HEALTH TECHNOLOGY ASSESSMENT

HOW DO YOU ASSESS OBSERVATIONAL RESEARCH, OUTCOMES RESEARCH, AND CLINICAL AUDIT?

MODULE 5

Workshop Manual
March 2006

Surgery Strategic Clinical Network: Evidence Decision Support Program

This Project was funded by:
Canadian Coordinating Office for Health Technology Assessment (CCOHTA)

WELCOME

Welcome to the **fifth** module of six in a series on Health Technology Assessment (HTA). The primary objective of this fifth module and workshop is to provide you with an overview of how to assess observational research, outcomes research and clinical audit for Health Technology Assessment (HTA).

We hope that the fundamentals presented in this module will not only assist you in your own search for a wide variety of health research information that may be considered in Health Technology Assessment, but also provide you with the tools required to critically evaluate research in a sound, objective, and appropriate manner.

We look forward to sharing this experience with you and your colleagues. Your feedback and comments on both the module and workshop will be greatly appreciated! Please send feedback and comments to the office of Surgical Research at osr@ucalgary.ca

Prepared By:

Elizabeth Oddone Paolucci

Tyrone Donnon

Paule Poulin

The HTA Education Team Members:

Paule Poulin, PhD
Office of Surgical Research, Associate Director, Calgary Health Region

Tyrone Donnon, PhD
Medical Education & Research Unit, Assistant Professor, Faculty of Medicine

Elizabeth Oddone Paolucci, PhD
Medical Education & Research Unit, Educational Consultant

Norm Schachar, MD, FRCSC
Office of Surgical Education, Director, Department of Surgery

Anita Jenkins, RN, BN, MN
Office of Surgical Education, Education Coordinator, Department of Surgery

David Sigalet, MD, PhD., FRCSC
Office of Surgical Research, Director, Department of Surgery

CONFLICT OF INTEREST

Conflict of interest is considered to be financial interest or non-financial interest, either direct or indirect, that would affect the research contained in a given report, or creation of a situation where a person's judgment could be unduly influenced by a secondary interest such as personal advancement. Indirect interest may involve payment which benefits a department for which a member is responsible, but which is not received by the member personally, e.g. fellowships, grants or contracts from industry, industry sponsorship of a post or a member of staff in the department, industry commissioning of work.

Based on the statement above, no conflict of interest exists with the author(s) and/or external reviewers of the fifth module.

ACKNOWLEDGMENTS

This Health Technology Assessment Module was prepared for the **Department of Surgery** by the **Office of Surgical Research,** the **Medical Education and Research Unit,** and the **Office of Surgical Education** and funded by the **Canadian Office for Health Technology Assessment (CCOHTA)**, a National-based organization with a mandate to facilitate the appropriate and effective utilization of health technologies within health care systems across Canada. CCOHTA's mission is to provide timely, relevant and rigorously derived evidence-based information to decision makers and support for the decision-making processes. The HTA Team wishes to thank Ms. Darlene Bridger and Ms. Linda Davis for their administrative support.

This Project was funded by:
Canadian Coordinating Office for Health Technology Assessment (CCOHTA)

TABLE OF CONTENTS

1.0 OBJECTIVES

Welcome to the fifth module in the Health Technology Assessment Workshop series.

This module includes an overview of how to assess observational and outcome research, and clinical audit.

The primary objectives of this workshop are to:

(1) Define observational research, outcomes research, and clinical audit;

(2) Examine types of observational research and the methods employed in conducting it;

(3) Describe outcomes information for decision-making and how to measure outcomes; and

(4) Identify the audit cycle and evaluate clinical audits.

After reading this module, you should be able to:

(1) Differentiate between observational and outcomes research, and clinical audit;

(2) Describe the types of observational research and the methods employed in conducting it;

(3) Identify outcomes information for decision-making and how to measure outcomes; and

(4) Describe the audit cycle and identify both the benefits and drawbacks to clinical audits.

2.0 INTRODUCTION

In the previous module (i.e., Module 4), we saw that conventional scientific wisdom favors large randomized controlled trials (RCTs) to collect customized data (Stevens, Raftery, & Roderick, 2005). As Stevens, Raftery and Roderick (2005) indicated, concealed randomized allocation of interventions prevents selection bias and controls for confounders, both known and unknown. Therefore, the randomized clinical trial occupies a central role in the assessment of treatment efficacy (Kestle, 1999). The most important principle of clinical trial design is to answer one question well, reserving secondary questions and analyses for the purpose of generating, not testing, hypotheses (Kestle, 1999).

However, there are also limitations to clinical trials. In many instances, RCTs are not the feasible approach to employ. In other cases, clinical trials are not the appropriate design to answer a research question. Consequently, depending on the research or HTA question, observational studies can provide a potential alternative or complement to RCTs, or in fact, can be the most appropriate research design for particular research or HTA question. Therefore, given the unique issues that arise when collecting primary data, as well as the considerable debate about what kinds of study designs are "good enough" for addressing important HTA questions of effectiveness, we explore the potential of observational studies, outcomes research, and clinical audits in health technology assessment.

3.0 WHAT IS OBSERVATIONAL RESEARCH, OUTCOMES RESEARCH, AND CLINICAL AUDIT?

Evidence-based medicine, evidence-based surgery, health services research, outcomes research, and practice guidelines are all terms of the 1990s that were virtually unknown to the practicing surgeon (McLeod, 1999). Clinicians are now forced to evaluate what they do, how they do it, and in whom, with increasing frequency and rigorousness. No longer is evidence from personal case series acceptable (McLeod, 1999).

Evidence-based medicine is defined as "the conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients" (Sackett as cited in McLeod, 1999). Although there is acceptance that individual practice must be more evidence-based, often it is difficult for the individual clinician to be aware of the best evidence and critically evaluate it (McLeod, 1999). Still, there continues to be increasing interest in the evaluation of surgical procedures. Coupled with the pressure of using the limited health care resources more wisely, it will be necessary that all surgeons have some knowledge about the fundamentals of health services research and related topics (McLeod, 1999).

Advocates of evidence-based medicine classify studies according to "grades of evidence" or "Level of evidence" on the basis of the research design, using internal validity (i.e., the correctness of the results) as the criterion for hierarchical rankings (Concato, Shah, & Horwitz, 2000). The highest grade is reserved for research involving "at least one properly randomized controlled trial", and the lowest grade is applied to descriptive studies (e.g., case series) and expert opinion; observational studies, both cohort studies and case-control studies, fall at intermediate levels (Concato, Shah, & Horwitz, 2000). Although the quality of studies is sometimes evaluated within each grade, each category is considered methodologically superior to those below it (Concato, Shah, & Horwitz, 2000).

Observational and outcomes research, and clinical audits are forms of non-experimental research. In these kinds of studies, it is not possible to determine a cause-effect relationship because the researcher does not control the variables under investigation. Instead it might be possible to find correlations between observed variables, in order to determine relationships or associations between them.

We begin our examination of these methods of collecting information with a definition of observational and outcomes research, and clinical audit. Following these definitions, the follow sections will include a more detailed description of each of the respective methods.

