8-18-1978

Statistical Measures for Admission Rates

Paula Diehr
University of Washington, pdiehr@u.washington.edu

Statistical Measures for Admission Rates

I. Introduction.

Frequently in health services literature a reader is asked to compare two groups of people, or two delivery systems, on the basis of hospital admission rates. Table 1 is an example of such rates, using data from the Seattle Prepaid Health Care Project which compared utilization under a prepaid group practice (PPG) to that of a fee-for-service system (FFS). We note that in all but one subcategory (males over 45) the admission rate is higher for the FFS enrollees, and may conclude that admission rates were higher under FFS. The intent of this paper is not to interpret these findings, which have been shown elsewhere [1]. Rather, we wish to illustrate the problems inherent in generalizing data presented in this manner to another setting or time.

[Table 1 about here]

First, note that the number of persons and the length of time they were studied are not shown. This is serious, since we would hesitate to generalize from findings based on small numbers. (The study discussed here followed about 8,000 people for an average of nearly two years, or about 16,000 person-years, which seems a fairly large base.) Second, as is often the case with studies involving admission rates, no measure of variability is shown. We thus do not know how much these rates might change if measured at another time or setting, or if the differences shown are statistically significant. This omission is frequently justified on the basis that large differences, which are of greatest interest, are surely statistically significant because of large sample sizes. Unfortunately, this is not usually the case.

In Table 1 the admission rates varied a good deal by age and sex categories. In Table 2, rates are shown for the two plans by calendar year. Across the four years the PPG rates varied by as much as 28.6 admissions, and

TABLE 1

ANNUALIZED ADMISSION RATES PER THOUSAND ENROLLEES PER YEAR BY AGE, SEX, AND PLAN. (EXCLUDING OBSTETRICAL ADMISSIONS)

SEX / AGE	MALE		FEMALE		TOTAL	
	PPG	FFS	PPG	FFS	PPG	FFS
0- 4	86.3	132.7	84.4	101.7	85.3	117.4
5-19	42.1	68.8	70.9 (32.1)	81.9 (59.8)	56.1 (37.2)	75.7 (64.2)
20-44	53.5	77.9	217.9 (95.1)	282.6 (198.0)	143.7 (76.3)	194.6 (146.4)
45+	243.8	141.1	170.9	239.0 (238.0)	197.8	201.4 (200.8)
TOTAL	72.0	89.5	143.6 (83.3)	182.3 (144.4)	110.1 (78.0)	140.4 (119.6)

PPG = Prepaid Group Practice

FFS = Fee for Service

the FFS by 10.1 admissions. Further, differences between the plans varied from 10.3 to 47.8 admissions, a range of 37.5 admissions per thousand, which is quite large relative to the differences observed between the plans in any year. Thus, even for relatively large patient groups, as in this study, there is a good deal of variation within and between plans, which must be considered in projecting current findings.

[Table 2 about here]

Standard errors, confidence intervals, and significance tests are probably omitted not because they are not needed, but more likely because the nature of the data used for computing admission rates often does not allow the computation of a "correct" measure of variability. However, in these situations, an ad hoc variance estimate would be of use if its properties were known. Here, we address this need, showing appropriate variance estimates in a variety of situations where they exist, and suggesting other methods for data which do not satisfy the requirements of the former methods. Project data are used to illustrate the methods as well as to compare competing approaches where there is not theoretically correct estimate of the variance. To the extent that other data are similar to ours, this empirical evidence should provide guidance for other researchers.

For the purposes of this paper the study group is considered to be a large random sample from an infinite population of similar people with similar access to care, at this or a future time. The study time for each person may be fixed or random. We assume that there is an underlying parameter of interest, namely the mean admission rate per unit of time. We rely on large sample theory and asymptotic normality for some of the results shown, which is probably justifiable in most studies using admission rates.

Notation to be used in the remainder of the paper is shown here. Consider person i, studied for t_i months, who had X_i admissions. The total number of people studied is N.

TABLE 2

ADMISSION RATES[1] PER 1,000 BY CALENDAR YEAR AND PLAN

YEAR[2]	PLAN		RATIO	DIFFERENCE
	PPG	FFS	FFS:PPG	FFS-PPG
1971	99.2	146.8	1.48	47.6
1972	93.5	141.3	1.51	47.8
1973	126.4	136.7	1.08	10.3
1974	122.1	140.0	1.15	17.9
TOTAL	110.1	140.4	1.28	30.3
RANGE	28.6	10.1		37.5

[1] Computed using ratio method, all 8,737 people, all admissions divided by all exposure

[2] "Calendar year" from February 1 through January 31.

For this study population the mean number of admissions and mean study time, $\bar{X}$ and $\bar{t}$, and the sample variances s^2_X and s^2_t are computed in the usual way. In addition we define the total number of admissions $A = \sum_{i=1}^{N} X_i$.

A subgroup of interest is the people with one or more admissions. In our notation there are n such people, with $\bar{x}$ and s^2 referring the sample mean and variance of their admissions. Finally, let f = n/N be the fraction of the study group with one or more admissions. In many studies, not all of this information is available.

The underlying admission rate per month, R, is estimated in some manner from the data available. The estimate, $\hat{R}$, has a variance $V(\hat{R})$. Assuming normality, we may test for the equality of rates R_1 and R_2 using the statistic:

$$Z = \frac{\hat{R}_1 - \hat{R}_2}{\sqrt{V(\hat{R}_1) + V(\hat{R}_2)}} .$$

A 95% confidence interval for the difference in two monthly rates is thus

$$\hat{R}_1 - \hat{R}_2 \pm 1.96\sqrt{V(\hat{R}_1) + V(\hat{R}_2)}$$

Similarly, the 95% confidence interval for a single rate is:

$$\hat{R} \pm 1.96\sqrt{V(\hat{R})}.$$

Although we discuss monthly rates in this paper, different units of time may be used. For instance, the 95% confidence interval for the difference in two "rates per thousand per year" is:

$$12000\ \hat{R}_1 - 12000\ \hat{R}_2 \pm 1.96\ (12000)\sqrt{V(\hat{R}_1) + V(\hat{R}_2)} .$$

The preceding notation is used in Section II, where rate and variance estimates are shown for various situations. Section III discusses patients days, and Section IV summarizes the need for such measures.

