In Search of a Solution to the Primary Care Denominator Problem

Daniel C. Cherkin, PhD, Alfred O. Berg, MD, MPH, and William R. Phillips, MD, MPH
Seattle, Washington

Before morbidity or utilization rates can be legitimately compared across practices or within a practice over time, it is necessary to know the number and age-sex distribution of individuals who are served by the practice in question. Estimating this "population at risk" has been referred to as the "denominator problem." Although a variety of methods for estimating practice denominators were proposed more than six years ago, none has been shown to be satisfactory, and no new ones have materialized. One method, however, has never been evaluated, and evidence is presented which suggests that this method may be capable of providing satisfactory estimates of practice denominators.

Data from the US Health Interview Survey, the British National Morbidity Study, and other sources suggest that it may be possible to derive age- and sex-specific correction factors which, when applied to the age-sex distribution of visiting patients, would provide reasonable estimates on the practice denominator. This "correction factor method" would require only that practices maintain age-sex registers of their visiting patients. Further investigation is required before this method can be considered a satisfactory means of estimating practice denominators.

Rates are the hallmark of epidemiology, for they form the basis of comparisons between populations. "Floating numerators" are anathema, for they cannot be interpreted.
G. Rose and D.J.P. Barker[1]

Although conscientiously gathered numerator data deriving from patients' visits to primary care practices are becoming more common, the number of persons served by each of these practices remains unknown in most North American practices. Such practice denominators are essential for calculating rates, which are in turn necessary for valid comparisons of the relative frequencies of events occurring in different practices. This lack of known denominators continues to limit the usefulness of numerator data from primary care practices.

The most commonly stated need for knowing practice denominators has been for the calculation and comparison of morbidity rates.[2,3] Others have suggested that the ability to calculate rates would be useful for planning, evaluation, and research,[4]

Portions of this paper were presented at the Ninth Annual Meeting of the North American Primary Care Research Group, Incline Village, Nevada, March 18, 1981. From the Department of Family Medicine, School of Medicine, University of Washington, Seattle, Washington. Requests for reprints should be addressed to Dr. Daniel Cherkin, Department of Family Medicine, Research Section JD-13, University of Washington, Seattle, WA 98195.

0094-3509/82/020301-09$02.25

and for the comparison of "rates of problems, procedures, outcomes, encounters, and other measures."[5] Hence, it is believed that the availability of practice denominators would permit a wide variety of potentially valuable comparisons that are not now possible. What is needed, then, is a practical method for estimating practice denominators. This article discusses methods that have been proposed in the past and identifies one method that merits further consideration.

Literature Review

All of the methods that have been proposed to date for estimating the denominator were enumerated by Garson in 1976.[3] The current status of each of these methods is summarized below.

Census Method

In the rare situations in which a medically isolated and well-defined community is served by a single practice, the denominator could be estimated precisely by a community census. Unfortunately, such communities are uncommon, and even where they exist, obtaining accurate information through a census may be difficult and expensive.

Registration by Intent Method

Pioneered in Canada, this approach requires that patients inform the practice about which members of their family consider the practice to be their regular source of care. There are a variety of drawbacks to this method. It seems likely that many persons, particularly if they are healthy and single, might never register because they have no need to seek care, even though they would register at the practice were they to become ill. On the other hand, persons filling out the registration form might indicate that other members of their family would be seeking care at a family practice when, in fact, family members may be receiving primary care elsewhere (eg, from a pediatrician, internist, or other family physician). A recent study using the "active family"* concept of registration found that only 53 percent of registered patients were seen in a practice during a 12-month period,[7] a figure well below the 68 percent with an annual visit found in British general practices[8] and the 77 percent found in a health maintenance organization in the United States.[9] This suggests that the denominator may have been overestimated through the inclusion of individuals not actually receiving primary care from the practice. Finally, Bass warns that the process of registration is time consuming and expensive as well as of questionable accuracy.[2]

De Facto Registration Method

This method requires determining the number of individuals who have visited a practice one or more times during a specified time interval (eg, one year), usually accomplished by maintaining an age-sex register.[10] Although this method would be simple to implement, the true population served by the practice would be considerably larger than the number who visit during a year as long as obviously transient patients were not registered. Despite the deficiency of this method when used alone, it may provide a useful basis on which to build (see Utilization Correction Factor Method below).

