ACCESS TO HEALTH CARE: A PRELIMINARY MODEL

Martin K. Chen, Bureau of Health Services Research and Evaluation

Access to health care is an important concept both in theoretical discussions on health services and in empiric research on existing and new health care delivery systems. It is rare, however, for this author to find a precise definition of the concept in the literature. Usually administrators and researchers use the term to mean quite different things and the reader is left with the task of inferring its meaning in the particular context in which the term is used. For example, Fox (1), in a discussion on the Federal Government's role in increasing access to medical care for the poor, uses the term to mean ability to get into a health care system. Shannon et al (2) appear to use the term to mean "ready availability" of health services. Other investigators have used such diverse measures as patient travel time and utilization of services as indicators of access (3, 4).

The first attempt to define the concept known to this author is made by Given et al (5) in a paper that defines access as "the social, psychological, economic and organizational factors that influence individual participation in the health services system given the availability of services." Conceptually, this definition, like the World Health Organization definition of health (6), is comprehensive but vague. As such it practically defies operationalization. As a matter of fact, the authors themselves disregarded this definition later in the paper when they operationalized access as the ratio of the total number of doctor-patient contacts to the total number of disability days per 1000 population in the past two weeks.

A much more practical definition of access is the "use-need discrepancy ratio" used in the household survey conducted by the National Center for Health Services Research and Development (7). Symbolically, this ratio is defined as:

$$R = 100 \sum_{i=1}^{n} V_i \Big/ \sum_{i=1}^{n} R_i \text{ , where}$$

R = discrepancy ratio,

V_i = number of physician visits made by individual i for two-week period,

R_i = number of days of restricted activities, including bed days, within the two-week period for individual i, and

n = number of individuals included in the computation of R.

This definition is, in effect, another version of the operational definition by Given et al, and as such it shares the problems of the other definition. First of all, either of the two ratios is by itself uninterpretable without some kind of norm. For instance, if the discrepancy ratio for Population A is .33 (1/3) and that for Population B is .40 (2/5), this information in itself cannot help a health administrator decide whether he should try to increase access of care to either or both populations.

This is not to say, however, that the use-need discrepancy ratio is not useful. Given the resources and time, one could collect data from a large number of populations on disability days, physician contacts, and some indicator of health status. With such data, one could then statistically correlate the discrepancy ratio and health status. If the correlation is reasonably high, one could determine a point or range in the value of the discrepancy ratio that corresponds to the population with the highest health status. Without implying any causal relationship between the discrepancy ratio and health status, one could use the point or range as a norm in comparative studies.

Another problem has to do with sample size. Unless the sample size is substantial, the discrepancy ratio will have to be based on rather scanty data because of the low probability of people being bedridden or restricted during the past two weeks unless it is during the height of an epidemic season. This shortcoming is overcome to some degree when the time frame is a year rather than a two-week period, but the problem remains that a proportion of the sample surveyed do not contribute any information to the "discrepancy ratio."

In spite of their arbitrariness, the operational definition of Given et al and the "discrepancy ratio" make it possible to quantify a concept that has largely been left undefined. There is little doubt that a valid quantitative index of access, however imperfect, can be a very useful tool to both health program administrators and researchers interested in the evaluation of new or innovative health delivery systems. Toward this end, an attempt is made in this paper to operationally define and quantify access as a composite measure of several parameters useful in selected situations.

Before we define what access is, we need to differentiate what may be termed perceived access and access as objectively derived. These two concepts are different and may or may not be statistically correlated. It is useful to keep the two concepts apart because both are real in their effects on consumer behavior and a merger of the two may mask the dynamics of complex interaction patterns of intra-person and interpersonal factors vis-a-vis utilization of health services. In this paper we have chosen to focus on access as objectively derived, not because we believe perceived access unimportant, but because for quantification purposes perceived access requires a different type of data, such as obtainable by the questionnaire items on barriers to access quoted in <u>Health Services Data System</u>:

the Family Health Survey (8).

