

Surgeons, American Medical Association, American Medical Group Association, American Society of Internal Medicine, California Association of Physician Organisations, California Medical Association, Illinois State Medical Society, IPA Association of America, Massachusetts Medical Society, Medical Group Management Association, Medical Society of the State of New York, Michigan State Medical Society, Minnesota Medical Association, National IPA Coalition, and Texas Medical Association.

Contributors: TGR, SMS, RRG, JCR, LC, TB, and NO contributed to the conceptual development of the study, study design, data collection and analysis, and writing the manuscript. MCW contributed to the study design, data collection and analysis, and writing the manuscript. JAS contributed to the data collection and analysis. DB, EW, and PG contributed comments and consultation. TGR is guarantor for the study.

Funding: This work was supported by the Robert Wood Johnson Foundation, grant No 038690.

Competing interests: None declared.

- 1 Partnership for solutions. *Better lives for people with chronic conditions*. Baltimore, MD: Johns Hopkins University, Robert Wood Johnson Foundation, 2002. (www.chronicnet.org/problem/index.htm)
- 2 National Center for Health Statistics. *New asthma estimates: tracking prevalence, health care, and mortality*. Hyattsville, MD: National Center for Health Statistics, 2001.
- 3 National Institute of Mental Health. *Mental disorders in America*. Washington, DC: NIMH, 2001.
- 4 Centers for Disease Control. *Diabetes fact sheet*. Atlanta, GA: CDC, 1998.
- 5 Jessup M. Mechanical cardiac support devices—dreams and devilish details. *N Engl J Med* 2001;345:1490-92.
- 6 Hoyert DL, Arias E, Smith BL, Murphy SL, Kochanek KD. Deaths: final data for 1999. *Natl Vital Stat Rep* 2001;49.
- 7 National Heart, Lung, and Blood Institute. *Congestive heart failure data fact sheet*. Washington, DC: National Institutes of Health, 1996.
- 8 Songer TJ, Ettaro L. *Studies on the cost of diabetes*. Atlanta, GA: Centers for Disease Control, 1998.

- 9 Weiss K, Sullivan S. The health economics of asthma and rhinitis: assessing the economic impact. *J Allergy Clin Immunol* 2001;107:3-8.
- 10 Wagner EH, Austin BT, Davis C, Hindmarsh M, Schaefer J, Bonomi A. Improving chronic illness care: translating evidence into action. *Health Aff* 2001;20(6):64-77.
- 11 Legoretta AP, Christian-Herman J, O'Conner RD, Hasan MM, Evans R, Leung KM. Variation in managing asthma: experience at the medical group level in California. *Am J Managed Care* 2000;6:445-53.
- 12 Young AS, Klap R, Sherbourne CD, Wells KB. The quality of care for depressive and anxiety disorders in the United States. *Arch Gen Psychiatry* 2001;58:55-61.
- 13 Clark CM, Fradkin JE, Hiss RE, Lorenz RE, Vinicar F, Warren-Boulton E. Promoting early diagnosis and treatment of type 2 diabetes. *JAMA* 2000;284:363-5.
- 14 Institute of Medicine. *Crossing the quality chasm*. Washington, DC: National Academy Press, 2001:28-30.
- 15 Wagner EH, Austin BT, Von Korff M. Organizing care for patients with chronic illness. *Milbank Q* 1996;74:511-43.
- 16 Wagner EH, Austin BT, Von Korff M. Improving outcomes in chronic illness. *Managed Care Q* 1996;4(2):12-25.
- 17 Wagner EH, Davis C, Schaefer J, Von Korff M, Austin B. A survey of leading chronic disease management programs: are they consistent with the literature? *Managed Care Q* 1999;7(3):56-66.
- 18 Wagner EH, Glasgow RE, Davis C, Bonomi AE, Provost F, McCulloch D, et al. Quality improvement in chronic illness care: a collaborative approach. *J Qual Improvement* 2001;27:63-68.
- 19 Miles MB, Huberman AM. *Qualitative data analysis, second edition*. Thousand Oaks, CA: Sage Publications, 1994.
- 20 Berg BL. *Qualitative research methods for the social sciences*. Boston, MA: Allyn and Bacon, 2001.
- 21 Bates DW. Using information technology to reduce rates of medication errors in hospitals. *BMJ* 2000;320:788-91.
- 22 Buntix F, Winkens R, Grol R, Knottnerus JA. Influencing diagnostic and preventive performance in ambulatory care by feedback and reminders. a review. *Fam Pract* 1993;10:219-28.
- 23 Degeling P, Kennedy M, Hill M, Carnegie M, Holt J. *Professional sub-cultures and hospital reform*. Sydney: Centre for Hospital Management and Information Systems Research, University of New South Wales, 1998.