3.1 DEFINITION OF OBSERVATIONAL RESEARCH

It is not always clear what is meant by observational research. It can refer to the use of observation as a technique for gathering data about behavior in a study, which might in general be referred to as an experimental design. On the other hand, "observational" might refer to the overall design of a study, in contrast to a controlled experiment. For harmful interventions, randomization is unethical so it is necessary to observe the results of "natural" treatment in observational studies. Therefore, observational studies are used to examine how exposure to risk factors influences the probability of developing a disease.

Essentially, an observation requires a clinician to enter a situation where some behaviour of interest is likely to take place, to watch the nature and frequency with which particular forms of behaviour occur, and to make a record of what is observed. Eventually, the record of observations is used to help answer a particular research question. However, observation does not just mean casually watching something, because issues such as definition of units of behavior or observations, structure, procedure, interpretation, and recording are crucial. In an observational study the emphasis is on the overall nature of the study being non-experimental and on simply observing the naturally and feely occurring behavior of people, with or without their knowledge.

Observational studies are used for detecting potential risk factors of health care problems and to examine how exposure to risk factors can influence the probability of developing disease. In this type of research there is no attempt by the research to manipulate any independent variable (risk factors), although it is still possible to test hypotheses. The researcher does not intervene in any way, but rather simply records data as unobtrusively as possible.

Some common reasons for conducting observational studies instead of experiments are that experiments may seem too artificial and lack ecological validity, and that sometimes observation is the only possible way and the most appropriate method to study certain research questions. Observational study designs are frequently used in epidemiology, but have other applications as well. Some drug treatments may have to be assessed over a long period of time and a clinical trial is not suitable. Therefore, the long-term performance of a drug can be assessed in this way. Also, safety assessment of a drug is often performed in an observational study design framework (White, Ashby, & Brown, 2000).

According to Black (1996), observational research refers to quantitative, epidemiological methods. Accordingly, the principal observational methods are non-randomised trials of cross-sectional studies (prevalence), cohort studies (prospective and retrospective longitudinal) and case-control methods.

3.2 DEFINITION OF OUTCOMES RESEARCH

Outcomes research refers to evaluation that focuses on: (1) the status of participants after receiving care, and (2) the process of care itself (http://www.dcri.duke.edu/patient/glossary.jsp).

Outcomes research seeks to understand the end results of particular health care practices and interventions. End results include effects that people experience and care about, such as change in the ability to function. In particular, for individuals with chronic conditions – where cure is not always possible – end results include quality of life as well as mortality. By linking the care people get to the outcomes they experience, outcomes research has become the key to developing better ways to monitor and improve the quality of care (http://www.ahrq.gov /clinic/outfact.htm).

The urgent need for outcomes research was highlighted in the early 1980s, when researchers discovered that "geography is destiny" (http://www.ahrq.gov/clinic/outfact.htm). Time and again, studies documented that medical practices as commonplace as hysterectomy and hernia repair were performed much more frequently in some areas than in others, even when there were no differences in the underlying rates of disease. Furthermore, there was often no information about the end results for the patients who received a particular procedure, and few comparative studies to show which interventions were most effective (http://www.ahrq.gov/clinic/outfact.htm). These findings challenged researchers, clinicians, and health systems leaders to develop new tools to assess the impact of health care services (http://www.ahrq.gov/clinic/outfact.htm).

3.3 DEFINITION OF CLINICAL AUDIT AND CLINICAL RESEARCH

"Audit" is a term that has acquired different meanings over time in relation to health care quality. In fact, a clinical audit can be used to systematically examine all aspects of patient care from assessment through to outcomes (Smith, 1992). Surgical audit has been defined as the "systematic critical analysis of the quality of medical care, including the procedures used for diagnosis and treatment, the use of resources, and the resulting outcome and the quality of life for the patient, carried out by those personally engaged in the activity concerned" (http://www.edu.rcsed.ac.uk/Lectures/Lt17.htm).

Sometimes clinical audit is confused with clinical research because the two disciplines have much in common. Research is about creating new knowledge; knowledge about whether new treatments work and whether certain treatments work better than others. Research

forms the basis of nationally agreed clinical guidelines and standards – it determines what best practice is (Smith, 1992). Clinical audit is a way of finding out whether we are doing what we should be doing. Are we following guidelines and are we applying best practice? (Smith, 1992). The borderline between audit for improvement of clinical practice and audit for research is thin. Many surgeons start auditing their practices, find deficiencies that then lead to different ways of doing things (i.e., research). Scientific research asks: "Are we doing the right operation", while audit research asks: "Are we doing the operation right?" (http://www.edu.rcsed.ac.uk/Lectures/Lt17.htm).

Clinical audit and clinical research both involve: answering a specific question relating to quality of care; can be carried out either prospectively or retrospectively; involve careful sampling, questionnaire design, and analysis of findings; and should be professionally led (Smith, 1992). However, in contrast to research, **clinical audit** answers the question: "Are we following best practice"; it measures against standards rather than based on a hypothesis; is usually carried out on a relatively small population over a short time span; never involves a completely new treatment; never involves anything being done to patients beyond their normal clinical management; never involves allocation of patients to different treatment groups; and depending on circumstances, may be pragmatically based on a sample size that is acceptable to senior clinicians (Smith, 1992).

Thus, clinical audit is not research, but does make use of research methodology in order to assess practice (Smith, 1992). Clinical audits ensure that the new knowledge created from research about whether treatments work or whether one treatment works better than another, is being used to best effect. Clinical audits usually look at processes (i.e., whether we are doing the things we should be doing), but can also look at outcomes (i.e., whether those processes are producing the "right" results (www.ubht.nhs.uk/Clinical Audit/questions.htm). The process of conducting clinical audit sometimes identifies areas where new research is needed.

Clinical audits are usually a multi-disciplinary activity and multi-sectoral (i.e., may involve health and social services, primary and acute care providers, education and health; Smith, 1992). Aspects of patient care (e.g., structure, processes, and outcomes) are selected and evaluated against explicit criteria (Jones & Cawthorn, 2002). Where necessary, changes are then implemented at an individual, team, or service level. Further monitoring can then be used to confirm the improvements in healthcare delivery (Jones & Cawthorn, 2002). When conducted well, clinical audit provides a way in which the quality of care can be reviewed objectively, within an approach that is supportive and developmental.

4.0 OBSERVATIONAL RESEARCH

As indicated in Module 3, in biomedical research, observational studies are used for detecting risk factors of health care problems and to examine how exposure to risk factors influences the probability of developing disease. In this type of research there is no attempt by the research to manipulate any independent variable, although it is still possible to test hypotheses. The researcher does not intervene in any way, but rather simply records data as unobtrusively as possible.

The main problem with the use of observation is that it remains difficult to establish causal links between variables. The lack of control over biases and potential confounding factors means there is always a possibility that some unknown factor is exerting an influence over the dependent variable. Observational studies are likely to be influenced by selection bias because of non-random sampling. The selection of a comparable control group is one of the most difficult decisions facing the authors of an observational (case-control or cohort) study. Few such studies succeed in identifying two groups of subjects who are equal in age, sex, mix, socioeconomic status, presence of coexisting illness, etc., with the single difference being their exposure to the agent being studied (Greenhalgh, 1997).

