ANNALS OF FAMILY MEDICINE

Online Supplementary Material

Katerndahl DA, Wood R, Jaén CR. A method for estimating relative complexity of ambulatory care. *Ann Fam Med.* 2010;8:(4):341-347.

http://www.annfammed.org/cgi/content/full/8/4/341/DC1

Supplemental Appendix. Computation of Complexity

Computation of Complexity of Each Input/Output

We used the National Ambulatory Medical Care Survey (NAMCS) for calculations. The NAMCS is an annual national probability sample survey of visits to the offices of physicians classified by the American Medical Association (AMA) and American Osteopathic Association (AOA) as "office-based, patient care." The complexity of each input/output is defined as the mean input/output per clinical encounter weighted by its inter-encounter diversity and variability. The mean of the input/output was chosen because it is the most unbiased estimator of central tendency,[1] and because the median and mode could be zero in meaningful but uncommon input/output measures, underestimating complexity. Thus, the complexity of diagnoses seen in family medicine (coded as "General/family practice" in the NAMCS database) would be the product of the mean number of diagnoses seen in family medicine encounters, the inter-encounter diversity of diagnoses weighting, and the inter-encounter variability of diagnoses weighting. Information theory also supports the weighting of information.[2] The NAMCS 2000 data set provides a patient weight that allows the sample of 27,369 visits to be "inflated" to represent the total of 832,541,999 visits that year in the United States. This patient visit weight was applied to the data set so that estimates of complexity parameters produced by re sampling techniques would better conform to national patterns of patient encounters.[3]

The diversity of an input/output is defined as the proportion of the number of categories needed to include 95% of the input/output reported out of the total possible categories. The 95% proportion was chosen to minimize the impact of a rare or miscoded input/output. The variability was defined as the coefficient of variation (COV) of the input/output, which is calculated as the standard deviation divided by the mean. The COV was chosen over other measures of variation because it is a unit-free measure.[4] Thus, diversity will and variability should typically range between 0 and 1. To standardize the weightings and limit the impact of low diversity or variability on complexity, the weightings used are the Z-transformations of the diversity proportion and the COV, and range between 0.5 and 1.0.

Using the 2000 NAMCS database,[2] the diversity of 95% of the diagnoses seen is 0.47 and the COV of diagnoses seen is 0.50. These Z-transform into weights of 0.68 and 0.69 respectively. Thus, the complexity of family medicine diagnoses is:

Mean Diagnoses per Encounter		Diversity Weighting		Variability Weighting		Complexity
1.70	×	0.68	×	0.69	=	0.80

Severity of illness was not included in this formula. First, severity of illness is distinct from complexity.[5] Second, a primary reason for including severity of illness when assessing complexity would be its impact on testing and outputs. Because these measures are already included, using severity of illness in the calculations would over-emphasize its impact on complexity. Similarly, the acuteness of illness was not included. Although it has been suggested that acute problems may represent higher complexity states due to their lack of equilibrium,[6] the fact that many are self-limited would suggest lower complexity across encounters. Thus, acuteness of illness was not included.

Because some inputs/outputs (ie patient characteristics, patient disposition) could not be represented in this manner, those variables were handled in a different but analogous way. For these variables, Z-transformations were performed on each component. Thus, patient characteristics are represented by 3 components (sex and race/ethnicity, age variability, and proportion of patients previously unknown to the physician). Sex and race/ethnicity were combined in a 2-way table. As with diversity, this combined sex-race/ethnicity is the proportion of possible categories that represents 95% of the patients seen. Similarly, age variability is measured as the COV for the ages of the patients seen. Finally, the proportion of patients previously unknown to the physician is also assessed. Previous work suggests that previously unknown patients represent situations of higher complexity.[7] Once these 3 components are represented by their proportions or COV, Z-transformation is performed to convert them to scores ranging from 0.5 to 1.0 and these scores are then summed to provide an estimate of patient characteristics complexity. Using the 2000 NAMCS data, this results in a patient characteristics complexity for family medicine as follows:

Component	Proportion/COV	Z-Transformation
Sex-race/ethnicity	0.63	0.74
Age variability	0.52	0.70
Previously unknown patients	0.09	0.53
PATIENT CHARACTERISTICS COMPLEXITY		1.97

COV = coefficient of variation

The number of components used in assessing patient characteristics complexity could be fewer or greater. Three components are used in this example so that the maximum possible patient characteristics complexity matches the maximum possible complexity of the other input components, such as diagnoses, thus providing similar weight across components when computing total input complexity. In a similar manner, patient disposition is measured as the Z-transformation of the proportion of possible patient disposition categories that represents 95% of the encounters. For family medicine, of the 5 possible patient disposition categories in the 2000 NAMCS data, 3 were used in 95% of the encounters (proportion = 0.60); this corresponds to a Z-transformation of 0.73.