Unwarranted variations in healthcare delivery: implications for academic medical centres

John E Wennberg

Everyday clinical practice is characterised by wide variations that cannot be explained by illness severity or patient preference. Professor Wennberg examines the causes for these variations and suggests ways to remedy the situation

Academic medicine has had only limited success in improving the scientific basis of everyday clinical practice, even within the walls of its own hospitals. Patterns of practice among academic medical centres—as among other institutions—are often idiosyncratic and unscientific, and local medical opinion and local supply of resources are more important than science in determining how medical care is delivered. In short, after nearly 100 years of academic medicine as we know it, much of medicine in the United States remains empirical.

The evaluative clinical sciences—those disciplines whose role in medicine is to evaluate medical theory, understand patient preferences, and improve systems—are capable of improving the scientific basis of clinical practice and warrant high priority in the national research agenda and full adoption into medical school curriculums. These sciences are essential to the development of organised healthcare systems in the 21st century, not least because they expose unwarranted variations in care and can be used to remedy them.

I will begin with a summary of the facts of unwarranted variations in clinical practice, derived

Summary box

Much of clinical medicine remains empirical, and everyday practice is characterised by wide variations that have no basis in clinical science

Patients served by even the best academic centres (teaching hospitals) experience unwarranted variations in health care and health outcomes

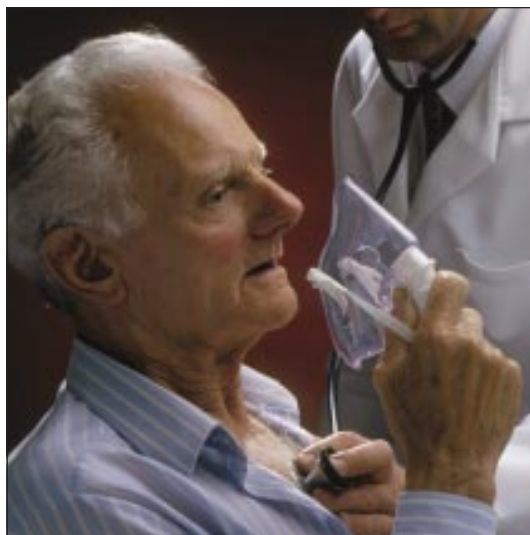
The evaluative sciences should be on national research agendas and medical school curriculums

Academic medical centres should start to lobby for this mandate and become advocates for reform

Center for the Evaluative Clinical Sciences, 7251 Strassenburgh, Dartmouth Medical School 03755-3863
John E Wennberg
director
john.wennberg@dartmouth.edu

BMJ 2002;325:961-4

from the Dartmouth Atlas of Health Care project, a US national study of traditional (fee for service) Medicare. The atlas project reports on the rates of use of resources and medical care by residents living in some



Patients with chronic conditions such as obstructive pulmonary disease experience unwarranted variations in care in different US regions

3436 hospital service areas, aggregated into 306 hospital referral regions, in which we examined unwarranted variations in three categories of clinical care.¹ The variations are unwarranted because they cannot be explained by type or severity of illness or by patient preferences. The categories are important because the causes of variation and their remedies differ according to category.

Unwarranted variations in clinical practice

Variations in effective care and patient safety

This category includes services whose effectiveness has been proved in clinical trials or well designed cohort studies and whose use does not involve substantial trade-offs that depend on patient preferences. In this category virtually all patients who are eligible for treatment should be treated; failure to treat thus represents underuse. The Dartmouth atlas documents systematic underuse, worse in some regions than in others, for each of 11 examples of effective care. A good example is the use of β blockers at time of discharge from hospital after a heart attack. The atlas shows variation among the 37 hospital referral regions containing one or more of the academic medical centres identified as the 50 best US hospitals for the treatment of cardiovascular disease by *US News and World Report*.² In the region with the best record only 83% of ideal candidates received a β blocker; in the region with the lowest record less than 40% did.

