Winter 2-1977

The Authors Respond

John E. Wennberg
Dartmouth College, john.e.wennberg@dartmouth.edu

Alan Gittelsohn
Johns Hopkins University

The Authors Respond

We appreciate the opportunity provided by the Editor of responding to Dr. Francis Moore's[1] preceding critique of our two recent articles[2,3] on small area variations in the utilization of surgical procedures and other health services. The language "capricious whim of hospitals, physicians and surgeons" is strictly Dr. Moore's, never having been employed by either of us. Our point by point response is roughly in the order of Dr. Moore's critique:

1. *Selection of procedures*

Dr. Moore's allegation that we look only at procedures which fit our argument is not true. The basis for selection of procedures for analysis was relative frequency and the types of surgery previously reported in the literature, as mentioned in the text. There is simply insufficient space in a journal article to deal with the hundreds of distinct procedures codes in the medical record. As to Dr. Moore's specific objection to our failure to consider fractures and obstetrical delivery because the "data failed to support their thesis," we indeed have examined and published data on these two categories and they support the thesis admirably (Figure 1). The hospitalization rate for deliveries and other pregnancy related conditions in Maine[4] and Vermont exhibit low variability over neighboring areas (as Dr. Moore's own data for Maine and SOSSUS clearly show). This is to be anticipated: Unlike such European countries as Holland, obstetrical traditions in the United States call for hospitalized delivery; and the birth rate among neighboring areas is approximately the same.

By contrast, admissions for fractures[5] show intermediate variability. Physicians familiar with treatment of trauma will recognize that not all fractures are admitted to hospitals for care. Judgment as to which cases should be hospitalized (as judgment concerning which should be treated with closed or open reduction techniques) may vary from physician to physician.

2. *Variation between years*

Dr. Moore suggests that "variations between successive years might be just as great as the variations between local areas." Only one year of data was available at the time of writing, 1969 for Vermont and 1973 for Maine and it was not possible to study annual trends. Since publication, we have examined and reported on trends between and within areas.[6,7,8,9] The general pattern is a marked consistency within areas over time (as exemplified by Table 1). Tonsillectomy procedures[8] and cesarean section[9] provide notable exceptions. Over the five-year period 1969-73 for Vermont, the overall tonsillectomy rate has declined by 40%. The proportion of deliveries by cesarean section nearly doubled in areas served by hospitals with fetal monitoring and remained fairly constant in areas served by community hospitals. Consistency is measured by the rank order correlation of areal rates at the beginning and end of the period and by the Kendall concordance coefficient.

3. *Control information*

Dr. Moore rightly points out that "control information" should be studied to provide meaningful comparison between areas. He undoubtedly is aware that only a limited amount of data is available by individual town from secondary sources including the U.S. Census. Indeed, we have examined a wide range of available socio-demographic variables in order to measure differences and similarities between areas. Based on Vermont vital statistics for the period 1968-72, one of us has prepared a report[6] on mortality differentials. In general, the 13 areas of Vermont exhibit low variability in total death rates, life expectation and cause specific mortality for such broad diagnostic classes as malignant neoplasms, heart disease, stroke and automobile accidents. Stillbirth, neonatal, perinatal and infant mortality rates similarly exhibit lower inter-area variability than all of the common surgical procedures except repair of inguinal hernia. In many instances, the vital events are of much lower frequency than the types of surgery under consideration. The latter should ease some of Dr. Moore's concern about rare events.

We also believe it is important to study comparative behavior in seeking medical care. An extensive household survey in six Vermont hospital service areas (including areas at the extremes in range of use of hospital and surgery) shows the populations to be well matched on socio-economic factors, including ethnic background, health insurance coverage and percent below poverty level; the distribution of reported illness and individual behavior in seeking care are similar. These studies, referenced in Maine I,[3] will be published in a subsequent issue of *The Journal of the Maine Medical Association.*[10]

4. *Size of numerators*

Dr. Moore need not fear that the importance of variations among small areas disappears because of "tiny integers in the numerators." He draws attention to our Figure 4.[3] The numerators for these rates are not small. The high and low rates for tonsillectomy are based on 652 and 1,042 cases, respectively. For the less common procedures, the high rate for hemorrhoidectomies is based on 94 cases, the low on 66. For varicose veins, the corresponding numbers are 60 and 65. Statistically, differences in rates between areas are highly significant. They are also of medical significance. For example, the number of hemorrhoidectomies in the high area ex-

Figure 1

Pattern of Variation for Congenital Anomalies, Pregnancy Related Admissions and Four Common Pediatric Surgical Procedures

