

A new learning environment: combining clinical research with quality improvement

Peter J. Pronovost MD¹ and Vahé A. Kazandjian PhD, MPH²

¹Departments of Anesthesiology and Critical Care Medicine and Surgery, Johns Hopkins School of Medicine, and Department of Health Policy and Management, Johns Hopkins University School of Public Health Baltimore, Maryland USA

²Adjunct Associate Professor, School of Public Health, Johns Hopkins University, and President Center for Performance Services, Inc., Lutherville, Maryland, USA

Correspondence

Peter Pronovost MD
Meyer 295
600 N. Wolfe Street
Baltimore
Maryland 21287-9106
USA

Keywords: case-control, clinical research, modified clinical trial, performance improvement

Accepted for publication:
17 July 1998

Abstract

The emphasis provided by quality improvement strategies on performance measurement and evaluation often results in our understanding of processes of care and, perhaps, better outcomes. There are different references for process evaluation: external peers, regional profiles of performance or a trending of one's own performance patterns. This paper proposes a methodology that enables learning from the daily practice of medicine by comparing alternative care processes and outcomes. Since it is estimated that 15–20% of medical practices are based on rigorous scientific data establishing their effectiveness, we have much to learn. We propose to learn from our daily practice by combining clinical research methods with quality improvement tools. The products comprise modified clinical trial and case-control studies. In a modified clinical trial, we would use a practice guideline as a control group and modify the guideline to create an experimental group. This method would maintain the internal validity of efficacy research while maintaining the external validity of effectiveness research. In the case-control method, it is possible to quantitate risk for a given outcome and focus improvement effort on factors associated with that outcome. We believe physicians will accept this learning approach because it is a more valid learning method than traditional quality improvement and, unlike randomized clinical trials, learning will occur in the daily practice of medicine.

Introduction

The goal of quality improvement in medicine is to create a learning environment where we use the daily practice of medicine to continually evaluate and improve care processes and, perhaps, outcome (Senge 1992; Berwick 1996). If learning is defined as the ability to make valid inferences between two or more alternatives, care providers seem to learn little from the routine practice of medicine because random variation in practice is greater than the effect

of most care processes (Wennberg *et al.* 1989; Wennberg 1995). To improve productivity, we must approach our daily work with the mental model: 'learning to learn'. This paper proposes a methodology that enables learning from the daily practice of medicine by comparing alternative care processes and outcomes. It is estimated that 15–20% of medical practices are based on rigorous scientific data establishing their effectiveness. Therefore, learning to improve quality of care implies learning about the science of medicine as much as its artful application

(Fink *et al.* 1987; Ropwe *et al.* 1988). Teaching institutions and university medical centres could champion this methodology and demonstrate its impact on performance.

TQM to control random variation

To reduce variation and improve quality, medicine has been experimenting with Total Quality Management (TQM) principles that have been successful in other industries (Shewart 1931; Demming 1986). Proponents of these principles state that if we reduce random variation, we will reduce waste and improve quality. TQM initiatives in medicine attempt to reduce random variation by creating standardized care processes such as practice guidelines or critical paths, which are common tools used in peer review and evaluation. Indeed, there are valid methods to create evidenced-based practice guidelines, and these guidelines have had some success (Field & Lohr 1992). The use of practice guidelines decreased mortality in patients with the adult respiratory distress syndrome, decreased deep wound infections after surgery and decreased length of stay (Horn & Hopkins 1994). Although TQM has reduced random variation in medicine, we propose that it has not created a learning environment. Evidence-based guidelines are rare and the evidence to support guidelines is often lacking (Blumenthal 1996). Additionally, TQM has generally used historical controls as a comparison group to evaluate treatment alternatives. The use of historical controls produces weak inferences about treatment alternatives, thus limiting our ability to improve practice guidelines.