Variations in preference-sensitive care

This category includes conditions where two or more medically acceptable options exist and choice should depend on patient preferences. One example is the use of lumpectomy or mastectomy for treating early stage breast cancer. Another example is the treatment of early stage prostate cancer, for which the treatment options include surgery, various forms of radiation, and watchful waiting. The Dartmouth atlas shows that treatment patterns for these and other preference-sensitive conditions vary strikingly among regions.

The essential feature common to all preference-sensitive conditions is that choice of treatment should belong to the patients. We know from recent trials of decision support systems designed to help patients understand their treatment options that informed patient choice (shared decision making) results in different patterns of practice than that found with patients who experience usual care.³ In real practice, however, medical opinion rather than patient preference tends to dominate the treatment choice.

Variations in medical opinion among the faculties of academic medical centres create wide variations in the risk for elective surgery in the populations they serve. In our original studies of Boston and New Haven we observed that, although the overall rate of surgery was nearly identical, the rates for individual procedures varied remarkably, even though the populations were well matched demographically. Residents of New Haven were more than twice as likely to receive coronary artery bypass surgery and 50% more likely to undergo hysterectomy than their counterparts living in Boston; by contrast, Bostonians were two times more likely to undergo carotid artery surgery and 50% more likely to have their hip joints replaced than residents of New Haven.⁴⁻⁶

Variations in supply-sensitive care

Variations in supply-sensitive services are perplexing because medical theory and medical evidence play virtually no role in determining the relative frequency of their use among defined populations. Indeed, medical texts and journals and the everyday discourse among clinicians rarely mention the subject. Most of this variation arises from the frequency with which patients with chronic diseases use consultations, diagnostic tests, referrals to medical specialists, hospitalisations, and stays in intensive care units. For each of these services, the per capita quantity of healthcare resources allocated to a given population largely determines the frequency of use. For example, patients with chronic diseases such as diabetes, congestive heart failure, cancers, and obstructive pulmonary disease who live in regions with more doctors per capita will have more consultations and diagnostic tests (which are associated with consultations). If they live in regions with more beds they will experience more hospitalisations and stays in intensive care units.

The phenomenon of supply-sensitive services is particularly perplexing to academic physicians who believe they are practising according to the dictates of medical science. They find it difficult to believe that something they can't observe and are unaware of—the relative capacity of the healthcare system in which they practise—has a profound effect on their practice patterns. Yet, from the systems perspective, the phenomenon makes sense: given the proclivity of doctors to fill their appointments to capacity, it is easy to see how, at the level of the population, a doubling of the supply of internists or cardiologists results in roughly a halving of the interval between repeat visits.

The relative frequency of visits to medical specialists, hospitalisations, and stays in intensive care units during patients' last six months of life is particularly striking. The 31 hospital referral regions served by the best 50 hospitals (those with the highest reputations for coordinated and compassionate geriatric and palliative care)

offer highly variable quality of care in the last six months of life. In the regions with the lowest rate the consultation rate is less than four per person; in the highest regions the rate is more than five times that—23 consultations per person. The use of intensive care also varies remarkably. In the regions with the lowest rate less than 23% of people are admitted one or more times to intensive care during the last six months of life; in the highest region more than 45% are.

The search for a remedy for unwarranted variation

Although research projects conducted over the past 20 years provides successful examples of reducing under-service, improving patient safety, reducing scientific uncertainty, and improving the scientific and ethical basis for informed patient choice, these successes have not yet had much impact on the patterns of practice in the United States, even among academic medical centres. Much less progress has been made in the critical evaluation of supply-sensitive services. Delivery of effective care requires a practice based infrastructure with reminder systems to ensure, for example, that patients with diabetes get immunisations and eye tests. Successful programmes along these lines depend on interdisciplinary research, the development and maintenance of clinically relevant registries for long term monitoring of clinical care and relevant clinical outcomes, and rigorous scientific methods to test theories on the causes and remedies of performance variation. Success also requires a long term view: converting everyday practice into an environment that learns from experience requires a stable infrastructure, sustained funding, and a research agenda based on the testing of sequential hypotheses as to how to improve care.