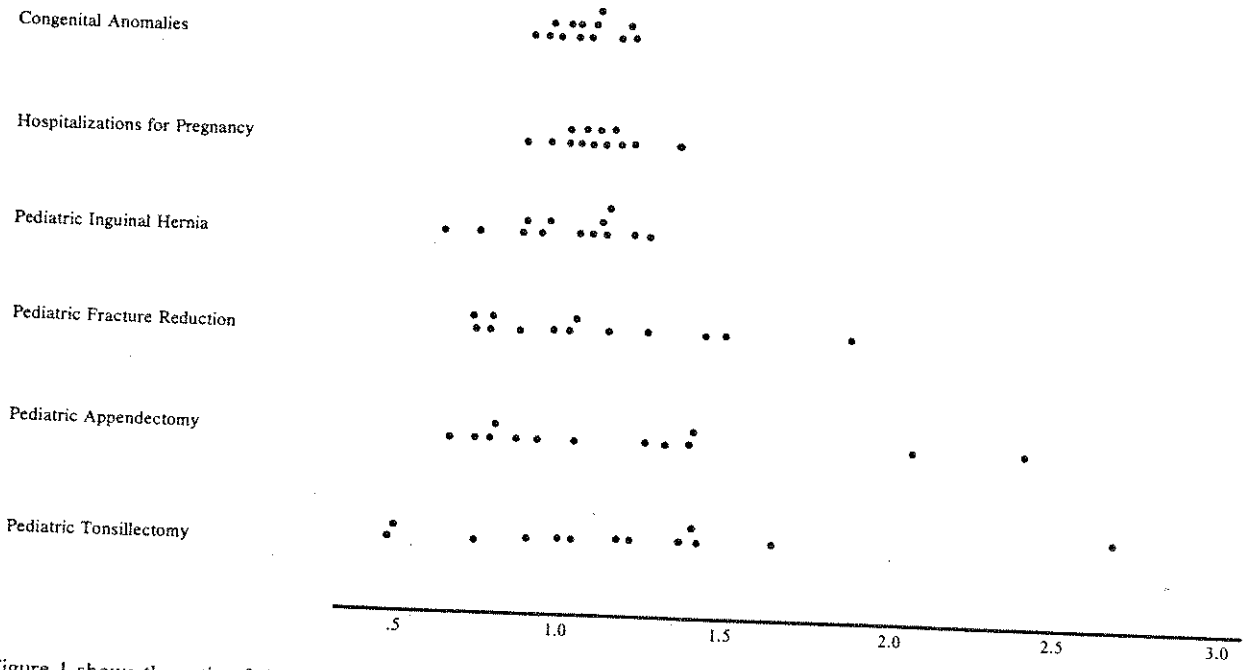


Figure 1 shows the ratio of the age-adjusted rate of admission to the state average rate for 13 hospital service areas. Admission for congenital anomalies and pregnancy associated conditions exhibit low variability. Of the four procedures, inguinal hernia shows low variability and none of the areas differ from the state average at the .01 level of significance. Fracture reduction demonstrates intermediate variability with four areas statistically different. Appendectomy and tonsillectomy show much greater variations. (The state average annual rate for congenital anomalies is 19 per 10,000; for conditions of pregnancy it is 189. The state average rate for the four procedures, in order, is 22, 55, 27, and 96 per 10,000 population per year. The data for pediatric procedures is for 13 Vermont hospital service areas and covers five years (1969-1973).[5] The data for congenital anomalies and pregnancy associated conditions is for the 13 largest Maine hospital service areas for the year 1973[4].)

ceeds the expected number by 60. As Dr. Moore will recognize, many inferences about therapeutic effectiveness often are based on smaller numbers of cases.

5. *Age corrections*

Dr. Moore remarks that the authors have applied "age" corrections to their data and have made the assumption that "the age group distribution of a disease will be the same throughout a large area." The first part of the statement is true and the second grossly false. As stated in the text, we have computed age adjusted rates using the direct method of adjustment which is standard demographic practice. Age adjustment is required because the conditions under study vary with age. The adjustment is applied to insure comparability between areas with differing age structures, a point which Dr. Moore has missed.

6. *Elective surgery*

Excluding appendectomy, Moore rightly states that the other eight surgical procedures involve an important interaction between patient, physician and surgeon. He then goes on to remark that "to assume that alterations in herniorraphy repair are always related to 'provider factors' " is a gross oversimplification. The statement errs badly in several major respects. We never made it. Had Dr. Moore looked at the data, he would have noted that herniorraphy, of all common types of surgery, exhibits the least variability between areas. Evidently, inguinal hernia, with its high apparency to the patient, is treated similarly across areas (See Figure 1). Dr. Moore's concluding 'platitude' in the same paragraph "no psychiatrist, no psychoanalysis, no surgeon, no surgery" is well supported by the SOSSUS study in which he is a participant. Although the SOSSUS report virtually ignores the issue," Dr. Moore's own data clearly demonstrate a higher incidence of surgery in the areas with the higher number of surgeons per capita (Figure 2).