Observational study designs range from relatively weak studies like descriptive studies to strong designs like case control and cohort studies (Pai, 2006). In the following section we examine various types of observational research in greater detail.

4.1 TYPES OF OBSERVATIONAL RESEARCH

There are different kinds of observational studies. The most common types of observational studies to assess risk factors for disease, are correlational studies, cross-sectional studies, the case report, case series, cohort, and case-control study designs. Each has distinctly different designs and differs in its advantages and disadvantages. Each of the study types will be described in the following sections.

(1) Descriptive or Correlational Studies (Case-report, Case-series)

Considered the weakest epidemiological design because they make no attempt to link cause and effect and thus, no causal association can be determined. The investigators simply describe the health status of a population or characteristic of a number of patients. Description is usually done with respect to time place and person (Pai, 2006). A case-report or case-series are examples of descriptive studies.

Correlational or descriptive studies are often the first step to a well-designed epidemiological study since they allow the investigator to define a good hypothesis which can then be tested using a better design (Pai, 2006). Therefore, these designs allow researchers to isolate possible causes for experiments to then assess. No matter how convincing data from descriptive and correlational studies may sound, because they have less control over the variables and environments that they study, these non-experimental designs cannot rule out extraneous variables as the cause of what is being observed.

(2) Case Report and Case Series

Case reports and case series are examples of descriptive studies as mentioned above. These studies describe a case or group of cases. Because they lack a comparison group, hypotheses cannot be tested (Peipert & Phipps, 1998). A case report involves the presentation of one case, while a case series refers to a collection of case reports that include details of the procedure used to treat the patients and the patient's outcomes. Case series are an extension of case reports, in that they are a description of a small number of individuals that have a similar experience with regards to a particular outcome and disease, rather than a documentation of just a single case as seen in case reports (Torabinejad & Bahjri, 2005). In both descriptive evaluations, however, there are no control groups, and so one cannot conclude with reasonable confidence that the observed outcome is because of the given exposure (Torabinejad & Bahjri, 2005).

These two types of descriptive observational studies are usually used to document unusual occurrences of outcomes and constitute the most common types of articles in medical journals (Torabinejad & Bahjri, 2005). They usually represent the first clues of new diseases or adverse effects of exposures. In fact, the surgical literature most commonly presents case series, which can be done by any physician who collects outcome information (Hartz & Marsh, 2003). These case series are useful for determining long-term outcomes of a given procedure. They are also useful for addressing complex questions of factors influencing outcome. A case series may show that patient outcomes are much better with a new procedure than historic results (Hartz & Marsh, 2003). However, we must remember that causality with the new procedure and outcome can not be established with descriptive studies.

Case series should include consecutive patients to avoid the possibility that only patients with the best results are kept in the data set (Hartz & Marsh, 2003). If the series is not consecutive, criteria for selecting patients should be defined carefully. Comparisons of case series may be misleading because patient risk and patient treatment may vary across studies in unknown ways (Hartz & Marsh, 2003). In particular, results from a case series may be better than historic controls because many

treatment techniques have improved, not just the technique described in the case series (Hartz & Marsh, 2003).

(3) Cross-Sectional Studies

Cross-sectional studies are sometimes described as a "snapshot" of time in that they measure exposure and outcome at the same (single) point in time, on the same individuals (Adamson, 2004; Peipert & Phipps, 1998). They are widely used to estimate the prevalence of disease or the prevalence of exposure to risk factors or both (thus, also called prevalence studies). Estimating the prevalence of conditions can be useful in predicting the need for health service use.

These studies are usually used before other analytical, observational, or experimental studies when little knowledge is available about the association of the particular outcome with the exposure under investigation (Torabinejad & Bahjri, 2005). Cross-sectional studies are useful for hypothesis generating, but cannot test hypotheses or provide evidence for causality (Peipert & Phipps, 1998). The main problem with respect to causality stems from the fact that both the exposure and the outcome are measured simultaneously. Thus, it is difficult to make any causal association (Pai, 2006).

These types of studies are easy to do and tend to be economical since repeated data collection is not done (Pai, 2006; Peipert & Phipps, 1998; Torabinejad & Bahjri, 2005). Measurements are made on a population at one point in time. The major disadvantage of cross-sectional studies is that the information about exposure and outcome (disease) are collected simultaneously, so it is difficult to sort out the temporal relationship (Peipert & Phipps, 1998). Since there is no longitudinal component, cross-sectional surveys cannot measure incidence of any disease (Pai, 2006). Another problem with cross-sectional studies is that this type of methodology selectively includes cases with more indolent diseases or people who live longer with the disease (Peipert & Phipps, 1998).

(4) Case-Control

In case-control studies, subjects who have a specified outcome (the cases) are compared with subjects who do not have this outcome (the controls) with respect to risk factors of interest (Hartz & Marsh, 2003; Peipert & Phipps, 1987; Torabinejad & Bahjri, 2005). As discussed in Module 3, in a case-control design (also known as retrospective studies), sampling starts with diseased (i.e., cases) and non-diseased (i.e., controls) individuals. The exposure status is then determined by looking backward in time (i.e., retrospectively), using documentation of exposures or recall of

historical events. The measure of association is called an Odds Ratio (OR), which is the ratio of the odds (chance) of exposure among cases in favor of exposure among controls (Pai, 2006). If the disease is rare, then the OR tends to be a good approximation of the Relative Risk (RR).

Case control studies are more cost-efficient, simple and easy to conduct than cohort studies. Moreover, they provide the only way of studying very rare disorders or those with a long time lag between exposure and outcome (Greenhalgh, 1997; Torabinejad & Bahjri, 2005). Case-control studies are an efficient method to evaluate what risk factors lead to adverse outcomes that occur infrequently (Hartz & Marsh, 2003). They rarely are used to evaluate treatment effectiveness. Case-control designs have been used to assess the benefit of screening (Davidov & Zelen, 2003). In a typical case-control study, cases and controls are matched on variables related to the outcome, and compared with respect to an exposure. Theoretically, the cases and controls are a random sample drawn from a hypothetical population of cases and controls. The key assumption underlying the case-control design is that both cases and controls are at risk for the outcome and the exposure (Davidov & Zelen, 2003). Case-control studies select subjects based on their clinical outcome and then determine the level of exposure to a risk factor or agent of interest (Peipert & Phipps, 1998).