There has been some progress in developing a scientific rationale for the frequency of consultations for patients with chronic illness. Experimental alterations in the way patients with chronic illnesses interact with their doctors have shown improved efficiency and healthcare outcomes compared with usual care (when follow up is scheduled by doctors). In one study doctors made periodic telephone calls to patients instead of scheduling regular follow up visits: patients in the study group experienced fewer hospitalisations and lower mortality than controls.⁷ In three other studies planned group visits to doctors were offered instead of one to one follow up visits for patients with chronic illness: those offered group visits achieved better control of diabetes, lower rates of use of emergency rooms and specialists visits, and improvement in other outcomes compared with controls.⁸⁻¹¹

While studies such as these would provide the basis for the redesign of clinical practice, they have been given little priority in US federal science policy and in the research agendas of academic medical centres. The lack of priority belies the importance of supply-sensitive services in determining variations in per capita spending in health care among regions. Among the 306 hospital referral regions, per capita Medicare spending varies more than twofold, even after adjustment of age, sex, race, illness, and price. Yet more spending does not necessarily result in more provision



RICHARD MELAP PHOTO



HENRYK T. KANSER/REX

Why are residents of New Haven (above) more than twice as likely to have coronary artery bypass surgery than residents of Boston (below), yet half as likely to undergo carotid artery surgery, when the two populations are well matched demographically?

of effective care or improved patient safety, as measured by 30 day mortality after bypass surgery. Nor does it result in more major surgery. It is the frequency of use of supply-sensitive services by chronically ill patients that distinguishes high cost regions from low cost ones, accounting, for example, for the more than twofold variation in per capita spending among the 31 hospital referral regions served by the best US hospitals for geriatric care.

Conclusion

The lack of priority given to the evaluative health sciences is depressing, given the growing evidence that, for populations, greater frequency of use does not improve health outcomes. Populations in high cost hospital referral regions with high frequency of use of services do not have longer life expectancy than those in low cost regions, even after controlling for differences in the prevalence of illness. Furthermore, more visits, more tests, more hospitalisations and

greater likelihood of dying in an intensive care unit mean that those living in high cost regions are unlikely to have better quality of life. For these reasons, disparities in spending should no longer be ignored. In the case of US Medicare, they signal gross inefficiency—more spending with no better, or possibly worse, global outcomes. Such variation is also unfair because workers and Medicare beneficiaries in low cost, more efficient regions subsidise the care of those in high cost regions. For example, the difference in lifetime Medicare spending for a typical 65 year old in Miami and one in Minneapolis is more than \$50 000 (£30 000, €50 000).¹

What must academic medical centres do to reclaim the authority of clinical science as the arbiter of the efficiency and effectiveness of supply-sensitive services? The first step is to acknowledge that their own patterns of practice in managing chronic illness provide no gold standard for excellence. Our studies of patterns of practice in teaching hospitals in Boston and New Haven provide the evidence.⁶ Hospitalisation rates for chronically ill patients using the Massachusetts General Hospital were 1.5 times higher than those for similar patients using the Yale-New Haven Hospital. Yet the Massachusetts General Hospital was the most efficient Boston teaching hospital: hospitalisation rates for patients using the New England Medical Center and Boston University Medical Center were 1.9 and 2.0 times that of Yale-New Haven Hospital.

An important second step would be the comparative study of different patterns of practice. Medicare claims, for example, could be used to assemble cohorts of patients with apparently similar diseases but treated at different centres, and these patients could be followed over time to establish the frequency of use of care. Case notes could be reviewed to record comorbidities and other patient factors relevant to predicting prognosis or measuring outcomes; this information would then be used to correct for possible differences in case mix when interpreting outcomes such as survival, complications from treatments, or quality of life.

