand editing is more flexible and, consequently, more expensive.[2] Once errors are found they can either be corrected or the data item(s) can be "thrown out." Corrections can be made either by imputing likely values for data elements from the sample as a whole (see Naus for details) or by actually going back to original documents or, if necessary, to the original respondent—an extremely expensive alternative.

This system provides an inexpensive method for dealing with the problem of data editing and correction, but cannot be applied to all situations. Its main limitations are that only certain types and sources of error are monitored and that errors, once made, cannot be corrected because personal identifiers are not available. The first limitation results because we chose machine, rather than hand, editing to save both money and time—close to 1,000 records are processed each week. Although the edit reports focus on specific errors, in the process of eliminating specific "known" sources of error, the overall quality of data collection will be improved. Because this system does not itself correct errors, improving data quality ultimately depends upon the cooperation and performance of clinic personnel. Thus, such a system would not have the desired effect under certain local conditions, such as understaffing or an uncooperative staff.* Since errors cannot be corrected, this system should not be used unless an initial "teaching" period, in which data may be sacrificed, can be tolerated. It takes time for the "lessons" obtained from the "feedback" system to be "learned."

We feel that we could tolerate both the loss of data during the initial "teaching" period and our ultimate error rate of one error per 100 entries because these data were being used for statistical analysis only and because our cost is only about $2 per record. Consequently, this system may be one answer to the problem of maintaining an acceptable level of data quality under the circumstances of anonymous data collection.

REFERENCES

1. Naus, J. I. Data Quality Control and Editing, New York: Marcel Dekker, Inc, 1975.
2. Pritzker, L., Ogus, J., and Hansen, M. H. Computer editing methods—Some applications and results. Bulletin of the International Statistical Institute, Proceedings 35th session, Belgrade, 41, pp. 442–65, 1965.

*We found, however, that if personnel were less than enthusiastic about the added burden of data collection, personal visits and phone calls could work wonders in improving morale. Many clinic staff members stated that they found their data collection tasks easier once they knew that someone at the CDC was taking a personal interest in their work.