Attributes of a Useful Objective Index

For an objective index of access to be useful, it must meet the following conditions:

1. It must be based on data that are readily available or can easily be collected by a health delivery system;

2. It must possess invariance from region to region or situation to situation; in other words, the parameters of the index should be given identical definitions and the data collection procedures standardized across regions and/or situations;

3. It must be easily computable given the required data; and

4. It should not include subjective elements based on feelings or perceptions.

The desirability of Conditions 1, 2, and 3 is self-evident and need no amplification. The importance of Condition 4 may not be appreciated unless it is remembered that the proposed index is objective in nature, and as such it should not be contaminated with subjective data. Furthermore, when a subjective index of access is developed, it will then be possible to study the statistical relation of the two indices and to determine their separate and joint contributions, if any, to the variance of health service utilization of a given population or group.

Definition of Access

In this paper access is defined as the degree of difficulty of a potential user in getting into the health system, and once in the system, the degree of efficiency of patient handling, given that the potential user has the resources for service and that he or she appreciates the value of service. The two given conditions are intended to isolate the health delivery system, be it clinic or hospital, from two important personal factors that are external to the system and that are postulated to have an effect on utilization. By insulating the care-providing institution from these factors in our formulations we hope to allow administrators and researchers to focus attention on the institution itself vis-a-vis the problem of access.

As defined, access may be affected by either or both factors: (1) inadequacy of the system in terms of personnel, facilities and equipment, and services; and (2) lack of efficiency in the utilization of the available supply of personnel, facilities and equipment, and services. The proposed index incorporates information about both factors. If a health facility has an adequate supply of personnel, equipment and services but its access index is low, then it may be assumed that these resources are not efficiently utilized, and the administrator should examine the system to ascertain the causes of the lack of efficiency.

Accordingly, the proposed index consists of two components, one component pertaining to a measure of adequacy of the physical facilities and personnel of the institution and the other component to efficiency in patient admission and handling. Let C_1 represent the first component. Then:

$$C_1 = \sum_{j=1}^{3} v_j \sum_{i=1}^{n_j} w_i \left|(I_i - R_i)\right| , \qquad (1),$$

where: I_i = ideal number of types of service, personnel or equipment; R_i = actual number of types of service, personnel or equipment; w_i = weight for the absolute difference between the ideal number and the actual number; v_j = weight for a given category (i.e., services, personnel or equipment); $i = 1,2, \ldots . n_j$; and $j=1, 2, 3$.

It is seen from Equation (1) that C_1 is a simple linear function of the absolute differences between the ideal number and the actual number for each type of service, personnel and equipment weighted in some manner. Although the absolute differences are used, positive and negative differences should be given different weights. This is so because the consequences of a positive difference (i.e., the ideal number exceeds the actual number) cannot be the same as those of a negative difference (i.e., the actual number exceeds the ideal number). For example, if the ideal number of ambulances for a service area is four, but the actual number is only two, many lives in the area may be threatened. This is not true if the numbers are reversed, although the situation is economically undesirable. The values of v_j reflect the relative importance of deficiencies in services, personnel and equipment.

If the economic factor of over-supply in equipment, services and personnel is ignored, Equation (1) can be reduced to:

$$C_1 = \sum_{j=1}^{3} v_j \sum_{i=1}^{n_j} w_i \; (I_i - R_i) \qquad (2)$$

Equation (2) is identical with Equation (1) except for the absolute sign. With Equation (2), $(I_i - R_i)$ is set to zero whenever its value is negative. Negative values are not allowed in the equation because they may neutralize the positive values, thus obscuring the different areas of deficiency in the system.