Reducing random variation may not mean increasing appropriateness (choosing the right patients to receive the guideline). Appropriateness is often measured by the rate at which health services are received in a population, such as the rates of carotid endarterectomy, upper gastrointestinal endoscopy and coronary angiography in a population (Brook *et al.* 1990). Indeed, studies that evaluated the appropriateness of care in patients with different rates of health services have shown that the proportion of inappropriate case management may be higher within the 'lower rate' category compared to the 'higher rate', thus high variation, category (Chassin *et al.* 1987). For example, patients with a

myocardial infarction treated at a hospital with higher rates of angiography had better outcomes than patients treated at hospitals with lower rates (Selby *et al.* 1996). Thus, unless reduction in random variation is associated with an evaluative component, it will lack the scientific rigour to create a learning environment, i.e. make valid comparisons between treatment alternatives (Chassin 1996). Additionally, the relationship between structure, process and outcome in industry is deterministic, but this relationship in health care is usually unknown, mainly probabilistic, and overwhelmingly untested. Because of this, most practice guidelines are opinion-based rather than science-based and should be thought of as 'doing the same thing' rather than 'doing the right thing' (Field & Lohr 1992; Kramer & Shapiro 1997).

Additionally, the TQM approach may not evaluate 'risk' adequately. While TQM can identify outliers (both good and bad), it does not weigh and quantify the risk associated with structure and process variables that lead to that statistically distinct (outlier) status. The inability to assign risk limits our ability to target and replicate resources efficiently, or eliminate the factors with the highest risk for the outcome. For example, through TQM we can identify a clinic with high patient satisfaction, but we may not measure the factors that lead to high satisfaction. Without such explanatory variables, the satisfaction survey becomes an exercise, not a strategy towards performance improvement. However, TQM does reduce variation and if combined with clinical research methods can create a learning environment. We propose to combine TQM with clinical trials and case-control studies.

Learning cannot be achieved without a framework within which epidemiological findings are translated to clinically applicable principles. We propose that an epidemiological approach fulfils the very definition of 'what befalls the people', and the 'people' need not always be the population at risk in its entirety, but can be patients; a segment of the population with a medical need who already has access to the care system. Thus, the universe upon which learning is based has well-defined risk factors (comorbidities, demographics, severity) and enablers (socio-economic factors, cultural expectations). Although focusing on patients rather than the population at large predisposes to selection bias, we

propose to minimize the effect of such biases through a case-control approach and randomization, which we discuss later (Sackett 1979). Additionally, patients who present for health services are often the denominator of interest.

The importance of a rapprochement between epidemiology, clinical practice and quality improvement is receiving renewed interest (Sackett *et al.* 1985; Proctor, 1994; Kazandjian 1996). The rationale for incorporating public health methodologies into randomized clinical trials (RCTs) and use of the observed outputs as an educational curriculum is attractive. A specific model has already been developed for select haematological diseases in the Northern Region of the United Kingdom (Charlton *et al.* 1997). This model, the population-adjusted clinical epidemiology (PACE) strategy, emphasizes well-controlled studies to obtain valid numerators, adjusts the estimates for the relevant denominator population, and uses these findings to inform clinical practice in a continuous process. Although methodologically elegant, PACE can be applied only where data are available about each patient, where data can be trended, where utilization and case management data can be linked to outcomes and where a formal auditing system assures complete, accurate and timely data. Such systems are rarely encountered. The alternative is to create a performance assessment, evaluation, distribution and monitoring scheme around a disease registry. The cost-effort analysis of such a model, limited to a handful of clinical conditions, may be prohibitive. The cost-usefulness analysis of focusing on a few conditions may render the education curriculum a mere seminar. We therefore propose to incorporate the epidemiological principles within a performance improvement mindset in the form of modified clinical trials (MCTs).

Modified clinical trials

In industry, learning occurs through multiple randomized experiments which often use a factorial design. A factorial design is a type of experiment that tests the effect of more than one treatment (Moen *et al.* 1991). The equivalent of these randomized experiments in health care are randomized clinical

trials (RCTs). A randomized clinical trial is the most powerful technique to evaluate alternative therapies but is rarely used in health care because a trial is rigid, expensive and often not generalizable (Berwick 1989). To maintain internal validity, clinical trials have strict entry criteria and treatment protocols which limit their ability to be applied to everyday clinical practice (external validity) (Horn 1995). For these reasons, few treatments have been evaluated with clinical trials and learning is divorced from the daily practice of medicine.