II. Estimates, Confidence Intervals, and Significance Testing

In this section we consider several methods of estimating an admission rate (per month) and its associated variance. One well-established estimate, the ratio estimate, is shown and used as a standard for comparison to two other methods which we call person-based ("P") estimators and FEHBPUS ("F") estimators. The ratio method is theoretically appropriate but requires a good deal of control over the data and may be inconvenient to use; P-estimates are not as well understood, and require equal control of the data, but may be more convenient than the ratio estimates; and, the F-method requires only totals, but is based on an assumption which is often injuriously incorrect. Based on empirical and theoretical analysis we conclude that with some adjustments all of these methods have similar performance, and might be used in various circumstances. Special methods are shown for the case when at most one admission is possible.

(1) Ratio Estimator

If the admission rate per month is estimated as:

$$\hat{R} = \frac{\bar{X}}{\bar{t}} \tag{1}$$

a good approximation for the variance of the estimate [2] is

$$V(\hat{R}) \sim \frac{s_X^2 + \hat{R}^2 s_t^2 - 2\hat{R}\, r_{Xt}\, s_X\, s_t}{N\bar{t}^2}\ , \tag{2}$$

where r_{Xt} is the sample correlation coefficient between X and t.

The variance estimate is biased and its distribution skewed, but both of these defects are unimportant for the large sample sizes usually considered. (In particular, the coefficient of variation for $\bar{t}$ and $\bar{X}$ should be below 0.1). Thus, this method is appropriate in most situations.

The ratio method has, however, several disadvantages. The first is the rather stringent data requirements: for each subgroup of people it must be

possible to compute the variance of the exposure, the variance of the number of admissions, and the correlation between exposure and admissions. This is often not possible due to the form of the data. Second, even if all of the data are available, the method is inconvenient if large numbers of rates are to be estimated. In sections (2) and (3) we present methods which, though less theoretically appropriate, do not have the drawbacks of the ratio method. Here we suggest some special cases where the ratio method may be more easily used.

(a) Constant exposure. Note that if all people are observed for the same period of time, t, the variance of t is zero and we have:

$$V(\hat{R}) = \frac{s^2_X}{N\bar{t}^2} \qquad (3)$$

(b) Positive Correlation. Unless the period of observation, t_i, is truncated when a person has an admission, the correlation between X and t will be positive. It is then conservative to delete the final term of (2), and correlation coefficients need not be computed.

(c) If the regression of X on t passes through the origin, then it is conservative to drop the two final terms of (2), leaving the form (3). (In this case, $\hat{R}$ is an unbiased estimate of R.)

(d) If an admission can occur at most once for each person, (X_i = 0 or 1), s^2_X may be computed knowing only the proportion of people with an admission:

$$s^2_X = f\ (1-f).$$

(e) If s^2_X is unknown, it may be calculated using only the people who had an admission, rather than requiring the entire study group. In many studies it would be relatively simple to compute $\bar{x}$ and s^2, the sample mean and variance for the n utilizers only. Then we have:

$$\bar{X} = \frac{n\,\bar{x}}{N} \quad \text{and} \quad s^2_X = (N-1)^{-1}\left[\frac{(N-n)\,n\,\bar{x}^2}{N} + (n-1)s^2\right] \tag{4}$$

or

$$\bar{X} = f\,\bar{x} \quad \text{and} \quad s^2_X \sim f(1-f)\bar{x}^2 + f\,s^2.$$

(f) Finally, if s^2_t is unknown it may be approximated. If the model in (b) above is used, it is conservative to use an extimate of s^2_t which is too large. If the largest and smallest observation times are known, we have:

$$s^2_t \leq \frac{(t_{max} - t_{min})^2}{4} \quad .$$

(2) Person-based estimates (P-estimates)

This method computes an estimate of the admission rate for each person i as

$$\hat{R}_i = X_i / t_i \ .$$

The over-all admission rate is then estimated as the mean of the individual estimates, or

$$\hat{R} = \bar{R} = \frac{1}{N}\sum_{i=1}^{n}\hat{R}_i = \frac{1}{N}\sum_{i=1}^{n}\frac{X_i}{t_i} \quad .$$

The variance of $\bar{R}$ is estimated in the usual way,

$$V_p(\bar{R}) \sim \frac{\Sigma(\hat{R}_i - \bar{R})^2}{N(N-1)} = \frac{s^2_{\hat{R}}}{N}$$

If t_i is constant, the ratio and P-method are equivalent.

The P-estimate is appealing because of its relative simplicity; once $\hat{R}_i$ is computed for each person, the mean and variance of admission rates for subgroups may be estimated directly simply by taking the mean and variance of this new variable, rather than requiring special calculations. In addition, the P-estimate provides an estimated admission rate for each person, which may be used in more detailed, multivariate analysis.

There are several drawbacks to the P-estimate. First, note that the data requirements are equally stringent to those of the ratio estimator, since X_i and t_i must be known simultaneously for each person. Second, the properties of the estimate are not known, and large bias is possible. Third, the rates computed for each person are poor estimators of that person's likelihood of admission, because of the low probability of admission and the short time the person is followed. This variability is greater for persons with low exposure. To deal with this problem we eliminate all people with 12 months or less of exposure, as these should have the highest variability. The person-based estimate using all people is referred to as the P-All estimate; that using only people with greater than one year of exposure is called the PG1 estimate.