As formulated, C_1 is inversely related to adequacy; that is, the higher the value of C_1, the less adequate the supply of personnel, services and equipment of the institution. A computational example based on Equation (1) and the data from Tables 1 and 2 is given below:

For Facility 1,

$$\begin{aligned} C_1 = & (.4)[5(0)+3(1)+1(2)+1(1)] \\ & + (.2)[2(4)+2(3)+4(1)] \\ & + (.4)[3(1)+3(2)+3(2)] \\ = & (.4)(6)+(.2)(18)+(.4)(15) = 12.0 \end{aligned}$$

For Facility 2,

$$C_1 = (.4)[2(1)+5(0)+2(1)] + (.2)[3(6)+3(2)+1(3)] + (.4)[5(4)+4(0)+2(2)+1(1)] = (.4)(4)+(.2)(27)+(.4)(25) = 17.0$$

It is noted that the lowest possible value of C_1 is zero, indicating exact correspondence between the ideal setting and the actual setting in terms of health services, personnel, and equipment. As the value of C_1 goes up, the deviation of the actual setting from the ideal becomes greater. Since the value of C_1 cannot be negative, its value cannot reflect the direction of the deviation; for that information one is referred to Table 2, where positive and negative differences are given.

Efficiency of Patient Handling

Now let C_2 be the second component of the index representing the degree of efficiency in patient admission and handling. Then

$$C_2 = f(A, T, W, P) \quad (3),$$

where A is appointment waiting-time in days, T is patient traveling time to the care delivery institution in minutes, W is waiting-room time in minutes, and P is throughput time from first contact with physician or other health professional to completion of visit, also in minutes. In words, this component of the index is a function of four parameters, all of which have to do with the duration of the patient visit. The smaller the value of any of the four parameters the shorter the duration of the visit and the more efficient the patient handling.

Intuitively, C_2 should be some kind of average of the four parameters. Since T, W, and P are in units of minutes and A is in units of days, we make the units commensurate by transforming an eight-hour day into minutes by simple multiplication, or 60 x 8 = 480. Then we derive C_2 as:

$$C_2 = \left[\left(\frac{a_1\sum_i^n T_i/n + a_2\sum_i^n W_i/n + a_3\sum_i^n P_i/n}{a_1 + a_2 + a_3}\right)\cdot 480 \sum_i^n A_i/n\right]^{1/2}$$

$$= \left[\left(a_1\sum_i^n T_i/n + a_2\sum_i^n W_i/n + a_3\sum_i^n P_i/n\right)\cdot 480 \sum_i^n A_i/n\right]^{1/2}$$

$$= \left(\overline{X}_1 \cdot 480\, \overline{X}_2\right)^{1/2} \quad (4)$$

where a_1, a_2 and a_3 are weights; $a_1 + a_2 + a_3 = 1$; and n is the number of patients sampled from a facility. We use the subscript i to represent sample patients in any facility and the n's need not be equal across facilities. It must be remembered, however, that the means of small n's are not reliable. Further, the weights for the different parameters must be the same across facilities to ensure invariance.

C_2 is actually two types of averages. The first term,

$$(1/n)(a_1\sum_i^n T_i + a_2\sum_i^n W_i + a_3\sum_i^n P_i)$$

is a weighted arithmetic mean of the values of T, W, and P for all sampled patients and the second term,

$$(480\sum_i^n A_i/n)$$

is the arithmetic mean of the values of A for all patients weighted by a constant. C_2 is simply the geometric mean of the two terms. The reason the geometric mean of the two terms is used rather than the arithmetic mean is that the two quantities are expected to be quite disparate, and when this is the case, the arithmetic mean tends to be automatically weighted toward the larger quantity and give negligible weight to the smaller quantity. For instance, the arithmetic mean of 1 and 100 is 50.5, but the geometric mean is 10. The two means will approximate each other as the two numbers approximate each other in magnitude, and will be identical if the two numbers are equal.

Mathematically, this phenomenon is explained by the fact that in using the geometric mean we have in effect done a logarithmic transformation of the original quantities to reduce their distance. In terms of logarithms, Equation (4) is written as:

$$\log C_2 = \log(\overline{X}_1 \cdot 480\overline{X}_2)^{1/2} = (1/2)(\log\overline{X}_1 + \log 480 + \log\overline{X}_2) = (1/2)(\log\overline{X}_1 + 2.68 + \log\overline{X}_2) \quad (5)$$

It is seen from Equation (5) that the second term, $480X_2$, becomes two additive quantities, the constant 480 now being 2.68. This property of logarithmic transformations makes them a valuable tool in dealing with averages of numbers that are widely disparate in magnitude.