However, case control studies are often criticized because of the possibility of various types of bias (e.g., recall bias; OR bias if the control group selected for comparison has very low odds for exposure; Pai, 2006). In fact, case-control studies are easy to do poorly and prone to many biases (Peipert & Phipps, 1998). Some biases include the difficulty in identifying comparable cases and controls resulting in selection bias; differential recall and report biases can also result as the information on prior exposure and other confounding variables are obtained (Torabinejad & Bahjri, 2005). Since case-control studies typically rely on records to determine exposure, it's often impossible to determine whether the exposure preceded the outcome (i.e., temporal sequence). Thus, unless the study involves choosing newly occurring incidents or cases, it is not certain that exposure occurred before the disease state (Peipert & Phipps, 1998). In addition, it is often difficult to select control groups. Accordingly, some have argued that the process that is most open to bias is not the assessment of outcome, but the diagnosis of "caseness" and the decision as to when the individual became a case (Greenhalgh (1997; Peipert & Phipps, 1998).

(5) Cohort

In the cohort design, outcomes are compared for persons receiving two or more treatments. It is considered the strongest of all observational designs. The idea is to measure and compare the incidence of disease in two or more study cohorts (i.e., group of people who share a common experience or condition). Usually there is one

cohort that is thought of as the exposed cohort, and another cohort is thought of as the unexposed cohort (Pai, 2006; Rational Pharmaceutical Management Plus Program, 2001). An attempt is made to match both cohorts with respect to age, sex, and other important variables, keeping the only difference between the two cohorts, the closure status.

A cohort study compares groups with respect to exposure to risk factors and to specific outcomes. Subjects are enrolled into a study group according to whether they have or have not been exposed to the risk factor being investigated. They then are observed for a period of time to determine whether they develop the outcome being investigated. When the follow-up period is complete, the two groups are compared to determine the incidence of the outcome among them (Peipert & Phipps, 1998).

Cohort studies have two primary purposes. One is descriptive: to describe the incidence of a disease or outcome in a group in a specified time. The other purpose is analytic; that is, to assess whether there is an association between a risk factor or exposures and an outcome (Peipert & Phipps, 1998). There are two basic types of cohort studies: the prospective (concurrent) and the retrospective (non-concurrent) cohort study.

In a **prospective** (or concurrent) cohort study, the group of subjects who have been exposed to the risk factor is identified at the onset of the study. They are followed longitudinally for a period of time to determine outcome. They are compared to a group of subjects who were not exposed to the risk factor at the onset of the study. This group also is followed for the same period of time to determine outcome (Peipert & Phipps, 1998). Well-designed cohort studies offer a clear and appropriate temporal sequence from exposure to outcome, providing strong support for causation (Torabinejad & Bahjri, 2005).

In a **retrospective** (non-concurrent) cohort study a group of patients are assembled based on a risk factor or exposure known to have occurred at some point in the past and the results or outcomes at the present time or at some point in the future are observed (Peipert & Phipps, 1998). The advantage of a retrospective cohort study is that one can assemble the cohort more efficiently than waiting several years for a large enough group of patients to be identified (Peipert & Phipps, 1998). However, retrospective cohort studies are relatively weaker because they rely on existing records (Torabinejad & Bahjri, 2005).

Cohort studies are usually prospective (i.e., forward looking) and are also called longitudinal studies. Disease-free cohorts are defined on the basis of the exposure status and then they are followed up for long time periods (Pai, 2006). New cases of the disease are picked up during follow-up and the incidence of the disease is computed on the basis of the exposure status. The incidence in the exposed cohort is then compared with the incidence in the unexposed cohort, known as the Relative Risk (RR) or Risk Ratio (RR; Pai, 2006). Relative Risk is a measure of association between the exposure and the outcome. The larger the RR, the stronger the association. The cohort study is the only study design in which the true incidence of a disease can be estimated (Pai, 2006).

Overall, the main advantages of cohort studies is that they can be used to determine prognostic factors related to specific diseases and may be more representative of the population with the disorder or disease in question (Peipert & Phipps, 1998). Data collection in the compared groups is usually thorough and systematic and is not as subjected to recall bias as in case-control studies (Peiper & Phipps, 1998). Good cohort studies provide better information for treatment comparisons than case series because data collection is standardized and patients are from the same period and in similar settings. Prospective cohort studies collect baseline data before treatment and again at follow-up (Hartz & Marsh, 2003). Treatments are compared with respect to patient outcomes after using statistical methods to take into account differences in patient risk before treatment. Retrospective cohort studies use data when information on the treatment, baseline risk, and patient outcome are abstracted from medical records (Hartz & Marsh, 2003). Cohort studies begin with the exposure and follow subjects for a period of time to assess the development of an outcome (Peipert & Phipps, 1998).

However, there are several disadvantages of a cohort study. One major disadvantage is the possibility of systematic error or bias. Because the risk factor, treatment, and other covariates may be chosen by the subject or the investigator based on some factor that may affect the outcome, cohort studies may be subject to bias (Peiper & Phipps, 1998). They cannot exclude unknown confounders, blinding is difficult and identifying a matched control group can be hard. They are difficult to use for rare events, large sample sizes or when long follow-up is necessary (Jones-Harris, 2003; Torabinejad & Bahjri, 2005). Another major disadvantage is that with retrospective cohort studies, one must rely on chart reviews or information from the past to determine level of exposure. Medical record review can lack important characteristics and variables of interest (Peipert & Phipps, 1998). Furthermore, cohort studies can be very time consuming and expensive. Since most diseases are rare, large cohorts have to be followed up for many years to get good estimates of incidence and this makes feasibility very difficult (Pai, 2006; Torabinejad & Bahjri, 2005). Nonetheless, the clear temporal (time) sequencing is extremely important while making causal inference.

4.2 EVALUATION OF OBSERVATIONAL RESEARCH

An evidence-based practice requires that health practitioners search for the best research evidence to answer clinical questions and appraise that research for its validity (i.e., is it true?), impact (i.e., how big is the effect?), and applicability (i.e., usefulness in clinical practice). To do this effectively requires an understanding of the fundamental principles of study design and biostatistics. Since observational studies dominate the surgical literature, it is essential that both the strengths and weaknesses of these studies be understood to determine both the quality and usefulness of the evidence deriving from them. The following section highlights some of the key strengths and weaknesses of observational studies. For information on how to critically appraise observational research studies, please refer to Appendix A which includes a checklist for reports of observational studies, as well as a summary of questions to assess the validity of studies of causation.

4.2.1 STRENGTHS

There is evidence that observational studies can be designed with rigorous methods that mimic those of clinical trials and that well-designed observational studies (e.g., cohort or case-control design) do not consistently overestimate the effectiveness of therapeutic agents (Concato, Shah, & Horwitz, 2000). Concato, Shah, and Horwitz (2000) found that the summary results of randomized controlled trials and observational studies were remarkably similar for each clinical topic examined. Viewed individually, the observational studies had less variability in point estimates (i.e., less heterogeneity of results) than RCTs on the same topic. Thus, the researchers concluded that observational cohort studies can produce results similar to those of RCTs when similar criteria are used to select study subjects. Furthermore, they argued that the popular belief that only RCTs provide trustworthy results and that all observational studies are misleading does a disservice to patient care, clinical investigation, and the education of health care professionals (Concato, Shah, & Horwitz, 2000).