However, clinical trials have many beneficial features and should not be abandoned. Since RCTs control for known and unknown confounders and treatment bias, they provide the strongest evidence for a treatment effect (Meinert 1986). A confounder is a variable that is associated with both a care process and outcome, and may distort the relationship between the process and outcome. For example, if we were evaluating mortality after coronary artery bypass surgery, age is a potential confounder because age is associated with the outcome (death) and age is also associated with risk of coronary artery bypass surgery. The most important confounders are severity of illness and comorbid diseases. Treatment bias is a systematic error that occurs when patients who are offered treatment are different from patients who are not offered treatment. For example, we enrol younger patients who are more likely to survive in a clinical study and exclude older patients who are more likely to die. Rather than abandon these methods, RCTs can be adapted and incorporated into TQM methodology in the form of modified clinical trials (MCTs).

Modified clinical trials allow learning from the daily practice of medicine. Indeed, alternative care processes can be compared through MCTs and inferences made about the effect of care process on outcome without the expense and burden of formal clinical trials. These modified clinical trials would increase external validity (generalizability) because we would include all patients who receive the practice guideline in the study. Additionally, randomization and the practice guideline (which would serve as a study protocol) would maintain internal validity. Thus, modified clinical trials would maintain the internal validity of efficacy studies, maintaining the external validity of effectiveness studies.

There are two steps to the conduct of a modified clinical trial. First, we must reduce variation in a process with a practice guideline or critical path. The receptivity of providers to the use of practice guidelines plays a critical part in the success of this step. Once interest in learning from a clinical trial is assured, and willingness to experiment with clinical guidelines is secured, TQM techniques can be used to identify, understand and manage variation. Secondly, we must compare alternative care processes and make inferences about the effect of these processes on outcome. To accomplish this, we would vary one (or more) of the care processes in the practice guideline and assign patients randomly to receive either routine (control group) or altered (treatment group) care processes. We would choose care processes that are not evidence-based and are likely to affect the outcome (Sackett 1989; Evidence-based Medicine Working Group 1992). To increase the rate of learning, we could alter multiple care processes at once (factorial design), especially if large numbers of patients are available.

The art of managing probabilities

As stated at the outset, most processes are causative in other industries but are correlative in medicine. Therefore, to make valid inferences about the effect of a care process, these modified clinical trials must be conducted with the same scientific rigour as larger clinical trials. Specifically, these MCTs should have the following characteristics: (1) be hypothesis-driven; (2) have a strict protocol for both treatment and control groups; (3) have a valid outcome measure; (4) have power and sample size calculations based on a primary outcome variable; (5) evaluate for potential confounders and interaction; (6) perform appropriate statistical analysis; and (7) draw valid conclusions. This stepwise methodology will allow us to learn from the daily practice of medicine.

For example, suppose a critical path for patients having a colectomy include H2 antagonists (Fig. 1). While this practice seems logical, there is little evidence to support it. We would randomize all patients in this critical path to receive H2 or no H2 antagonist.