Examination of Table 3 reveals several things worthy of note. First, regardless of which of two sets of weights is used, C_2 is larger for Facility 2 than it is for Facility 1, indicating that Facility 1 is more efficient than Facility 2. Second, it is seen that when a new set of weights is used, the value of C_2 tends to increase in Facility 2, but decrease in Facility 1, although there is no change in the rank order of the two facilities.

Index of Access

With the values of C_1 and C_2 known or computable, we now propose an index of access:

$$I_x = \arcsin(10^k C_1/N)^{1/2} + \arcsin(10^r C_2/N)^{1/2} \quad (6),$$

where I_x is the index of access, k and r are single-digit integers, and N is the population of users of the facility in a catchment area.

Equation (6) is actually an angular transformation of two ratios, the first being the ratio of the number of deficiencies to the size of the population and the second the ratio of the number of wasted minutes to the same population. This transformation serves two functions; namely, to make the two components of the index additive and to stabilize the variances of the two ratios because of the known statistical relation between the mean and the variance of ratios or proportions. The constants k and r are intended to adjust the values of the ratios such that the ratios are not too small or too large. Their values range between -9 and +9. In any comparative study, the constants must have identical values across different facilities.

Computationally, Equation (6) may look complex, but in reality it is simple. A table such as Table 4 with given values of C_1 and C_2 as well as N would help. The values of k and r can be easily assigned after looking at the ratios of C_1 or C_2 to N. In this case it is seen that if we add two zeroes to C_1 and one zero to C_2, the ratios should be just right. Accordingly, we give the value of 2 to k and the value of 1 to r. Once the ratios are computed, we take their square roots, which can be done on a desk calculator or by referring to a table of square roots. Then the square roots are converted to degrees by using a standard table available in most statistical textbooks.

I_x has an inverse relationship with access; that is to say, the higher the value of I, the less accessible the facility is. Table 4 shows that the I_x value for Facility 1 is 97.98 and that for Facility 2 is 115.70, indicating that Facility 1 is more accessible than Facility 2. Note that I_x is a function of three parameters, C_1, C_2 and N. By holding two of the parameters constant, it is seen that the value of I_x will be high if either C_1 or C_2 is high. On the other hand, the value of I_x will be low if N is high. This phenomenon makes sense in that if a facility can keep the values of C_1 and C_2 comparatively small while it has a larger population to serve, it is bound to be more accessible than another facility with the same values of C_1 and C_2, but with a smaller population to serve.

Weighting Problems

For computing both C_1 and C_2, basic components of the index, weights are assigned to the various parameters. What should these weights be? In the case of C_1, two different weighting scales are used, the v_j's and w_i's. The weights are entirely arbitrary for demonstration purposes. Where external criterion or criteria are available, it is theoretically possible to collect enough data to statistically determine the optimal weights for the different categories of personnel, equipment and services to minimize the error of predicting the criteria by a linear combination of these categories. For example, the degree of adequacy of personnel, equipment and services may be statistically related to patient satisfaction as an external criterion. If so, different least squares techniques can be used to determine the optimal weights for personnel, equipment and services such that the error of predicting patient satisfaction by a linear combination of the three parameters is minimum.

In the case of C_2, the relative importance of T or travelling time, W or waiting-room time, and P or patient processing time, must be somehow determined. Here again, patient satisfaction may be used as an external criterion to be predicted by a linear combination of the three parameters such that the error of prediction is minimum. The weights so determined will be objective based on available data.

There are two basic problems with this approach. One is that cumulation of the data necessary for least squares solutions, if at all feasible, will be time-consuming because of the lack of valid instruments for the measurement of external criterion, such as patient satisfaction. The developmental work alone may take six months or longer. Validating the instruments may require another six months.