7. *Population coverage*

Dr. Moore, in his comments on population coverage, misrepresents the concept of a population rate, the most fundamental idea in epidemiology and demography. In the numerator of the rate, we require a complete count of the events occurring to the members of the population at risk constituting the denominator. Both of these conditions have been fulfilled in the Vermont and Maine studies, not to mention the Lewis[12] report on Kansas Blue Cross/ Blue Shield subscribers which Dr. Moore also

TABLE 1

PRIMARY APPENDECTOMIES AND VARICOSE VEIN STRIPPING IN TWO SMALL VERMONT HOSPITAL SSERVICE AREAS*

	Observed and Expected Number of Cases by Year					
	1969	*1970*	*1971*	*1972*	*1973*	*All Years*
Appendectomy						
Area 1						
Observed	29	34	37	46	31	177
Expected	22.1	20.7	21.1	19.5	19.7	103.1
Area 2						
Observed	14	15	22	17	11	79
Expected	22.5	21.0	21.9	20.1	19.9	105.4
Varicose Veins						
Area 1						
Observed	11	12	9	18	10	60
Expected	9.8	9.7	9.7	8.2	8.9	46.1
Area 2						
Observed	29	22	22	13	23	109
Expected	10.7	10.6	10.6	9.0	9.4	50.3

*Population circa 10,000

The data are for two Vermont hospital service areas with populations of about 10,000 (1970 census). Area 1 has a high appendectomy rate, Area 2 a high rate of varicose vein procedures. The numbers are consistent between years. By the end of five years, Area 1 has about 75 more appendectomies than expected (based on STATE average); Area 2 has about 60 more examples of varicose vein stripping than expected.

criticizes. It is irrelevant whether or not the population of an entire state is covered. The issue is that all of the events under consideration be included in the numerator, a point explained in the text and in all basic works on epidemiology.

8. *Population size*

We agree with Dr. Moore on the difficulty of drawing inferences about rare events in small populations. In part, this is a classical statistical problem accounted for directly in the standard deviations of the population rates. To argue that the four SOS-SUS study areas average about 1,000,000 persons and that the total surgical rates range narrowly between 58 and 91 is not relevant. Such information tells us nothing about intra-regional variability which is the central point of our two reports. Each hospital service area in Maine and Vermont is an aggregation of adjoining towns whose residents receive most of their hospital care primarily at local facilities. Most hospital service areas contain only one facility staffed by one group of physicians. An important thrust of our effort has been to study local differences in the delivery of care and to examine the relationships between morbidity, care, and community and provider characteristics. To lump the experience on a regional basis misses the entire point.