Indicator Disease Method

This approach would require the identification of at least one disease that occurs with relatively constant frequency across all patient populations

*An "active family" is defined as a registered family containing at least one member who has received health care at least one time in the preceding two years.[6]

and for which a relatively constant proportion of individuals seek medical care. If such a disease were identified, the practice denominator could be estimated from the number of cases of the disease seen by the practice. For example, if it were known that 48 of every 1,000 people seek care for a particular disease each year, then a practice which recorded 96 cases of the disease during a year would be serving an estimated 2,000 people. In order for this approach to provide reasonable estimates, it would be necessary for the disease to be fairly common (ie, a rare disease might never be seen in a practice during a year), and it must be readily and uniformly diagnosed and recorded. These requirements, however, are somewhat academic, since the incidence of virtually all diseases varies by age, sex, race, and other factors, and a candidate for this indicator disease remains unidentified.

Episodes of Illness Method

Working with data from the British Second National Morbidity Study, Kilpatrick noted that "the frequency of episodes of illness which 315,000 people brought to their doctors' attention in one year" followed a negative binomial distribution.[11] Kilpatrick suggested that knowledge of the distribution of episodes of illness for patients visiting a practice could be used to estimate the number of people who were served by the practice but who had not visited. This could be accomplished by fitting a truncated negative binomial distribution to frequency of episode data for visiting patients, and then using the distribution's parameters derived in this manner to estimate the number of nonvisiting patients (ie, those with zero visits during the year). The total practice denominator could then be calculated by adding the estimated number of nonvisiting patients to the number of patients who had one or more visits during the year. Hence, if community practices in North America were able to determine the annual number of episodes of illness for each of their patients, this method could be used to derive estimates of the practice denominators.

Unfortunately, evaluations of this approach in North American practices have been discouraging. After extensive investigation of this method using data from practices in New York, Pennsylvania, and Virginia, Kilpatrick concluded that "a practice population is a nebulous concept: it is undefinable and cannot consistently be estimated," and that "under the present health care system, we cannot use encounter records to do population based research."[12] Hence, the episode of illness approach, at least when used in conjunction with the negative binomial distribution, has not proven useful.

Utilization Correction Factor Method

This approach assumes that relatively constant proportions of the populations served by primary care physicians will visit these physicians during a particular time period.[3] Therefore, if one knows the proportion of a population in a region that has visited a primary physician in the past year (as in some Canadian provinces where utilization statistics are available from government operated, prepaid medical care insurance plans), then the number served by any particular physician could be calculated. For example, if it were known that 75 percent of individuals in a region visited a primary care physician during a particular year, a general practitioner in the region who saw 750 individuals during that year would be serving an estimated 1,000 patients. The correction factor in this example would be the number which, when multiplied by 750, increases it by 250 to 1,000, ie, 1.33. In general, the denominator would be estimated by multiplying the number of individuals visiting during a time interval by an appropriate correction factor.* Since certain age and sex groups are more likely to visit than others, this approach could be refined by determining appropriate utilization correction factors for several age-sex categories. Although Garson first proposed this method more than five years ago,[3] it has never been evaluated.

*The correction factor is, in fact, simply the reciprocal of the estimated proportion visiting. Hence, in the above example 1/.75 = 1.33.

In summary, although several methods for estimating practice denominators have been proposed, none have been shown to be capable of providing satisfactory estimates of the denominator. There is evidence, however, that the correction factor method may represent a practical and satisfactory approach to estimating the denominator. The reasons for this belief are discussed below.

Rationale for Consideration of Correction Factor Method

The correction factor method requires two types of information: (1) an accounting by age and sex of the number of individuals who visited the practice during a particular time period (eg, one or two years) and (2) a set of age- and sex-specific correction factors that relate the number of individuals in particular age-sex groups who visited a practice during the time period to the number who were "at risk" of visiting during the time period.* The age-sex accounting is precisely the product of an annually updated age-sex register, something that physicians who carefully record morbidity data would be capable of and likely to be maintaining. Identification of widely applicable correction factors is not a simple matter, however, and represents the basic challenge to the success of this method.