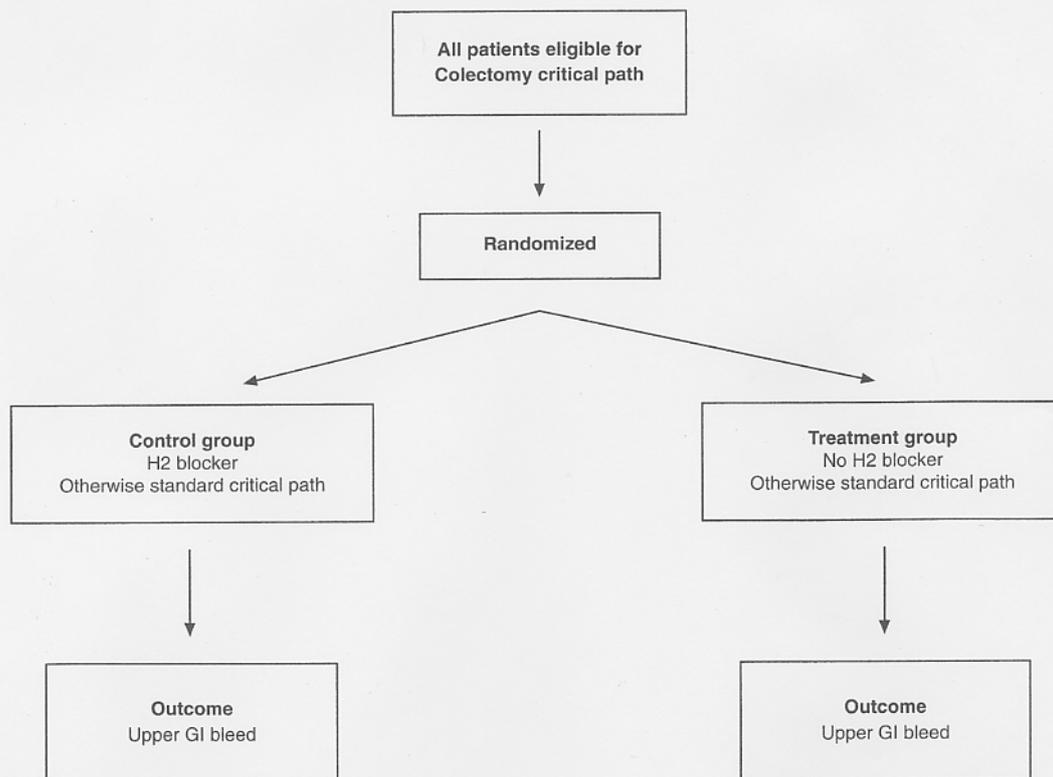


Figure 1 Model of modified clinical trial in colectomy patients.

In this example, randomization would occur in the pharmacy and patients would receive either an H2 antagonist (control group) or a placebo (experimental group). Our outcome measure would be an upper gastrointestinal (GI) bleed that requires a blood transfusion. In patients on this critical path, the incidence of upper GI bleed that requires a transfusion ($P1$ – probability in control group) is 1%, and we would calculate our sample size to demonstrate that $P2$ (probability in treatment group) is also 1%. We are already collecting data on adherence to the critical path and complications after surgery which include an upper GI bleed that requires a transfusion. Therefore, with little additional effort, we will be able to evaluate the effect of H2 antagonist in patients having colectomies and learn from the daily practice of medicine. The need for informed consent with this methodology must be addressed but is not the focus of this essay.