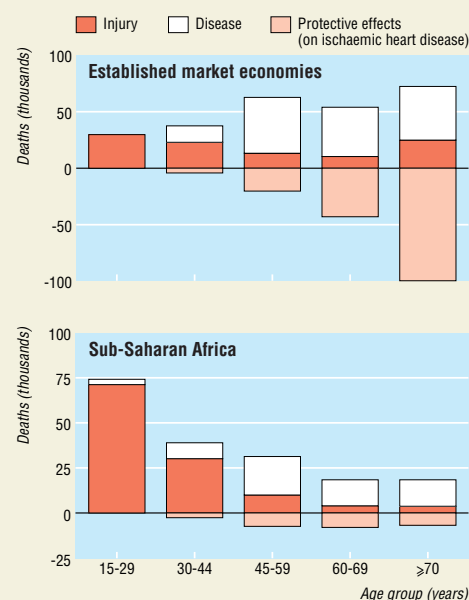
Although studies such as these would not prove definitive evidence (none ever do), they would put the issue of variation in supply-sensitive services on the agenda. Further inquiry would lead to the development and testing of hypotheses and the design and testing of remedies for unwarranted variation. Since the bulk of supply-sensitive services are used to treat a few chronic illnesses such as cancer, congestive heart failure, and chronic pulmonary obstructive disease, disease management interventions designed to improve the efficiency and effectiveness of such care would provide an important outlet for the development of a relevant clinical science. Clinical trials of such interventions are a promising development.

If academic medical centres succeed in meeting these challenges, new models of practice incorporating continuous improvement in process of care, shared decision making, and redesigned systems of care for chronically ill patients should rightly dominate clinical practice.

This article is an abridged version of the 13th annual Coggeshall lecture given by Professor Wennberg at the University of Chicago on 30 April 2002 (for more information, see www.dartmouthatlas.org).

- 1 Wennberg JE, Fisher ES, Skinner JS. Geography and the debate over Medicare reform. *Health Affairs* www.healthaffairs.org/WebExclusives/Wennberg_Web_Excl_021302.htm (accessed 11 Oct 2002).
- 2 America's best hospitals. *US News and World Report* 2001;131(3):44-58.
- 3 O'Connor AM, Rostom A, Fiset V, Tetroe J, Entwistle V, Llewellyn-Thomas H, et al. Decision aids for patients facing health treatment or screening decisions: a Cochrane systematic review. *BMJ* 1999;319:731-4.
- 4 Wennberg JE, Freeman JL, Culp WJ. Are hospital services rationed in New Haven or over-utilized in Boston? *Lancet* 1987;i:1185-8.
- 5 Wennberg JE, Freeman JL, Shelton RM, Bubolz TA. Hospital use and mortality among Medicare beneficiaries in Boston and New Haven. *N Engl J Med* 1989;321:1168-73.
- 6 Fisher ES, Wennberg JE, Stukel TA, Sharp SM. Hospital readmission rates for cohorts of Medicare beneficiaries in Boston and New Haven. *N Engl J Med* 1994;331:989-95.
- 7 Wasson J, Gaudette C, Whaley F, Sauvigne A, Baribeau P, Welch HG. Telephone care as a substitute for routine clinic follow-up [see comments]. *JAMA* 1992;267:1788-93.
- 8 Wagner EH, Grothaus LC, Sandhu N, Galvin MS, McGregor M, Artz K, et al. Chronic care clinics for diabetes in primary care: a system-wide randomized trial. *Diabetes Care* 2001;24:695-700.
- 9 Trento M, Passera P, Tomalino M, Bajardi M, Pomero F, Allione A, et al. Group visits improve metabolic control in type 2 diabetes: a 2-year follow-up. *Diabetes Care* 2001;24:995-1000.
- 10 Coleman EA, Eilertsen TB, Kramer AM, Magid DJ, Beck A, Conner D. Reducing emergency visits in older adults with chronic illness. A randomized, controlled trial of group visits [see comments]. *Effect Clin Pract* 2001;4(2):49-57.
- 11 Beck A, Scott J, Williams P, Robertson B, Jackson D, Gade G, et al. A randomized trial of group outpatient visits for chronically ill older HMO members: the Cooperative Health Care Clinic [see comments]. *J Am Geriatr Soc* 1997;45:543-9.

The burden of disease Deaths among men, attributable to and averted by alcohol, 1990



Reproduced with permission from Murray CJL, Lopez AD. *The global burden of disease*. Cambridge, MA: Harvard University Press, 1996

The figure at the top shows the effects of alcohol on overall mortality in the world's rich countries, those with established market economies. Here heart disease is a common killer, but deaths from injuries and violence are relatively rare, and alcohol may prevent about as many deaths as it causes. This figure doesn't, however, show the big disease burden caused by alcohol in rich countries through disability and premature death. In sub-Saharan Africa (bottom figure) the protective effect of alcohol is far outweighed by the injuries it causes.