The fundamental assumption underlying the correction factor method is that within any age-sex category, the proportion of individuals served by a practice who make one or more visits during a time interval does not vary substantially from practice to practice. If this assumption were shown to be true, the age-sex composition of the population served by the practice, ie, age-sex specific denominators, could be estimated directly from age-sex register data. For example, if it were known that 70 percent of male patients aged 25 to 44 years visit a primary care physician at least once during a 12-month period, then a primary care practice that encountered 70 male patients aged 25 to 44 years during a one-year period would be serving an estimated 100 men in this age group. While it would be very difficult to test this fundamental assumption for North American practices directly, existing data from a variety of sources suggest that the variation in the annual percent of individuals who visit their primary care physician may be small, particularly after adjusting for differences in age-sex composition.

The US Health Interview Survey (HIS) is possibly the most widely available source of data on variation in the percent of individuals who visit a physician during a particular time interval.[13] The HIS is a continuing nationwide survey conducted by household interviews of a probability sample of households of the civilian noninstitutionalized population of the United States. It gathers data on illness, injury, disability, and use of hospital and medical services. In 1977, 41,000 households with 111,000 persons were sampled, and 97 percent of the households were successfully interviewed. Of greatest relevance are the HIS data on the percent of individuals in various population subgroups who had visited a physician within the past 12 or 24 months. HIS data on the percent reporting one or more visits to a physician within the previous 24 months for several population subgroups are shown in Table 1.[13-15] Although there is some variability in percent visiting across subpopulations, it is the relative uniformity that seems most noteworthy.

Overall, 86 percent of the people in the United States reported they had visited a physician within the previous 24 months. As would be expected, there was some variation in the proportion visiting by age and sex. Relatively small differences, however, were noted between blacks and whites and among those in various income groups. Furthermore, visit rates by place of residence in terms of SMSA (Standard Metropolitan Statistical Areas) and non-SMSA (nonfarm) were very similar, as was the proportion visiting in each of the four regions of the country. Not surprisingly, those with some activity limitation were more likely to have visited than were those who had none. The most variation was noted for the education level of the family head of household, ranging from 80 percent visiting in the previous two years for those with least education, up to 91 percent for those with the

*The term "at risk" is used in its epidemiological sense and refers to those individuals who would seek primary care services from a particular practice were they to seek primary care services at all.

Table 1. Percent of Persons Reporting That They Had Visited a Physician Within Previous 24 Months, According to Selected Characteristics

Characteristics	Percent Reporting Physician Visit in Previous 24 Months
Sex (1978)	
Female	89.3
Male	83.0
Age, female (1978)	
Under 17 years	89.1
17-24 years	91.3
25-44 years	91.6
45-64 years	86.6
65 years and over	87.7
Age, male (1978)	
Under 17 years	89.5
17-24 years	80.8
25-44 years	78.8
45-64 years	80.5
65 years and over	83.7
Race (1977)	
Black	86.6
White	86.4
Family income (1977)	
Less than $5,000	86.3
$5,000-$9,999	85.2
$10,000-$14,999	86.5
$15,000-$24,999	87.3
$25,000 or more	88.7
Education of head of family (1975)	
Less than 9 years	80.3
9-11 years	84.0
12 years	87.2
13-15 years	89.4
16 or more years	91.0
Activity limitation (1975)	
Unable to carry on major activity	93.2
Limited in amount or kind of major activity	92.5
Limited, but not in major activity	90.9
Not limited in activity	85.0
Geographic region (1977)	
Northeast	87.1
North Central	86.3
South	86.2
West	86.1
Place of residence (1975)	
SMSA	86.5
Outside SMSA	
Nonfarm	85.5
Farm	80.2

Data for 1975 are from National Center for Health Statistics,[13] for 1977 from the Public Health service,[14] and for 1978 from National Center for Health Statistics.[15]

