Should the NHS strive to eradicate all unexplained variation?

Variation exists in all aspects of health care. **Stephen Richards** argues that it is damaging to both quality of care and finances, but **Richard Lilford** believes that imposing uniformity risks stifling medical progress

Stephen Richards medical director, Oxfordshire Primary Care Trust, Oxford OX4 2LH
lesley.corfield@oxfordshirepct.nhs.uk

YES Medicine is an art informed by science, and the best medicine requires both the art to be well done and the science to be accurate. Every person is different, and variation is inherent to all things. But the NHS should strive to eradicate all unwarranted, unintended, and inappropriate variation.

Donald Berwick wrote the seminal paper on controlling variation in health care in 1991, building on the work of Walter Shewart from the 1930s.[1] Berwick refers not to unexplained variation but to unintended variation. Deming recognises two types of variation: common cause (the random variation that occurs in all processes) and special cause (arising from unexpected events or unplanned situations).[2] The NHS Institute for Innovation and Improvement refers to "natural variation" and "artificial variation"—again there is no use of the word unexplained.[3]

"Only by explaining the variation can we expect to improve quality and cost effectiveness"

Economic implications

The King's Fund has exposed large variation in how primary care trusts (PCTs) spend the £15.5bn (€17.2bn; $26bn) that they are responsible for.[4] In 2004-5 Islington PCT spent seven times more per head on mental health services than did Bracknell Forest PCT. What the paper cannot say is if this resulted in better outcomes. The paper highlights the variation and the need to understand that variation. There is a fourfold variation in spending by PCTs on cancer and an eightfold variation on musculoskeletal services.

Unwarranted variation should be eradicated for both financial and health reasons. This need has led to the establishment of the quality, innovation, productivity prevention agenda led by Jim Easton.[5] Analysis of the NHS's Better Care Better Value (BCBV) indicators suggest that "if every organisation improved its performance to match the top quartile in each BCBV indicator NHS England could realise £3.6bn."[3] Also from the economic stance, studies undertaken by

JETTA PRODUCTIONS/ICONICA/GETTY

Richard J Lilford professor of clinical epidemiology, School of Health and Population Sciences, University of Birmingham B15 2TT
r.j.lilford@bham.ac.uk

NO I assume that most people would agree that variation is acceptable when patients make different personal choices—for example, whether to have screening for Down's syndrome or prostate cancer. Likewise, I take it that variation is not acceptable when compelling evidence of cost effectiveness is available. These points I take as "self-evident truths" and turn to the interesting question that relates to the acceptability of eclectic practice when the evidence is not compelling.

When I started out in practice, some clinicians recommended bed rest for threatened miscarriage while others advised women to continue with normal activities. Some mothers were encouraged to nurse their babies prone, others supine. Should mandatory compliance with one or the other practice have been imposed in each case? If so, how could the decision have been made? At the time little salient empirical evidence was available so the opinion of the majority would have prevailed. In that case bed rest would have been recommended for threatened miscarriage and the prone position advised (or perhaps even mandated) for infants. Yet we now have strong evidence that bed rest does not reduce the probability of miscarriage and that risk of infant death is increased in the prone position.

In both scenarios premature enforcement of a standard that was poorly supported by the prevailing evidence would have thwarted assembly of the necessary data. In the event, lack of an enforced policy on sleeping positions generated two sufficiently large groups of infants that enabled rates of unexpected infant death to be compared with a high degree of precision. Likewise, Everett and colleagues perturbed previously entrenched beliefs concerning management of threatened miscarriage by showing non-uniform practice and opinion,[1] thereby generating sufficient equipoise to mount a clinical trial.[2 3]

Value of variation

The cost effectiveness of new treatments should be established before they are widely

All references are in the version on bmj.com

WHERE DO YOU STAND ON THE ISSUE?
Tell us on bmj.com

Rand Health suggest that 20-40% of health care is inappropriate.[6] To quote Berwick, "Unintended variation is stealing health care blind."[1]

Quality of care

Although the global financial downturn has greatly increased the need for cost effective use of health resource, that same pressure must also be used to improve the quality and safety of patient care.

In *Crossing the Quality Chasm* the US Institute of Medicine stated that "patients should receive care based on the best available scientific knowledge. Care should not vary illogically from clinician to clinician or from place to place."[7] The paper describes poor quality of care as overuse, misuse, and underuse (each of which can represent waste).