Case-control study

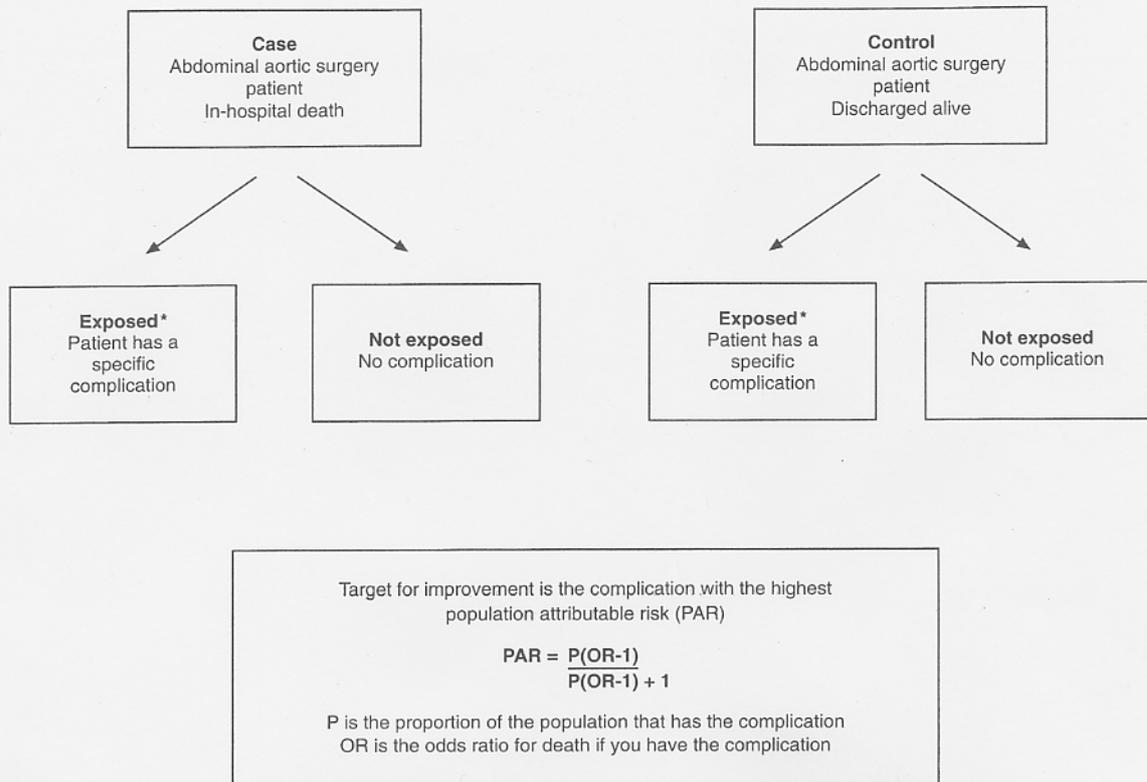
The case-control study is a well-tested epidemiological method which allows estimation of risks between care processes and outcome. In a case-control study, we compare a group of people with an outcome of interest (cases) to a group of people without the outcome of interest (controls). For example, we may identify a hospital that has a low mortality rate after aortic surgery, but this information does not tell us what care processes are associated with the low mortality. Therefore, we are unable to focus improvement efforts on care processes that are associated with low mortality. However, we can use the case-control method to evaluate the strength of the factors (risks) that lead to the good outcome. We found that patients who had a post-operative infection had a 2.8 times increased risk of death and an increase in total charges of US\$40 000. Additionally, patients who had a post-operative myocardial infarction had a 10 times increased risk of death (Pronovost 1997). We can use these data to focus improvement efforts by calculating the attributable fraction (population attributable risk) for each risk factor. Attributable fraction is the proportion of disease occurrence that could potentially be eliminated if the risk factor were prevented (Kelsey *et al.*

1994). We could focus improvement on the factors with the highest attributable fraction (Fig. 2).

Additionally, the case-control methodology can be used to identify risk factors for errors in clinical practice (Lilienfield & Lilienfield 1979) For example, to identify factors associated with medication errors, we could identify patients who had a medication error (cases) and those without medication errors (controls). The number of controls will depend on the number of cases and the desired power of the study. At the time the case is identified, we could choose 4–5 patients who were receiving medication but did not have an error. We choose 4–5 controls for every case to increase power, which is the ability to detect a difference if one truly exists (Cochran & Cox 1957). We will then choose exposure factors from the available literature and expert opinion. In this example, exposure factors could include ordering the medication, transcribing the order, processing the order and dispensing the medication. Using a case-control method we can assign risk (odds ratio) to the exposure variables that are associated with a medication error. This method moves beyond traditional quality assurance because it identifies risk factors on which we can focus improvement efforts. The immediate benefits of this case-control design, contrasted to a traditional design to study errors, is in the smaller sample size needed for the former approach. Indeed, the case-control design achieves similar power with fewer cases than the traditional quality assurance approach, which often does not select the cases with strict criteria for risk stratification. Ultimately, through the case-control method, we can eliminate (or replicate) the care processes strongly associated with the outcome.