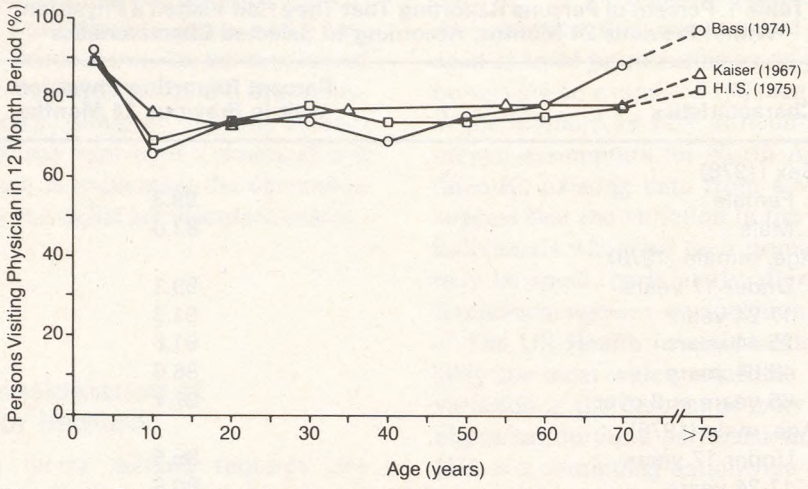

Figure 1. Percentage of persons visiting physician in 12-month period reported in three data sources (Bass,[2] Greenlick et al,[9] and Health Interview Survey[13])

most. Similar utilization data for different occupational groups and across states also evidence remarkably little variation.[16,17] It should be kept in mind, however, that the data from the Health Interview Survey are for a random sample of the US population. The question remains: Does this type of data mirror the visit rates in physician practices?

There are data from North American practices with essentially known denominators which, when juxtaposed to the HIS data for proportion visiting, suggest that the HIS data may indeed reflect practice data. It can be seen in Figure 1, that when the estimates for the percent visiting within the previous 12 months from the Health Interview Survey are compared with the percent visiting reported by a Canadian practice with a carefully registered patient population[2] and by the Kaiser-Permanente Medical Care Program in Portland,[9] there is remarkable similarity across all age groups, except possibly the oldest.

Finally, the British Second National Morbidity Study provides information on the variation in the percent of patients visiting 60 general practices in Great Britain.[8] Under the British National Health Service, persons are assigned to the care of specific general practitioners; hence, practice denominators are known.* Despite some curious outliers, 90 percent of the 60 practices providing data saw between 61 percent and 74 percent of their patients, and 50 percent of the practices saw between 67 percent and 72 percent of their patients during a 12-month period. It is believed that had a 24-month interval been examined, the interpractice variation would have been considerably lower (see Discussion). In addition, adjustment for age and sex differences among the practices might decrease the variation in percent visiting.

*There is some question about how accurately practice lists reflect the true practice denominators. Inflation rates as high as 22 percent have been reported, though over 90 percent of the practices participating in the 1971 National Morbidity Study were found to experience less than 5 percent inflation.[18,19]

Table 2. Annual Percentage of Registered Persons Visiting British General Practices by Practice Characteristics

Characteristics	Practices Reporting	Percent of Patients Visiting
Overall	55	67.7
Region		
North	21	68.9
Midlands and Wales	12	67.7
South	22	67.0
Urban/Rural		
Urban	40	67.2
Rural	15	69.6
Physicians in practice		
1	21	69.7
2	13	67.0
3	10	65.9
4 and over	11	68.2
Average practice population		
Under 1,900	7	79.2
1,900-2,499	19	67.7
2,500-2,999	10	68.0
3,000 and over	19	67.3
Physician's age (practice mean)		
Under 40 years	11	69.4
40-49 years	32	68.0
50 years and over	11	65.9

Data from Office of Population and Census Surveys[8]

When British practices with similar characteristics are grouped, the variation in the percent visiting all but disappears (Table 2). Hence, at least in Great Britain, the percent of registered patients who make one or more visits in a year appears to be unrelated to region of country, urban vs rural location, number of physicians in the practice, average age of the physicians, or practice size (with the notable exception of the very smallest practices). The percent visiting was also found to be unrelated to practices' ratio of nurses to population, access to hospital beds, existence of special clinics run by the practice, or access to facilities outside the practice.[8]

Discussion

The evidence presented above, though suggestive, does not prove that the correction factor method can satisfactorily estimate practice denominators. Two types of error could render the correction factor estimates of the denominator unsatisfactory for particular practices: random and systematic error.