Utilization data from the Seattle Prepaid Health Care Project are used to examine the efficiency of the P-all and PG1 methods, relative to the ratio method. Table 3 shows the distribution of the study group by age, sex, and plan, also showing the number with more than 12 months of exposure.

[Table 3 about here]

Table 4 shows the difference in admission rates for the two plans using the ratio, P-all, and PG1 methods. Note that there are large discrepancies between the P-all and the ratio estimates, but that the agreement between the PG1 and ratio estimates is quite good. This suggests that the people with lower exposure were indeed contributing to the variability of the P-all estimate.

[Table 4 about here]

In Table 5, the variance estimates are compared by examining the standard error of the plan difference for the three methods. Under the column labelled "All," including all people, the P-all standard error is seen to be considerably larger than the ratio standard error in most cases, again suggesting that people with low exposure cause high variability. When only people with above 12 months are considered, however, the PG1 and ratio standard errors are quite similar. The greatest discrepancies occur in categories with small numbers which is to

TABLE 3

NUMBER OF PEOPLE USED TO CALCULATE ESTIMATES BY AGE, SEX, AND PLAN: ALL PEOPLE, AND ONLY THOSE WITH MORE THAN TWELVE MONTHS OF EXPOSURE

Sex	Age	All People		>12 Months Exposure	
		PPG	FFS	PPG	FFS
Male	0-4	121	275	80	163
	5-19	471	973	396	814
	20-44	475	909	378	677
	45+	128	378	101	328
	Total	1195	2535	955	1982
Female	0-4	143	262	89	157
	5-19	425	1012	357	863
	20-44	581	1147	459	867
	45+	189	532	157	438
	Total	1338	2953	1062	2325

TABLE 4

PLAN DIFFERENCES IN ADMISSIONS PER THOUSAND BY AGE AND SEX

ESTIMATED BY THE RATIO P-ALL AND PGl METHODS

Sex	Age	All People		>12 Months Exposure	
		Ratio	P-All	Ratio	PGl
Male	0-4	-57.4	-94.0	-60.1	-58.3
	5-19	-26.0	-26.1	-25.5	-21.4
	20-44	-26.6	-43.6	-21.6	-24.4
	45+	110.9	73.9	128.6	98.5
	Total	-16.9	-34.0	-11.8	-16.1
Female	0-4	22.6	76.8	12.1	33.1
	5-19	-29.3	-45.1	-27.8	-29.5
	20-44	-44.8	-27.8	-56.1	-58.6
	45+	-64.5	-103.0	-49.8	-51.2
	Total	-34.4	-31.6	-36.1	-34.5

TABLE 5

STANDARD ERROR[1] OF PLAN DIFFERENCE IN ADMISSIONS PER THOUSAND BY AGE AND SEX: ALL PEOPLE AND ONLY THOSE WITH MORE THAN TWELVE MONTHS OF EXPOSURE

Sex	Age	All		>12 Months Exposure	
		Ratio	P-all	Ratio	PGl
Male	0-4	53	58	60	51
	5-19	11	11	11	12
	20-44	13	16	13	12
	45+	49	62	50	53
	Total	10	12	10	10
Female	0-4	36	71	38	51
	5-19	11	16	12	13
	20-44	25	29	26	28
	45+	43	48	44	45
	Total	14	17	14	15

[1] Standard Error Difference = $\sqrt{s.e._{GH}^2 + s.e._{KCM/BC}^2}$

be expected.

Based on these empirical results, we suggest that PG1 estimates may be used as well as ratio estimators, with some gain in the ease of calculation and further analysis, and without a good deal of loss in terms of variability. If, however, it is not practical to discard persons with one year or less of exposure, the ratio estimator is probably preferable, as rates including these people appear too variable to use.

[Table 5 about here]

(3) The F-method

In the preceding sections we showed two rate estimates which were appropriate in many situations but which required that a record must exist for each person with his exposure and number of admissions. Unfortunately, this is not always the case; frequently only A, the total number of admissions, and $\bar{Nt}$, the total months of exposure, are known. In this situation researchers commonly present no measures of variability. One other approach, however, was suggested in the FEHBPUS study [3]. We shall describe this "F-estimate", show some problems with its use, and provide recommendations for solving these problems.

Here, a person-week of exposure is considered to be a unit of time with a certain probability of admission, but with virtually no probability of two admissions. (This may not be a good model for maternity admissions, where we have frequently seen two in the same week.) If p is the probability of an admission in one week, and t_i is the number of months studied, then $4.33\ t_i$ is the number of person-weeks for person i, and

$$\hat{p} = \frac{\text{number of admissions}}{\text{person-weeks observation}} = \frac{A}{4.33\ \bar{Nt}}$$

or,

$$\hat{R} = 4.33\ \hat{p}$$

because of the change from months to weeks. Note that $\hat{R}$ is identical to the Ratio

estimate; only the variance computation is different. If the number of admissions in a person-week is the outcome of an independent binary trial with probability p, then

$$V(\hat{p}) = \frac{\hat{p}(1-\hat{p})}{4.33\, N\, \bar{t}}$$

and

$$V_F(\hat{R}) = (4.33)^2 V_F(\hat{p}) = \frac{A}{N^2\bar{t}^2}\left(\frac{4.33\, N\bar{t}-A}{4.33\, N\bar{t}}\right) \qquad (5)$$

which may be conservatively simplified to:

$$V_F(\hat{R}) \sim \frac{A}{N^2\bar{t}^2} \qquad (6)$$

As mentioned, the variance estimate of (5) was used to test for significant differences in the FEHBPUS study.

Unfortunately, this model assumes that the person-weeks are independent, when it is intuitively clear that multiple observations on the same person are correlated. As we shall show, this correlation leads to an under-estimate of the variance in many cases and hence to confidence intervals which are too small, and to too many "significant" results. Further, for surgical admissions it may over-estimate the variance. Still, because of its simplicity, the F-estimate would be attractive if the amount by which it underestimates the variance were known. For this reason we examine the relation of the F-variance to the Ratio variance, which is theoretically appropriate, and study its properties empirically.