Requirements for this new methodology

The combination of TQM and clinical research will require an integrated clinical information system, or health informatics. The specific requirements of the informatics system will depend on the health-care delivery level where the methodology is to be applied: clinic, hospital or integrated health systems level. To use this methodology at the clinic or hospital level would require either chart abstractors or a clinical information system. Many hospitals are already monitoring adherence to practice guidelines



* In this example, exposure is defined as having an in-hospital complication. For example, nosocomial pneumonia, postoperative infection, or line sepsis. Exposure can be any structure or process of care.

Figure 2 Case control method.

and have clinical information systems. To use this methodology at the health system level would require a large clinical information system. Many managed-care organizations and health systems are collecting process and outcome data on patients who are treated according to a practice guideline. On either level, we must maintain the same data quality standards as clinical research. Nevertheless, little is written about quality control in clinical information systems. However, as the demand for information about performance and quality amplifies within the realm of social accountability (e.g. the recent deaths of 29 babies in the Bristol Royal Infirmary) the need for performance indicators and clinical information systems will increase (National Committee for Quality Assurance 1995).

The combination of clinical research with TQM

will also require a multi-disciplinary team of administrators, clinicians, clinical trial methodologists and biostatisticians. While large managed-care organizations do not currently have the resources to perform these studies, university medical centres are uniquely qualified to implement this methodology and become centres of excellence in quality of care research (Lawrence 1998). In the competitive health-care market, university medical centres are seen as a liability because their average costs are 15–30% above competitors. This parochial view fails to acknowledge university medical centres' intellectual capital and value for improving quality. This new methodology will also require capital investment. The required capital could come from the following sources: (1) management in an effort to improve quality; (2) consultation fees for quality of care

research from managed-care organizations, disease management companies or integrated delivery systems; (3) university medical centre (acting as an insurer); and (4) government.

Summary

Through the union of clinical research with TQM, we can learn continually by evaluating the association between care processes and outcome. This strategy will require knowledge of both clinical research and TQM and will allow us to learn from the daily practice of medicine. This paper proposes the usefulness of MCTs and case-control studies to the learning methodology based on daily practice analysis (Fig. 3).

We believe that physicians will accept this learning approach because it is a more valid learning method than TQM and, unlike RCTs, learning will occur in the daily practice of medicine. The methods discussed in this paper capture aspects of performance and daily practice as they take place, without the effect of observer-over-the-shoulder, or the resulting Hawthorn effect (Lied & Kazandjian, in press).

Acknowledgements

The authors thank Dr Haroutune Armenian for his thoughtful comments on this manuscript and Abbott Pharmaceuticals for an educational grant in partial support of Dr Pronovost.