Random errors result from chance fluctuations in the number visiting. For example, even a stable practice would be expected to experience some variation in the number of patients visiting from month to month or year to year. The variance of

the correction factor method estimate of the practice denominator (see Appendix) indicates that among practices of a particular size, random variation will be lower for higher values of the proportion visiting. Hence, other factors being equal, use of the proportion visiting during a two-year interval (.86) should provide more accurate estimates of the practice denominator than use of the proportion visiting during a one-year interval (.75).

Of greater concern, but less readily predictable, is the possible effect of systematic error which would result if the age- and sex-specific national estimates of the proportion visiting did not reflect the true situation in particular practices. For example, even though age- and sex-specific national estimates may resemble the true situation in many practices, it is possible that certain practices with peculiar characteristics may have higher or lower than expected percentages visiting. Data from the Health Interview Survey shown in Table 1 suggest that practices which serve mostly people with little education might expect different percentages visiting than those who serve populations with extensive education. Indeed, it is possible that education level as well as age and sex should be incorporated into any correction factors that estimate practice denominators using percent visiting data.

There are a number of implicit assumptions upon which the correction factor method rests and which may influence the amount of systematic error contained in estimates of practice denominators. First is the assumption that virtually all persons have one and only one regular source of care. Data from a study of a representative sample of the United States population suggest that 88 percent of the population could identify a regular source of care, and that most of those without a regular source of care either felt no need (5 percent) or were only temporarily without a regular source of care, since they or their physician had recently moved (3 percent).[20] Only about 1 percent of the population reported using more than one physician regularly. Hence, to the extent that these data reflect actual behavior, this first assumption is largely justified.

The second implicit assumption is that people who have visited a physician other than their regular physician within the previous one or two years will also have made at least one visit to their regular physician. While this assumption sounds reasonable, the authors are unaware of evidence to document this.

A third implicit assumption is that the effects of migration into the practice will balance the effects of migration out of the practice. Clearly, practices that are either rapidly losing or gaining patients would not be suitable candidates for estimating their denominators or, in all likelihood, for carrying out population based research. Hence, the main concern here is with those practices that experience a relatively small amount of net growth or loss. A major question then becomes whether the probability of a visit by an individual who is about to leave the practice differs significantly from that of an individual who has just entered the practice population. This question requires further investigation.

In summary, preliminary evidence suggests that the correction factor method may be able to provide reasonable estimates of practice denominators. Should this approach prove satisfactory, primary care physicians interested in calculating rates of events occurring in their practices need only maintain an age-sex register and a morbidity register. These registers have been successfully incorporated into the practices of a growing number of primary care physicians and should not represent an unacceptable burden to physicians truly committed to an involvement in research.

The question of the precision of estimates of the practice denominator has received little attention. Is it important to be able to detect small differences in the morbidity rates between practices, or is it only the more striking differences that are worth noting? In view of the considerable *numerator* problems that exist,[21] it seems that focusing on small differences in rates would be very hazardous. Furthermore, due to the ill-defined nature of a "practice population" in North America, precise estimates of practice denominators may not be possible.

Further investigation is required before the correction factor method can be used with confidence. Do age and sex alone adequately explain the variation in proportion visiting, or is it necessary to incorporate measures of educational level, activity limitation, or other variables into the correction factors? Do the Health Interview Survey data satisfactorily mirror the percent visiting physicians' practices, or is a better source of correction factors required? When applied to data from a

health maintenance organization, will the method correctly estimate the known denominator? The answers to these questions are being pursued by the authors.