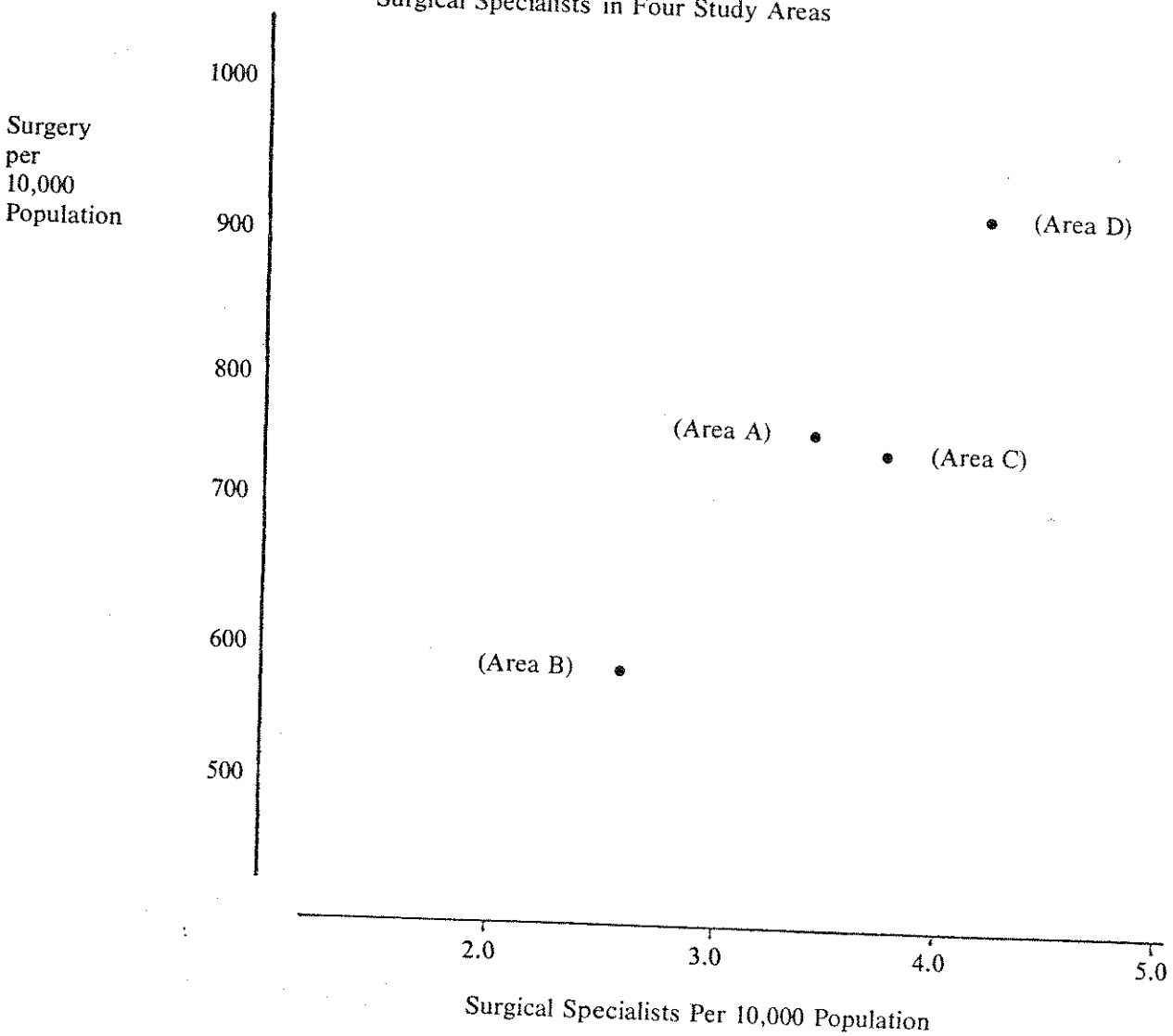
Dr. Moore should be well aware that major differences in the incidence of surgery exist between large population groups. For years, the Canadians have recorded rates for gynecological and abdominal surgery far in excess of those for the United States. Similarly, the rates in Great Britain and among members of group health plans in the United States are less than half those for subscribers of Blue Cross/Blue Shield type plans. Canadian cholecystectomy rates are more than five times the British rates. The West Germans experience remarkably high-appendectomy rates. It is clear that these differences between these areas, based on millions of persons, are real and statistically significant. In no instance has it been possible to relate variations in common surgery rates to variations in the incidence and prevalence of the conditions for which the surgery is performed. There is no evidence suggesting that group health subscribers have half as much gynecological disease as subscribers of other health insurance plans or that English females have significantly less cholelithiasis than Californians or that the incidence of acute appendicitis varies by a factor of 2 between blue and white collar workers in Hanover. By the same token, it should not be surprising that practice differences and similarities occur between neighboring Vermont or Maine communities served by different groups of physicians. In one area, served by a single hospital, over half the deliveries are induced while in another the cesarean section rate is about double that recorded in the rest of the State. The same situation pertains in varying degrees to the types of common surgery under discussion. The issue will not disappear by resorting to innuendo, ad hominums and distortion. Rather, the responsible approach is to study current practice and to ascertain which level of care results in betterment of the health, tranquility and general well-being of the population.

9. *International comparisons*

Dr. Moore has misread our quote of the Bunker report[13] on the incidence of common surgical procedures in Great Britain and the United States. We have quoted Bunker accurately, directly and without error. Dr. Moore's allegation that we have sup-

Figure 2

Operations Performed and Supply of Surgical Specialists in Four Study Areas



Data is from Table 11, Chapter IV, *Surgery in the United States: A Summary Report of the Study on Surgical Services for the United States*. Sponsored Jointly by The American College of Surgeons and The American Surgical Association, 1975. There is a concommitant variation between surgeon supply and rate of surgery.

pressed similarities in appendectomy, cataract extraction, thyroidectomy and circumcision rates is beside the point; by the same token Dr. Moore has suppressed the 3 to 1 differential in cholecystectomy rates and the 2 to 1 differential in hysterectomy rates. It was not our intention to reproduce the Bunker article in full since it was published in a widely available journal.