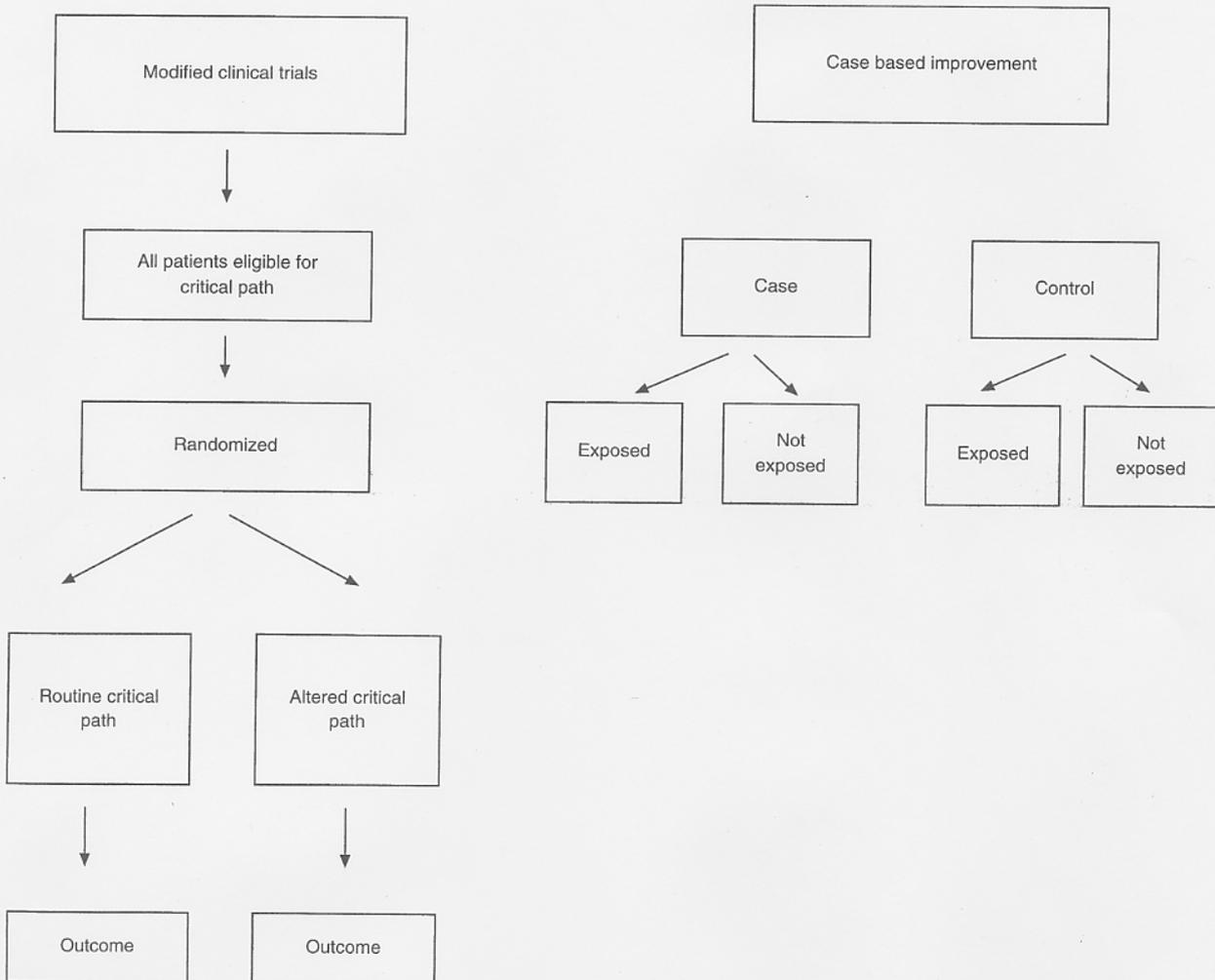


Figure 3 Epidemiologic methods to learn from the daily practice of medicine.