According to Concato, Shah, and Horwitz (2000), a specific method used to strengthen observational studies is the "restricted cohort" design. This design adapts principles of the design of RCTs to the design of an observational study by identifying a "zero time" for determining a patient's eligibility and base-line features, using inclusion and exclusion criteria similar to those of clinical trials, adjusting for differences in base-line susceptibility to the outcome, and using statistical methods (e.g., intention-to-treat analysis) similar to those of RCTs (Concato, Shah, & Horwitz, 2000).

Observational studies can be more challenging to design and execute in terms of controlling bias, but they have several advantages compared with RCTs. Perhaps the most important advantage is that they do not interfere with patient and physician

choices (Hartz & Marsh, 2003). Observational studies are "natural" in that the researchers do not interfere with what happens, they simply collect and analyze relevant data that can reduce the costs and improve the generalizability of studies (Sheldon, 2001). They can be designed and implemented sooner and with less cost than RCTs, and they more often test treatment in relevant settings and with relevant patients. Historically, observational studies have more biases than RCTs, but most of these could be overcome with careful planning and execution (Hartz & Marsh, 2003).

Questions regarding diagnosis, prognosis, and causation may be best addressed by observational (i.e., epidemiological) studies. Moreover, many studies employ the observational design because it is less expensive, more timely, more relevant, and leaves medical decision-making under the control of the physician and patient (Hartz & Marsh, 2003). Well-designed observational studies can report on rare diseases and novel interventions, prevent unethical allocation of treatments, and their findings can be closer to representation of the general population (Torabinejad & Bahjri, 2005). Furthermore, careful study design, data collection, and statistical analysis can reduce much of the selection and information bias that has been associated with observational studies. With improved standards for reporting observational studies, it may be possible to identify those studies most likely to be valid (Hartz & Marsh, 2003).

4.2.2 WEAKNESSES

As already indicated in the previous sections, causal inference is difficult, if not impossible with the weaker observational study designs. Since the clinician or researcher may not be able to control and hence detect all the differences between the groups studied, any difference detected may consequently be attributed to the wrong cause. Moreover, without establishing clear temporal (time) sequencing in the study, it is unlikely that any causal inferences can be made.

Related to this is the primary inherent disadvantage of observational studies; that they are vulnerable to confounding from unrecognized or unrecorded risk factors (Hartz & Marsh, 2003). Errors in observational studies that cause a study to be invalid often are referred to as bias. **Selection bias** is preferential inclusion of subjects with certain treatment outcomes. For cohort studies this bias occurs when follow-up information is less likely to be collected on subjects who have better (or worse) outcomes (Hartz & Marsh, 2003).

Another type of bias related to measurement error is **information bias**. This bias may be attributable to either imperfect definition of a study variable or a flawed data collection procedure (Hartz & Marsh, 2003). Well-specified outcome measures will not only reduce differential information bias that favors one of the treatments but also will

reduce random misclassification that makes it more difficult to detect a true difference in treatments (Hartz & Marsh, 2003).

With confounding, apparent effects of treatment are altered by effects of risk factors. It occurs if patients who have Treatment A are at higher risk for a bad outcome before treatment than patients who have Treatment B. Because of this confounding, Treatment A may appear worse than Treatment B if it actually is equally effective (Hartz & Marsh, 2003). Sometimes observational studies of surgical procedures select one type of patient for one treatment and a second type of patient for another treatment (Hartz & Marsh, 2003).

Related to this is the confound known as **confounding-by-indication**. This weakness of observational studies that compares one therapeutic strategy with another involves the treating physicians selecting patients for a given therapy based on clinical features that are also related to the outcomes of interest (Siderowf, 2004). If it were possible to completely control for the factors upon which patients were selected, confounding-by-indication might not be a problem. However, because there are often unmeasured as well as measured factors related to selection of patients for a given treatment, confounding-by-indication may be accompanied by bias (Siderowf, 2004). It is this associated bias that threatens the validity of observational studies.

5.0 OUTCOMES RESEARCH

Outcomes and effectiveness research have been terms used to refer to a wide range of studies, and there is no single definition for either that has gained widespread acceptance. However, one definition of outcomes research refers to evaluation that focuses on: (1) the status of participants after receiving care, and (2) the process of care itself (http://www.dcri.duke.edu/patient/glossary.jsp). In this regard, they are used to measure a variety of effects on diverse populations of patients. For example, outcome indicators that measure the impact of care include traditional measures such as mortality, complications, and costs, as well as less traditional measures such as function, quality of life, and patient satisfaction (Hubbard, Walker, Clancy, & Stryer, 2002).

Outcomes research is said to possess several distinguishing characteristics: (1) It focuses on conditions and treatments for those conditions rather than on individual treatment; (2) outcomes are not limited to physiologic measures, but also encompass health-related quality of life outcomes; and (3) the influence of non-clinical factors on the ultimate result of treatment is recognized (Mark & Salyer, 1999).

Ultimately, outcomes research seeks to understand the end results of particular healthcare practices and interventions. Evidence is needed to inform practice, and outcomes research provides evidence about benefits, risks, and results of treatments so that both clinicians and patients can make more informed decisions (Hubbard, Walker, Clancy, & Stryer, 2002). It is believed that outcomes research can identify potentially effective strategies for purchasers and healthcare managers that can be implemented to improve the quality and value of care (Hubbard, Walker, Clancy, & Stryer, 2002).

5.1 MEASURING OUTCOMES

With the outcomes movement, came recognition that outcome, as measured from the patient perspective, not just the process of health care delivery is important (McLeod, 1999). Historically, clinicians relied primarily on traditional biomedical measures such as the results of laboratory tests, to determine whether a health intervention is necessary and whether it is successful. However, researchers discovered that the sole use of these measures resulted in their missing many of the outcomes that mattered most to patients (http://www.ahrq.gov/ clinic/outfact.htm). Consequently, outcomes research included how people function, their experiences with care, and health-related quality of life issues that are of great clinical relevance.

Today, there is a range of instruments available to measure quality of life, including both generic and disease-specific instruments (McLeod, 1999). While the generic instruments have the advantage that they have usually been used extensively, have been tested for their reliability and validity, and can be used to compare patients with different disease processes, disease-specific instruments tend to be more sensitive for detecting small but clinically important changes in patients with the same disease process (McLeod, 1999). Depending on the study design and question being posed, the choice of an instrument to measure outcome may vary (McLeod, 1999).

Outcomes research has altered the culture of clinical practice and health care research by changing how we assess the end results of health care services. In doing so, it has provided the foundation for measuring the quality of care. The results of outcomes research are becoming part of the "report cards" that purchasers and consumers can use to assess the quality of care in health plans (http://www.ahrq.gov/clinic/outfact.htm). Studies have demonstrated that the difference between traditional clinical measures for a disease and the outcomes that matter to patients can, in fact, be dramatic.

Outcomes research is believed to provide policymakers with the tools to monitor and improve quality both in traditional settings and under managed care. Moreover, it has been described as the key to knowing not only what quality of care can be achieved, but how it can be achieved (http://www.ahrq.gov/clinic/outfact.htm). However, just as was the case with the other approaches examined, inappropriate attention to a variety of methodological issues can invalidate the findings of outcomes studies. In the following section we briefly describe some of the more pressing methodological considerations that should be taken into account when the findings in outcomes research are evaluated.

