

Heterogeneity: we can't live with it, and we can't live without it

Frank Davidoff

Correspondence to

Dr Frank Davidoff, Institute for Healthcare Improvement, 143 Garden Street, Wethersfield, CT 06109, USA; fdavidoff@cox.net

Accepted 13 October 2010

As hard as it is to believe now, it was permissible not long ago to sell a prescription drug (in the USA) without providing evidence that it worked; all that was needed was evidence it was safe. Then, Kefauver–Harris changed everything. Acting largely in response to the thalidomide tragedy, Estes Kefauver, a senator from Tennessee, and representative Oren Harris from Arkansas introduced the amendment that now bears their names. The amendment required convincing evidence of efficacy, as well as safety, before a drug could be brought to market.¹ It was signed into law by John F Kennedy on 10 October 1962.

The effect of Kefauver–Harris on clinical practice is unequivocal; its effect on the shaping of clinical evidence has been less obvious. Although it would be hard to prove a strong cause–effect relation between enactment of the amendment and emergence of controlled trials as the evidentiary ‘gold standard,’ the timing is right. Reports of controlled trials that meet reasonable methodological standards began to enter the medical literature in 1948, but their numbers were small until 1966—just a few years after Kefauver–Harris—at which point their numbers soared, increasing about fivefold over the next 14 years.² (The Cochrane Collaboration, which has been a powerful force in establishing the dominance of controlled trial evidence, clearly did not play a role in this surge, since it came into existence only in 1993.) The rapid rise in healthcare spending in the 1960s and 1970s almost certainly reinforced the need for hard evidence of efficacy to avoid wasting medical resources. However, it is hard to avoid the conclusion that after 1962, clinical research went increasingly where the money was, and the money was flowing increasingly from industry as it set about meeting the new legislative requirement for strong experimental evidence that its products actually work.

Since biological and social heterogeneity among trial participants interferes with the detection of true cause–effect relationships, controlled trials are carefully designed to minimise or eliminate heterogeneity’s effects, primarily by aggregating the outcomes from many participants rather than drawing inferences from outcomes in individual persons.³ This intentional ‘heterogeneity blindness’ in controlled trials has brought with it a progressive shift from qualitative documentation of the concrete—meticulous narrative case reports and case series describing illness in individual patients, which had dominated the medical literature for decades—to quantitative assessment of the abstract—sophisticated statistical inferences about summary effect size in average patients.

Determining effect sizes by merging outcome data from groups of study participants obscures the reality that clinical interventions rarely work for everyone and under all circumstances. In large part, clinicians are unable to face that reality because they lack information about the biological, experiential and environmental sources of that heterogeneity, and the impact of that heterogeneity on clinical responses. When an intervention known to be effective in study populations fails to work in a particular patient, clinicians therefore have little choice but to move to alternative interventions on the basis of pragmatic rules of thumb (heuristics).^{4 5}

The shift in medical evidence from fine-grained narrative to sweeping statistical inference calls to mind analogous shifts in perspective that have followed certain other landmark changes in information systems. For example, when satellite cameras became available, they aggregated billions of bits of information from the earth’s surface into images from space, which have created a new, emergent reality about land use, oceans,



This paper is freely available online under the BMJ Journals unlocked scheme, see <http://qualitysafety.bmj.com/site/about/unlocked.xhtml>

climate and human activity. Curiously, although those satellite images have created a new gold standard of evidence, we continue to value ground-level knowledge very highly, in contrast with the disdain we exhibit towards clinical evidence from case reports and case series; in fact, we easily accept the idea that these two levels of evidence about the physical world are complementary.

What will it take to bring the focus of clinical evidence back down from the homogenised ‘outer space’ of controlled trials to the ground level of individual patient experience, with its inherent variability? Two powerful forces could move us in that direction. First, the explosion in molecular genetics has reawakened interest in the possibility of ‘personalised medicine’ in which therapies are tailored to the specific biological and clinical circumstances of individual patients.⁶ The implications of building heterogeneity from genetic and other sources formally into clinical decision-making are at least as profound for the business model of the pharmaceutical industry as they are for clinical research, not least because such individuation would probably spell the end of the era of ‘blockbuster’ drugs. Second, it is increasingly clear that improving the performance of healthcare systems inherently involves social change. Since the social structure of individual care systems is at least as heterogeneous as the biology of individual patients, it is rapidly becoming clear that anyone who works to change the performance of healthcare systems ignores that heterogeneity at their peril.^{3 7} The epistemological value of case reports has in fact received increased scholarly

attention in recent years^{8–10}; moreover, twice as many clinical case reports are published annually as randomised clinical trials, and the numbers of both are increasing at about the same rate¹¹—distant rumblings that suggest the forces of individuation may be poised to modulate the homogenising influence of experimental study methods.

Competing interests None.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. FDA Overview. http://www.emedicinehealth.com/fda_overview/article_em.htm (accessed 8 Jun 2010).
2. McDonald S, Westby M, Clarke M, *et al*; Cochrane Center’s Writing Group on 50 Years of Randomized Trials. Number and size of randomized trials reported in general health care journals from 1948 to 1997. *Int J Epidemiol* 2002;31:125–7.
3. Davidoff F. Heterogeneity is not always noise. Lessons from improvement. *JAMA* 2009;302:2580–6.
4. McDonald CJ. Medical heuristics: silent adjudicators of medical practice. *Ann Intern Med* 1996;124:56–62.
5. Pauker SG, Wong JB. How (should) physicians think? A journey from behavioral economics to the bedside. *JAMA* 2010;304:1233–5.
6. Fraenkel L, Fried TR. Individualized medical decision making: necessary, achievable, but not yet attainable. *Arch Intern Med* 2010;170:566–9.
7. Bate P, Mendel P, Robert G. *Organizing for Quality. The Improvement Journeys of Leading Hospitals in Europe and the United States*. New York: Radcliffe Publishing. 2008.
8. Vandenbroucke JP. In defense of case reports and case series. *Ann Intern Med* 2001;134:330–4.
9. Vandenbroucke JP. Observational research, randomized trials, and two views of medical science. *PLoS Med* 2008;5:e67.
10. Glasziou P, Chalmers I, Rawlins M, *et al*. When are randomized trials unnecessary? *BMJ* 2007;334:349–51.
11. Bastian H, Glasziou P, Chalmers I. Seventy-five trials and eleven systematic reviews a day: How will we ever keep up? *PLoS Med* 2010;7:e1000326.