References

- Berwick D.M. (1989) Continuous improvement as an ideal in health care. *New England Journal of Medicine* **320**, 53–56.
- Berwick D.M. (1996) Harvesting knowledge from improvement *Journal of the American Medical Association* **122**, 877–878.
- Blumenthal D. (1996) Quality of health care: part 4. The origins of the quality-of-care debate. *New England Journal of Medicine* **335**, 1146–1149.
- Brook R.H., Park R.E., Chassin M.R. *et al.* (1990) Predicting the appropriate use of carotid endarterectomy, upper gastrointestinal endoscopy, and coronary angiography. *New England Journal of Medicine* **323**, 1173–1177.
- Charlton B.G., Taylor P.R., & Proctor J.S. (1997) The PACE (population-adjusted clinical epidemiology) strategy; a new approach to multicenter clinical research. *Quarterly Journal of Medicine* **90**, 147–151.
- Chassin M.R., Kosecoff J., Park R.E. *et al.* (1987) Does inappropriate use explain geographic variation in the use of health services? A study of three procedures. *Journal of the American Medical Association* **258**, 2533–2537.
- Chassin M.R. (1996) Quality improvement nearing the 21st century: prospects and perils. *American Journal of Medical Quality* **11**, s4–7.
- Cochran W.G., & Cox G.M. (1957) *Experimental Designs*, 2nd edition. John Wiley and Sons Inc., New York.
- Demming W.E. (1986) *Out of the Crisis*. Massachusetts Institute of Technology, Center for Advanced Engineering Study, Cambridge, MA.
- Evidence-based Medicine Working Group (1992) Evidence-based medicine: a new approach to teaching the practice of medicine. *Journal of the American Medical Association* **268**, 2420–2425.
- Field M.J., & Lohr K.N., eds. (1992) *Guidelines for Clinical Practice: From development to use*. National Academy Press, Washington, DC.
- Fink A., Brook R.H., Kosecoff J. *et al.* (1987) Sufficiency of the medical literature on the appropriate use of six medical and surgical procedures. *Western Journal of Medicine* **147**, 609–614.
- Horn S.D., & Hopkins D.S. (1994) *Clinical Practice Improvement: A new technology for developing cost-effective quality care*. Faulkner and Gray, Washington, DC.
- Horn S. (1995) Using clinical practice improvement: an interview with Susan Horn. *Journal Quality Improvement* **21**, 301–308.
- Kazandjian V.A. (1996) *The Epidemiology of Quality*. Aspen Publishing.
- Kelsey J.L., Whittemore A.S., Evans A.S. *et al.* (1996) *Methods in Observational Epidemiology*. Oxford University Press, Oxford.
- Kramer M.S. & Shapiro E.D. (1997) Management of the young febrile patient: a commentary on recent practice guidelines. *Pediatrics* **100**, 128–134.
- Lawrence, J. 'How can we watch the doctors?' *Independent*, Sunday, 31 May, 1998.
- Lied T.R. & Kazandjian V.A. 'A Hawthorne strategy for performance measurement and improvement'. *Clinical Performance and Quality in Health Care*, in press.
- Lilienfeld A.M. & Lilienfeld D.A. (1979) A century of case-control studies. *Journal of Chronic Disease* **32**, 5.
- Meinert C.L. (1986) *Clinical Trials: Design, conduct, and analysis*. Oxford University Press, New York.
- Moen R.D., Nolan T.W. & Provost L.P. (1991) *Improving Quality Through Controlled Experimentation*. McGraw-Hill Inc., New York.
- National Committee for Quality Assurance (1995) *Health Plan Employer Data and Information Set and User's Manual*, Version 2.0. National Committee for Quality Assurance, Washington, DC.
- Proctor S.J. (1994) From research to practice. *Lancet* **344**, 417–418.
- Pronovost P.J. (1997) Unpublished data.
- Ropwe W.L., Winkenwerder W. & Hackbarth G.M. (1988) Effectiveness in health care: an initiative to evaluate and improve medical practice. *New England Journal of Medicine* **319**, 1197–1202.
- Sackett D.L. (1979) Bias in analytic research. *Journal of Chronic Disease* **32**, 51–63.
- Sackett D.L., Haines R.B. & Tugwell P. (1985) *Clinical Epidemiology*. Little Brown and Co., Boston.
- Sackett D.L. (1989) Rules of evidence and clinical recommendations on the use of antithrombotic agents. *Chest* **95**, 2S–4S.
- Selby J.E., Fireman B.H., Lundstrom R.J. *et al.* (1996) Variation among hospitals in coronary-angiography practices and outcomes after myocardial infarction in a large health maintenance organization. *New England Journal of Medicine* **335**, 1888–1896.
- Senge P. (1992) Building learning organizations *Journal of Quality Participation* **March**, 30–38.
- Shewart W.A. (1931) *Economic Control of Quality Manufactured Product*. Van Nostrand Co Inc., New York.
- Steinwachs D.M., Wu A.W. & Skinner E.A. (1994) How will outcomes management work? *Health Affairs* **Autumn**, 153–162.
- Wennberg J.E., Freeman J.L., Shelton R.M. *et al.* (1989) Hospital use and mortality among Medicare beneficiaries in Boston and New Haven. *New England Journal of Medicine* **321**, 1168–1173.
- Wennberg J.E. (1995) Variation in medical practice and hospital costs. *Connecticut Medicine* **49**, *Journal of the American Medical Association* 444–453.