The two estimates may be compared theoretically in the case where t_i is constant. Here, the ratio variance, V_R, is obtained from equations (3) and (4), and V_F from (5). Their ratio,

$$\frac{V_R}{V_F} = \left[(1-f)\,\bar{x} + \frac{s^2}{\bar{x}}\right] \cdot \left[\frac{4.33\, t}{4.33\, t - f\bar{x}}\right] \sim (1-f)\,\bar{x} + \frac{s^2}{\bar{x}}$$

should be approximately 1.0 if V_R and V_F estimate the same quantity. However, this ratio is clearly a function of the proportion of people with an admission, f, and the readmission rate. If there are no readmissions (as in surgical cases), then $\bar{x}$ is identically one and s^2 is zero; if the period studied is greater than 1 week then the ratio tends to (1-f), showing that V_F is too conservative, or small relative to V_R. If the probability of readmission is high, however, $\bar{x}$ and s^2 will increase, and V_F will be too small, or an anti-conservative estimate of the variance. Thus, for a fixed time period, the appropriateness of the F-method depends on the probability of readmission, which is similar to the independence of person-weeks. We next consider empirical evidence on the performance of the F-method when t may vary, and also study how large $\bar{x}$ and s^2 may be in several situations.

Table 6 shows V_R/V_F, the variance ratio, calculated from self-reported data on the number of admissions in the 12-months preceding project enrollment. Approximately ten percent of the people reported admissions; 16% of those admitted were readmitted at least once, with $\bar{x} = 1.32$. Since the period of observation is constant for these data, the ratio of the two variances reduces to $s^2_X / \bar{X}$, which would be about 1.0 if the number of admissions per year had a Poisson distribution. However, in Table 6 most of the ratios shown are on the order of 2.0, showing that the F-estimate is only half as large as it should have been, and that the Poisson model does not hold.[1]

The groups of Table 6 are heterogeneous, suggesting a large value of s^2_X. When the population was subdivided into 16 subgroups by age, sex, and plan, only four of the variance ratios calculated were as large as 2, showing that the F-estimate is more conservative for homogeneous subgroups,

---

[1]The number of admissions seems rather to have a negative binomial distribution, which results when admissions have a Poisson distribution but individuals do not all have the same parameter value.

TABLE 6

Variance Ratio $\frac{V_R}{V_F}$

Computed from self-reported data* prior to Program Enrollment

Group	n	$\frac{V_R}{V_F}$
All	6083	2.2
Male	2796	2.3
Female	3287	2.1
PPG	1807	1.2
FFS	4276	2.5
Age < 20	2606	1.7
Age >20	3477	2.3

* 12 month recall. All admissions

where $s^2_{\bar{X}}$ is smaller. The variance ratio was very nearly a function of the maximum number of admissions per person observed. Based on our data, it appeared that a useful approximation was:

$$\frac{V_R}{V_F} = h = .8 + .2(X_{maximum}), \qquad (7)$$

which gives h = 1 in the case of no readmissions, when $X_{maximum}$ is 1.0.

[Table 6 about here]

The previous example used self-reported data over a fixed time period, where we had predicted that the F-estimate of the variance would be too low if there were readmissions. We next consider a set of data generated by people enrolled in the project which is provider reported, and where the length of time each person was studied varied from 1 to 48 months. Approximately 13% of the PPG and 16% of the FFS enrollees had an admission. The mean number of admissions per admitted person was about 1.5; and, about 27% of those admitted were readmitted. Since exposure time is not constant, and the information source and access to care are different, we would not be surprised to find very different results for these data; however, Table 7 shows that the ratio of the two variance estimates again appears to be near to 2.0 in most of the comparisons. In addition, when more homogeneous age/sex/plan subgroups were examined it was seen that these, too, had smaller variance ratios and that these, too, were fairly well predicted by equation (7).

[Table 7 about here.]

The similarity in the results of these two very different data sets suggests that it may be safe to use the F-estimate, but that the variance estimate should be adjusted in the case where multiple admissions are possible. We suggest using:

$$\hat{R} = \frac{A}{N\bar{t}} \qquad (8)$$

and

$$V_{adjusted}(\hat{R}) = \frac{hA}{N^2\bar{t}^2}$$

TABLE 7

Variance Ratio $\frac{V_R}{V_F}$

Computed from Provider Reported Data
During Program Enrollment

	Group				
	FFS			PPG	
Total months enrollment	All	Male	Female	Male	Female*
1-12	1.6	1.8	1.4	1.0	1.5
13-24	1.9	1.3	2.0	2.0	2.0
25-36	1.7	1.4	1.7	1.9	3.0
37-48	2.5	2.9	2.2	1.8	1.9
1-48	2.0	1.9	2.0	1.8	2.4
N	5488	2534	2954	1195	478

*First 24 months include obstetrics. "Total" for 25-48 months only.

where it is probably safe to use the value of 2 for h; or if the maximum number of admissions per person is known, one might use equation (7) to choose a better value for h.

Admittedly, these empirical results may vary considerably for other data; however, in the case where only totals are known for admissions and for exposure the F-estimate (adjusted) may be the only way of providing a measure of variability for the rate estimates.

Table 8 shows the standard error of the estimated plan differences calculated by three methods: the ratio method, the F-method, and the F-method adjusted, with h = 2.0. These are shown in homogeneous age/sex categories as well as over-all by sex. Note first that, as discussed above, the F-estimates are much smaller than the ratio estimates due to the number of readmissions. These differences are large enough to have caused spuriously significant test statistics.[1] In the final column, however, we note that the F-adjusted estimates compare quite favorably with the ratio estimates, suggesting that these adjusted variance estimates are useful. Further investigation might have shown that a smaller value of h could be used for the homogeneous subcategories.