5.1.1 MEASUREMENT ISSUES

Successful measurement of health outcomes and changes in those outcomes is critical to the validity of any treatment effectiveness or outcomes research. Two major frameworks guide the design and interpretation of measurement:

(1) Norm-Referenced Measurement

Discriminates among individuals on the basis of the extent to which a particular concept (measure) is present. Health status and quality of life measures are examples (Mark & Salyer, 1999).

(2) Criterion-Referenced Measurement

Useful in determining what a person knows or can do in relation to a fixed standard. How one individual compares to another is irrelevant in this framework (Mark & Salyer, 1999). Measures developed according to this framework produce classifications or judgments such as satisfactory or unsatisfactory; met or not met. An example is the National Cancer Institute's criteria for correct performance of breast self-examination (Mark & Salyer, 1999).

Regardless of the framework adopted, the following 3 key measurement concerns apply:

(1) Reliability

(2) Validity

(3) Sensitivity to chance or responsiveness; the ability to capture inter-individual differences in the outcomes among participants who received an intervention and those who did not, at both pre- and posttest.

According to Mark and Salyer (1999), measures that do not have documented ability to capture these differences and that demonstrate change over time may have less use in outcomes research than those in which sensitivity to change has been evaluated. As indicated by Lipscomb, Donaldson, and Hiatt (2004), for outcomes research to achieve its potential to enhance care delivery, 3 prerequisites must be met: (1) technically sound and decision relevant outcome measures; (2) persuasive evidence about the impact of interventions on those outcomes; and (3) the capacity, determination, and ingenuity to translate findings into useful information for decision making.

6.0 SURGICAL CLINICAL AUDIT

Clinical audit is an important activity for various reasons. The main reason is that it helps to improve the quality of the service being offered to users (Jones & Cawthorn, 2002; Smith, 1992). Additional benefits of clinical audit (as presented by Smith, 1992 and (http://www.edu.rcsed.ac.uk/Lectures/Lt17.htm) are, it:

(1) Identifies and promotes good practice and can lead to improvements in service delivery and outcomes for users. In order to investigate the avoidable complications, mortality or morbidity data can be made available through assessments. Moreover, assessments of patients' stay in the hospital might bring to light the misuse or inefficient use of diagnostic services, which can then be improved. On the other hand, information may come to light indicating good practices are already in effect and thus should simply be sustained.

(2) Can provide the information you need to show others that your service is effective (and cost-effective) and thus ensure its development. Assessments of newer technologies with an audit can help determine whether these additions have improved the delivery and quality of health care services. Moreover, in case of accusation of malpractice, audit data can help to establish that the rate of complications compares favorably with that of the accepted standards.

(3) Provides opportunities for training and education. Monitoring the performance of the staff is one essential outcome of an audit.

(4) Helps to ensure better use of resources and, therefore, increased efficiency. A surgeon must know how (s)he is spending his or her time and the resources of the hospital before attempting to improve on time utilization.

(5) Can improve working relationships, communication and liaison between staff and service users, and between agencies. For example, the patient's view of health care delivery can be assessed to obtain information on satisfaction measures. This information can be used to strengthen or repair rapport with patients and staff, wherever necessary.

According to Jones and Cawthorn (2002), clinical audit should be an integral part of clinical practice. Given that all clinicians want to provide the best possible care for patients, clinical audit is one tool that can systematically facilitate this. It can be a powerful tool for positive change, resulting in improved practice and outcomes for patients (Jones & Cawthorn, 2002)

6.1 TYPES OF AUDIT

There are various types of audit that may be conducted depending on the desired objectives.
Below is an overview of 5 types of clinical audits:

National audit – are vital to study the overall health trends in the country and to be meaningful, they should:

(a) Be open to debate or self-evaluation,
(b) Be interesting,
(c) Maintain confidentiality of the surgeon and patient,
(d) Demonstrate change with improvement of patient care,
(e) Keep resources spent at a minimum,
(f) Set standards and review periodically, and
(g) Include priority topics for audit (http://www.edu.rcsed.ac.uk/Lectures/Lt17.htm).

Standards-based audit – A cycle that involves defining standards, collecting data to measure current practice against those standards, and implementing any changes deemed necessary (Jones & Cawthorn, 2002). For instance, a retrospective analysis of case records may be made to judge against a set of chosen criteria like assessment of quality of writing of operating theatre records, quality of discharge summary, appropriateness of investigators for a particular diagnosis, and appropriateness of treatment (http://www.edu.rcsed.ac.uk/Lectures/Lt17.htm)

Adverse occurrence screening and critical incident monitoring – Often used to peer-review cases which have caused concern or from which there was an unexpected outcome. The multidisciplinary team discusses individual, anonymous cases to reflect upon the way the team functioned and to learn for the future. In the primary care setting, this is described as a "significant event audit" (Jones & Cawthorn, 2002). For instance, details of adverse occurrences such as wound infections, unplanned readmissions, delay or error in diagnosis are reviewed to identify trends and perform comparative analysis (http://www.edu.rcsed.ac.uk/Lectures/Lt17.htm).

Peer review – "An assessment of the quality of care provided by a clinical team with a view to improving clinical care." Individual cases are discussed by peers to determine, with the benefit of hindsight, whether the best care was given. This is similar to the method described above, but might include 'interesting' or 'unusual' cases rather than problematic ones. Unfortunately, recommendations made from these reviews are often not pursued as there is no systematic method to follow (Jones & Cawthorn, 2002).

Patient Surveys and focus groups – These are methods used to obtain users' views about the quality of care they have received. Surveys carried out for their own sake are often meaningless, but when they are undertaken to collect data they can be extremely productive (Jones & Cawthorn, 2002). In a global audit, the entire process of health care delivery during a patient's stay in the hospital, including the spectrum of administration, nursing staff, para-clinical staff and doctors is assessed as outcome which is an important measure of the quality of care (http://www.edu.rcsed.ac.uk/Lectures/Lt17.htm).

6.2 COMPONENTS OF THE AUDIT CYCLE

There is only one clinical audit method – the clinical audit cycle (Smith, 1992). As can be seen in the figure in Appendix B, the audit method can be described as a cycle or a spiral. In the following sections, we identify and provide further details about a typical audit process.

Stage 1: Preparing for Audit

In practical terms, preparing for audit can be broken down into five elements: (1) Involving users (i.e., patients, carers, and organizations representing their interests) in the process, (2) Selecting a topic, (3) Defining the purpose of the audit, (4) Providing the necessary structures, (5) Identifying the skills and people needed to carry out the audit, and training staff and encouraging them to participate (Jones & Cawthorn, 2002).

Stage 2: Selecting Criteria/Standards

In clinical audit, criteria and standards are used to assess the quality of care provided by an individual, a team or an organization. These criteria are explicit statements that define what is being measured and represent elements of care that can be measured objectively (Jones & Cawthorn, 2002). Criteria can be classified into those concerned with (1) structure (what you need), (2) process (what you do), and (3) outcome (what you expect) (Jones & Cawthorn, 2002).