As a general practitioner I have a special interest in variation in prescribing. Every bar chart I refer to shows a similar pattern of variation across the 82 general practices in Oxfordshire. Bar charts do not distinguish common cause from special cause variation. However, using statistical process control can expose special cause variation for investigation in many therapeutic areas.[8] Once special cause variation has been explained or eliminated work can begin to reduce the range of common cause variation. Nearly every practice has one indicator where special cause variation is found. This does not necessarily imply poor quality, indeed the reverse may be true. If the range of variation was controlled there would be a saving of nearly £2m on just the four BCBV prescribing indicators.

Variation has also been found in performance in the Quality and Outcome Framework, which is used to determine performance related pay to general practices in England.[9] The clinical guidelines are not intended to apply to every patient, and practices are allowed to report exceptions. The average rate of exception reporting is 6%, but in the framework's first year one practice had an exception reporting rate of 89%. Subsequent investigation and support narrowed the gap. In the second year no practice had an exception rate above 27%, yet performance in both affluent and non-affluent areas continued to improve. PCTs have a mechanism for investigating practices with unusually high rates of exception reporting and they should use it.

A large amount of current variation remains unwarranted. Only by explaining the variation can we expect to improve quality and cost effectiveness. Yet knowing the cause of variation is only the first battle; the second is to convert that knowledge into action. Even when evidence has been accepted and adopted there is a third critical phase—that of informing the patient and the public. More control should rest with patients—for example, the decision to have a prostatectomy or not.

Applying statistical process control to all key examples of variation will be a challenge. Educating the profession will be a greater challenge. Embedding that knowledge into the decisions made with patients will be the hardest of all but critical to improving quality and driving down costs.

In the words of Donald Berwick, "The need is not so much for payers and regulators to force the medical system into uniformity . . . but rather for the profession and its leaders to recognize that there is embedded in this cacophony of practice so much waste and hazard that physicians simply owe it to themselves to reduce the variation wherever they can."[1]

Competing interests: None declared.

Cite this as: *BMJ* 2009;339:b4811

adopted in health services, but no matter how many randomised trials are done observational studies will still be needed .[4] There will always be more questions than answers, in part because for every piece of new knowledge, a string of secondary questions arises. For example, the National Institute for Health and Clinical Excellence (NICE) appraisal committee had plenty of evidence to recommend clopidrogel for prevention of recurrent stroke, but subsidiary questions, such as recommended duration of treatment and its use in patients with various comorbidities, were harder to answer.

Historically controlled studies are more biased than concurrently controlled observational studies,[5] but concurrent studies can contribute data on effectiveness and safety only if treatment is allowed to vary. The epistemic advantages of non-uniform practice do not entail any loss for patients. Different clinician-patient pairs are likely to reach different judgments concerning the effects of treatment. More than one treatment option is therefore acceptable when evidence is weak. Prior opinions about the effects of treatment can be captured in the form of prior probability estimates under a bayesian model. However, prior beliefs are updated as evidence accrues so that an individual becomes progressively less uncertain while the beliefs of different individuals converge.[6] Mothers are no longer sent routinely to bed for threatened miscarriage, and a much higher proportion of the world's babies are now nursed supine. It has been shown that variation in practice diminishes as uncertainty is reduced.[6 7] However, in the absence of compelling evidence different prior beliefs are rational and differences in practice do not offer a disservice to patients. They are therefore not inequitable.

A Boston physician was censured in the 19th century for not using blood letting to treat pneumonia.[8] Semmelweis was ridiculed for insisting on antisepsis in maternity wards.[9] There is a real risk that practice might ossify if uniformity is imposed when the evidence is insufficient to generate consensus (that is, before convergence of probability estimates has occurred). Eclectic practice uninformed by evidence is obviously undesirable. But different decisions in a given scenario are entirely justifiable and indeed desirable when the jury is still out.

"In the absence of compelling evidence different prior beliefs are rational and differences in practice do not offer a disservice to patients"

People with the power to impose their beliefs on others will always be tempted to do so.[10] However, post-enlightenment societies have succeeded because they allow the freedom to use individual judgment when there is no good reason to impose conformity. Competition between ideas can be resolved empirically when more than one option can be experienced. A randomised trial is but one method under which treatment may vary so that safety and effectiveness can be assessed.

Competing interests: None declared.

Cite this as: *BMJ* 2009;339:b4809