[Table 8 about here]

(4) Special Method for "Surgical Admissions."

Occasionally we are interested in a type of admission which can occur at most once for each person. One example is a surgical procedure where an organ is removed. Or, we might be interested only in a person's first admission. Estimation of these admission rates is comparable to the estimation of a proportion, where the "denominator" is the number of persons initially eligible for the procedure. We first assume that the number of people at risk is known.

[1] The data of the FEHBPUS report were re-examined using (8) with h = 2.0. In only two diagnostic categories did differences reported significant fail to be significant with the adjusted variance. It was felt that for these two categories it was virtually impossible to have a readmission, and thus that h = 1 was more appropriate. Thus, if they may be compared to our data, the FEHBPUS results stand.

Table 8

Standard Error of Plan Difference[1] in Admissions per Thousand per Month by Age and Sex: Ratio, F-method, and F-method (adjusted).

Sex	Age	Ratio	FEHBPUS	F-adjusted[2]
Male	0-4	53	36	51
	5-19	11	10	14
	20-44	13	11	16
	45+	49	30	42
	Total	10	8	11
Female	0-4	36	27	38
	5-19	11	10	14
	20-44	25	20	28
	45+	43	28	40
	Total	14	10	14

[1]Standard Error Difference = $\sqrt{V(\hat{R}_1) + V(\hat{R}_2)}$

[2]Adjusted by $\sqrt{h} = \sqrt{2}$

If a person may have at most one admission (X_i = either 0 or 1), variance estimates are readily available in three special cases, again depending on the organization of the data. In the first case, if each person is observed for exactly t months, we are in a binomial situation and can easily estimate the probability of an admission in t months, p_t:

$$\hat{p}_t = A/N \; ; \quad V(\hat{p}_t) \sim \hat{p}_t (1-\hat{p}_t)/N.$$

For other time units, say ct months, if p_t is small we have

$$\hat{p}_{ct} = c\hat{p}_t \; ; \quad V(\hat{p}_{ct}) \sim c^2 V(\hat{p}_t).$$

A second special case occurs when a person with an admission is not observed further, since he is no longer at risk. If all people without an admission are studied exactly m months, then the probability of an admission in one month, p_1, may be estimated as:

$$\hat{p}_1 = A/PMR \; ; \quad V(\hat{p}_1) = \hat{p}_1(1-\hat{p}_1) \; / \; Nm,$$

where PMR is the total "person months at risk." If m = 1 this is the simple binomial case. In other cases, the variance estimate should be multiplied by $[1 + ((m-1)(1+3\, p_1(m-2)/Nm(1-p_1)]$, but this factor may be omitted unless, say, N is below 100, p is above .10, or m is above 24 months.

The above assumes that all people were studied for exactly m months. More commonly, people without an admission are studied for a variable number of months, t_i. In this case, life table methods may be used to show the proportion with an admission over time. This is felt by some [4] to be more appropriate than the previous method since the probability of admission in one month is probably not constant for all of the people studied. In life table terminology people are studied until they are lost to follow-up (drop out), censored (the study period ends before the person has an admission) or they die (have an admission). If we know for each person the month he entered the study, the month of the admission (if not admitted, the month in which he left

the study), available computer programs will estimate the proportion of a cohort with an admission by length of time followed. As an example, Figure 1 was used to compare the proportion of people with one or more admissions in the two systems over time, showing that consistently more people in the FFS system had admissions. Standard errors are available for each point on the curve, which allowed us to show that the curves were significantly different except at month 1. An introduction to life table methods may be seen in [5].

[Figure 1 about here]

Finally, we consider the problem of an unknown denominator. In "surgical" admissions the correct denominator for a rate is the number of persons eligible for a certain procedure. This number is often unknown and may be substantially smaller than the number of persons available for observation, depending on the number who had the procedure prior to the study period. For example, as many as 50% of women over age 50 may not have a uterus and so are not eligible for a hysterectomy; for older persons the probability of having intact tonsils or an appendix may also be quite low. Since frequently occurring operations are the most often studied, accurate knowledge of the denominator may be an important problem.

Here, we note that two groups may be compared, even though the actual rates cannot be estimated. We need know only the approximate ratio, K, of eligibles in the 2 groups. If the groups under consideration are comparable on other measures, K may be taken as the ratio of the number of people in the two study groups (or, of the person-months of exposure in the groups if these differ). If the groups are not similar, it may be possible to find subclasses of both groups (e.g., by age, sex, health status) where the assertion can be supported that the proportion of people at risk, though unknown, is approximately the same for both groups. Let K = the ratio of the number of study persons (or person months) in the two groups:

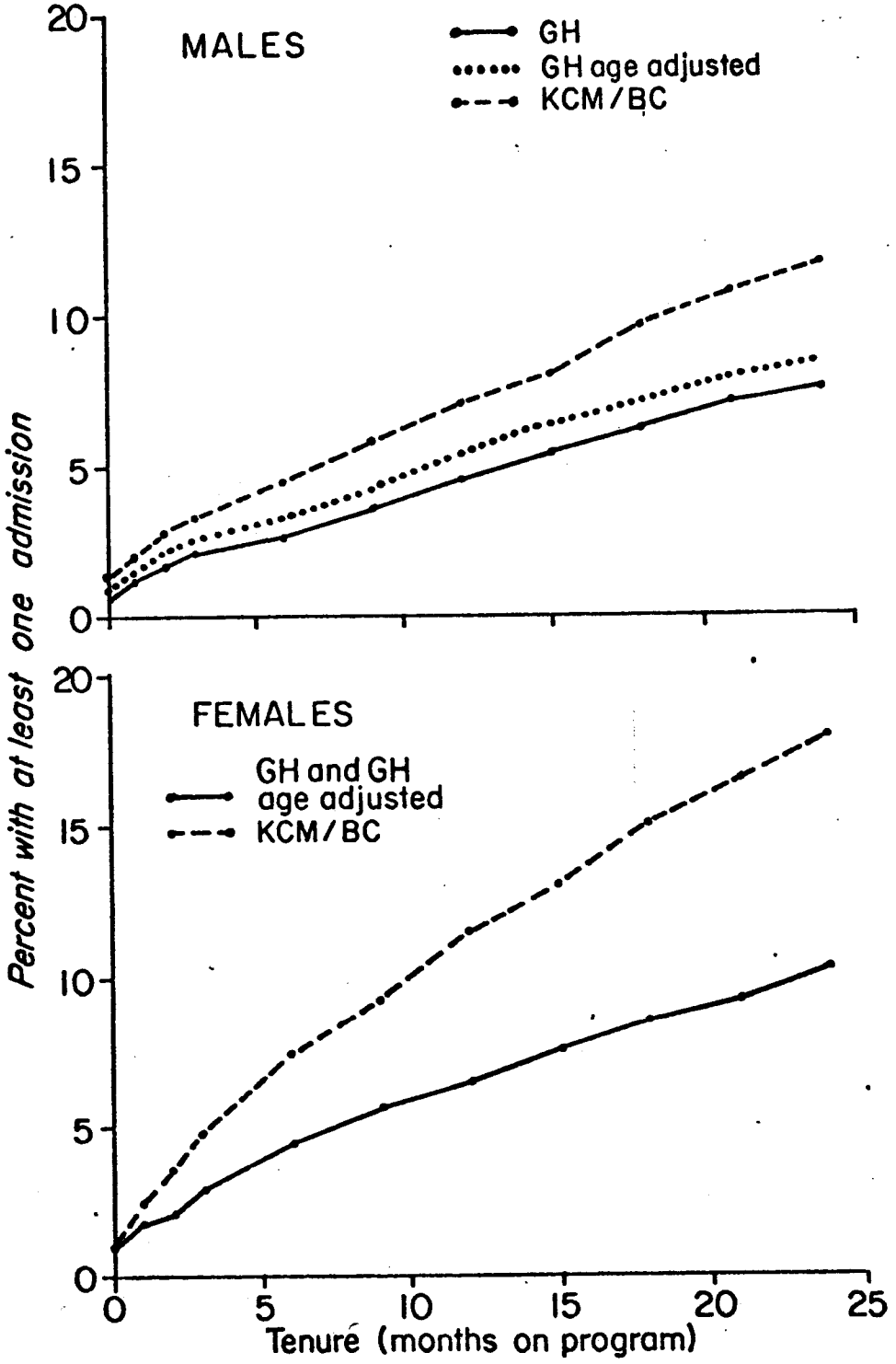

FIGURE 1

PERCENT OF ENROLLEES WITH ONE OR MORE ADMISSIONS BY SEX, PLAN, AND TENURE

$$K = \frac{n_1}{n_2} \qquad \text{or} \qquad K = \frac{n_1 \bar{t}_1}{n_2 \bar{t}_2}$$

Let A_1 and A_2 be the number of admissions in groups 1 and 2, respectively.

We wish to estimate the difference between the (unknown) rates. One possible measure is a "95% prediction interval," where the number of admissions in one group is used to predict a probable range for the number of admissions in the other group. For example, given A_1, we are 95% confident that A_2 will lie in the interval

$$\frac{A_1}{K} \pm \frac{1.96}{K} \sqrt{1 + K} \sqrt{A_1}$$

if the two groups have the same admission rate. If A_2 is outside this prediction interval, this is equivalent to rejecting (at the .05 level) the null hypothesis of equal admission rates. In addition, the prediction interval provides a measure of the number of "excess" admissions in the second group. Since K, the ratio of eligibles in the two groups, is known only approximately, sensitivity analysis should be used to investigate whether small changes in K would change the conclusions. This is illustrated in the following example.

[Table 9 about here]

Table 9 shows the admissions for the two plans for four types of surgery: appendectomy, cholecystectomy, hysterectomy, and tonsillectomy. The number of people in the population who were "at risk," still having the relevant organ, was unknown. For cholecystectomy the denominator is probably near to the number of persons observed, since this procedure is not frequently performed at young ages. For T&A, appendectomy, and hysterectomy, however, the number of people at risk may be considerably smaller than the number of persons under observation, since many of the study subjects may have had the procedure prior to this study. Prediction intervals were computed to estimate plan differences.

The enrollment ratio (K) between the two plans is 2.2:1, suggesting that

Table 9

95% Prediction Intervals for the Number of PPG Admissions in Four Surgical Categories

	Appendectomy	Cholescystectomy	Hysterectomy	Tonsillectomy
Number of Admissions				
FFS A_1	12	22	44	31
PPG A_2	4	6	3	5
95% Prediction Interval*				
K = 2.0	(0.1,11.9)	(3.0,19.0)	(10.7,33.3)	(6.0,25.0)
K = 2.2	(0,11.0)	(2.5,17.5)	(9.4,30.6)	(5.2,23.0)
K = 2.4	(0,10.2)	(2.1,16.2)	(8.3,28.3)	(4.5,21.3)

* $\frac{A_1}{K} \pm \frac{1.96}{K} \sqrt{1+K} \sqrt{A_1}$

there should be twice as many procedures performed in the FFS group as for the PPG. The observed number of cases for the PPG is thus considerably lower than would have been expected for all four categories. The size or significance of this difference is next of interest.