Stage 3: Measuring Performance

To ensure the data collected are precise and only essential data are collected, certain details of what is to be audited must be established from the outset. The user group included and/or excluded should be noted, the healthcare professionals involved in the user's care, and the period over which the criteria apply (Jones & Cawthorn, 2002). It is also necessary to define the population to which the audit applies and obtain a representative sample from it. Data are collected to measure current practice against agreed standards. Some data may already be available in computerized clinical information systems. In other cases, data may have been collected routinely by other methods (Jones & Cawthorn, 2002). The data collected should relate only to the objectives of the audit; Jones and Cawthorn (2002) recommend not being tempted to collect additional "interesting" information. Finally, simple statistical analysis is usually all that is required of audit data. For example, actual length of stay can be calculated as the number of occupied beds multiplied by the number of days in the study period, all divided by the number of patients discharged or dead during the study period (http://www.edu.rcsed.ac.uk/Lectures/Lt17.htm). Other simple performance indicators include: expected length of stay, turn over interval, actual throughput, expected throughput, expected throughput, percentage of day cases, percentage of cases not operated, pre-operative stay, post-operative stay, and waiting list per 1000 population.

Stage 4: Making Improvements

Once the results of the audit have been published and discussed, an agreement must be reached about the recommendations for change. Using an action plan to record these recommendations is good practice and should include who has agreed to do what and by when (Jones & Cawthorn, 2002). Aside from those who will never agree to the need for change, there are potential barriers to change in terms of resources, politics, or environment. Change needs to be implemented in a systematic way, ensuring that communication and dissemination are sustained throughout the process (Jones & Cawthorn, 2002).

Stage 5: Sustaining Improvements

After an agreed period, the audit should be repeated. The same strategies for identifying the sample, methods and data analysis should be used to ensure comparability with the original audit (Jones & Cawthorn, 2002). The re-audit should demonstrate that the changes have been implemented and that improvements have been made. Further changes may then be required, leading to additional re-audits (Jones & Cawthorn, 2002).

6.3 STRUCTURE OF AUDIT

The structure of an audit refers to the organization within which surgical practice is carried out. For instance, it could be the hospital, the department, or the operating theatre. Standards to change the working environment for the better can be set to improve patient care and encourage the staff to function to the best of their abilities (http://www.edu.rcsed.ac.uk/Lectures/Lt17.htm).

Although there is only one audit method, there are apparent variations on it. Thus, in an attempt to maintain some consistency, the Royal College of Physicians (1989) suggested that all audits should include the following:

(1) The purpose of the audit should be educational and relevant to patient care.

(2) Control should be maintained by clinical peers with voluntary participation.

(3) Standards should be set locally by the participating clinicians.

(4) The methods employed should be non-threatening, interesting, objective, and repeatable.

(5) Resources should be cheap and simple.

(6) Records should contain adequate clinical content and be easily retrieved.

All clinical audits need to be conducted within an ethical framework. Since clinical audit by definition does not involve anything being done to patients beyond their normal clinical management, it does not require formal ethical approval. Clinical audit aims to improve patient care through systematic review of care against explicit criteria and the implementation of change (Gloucestershire Health Community, 2005). Decisions about whether "audit" projects need ethics approval often hinge on the question of whether they really are an audit, or whether they are actually research. For instance, research always requires ethical approval. In the "hard-to-call" cases, the decision about whether a project is an audit or research is ultimately a matter for the Local Research Ethics Committee (Gloucestershire Health Community, 2005).

6.4 EVALUATION OF CLINICAL AUDIT

Increasing multi-professional participation is the key to a successful audit. Successful audit means good-quality studies that are based on agreed, evidence-based standards of care, that have agreed outcomes and that achieve sustained improvements in care for patients (Jones & Cawthorn, 2002).

6.4.1 ADVANTAGES

(1) Clinical audit has the advantages of being an excellent educational tool, and when properly carried out, is non-punitive (Wagaarachchi, Asare, Ashley, Gordon, Graham, Hall, Henneh, McCaw-Binns, Penney, Antwi, & Bulloughv, 2001).

(2) It can be initiated locally and results in the production of locally relevant and immediately actionable information (Wagaarachchi, Asare, Ashley, Gordon, Graham, Hall, Henneh, McCaw-Binns, Penney, Antwi, & Bulloughv, 2001).

(3) The participatory element of audit cannot be overemphasized. The process of involving local staff in reflecting on their current practice and setting targets is believed to be an effective mechanism for bringing about improvements in health care (Wagaarachchi, Asare, Ashley, Gordon, Graham, Hall, Henneh, McCaw-Binns, Penney, Antwi, & Bulloughv, 2001).

(4) Since non-medical personnel are capable of doing the necessary data extraction, it can be less expensive than other forms of audit (Wagaarachchi, Asare, Ashley, Gordon, Graham, Hall, Henneh, McCaw-Binns, Penney, Antwi, & Bulloughv, 2001).

(5) It provides a structured framework for gathering information and involves less subjective assessment of case management (Wagaarachchi, Asare, Ashley, Gordon, Graham, Hall, Henneh, McCaw-Binns, Penney, Antwi, & Bulloughv, 2001).

(6) The review process can also help to highlight deficiencies both in recording in-patient records and in records storage (Wagaarachchi, Asare, Ashley, Gordon, Graham, Hall, Henneh, McCaw-Binns, Penney, Antwi, & Bulloughv, 2001).

6.4.2 DISADVANTAGES

(1) Clinical audit is limited to clinical care in the facility in which it is carried out and cannot deal with community issues (Wagaarachchi, Asare, Ashley, Gordon, Graham, Hall, Henneh, McCaw-Binns, Penney, Antwi, & Bulloughv, 2001).

(2) Having always based their clinical practice on their own personal experience and preferences, some clinicians may find the concepts of evidence-based practice and audit both difficult to understand and threatening (Wagaarachchi, Asare, Ashley, Gordon, Graham, Hall, Henneh, McCaw-Binns, Penney, Antwi, & Bulloughv, 2001).

(3) Either an appropriate set of criteria needs to be available, or local criteria must be developed if a criterion-based clinical audit (Wagaarachchi, Asare, Ashley, Gordon, Graham, Hall, Henneh, McCaw-Binns, Penney, Antwi, & Bulloughv, 2001).

(4) It relies on being able to identify relevant cases from facility registers and retrieve the case notes (Wagaarachchi, Asare, Ashley, Gordon, Graham, Hall, Henneh, McCaw-Binns, Penney, Antwi, & Bulloughv, 2001).

(5) Non-medical audit assistants (usually records staff) must be available to find patient records and undertake the extraction of information (Wagaarachchi, Asare, Ashley, Gordon, Graham, Hall, Henneh, McCaw-Binns, Penney, Antwi, & Bulloughv, 2001).

(6) There must be a willingness to close the audit loop with at least one further round of reviewing practice (Wagaarachchi, Asare, Ashley, Gordon, Graham, Hall, Henneh, McCaw-Binns, Penney, Antwi, & Bulloughv, 2001).