10. *Social responsibility*

In his concluding paragraph, Dr. Moore's gratuitous comment on the social responsibility of sociologists is not applicable. Neither of us has ever been a sociologist; one of us is a physician and the other is a biostatistician. Dr. Moore's term "uncontrolled studies" is correct only in the sense that all observational studies including SOSSUS are uncontrolled. The "glaring deficiencies" of our studies remain unknown to us, our colleagues or referees. We have attempted to insure that the data are correct within the limits imposed by the medical record and other data systems available to us. It has been our objective to let the data tell their own story. We feel gratified that our work has received national attention and regret that Dr. Moore finds the situation "especially unfortunate."

REFERENCES

1. Moore, F.: Small Area Variations in Health Care Delivery: A Critique. Journal of the Maine Medical Association, 68: *to come*, 1977.

2. Wennberg, J. E., Gittelsohn, A.: Small Area Variations in Health Care Delivery. A population-based health information system can guide planning and regulatory decision-making. Science, 182: 1102-1108, 1973.
3. Wennberg, J. E., Gittelsohn, A.: Health Care Delivery in Maine I: Patterns of Use of Common Surgical Procedures. Journal of the Maine Medical Association, 66: 123-130, and 149, 1975.
4. Wennberg, J. E., Gittelsohn, A., and Soule, D.: Health Care Delivery in Maine II: Conditions Explaining Hospital Admission. Journal of the Maine Medical Association, 66: 255-261, and 269, 1975.
5. Wennberg, J. E., Grufferman, S. K.: Common Uses of Hospitals by Vermont Children. To be published as a chapter in a forthcoming book by The Child Health Project, Harvard University. (Ballinger Press, Cambridge, Massachusetts).
6. Gittelsohn, A.: Vermont Surgery Report. Technical Report, Cooperative Health Information Center of Vermont, Burlington, 1974.
7. Gittelsohn, A., Wennberg, J. E.: "On the Incidence of Tonsillectomy and Other Common Surgical Procedures" Chapter I-4 IN Bunker, J.P., Mosteller, F., and Barnes, B. A. (eds.), *An Inquiry into the Efficacy of Some Common Operations*. Oxford University Press, New York City, (In Press) 1976.
8. Wennberg, J. E., *et al*: Changes in Tonsillectomy Rates Associated with Feedback and Review. The Journal of Pediatrics (In Press).
9. Wennberg, J. E.: Issues and Recommendations of a Seminar of National Health Planning Goals. Boston, 1976. National Technical Information Service, Springfield, Virginia (In Press).
10. Wennberg, J. E., Fowler, F. J.: A Test of Consumer Contributions to Small Area Variations in Health Care Delivery. Journal of the Maine Medical Association, forthcoming.
11. *Surgery in the United States: A Summary Report of the Study on Surgical Services for the United States*. Sponsored Jointly by The American College of Surgeons and The American Surgical Association, 1975.
12. Lewis, C. E.: Variations in the Incidence of Surgery. NEJM, 281: 880-884, 1969.
13. Bunker, J. P.: Surgical Manpower, A Comparison of Operations and Surgeons in the United States, England and Wales. NEJM, 285(3): 135-144, 1970.

Alan Gittelsohn, Ph.D.
Department of Biostatistics
School of Hygiene and Public Health
Johns Hopkins University
Baltimore, Maryland 21205

John E. Wennberg, M.D.
Department of Preventive and Social Medicine
Harvard Medical School;
Center for the Analysis of Health Practices
Harvard School of Public Health
Boston, Massachusetts 02115

COMMENT BY DOCTOR MOORE — (JANUARY 12, 1977)

I am delighted that Wennberg and Gittlesohn have been stimulated to present the obstetrical and fracture data. The obstetrical variation (0.9 to 1.4) and the fracture variation (pediatrics only, 0.6 to 1.8) — if we may interpolate from their graph — clearly indicates that "provider behavior" is not the only source of "small area variation in health care delivery" with which they are so concerned.

Unanswered remains the plea that they broaden their view and document small area variations in other social phenomena that reflect both local custom (fraction of registered voters voting, for example) or market forces (auto, television or refrigerator sales, movie house attendance). Without this we are left with the message that medical variability is due to the behavior of doctors only, and the conclusion drawn by DHEW that the best way to cut costs is to reduce services.

For any "ad hominums" (sic) I apologize; but if these two authors are going to make a career of "small area variations"; let us hope these will put their studies on a sounder scientific basis. We all owe them a debt of gratitude for their concern about this especially difficult statistical and interpretive problem: local variability.

FRANCIS D. MOORE, M.D.