Prediction Intervals are shown for the number of PPG admissions using K = 2.0, 2.2, and 2.4 for sensitivity analysis, as suggested. Note that for K = 2.2, the most likely ratio of eligibles, A_2 falls within the prediction interval for appendectomy and cholecystectomy, but is lower than expected for hysterectomy and tonsillectomy. We thus have significant difference for the final two categories. Note, however, that the number of PPG hysterctomies is 6.4 visits below the prediction interval, but that the number of tonsillectomies is only 0.2 from the interval endpoint. When other values of K are used, the results in the first three categories remain unchanged, but for tonsillectomies the number of admissions for the PPG would fall within the prediction interval if the true ratio of eligibles were 2.4:1, rather than the estimated 2.2:1. We thus have a feeling for the size of the differences in the admitting rates in the plans, even though we have not estimated the rates themselves.

(5) Hypothesis Testing

The preceding material on rate estimates and their variances makes it possible to construct confidence intervals and test hypotheses, as noted in the introduction. Since, however, research emphasis is initially on testing rather than estimation, we here provide two "quick" significance tests which can be used in a variety of situations. We consider here that there are two groups whose rates are to be compared, that they have A_1 and A_2 admissions respectively, and that the relative sample size is K, as in the previous section. Using this notation, the F-method provides a test statistic:

$$Z = \frac{A_1 - K A_2}{\sqrt{h}\sqrt{A_1 + K^2 A_2}} \qquad (9)$$

where h can be taken as 2.0 unless there is evidence for a smaller value, from the material in section (3). Z is approximately normally distributed with mean zero and variance one, and the standard critical values of 1.645 and 1.96 apply here for one and two-tailed tests, respectively.

A second test statistic arises from the preceding section on surgical admissions. The use of K is particularly valuable here since, as was pointed out earlier, the actual numbers at risk may not be known in this case, but the ratio may be assumed to be known. If each person has at most one admission, then:

$$X_1^2 = \frac{(A_1 - KA_2)^2}{K(A_1 + A_2)} \qquad (10)$$

has a chi square distribution with 1 degree of freedom if both $(A_1 + A_2)/(1 + K)$ and $(A_1 + A_2)/(1 + K^{-1})$ are greater than 5. In the surgical case, where h = 1, statistics (9) and (10) are virtually the same, differing in whether the estimate of the variance for the difference in two proportions should be pooled or summed.

III. Patient Days

Measures of variability may be needed for patient days per thousand, or the number of hospitalized days divided by person years of exposure. The first two methods of section II, the ratio and person-based estimators, are also applicable to patient days data. As before, we compared the methods empirically on the Seattle Prepaid Project Data. In tables 10-12 the results are shown for the Ratio ., the P-all, and the PG1 estimates for various age, sex, and plan categories: rates for both plans, their difference D, the Standard error of the difference, the Z-statistic testing the differences, and 95% confidence intervals are shown. The general conclusion to be reached is that the Ratio and the PG1 estimators have similar properties but that, if persons with under a year of exposure must be included, the P-all estimate is much too variable.

There is no counterpart to the F-estimate for patient days. If it is not feasible to compute Ratio or PG1 estimates we suggest restricting the study to length of stay per admission rather than per person-year, since measures of variability are readily obtainable for this comparison, and analysis of the admission rates and length of stay separately are essentially equivalent to a study of patient days.

TABLE 10

PATIENT DAYS PER THOUSAND BY AGE WITH THREE ESTIMATION METHODS: MALES

Age	Method	KCM/BC Rate	GH Rate	D	Se (D)	z Statistic	95% Confidence Interval	
0-4	PGl	1220.2	311.2	909.0	464.0	1.96	1818.0	.4
	P-all	2101.8	329.8	1772.0	1176.0	1.50	4076.0	-532.0
	R	1202.6	345.2	857.4	342.05	2.50	1527.81	186.96
5-19	PGl	168.6	128.8	39.8	45.0	.88	128.0	-49.0
	P-all	162.0	112.5	49.5	39.0	1.25	127.0	-28.0
	R	184.4	160.8	23.63	55.58	.42	132.58	-85.32
20-44	PGl	271.6	250.3	21.3	90.0	.24	198.0	-155.0
	P-all	298.8	208.2	90.6	88.0	1.02	263.0	-82.0
	R	301.8	281.3	20.51	97.39	.21	211.41	-170.38
45+	PGl	785.1	1552.5	-767.4	453.0	-1.69	121.0	-1656.0
	P-all	1097.1	1585.2	-488.1	455.0	-1.09	384.0	-1360.0
	R		1574.4	-694.86	403.65	-1.72	96.29	6.01
Total	PGl	392.4	342.7	49.7	77.0	.64	200.0	-101.0
	P-all	561.1	330.3	230.8	144.0	1.60	512.0	-51.0
	R	422.6	377.8	44.82	72.25	.62	186.43	-96.78

TABLE 11

PATIENT DAYS PER THOUSAND BY AGE WITH THREE ESTIMATION METHODS: FEMALES

Age	Method	KCM/BC Rate	GH Rate	D	Se (D)	z Statistic	95% Confidence Interval	
0-4	PGl	304.5	561.8	-257.0	388.0	-.66	503.0	-1018.0
	P-all	2237.4	2000.0	237.4	1755.0	.13	3678.0	-3203.0
	R	661.7	777.9	-116.21	428.84	-.27	724.31	-956.74
5-19	PGl	221.8	79.6	142.0	45.0	3.18	229.8	55.0
	P-all	276.6	89.4	187.2	64.0	2.90	313.0	61.0
	R	208.2	91.9	116.34	37.44	3.10	189.72	42.96
20-44	PGl	1205.5	783.9	422.0	217.0	1.94	847.0	-4.2
	P-all	1139.3	771.8	367.5	189.0	1.94	738.0	-3.3
	R	1276.3	762.7	513.64	176.49	2.91	859.57	167.71
45+	PGl	1876.5	1128.7	748.0	421.0	1.78	1573.0	-78.0
	P-all	2371.8	1011.6	1360.2	474.0	2.86	2290.0	430.0
	R	2124.5	1175.3	966.02	426.27	2.26	1801.52	130.52
Total	PGl	905.9	579.5	326.0	122.0	2.67	566.0	86.0
	P-all	1162.8	720.2	442.6	205.0	2.16	844.0	41.0
	R	1008.8	602.4	406.38	111.93	3.63	625.77	187.0