Another problem is that, even if a valid external criterion were readily available, the weights for the different parameters determined on the basis of data from one sample of patients would have to be cross validated with data from other samples. The patient population of a health care institution may change character over time, thus aggravating the already serious problem of sampling variations of the least squares weights. Furthermore, once a stable set of weights has been determined, periodic updating and perhaps revision with new data must be undertaken to ensure the continued validity of the weights.

An alternative to the statistical technique of minimizing the error of prediction is the psychometric technique of scaling. The values of the weights for the different parameters, say T, W, and P for C_2, can be scaled by a variety of techniques, such as pair comparison and successive categories discussed in Guilford's work (9). With proper sampling procedure, the scale values so derived represent the values to which a given population subscribes. Although still subjective in nature, the weights or values so determined may in fact have greater intrinsic validity than objective weights because it is well known that people's behavior is governed by their perceptions of reality than by reality per se.

The above discussions may give the impression that the utility of the model depends on the proper determination of the weights in C_1 and C_2. That is not true. Initially, one could give the parameters equal weights by assigning the value of one to them. Or, one could use a panel of judges to determine the rank order of the parameters and use the ranks as weights. The comparisons of the health care institutions in terms of access will be valid if the weights are accepted by these institutions.

Some Cautionary Remarks

It is stated earlier in this paper that the index should be useful in selected situations, implying that certain requirements must be met for the valid use of the index. What are these requirements? One is that the care providing institutions have available data or are willing to collect such data on new patients or old patients seeking appointments of their own accord. This requirement is intended to rule out physician-ordered or recall visits that are prescheduled and that would make appointment waiting time meaningless. If the physician orders a patient to return for a checkup in three months, it is not fair to the physician or clinic to use three months as A or appointment waiting time.

Another requirement is that the data are limited to non-emergency cases. In true emergency cases there usually is no appointment necessary and appointment waiting time is the time interval between the call to the ambulance service and the arrival of the ambulance, probably a matter of minutes. Since true emergency cases constitute only a small fraction of the total caseload of an institution, it is better not to apply the index to these cases.

The third requirement is dictated as much by validity as by common sense. That is, for comparative purposes the institutions should be similar in nature and serving similar types of populations. One would not compare a fee-for-service multi-specialty group with a public-supported hospital, such as a U.S. Public Health Service hospital. They are different in nature and they serve different types of populations. While the question, "How similar is similar?" may be legitimately raised, one need not be slavish in matching the institutions to be compared or the populations to be served. Nonetheless, the closer the match of the institutions compared, the easier it is to pinpoint the causes of their relative efficiency as measured by I_x.

The fourth and last requirement is that for each institution there is a well-defined catchment area. Without a well-defined area, it would be futile to talk about "ideal numbers" of personnel, equipment and services. The numbers are "ideal" in relation to the population served; if this population were unknown or only vaguely known, then it would be impossible to come up with an ideal that has any meaning. This requirement may be difficult to meet because of the plurality of service institutions in a community, except for prepaid group practices.

The first component of the index, C_1, deals with quantity only. An additional weight for quality could easily be incorporated into the formula, but this would make the index computationally complex. Besides, quality, particularly of professional people, is difficult to measure. For this reason the parameter of quality is ignored in the present model.

Finally, it is to be remembered that this index is not a precise indicator of access. It is some number that has no meaning in itself, but that it assumes meaning only when the same set of operations are applied to selected institutions that meet certain requirements.

REFERENCES

1. Fox, Peter D., "Access to Medical Care for the Poor: The Federal Perspective," Medicare Care 10:272-277, May-June 1972

2. Shannon, G. W., L. Bachshur, and C. A. Metzner, "The Concept of Distance as a Factor in Accessibility and Utilization of Health Care," Medical Care Review 26:143-161, 1969

3. Drosness, D. L. and J. W. Lubin, "Planning Can Be Based on Patient Travel," The Modern Hospital 106:92-94, April 1966

4. Willie, C. V., "Health Care Needs of the Disadvantaged in Rural-Urban Area," HSMHA Health Reports 87:81-86, January 1972