Computation of Total Complexity

Once the complexity of each component has been calculated, the total input and total output complexities are calculated by summing the component complexities.[6] However, calculation of the total specialty complexity is not merely the sum of the input and output complexities. A fundamental principle of complex systems is that there is a logarithmic relationship between input and output, so that, as the information in the input increases linearly, the complexity of the system increases exponentially. Thus, for binary data, the total system complexity is determined by the following formula[6]:

$$\text{System Complexity} = \text{Output Complexity} \times 2^{(\text{Input Complexity})}$$

In this case, we accept the assumption of binary data for 2 reasons. First, the components used generally represent the presence or absence of an entity (ie, a particular diagnosis, a particular medication). Second, biological systems behave as if they are binary no matter what system we examine.[8] Thus, total system complexity depends heavily upon the complexity of the input.

Table 1 presents the complexity of family medicine using the 2000 NAMCS data. The total input complexity is the sum of the complexities of reasons-for-visit (0.77), diagnoses (0.80), examination/testing (0.83), and patient characteristics (1.97). Using the formula presented above, we calculate the total specialty complexity as:

$$\begin{array}{ccccc} \underline{\text{System Complexity}} & = & \underline{\text{Output Complexity}} & \times & \underline{2^{(\text{input Complexity})}} \\ 44.04 & = & 2.15 & \times & 2^{(4.36)} \end{array}$$

For the purposes of this study, we used the above procedure to estimate complexity of ambulatory care for family medicine, cardiology, and psychiatry.

Limitations

The process of estimating complexity has several limitations. In addition to the difficulty in counting events, the lack of knowledge about the full behavior of the system, and the appropriateness of the framework of estimation,[6] these measures have no gold standard. Units in this complexity measure are purely abstract, without any concrete meaning. This method is only useful in comparisons. In addition, if we could account for the full range of decision-making strategies, the gap between generalist and specialist relative perceived complexity may be even greater. Finally, the database used for any such estimate will almost certainly be limited. For example, NAMCS only allowed physicians to report a maximum of 3 problems per visit; previous work suggests that the average number of problems address in a brief visit to a family physician is 3-4.[9,10] Hence, such estimates will tend to underestimate complexity. Their value in relative comparisons lies in the unbiased limitations of measurement across specialties.

References

1. Sokal RR, Rohlf FJ. *Introduction To Biostatistics.* 2nd ed. New York, NY: W. H. Freeman & Company; 1987.
2. Reza FM. *Introduction To Information Theory.* New York, NY: Dover Publications, Inc; 1994.
3. National Center for Health Statistics. *National Ambulatory Medical Care Survey.* Hyattsville, MD: US Department of Health and Human Services; 2000.
4. Armitage P, Berry G. *Statistical Methods In Medical Research.* 2nd ed. Boston, MA: Blackwell Scientific Publications; 1987.
5. Horn SD, Buckle JM, Carver CM. Ambulatory severity index: development of an ambulatory case mix system. *J Ambul Care Manage.* 1988;11(4):53-62.
6. Bar-Yam Y. *Dynamics of Complex Systems*. Reading, MA: Perseus Books; 1997.
7. Fraser SW, Greenhalgh T. Coping with complexity: educating for capability. *BMJ*. 2001;323(7316):799-803.
8. Kauffman SA. Origins of Order. New York, NY: Oxford University Press; 1993.
9. Beasley JW, Hankey TH, Erickson R, et al. How many problems do family physicians manage at each encounter? A WReN study. *Ann Fam Med.* 2004;2(5):405-410.
10. Flocke SA, Frank SH, Wenger DA. Addressing multiple problems in the family practice office visit. *J Fam Pract.* 2001;50(3):211-216.