TABLE 12

PATIENT DAYS PER THOUSAND BY AGE WITH THREE ESTIMATION METHODS:

FEMALES EXCLUDING OBSTETRICAL

Age	Method	KCM/BC Rate	GH Rate	D	Se (D)	z Statistic	95% Confidence Interval	
0-4	PGl	304.5	561.8	-257.1	388.0	-.67	503.0	-1018.0
	P-all	1115.6	1883.9	-768.3	714.4	-1.07	632.0	-2168.0
	R	661.7	777.9	-116.21	428.84	-.27	724.31	-956.74
5-19	PGl	163.0	40.3	123.0	39.0	3.12	199.7	45.7
	P-all	651.2	321.9	329.3	134.9	2.44	594.0	65.0
	R	152.4	41.0	111.4	30.13	3.69	170.46	52.34
20-44	PGl	934.7	442.0	493.0	195.0	2.53	874.1	111.2
	P-all	1664.4	903.1	761.3	206.4	3.68	1166.0	357.0
	R	1011.6	391.6	619.94	164.46	3.76	942.3	297.58
45+	PGl	1873.7	1128.7	745.0	421.0	1.76	1570.3	-80.3
	P-all	2615.2	1596.8	1018.4	476.8	2.13	1953.0	84.0
	R	2120.9	1175.3	945.57	431.78	2.18	1791.87	99.26
Total	PGl	782.6	418.5	364.0	115.0	3.17	589.1	139.1
	P-all	1439.5	921.3	518.2	145.1	3.57	802.0	234.0
	R	890.1	433.9	456.15	108.83	4.19	669.47	242.83

IV. Wise Sayings and Conclusions

We noted in the introduction to this paper that admission rates were often presented and interpreted without the benefit of statistical testing and estimation procedures. This is, of course, appropriate if one wishes only to assess past performance. However, since one usually wants to generalize to other cases, or to these groups at a future time, some inferential techniques should be used. In the body of the paper, we suggested several ways of estimating variances, even in situations where they are not usually calculated. We now consider the importance of their application.

Table 1 presented annualized admission rates in the usual form: admissions per thousand, with no associated measure of variability. With the exception of the oldest male category, the admission rate was higher in the FFS system than in the PPG system. We later noted that these rates are based on a large number of people and months, which strengthens this conclusion. We next consider statistical significance and importance, using the measures of variability developed above.

Table 13 shows the plan differences (ratio estimate), which appear substantial, on the order of 17 admissions per thousand for males and 34 admissions per thousand for females. The second column of the table shows the standard errors of the plan differences, which also seem rather large. The third column shows the ratio of the first two, labelled Z, which is normally distributed with mean zero and variance 1 if there is no plan difference. For a two-tailed test at the .05 level, only five of the ten differences are significantly different from zero. Further, two of the three largest differences (males 0-4, females over 45) are not significantly different from zero, despite the hopeful rule of thumb presented in section I.

Even where the differences are significant, we can not be sure that they are large or important because of the large confidence intervals afforded by

these data. The two shortest intervals estimate the underlying difference for males aged 5-19 as between 4.4 and 47.6 admissions per year, and for females aged 5-19 the difference is 7.7 to 50.9 admissions per thousand. We may feel confident that FFS has a higher admission rate than PPGP for these two categories, but the difference may still be too small for policy decisions.

We do not claim here that admission rates are not higher in the FFS group, since there is a large literature supporting the opposite contention. Rather, we note that even with large numbers of people and evident differences, statistical significance and importance are difficult to show, and can be demonstrated conclusively only with larger samples. Thus, the inherent variability in admission rates is very large, and casual readers should be deterred by this fact, and demand measures of this variability before making conclusions or decisions based on such data.

TABLE 13

TESTING AND ESTIMATION OF DIFFERENCES IN ADMISSION RATES

AGE/SEX	PLAN DIFFERENCE (FFS—PPG)[1]	STANDARD ERROR OF RAW DIFFERENCE[1]	Z	95% CI FOR DIFFERENCE
Male				
0- 4	+ 57.4	53	+1.08	(- 46.5 161.3)
5-19	+ 26.0	11	+2.36	(4.4 47.6)
20-44	+ 26.6	13	+2.05	(1.1 52.1)
45+	-110.9	49	-2.26	(-206.9 -14.9)
TOTAL	+ 16.9	10	+1.69	(- 2.7 36.5)
Female				
0- 4	- 22.6	36	- .63	(- 93.2 48.0)
5-19	+ 29.3	11	+2.66	(7.7 50.9)
20-44	+ 44.8	25	+1.79	(4.2 93.8)
45+	+ 64.5	43	+1.50	(- 19.8 148.8)
TOTAL	+ 34.4	14	+2.46	(7.0 61.8)

[1] Admissions per thousand per year

References

1. Diehr, et al.: Utilization: Ambulatory and hospital. In the Seattle Prepaid Health Care Project: Comparison of Health Services Delivery, Chapter II, W. C. Richardson, et a. (Eds.) National Technical Information Service, PB #267488-SET, 1976.

2. Cochran, William G.: Sampling Techniques Second edition, p.31, John Wiley and Sons, Inc., New York 1963.

3. Riedel, D. C., Walden, D. C., et al.: Federal Employees Health Benefit Plan Utilization Study. DHEW Publication No. (HRA) 75-3125, 1975.

4. Sheps, M. C.: On the Person Years Concept in Epidemiology and Demography. Milbank Memorial Fund Quarterly, 44, 69-91, 1966.

5. Armitage, P.: Statistical Methods in Medical Research, John Wiley and Sons, New York 1973.