5. Given, C., C. Akpom, and S. Katz, "Management Information for Improvement of Access to Ambulatory Care Facilities," unpublished report

6. World Health Organization, Measurement of Levels of Health, Technical Report Series No. 130, Geneva, World Health Organization, 1957

7. National Center for Health Services Research and Development, Health Services Data System: The Household Survey, unpublished report, August 1972

8. Eichhorn, R. L., Health Services Data System: The Family Health Survey, unpublished report

9. Guilford, J. P., Psychometric Methods, New York:McGraw-Hill Book Co., 1954

Table 1

Values of w_j and v_j for the Computation of

C_1 for Two Health Facilities*

Facility 1					Facility 2				
I. Types of Service (v_1 = .4)	1	2	3	4	I. Types of Service (v_1 = .4)	1	2	3	
$w_i(+)$	5	3	3	2	$w_i(+)$	2	5	3	
$w_i(-)$	2	2	1	1	$w_i(-)$	1	3	2	
II. Types of Personnel (v_2 = .2)	1	2	3		II. Types of Personnel (v_2 = .2)	1	2	3	
$w_i(+)$	2	4	4		$w_i(+)$	3	3	2	
$w_i(-)$	1	2	3		$w_i(-)$	2	2	1	
III. Types of Equipment (v_3 = .4)	1	2	3		III. Types of Equipment (v_3 = .4)	1	2	3	4
$w_i(+)$	3	3	3		$w_i(+)$	5	4	2	2
$w_i(-)$	2	2	1		$w_i(-)$	3	2	1	1

* For v_j a scale from 0 to 1 is used, whereas for w_i a scale from 1 to 5 is used. The scales are entirely arbitrary. $w_i(+)$ refer to weights to be used if (I_i-R_i) is positive, and $w_i(-)$ to weights to be used if (I_i-R_i) is negative.

Table 2

Comparison of Two Health Facilities in Access
With Fictitious Data

Facility 1					Facility 2				
I. Types of Service	1	2	3	4	I. Types of Service	1	2	3	
Ideal (I_i)	1	3	2	0	(I_i)	4	3	7	
Actual (R_i)	1	2	4	1	(R_i)	3	3	8	
Diff.	0	+1	-2	-1		+1	0	-1	
II. Types of Personnel	1	2	3		II. Types of Personnel	1	2	3	
Ideal (I_i)	20	12	8		(I_i)	24	15	5	
Actual (R_i)	16	15	7		(R_i)	18	13	8	
Diff.	+4	-3	+1			+6	+2	-3	
III. Types of Equipment	1	2	3		III. Types of Equipment	1	2	3	4
Ideal (I_i)	3	2	5		(I_i)	7	2	5	0
Actual (R_i)	2	0	3		(R_i)	3	2	3	1
Diff.	+1	+2	+2			+4	0	+2	-1

Table 3 1/

Hypothetical Data for Computing C_2

	Facility 1				Facility 2			
User	T	W	P	A	T	W	P	A
1	5	12	30	4.5	10	21	8	7
2	13	5	26	6	7	15	10	3.5
3	17	20	15	3	20	21	30	6
4	32	4	7	8	23	20	28	7.5
5	12	15	25	4	18	19	15	9.5
6	10	8	11	10				
Total	89	64	114	35.5	78	96	91	33.5
Mean	14.8	10.6	19	5.9	15.6	19.2	18.2	6.7
C_2* = 207.6		C_2** = 198.8			C_2* = 238.7		C_2** = 241.8	

* Computations based on following weights: $a_1 = .3$, $a_2 = .3$, $a_3 = .4$

** Computations based on following weights: $a_1 = .2$, $a_2 = .5$, $a_3 = .3$

1/ To facilitate computation, a table such as Table 3 may be used.

Table 4

Computational Table for Comparing Two Facilities in Terms of Access

	Facility 1	Facility 2
C_1	12.0	17.0
C_2	207.6	238.7
N	5000.0	4000.0
k	2	2
r	1	1
I_x	97.98	115.70