References

1. Rose G, Barker DJP: Epidemiology for the uninitiated: Rates. Br Med J 2:941, 1978
2. Bass M: Approaches to the denominator problem in primary care research. J Fam Pract 3:193, 1976
3. Garson JZ: The problem of the population at risk in primary care. Can Fam Physician 22:71, 1976
4. Boyle RM, Rockhold FW, Mitchell GS, VanHorn S: The age/sex register: Estimation of the practice population. J Fam Pract 5:999, 1977
5. White KL: Primary care research and the new epidemiology. J Fam Pract 3:579, 1976
6. A glossary for primary care. Report of the North American Primary Care Research Group (NAPCRG). J Fam Pract 5:633, 1977
7. Robertson D, Schofield TPC: Comparison of methods of determining the denominator for morbidity in primary care. Paper presented at Ninth Annual Meeting of the North American Primary Care Research Group, Incline Village, Nevada, March 18, 1981
8. Morbidity statistics from general practice: Second national study 1971-72. In Office of Population Censuses and Surveys: Studies on Medical and Population Subjects, No. 36. London, Her Majesty's Stationery Office, 1979
9. Greenlick MR, Hurtado AV, Pope CR, et al: Determinants of medical care utilization. Health Serv Res 3:296, 1968
10. Froom J: An integrated medical record and data system for primary care: Part I. The age-sex register: Definition of the patient population. J Fam Pract 4:951, 1977
11. Kilpatrick SJ: The distribution of episodes of illness: A research tool in general practice? J R Coll Gen Pract 25: 686, 1975
12. Kilpatrick SJ: Successes and failures in the analysis of family medicine data bases. In New Challenges for Vital and Health Records: Proceedings of the 18th National Meeting of the Public Health Conference on Records and Statistics. National Center for Health Statistics (Hyattsville, Md). DHHS publication No. (PHS) 81-1213. Government Printing Office, 1980
13. Physician visits: Volume and interval since last visit, United States, 1975. In National Center for Health Statistics (Hyattsville, Md): Vital and Health Statistics series 10, No. 128. DHEW publication No. (PHS) 79-1556. Government Printing Office, 1979
14. Health, United States, 1979. US Department of Health, Education, and Welfare, Public Health Service. DHEW publication No. (PHS) 80-1232. Government Printing Office, 1980
15. Current estimates from the health interview survey, United States, 1978. In National Center for Health Statistics (Hyattsville, Md): Vital and Health Statistics, series 10, No. 130. DHEW publication No. (PHS) 80-1551. Government Printing Office, 1979
16. Selected health characteristics by occupation, United States, 1975-76. In National Center for Health Statistics (Hyattsville, Md): Vital and Health Statistics, series 10, No. 133. DHHS publication No. (PHS) 80-1561. Government Printing Office, 1980
17. State estimates of disability and utilization of medical services: United States, 1974-76. National Center for Health Statistics (Hyattsville, Md). DHEW publication No. (PHS) 78-1241. Government Printing Office, 1978
18. Fraser RC: The reliability and validity of the age-sex register as a population denominator in general practice. J R Coll Gen Pract 28:283, 1978
19. Lambert PM, Birch D: Errors in the morbidity study, letter. J R Coll Gen Pract 27:185, 1977
20. Aday L, Andersen R, Fleming G: Health care in the United States: Equitable for whom? Beverly Hills, Sage Publications, 1980
21. Anderson JE: Reliability of morbidity data in family practice. J Fam Pract 10:677, 1980

Appendix

Assume the proportion of individuals (p) in a practice population who visit a physician during a certain time interval is known. Let n represent the number of individuals with one or more visits during the interval; n is binomially distributed with parameters N (the true population denominator) and p (the proportion visiting during the interval). The following formulas yield the expected value of n: $E(n) = Np$; and its variance: $Var(n) = Np(1-p)$.

The true denominator N can be estimated by $\hat{N} = n/p$. $\hat{N}$ is an unbiased estimate of N since $E(\hat{N}) = 1/pE(n) = 1/p(Np) = N$. The variance of $\hat{N} = 1/p^2 Var(n) = Np(1-p)/p^2 = N(1-p)/p$, which for a fixed N will decrease as p increases.