7.0 CONCLUDING SUMMARY

The intuitive assumption that only evidence from RCTs counts in evidence based practice is understandable. However, this position is now being challenged, and other designs, such as observational and outcomes research, are being considered legitimate providers of the evidence in evidence-based practice.

As discussed in this module, observational studies are important methodologies to evaluate exposures and risk factors that are not amenable to experimental trials (Peipert & Phipps, 1998). They offer the advantage of being more generalizable to patients, as these studies may have more liberal inclusion criteria than the typical randomized trial. Their disadvantage is their susceptibility to biases and their inability to control for unknown factors that may impact on the outcome of interest (Peipert & Phipps, 1998). Establishing causality of an association noted in observational studies is an intricate process that requires careful assessment. Clinicians and researchers should be familiar with observational studies so they may better evaluate a proposed causal relationship and the quality of reports claiming such relationships. Only then can they determine if the findings are valid and applicable to their patient population (Peipert & Phipps, 1998). No observational study (or randomized controlled trial for that matter) should be considered definitive (Hartz & Marsh, 2003). More information can be obtained by comparing results from several observational studies. Of particular interest are results of good observational studies with different designs that consistently support the superiority of one treatment over another even though the studies are unlikely to have similar biases (Hartz & Marsh, 2003).

We also saw that outcomes measurement and research is aimed at improving the quality of interventions and policies governing interventions. Consequently, associating differences in the process of care with differences in patient outcomes can assist in clarifying which services are worth providing and which require improvement. Similarly, clinical audit was described as a cycle, wherein there are stages that follow a systematic process aimed at reviewing the quality of everyday care provided to patients with common conditions (http://www.nelh.nhs.uk/BestPracticeClinicalAudit.pdf).

Information gained through observational and outcomes studies, and clinical audit should be readily accessible to clinical practitioners to assist in decision-making. Instead of being frightened by the use of evidence (or lack of it), Kreder (1999) suggests surgeons seize the opportunity to steer policymakers and payors in the right direction. If the evidence is lacking, the challenge to generate it should be undertaken (Kreder, 1999). According to Kreder (1999), surely the art of surgery can only be enhanced by the thoughtful

generation and application of the best possible scientific evidence. Researchers and practitioners alike should be united in their quest for scientific rigour in evaluation, regardless of the method used (Black, 1996).

7.1 REVIEW OF MODULE OBJECTIVES

At the end of this module, participants should now be able to:

(1) Differentiate between observational studies, outcomes research, and clinical audit;

(2) Describe the types of observational research and the methods employed in conducting it;

(3) Identify outcomes information for decision-making and how to measure outcomes; and

(4) Describe clinical audits, research and the audit cycle.

8.0 REFERENCES

Adamson, J. (2004). Evaluation of studies of causation (aetiology). *Evidence Based Nursing, 7*; 36 – 40. Downloaded from ebn.bmjjournals.com on 1 April 2006.

Agence d'evaluation des technologies et des modes d'intervention en sante' (AETMIS) & International Society of Technology Assessment in Health Care (ISTAHC). (2000 – 2001). Health Technology Assessment: Decision-Making for Health. Distance Learning Course. Units 1 to 12. Course content developed by a team of researchers and instructors at McGill University, University of Montreal and AETMIS in Montreal, Canada. http://aetmis.gouv.qc.ca/e-learning/mod/forum/discuss.php=73

Alberta Heritage Foundation for Medical Research; AHFMR. (2003). *Health technology assessment: Alberta's needs and future directions. Summary of results of a consultation workshop.* Edmonton, AB. Canada.

Australian Government. (1996). "Measuring and improving outcomes: A note from the editor." *Better Health Outcomes Newsletter, 2(1)*. Department of Health and Aging. http://www7.health.gov.au/pq/bho/1996/v2n1/outcomes.htm#definitions

Black, N. (1996). Why do we need observational studies to evaluate the effectiveness of health care. *British Medical Journal, 312*: 1215-1218.

Concato, J., Shah, N., & Horwitz, R.I. (2000). Randomized, controlled trials, observational studies, and the hierarchy of research designs. *The New England Journal of Medicine, 342*; 1887 – 1892.

Danish Institute for Health Technology Assessment; DIHTA. (2001). Health Technology Assessment Handbook. 1st Edition. Kristensen, F.B., Horder, M., & Poulsen, P.B. (Eds.). National Board of Health.

Davidov, O. & Zelen, M. (2003). The theory of case-control studies for early detection programs. *Biostatistics, 4(3)*; 411 – 421.

First International Conference on Health Promotion Ottawa. (1986). Ottawa Charter for Health Promotion. WHO/HPR/HEP/95.1. http://www.who.int/hpr/NPH/docs/ottawa_charter_hp.pdf

Gloucestershire Health Community. Do clinical audit projects ever need formal ethics approval? Accessed on December 8, 2005.

Hartz, A. & Marsh, J.L. (2003). Methodologic issues in observational studies. *Clinical Orthopaedics and Related Research, 413*; 33 – 42.

Hubbard, H., Walker, P.H., Clancy, C.M., & Stryer, D. (2002). Outcomes and effectiveness research: Capacity building for nurse researchers at the agency for healthcare research and quality. *Outcomes Management for Nursing Practice, 6(4)*; 146 – 151.

Jones, T. & Cawthorn, S. (2002). What is clinical audit? *"What is …", Volume 4(1)*; 1-8. www.evidence-based-medicine.co.uk

Kestle, J.R.W. (1999). Clinical Trials. *World Journal of Surgery, 23*; 1205 – 1209.

Khan, K.S., ter Riet, G., Glanville, J., Sowden, A.J., & Kleijnen, J. (Eds.) (2001). Undertaking systematic reviews of research on effectiveness. CRD's Guidance for those Carrying out or Commissioning Reviews. CRD Report Number 4 (2nd Edition). NHS Centre for Reviews and Dissemination, University of York. York Publishing Services Ltd

Kreder, H.J. (1999). Evidence-based surgical practice: What is it and do we need it? *World Journal of Surgery, 23*; 1232 – 1235.

Lawlor, D.A., Smith, G.D., & Shah, E. (2004). Commentary: The hormone replacement-coronary heart disease conundrum: Is this the death of observational epidemiology? *International Journal of Epidemiology, 33(3)*; 464 – 467.

Mark, B.A. & Salyer, J. (1999). Methodological issues in treatment effectiveness and outcomes research. *Outcomes Management for Nursing Practice, 3(1)*; 12 – 19.

McCulloch, P., Taylor, I., Sasako, M., Lovett, B., & Griffin, D. (2002). Randomised trials in surgery: Problems and possible solutions. *British Medical Journal, 324*; 1448-1451.

McLeod, R.S. (1999). Clinical trials, outcomes research, and practice guidelines in surgery: Introduction. *World Journal of Surgery, 23*; 1203 – 1204.

Outcomes Research. Fact Sheet. AHRQ Publicatio No. 00-P011, March 2000. Agency for Healthcare Research and Quality, Rockville, MD. http://www.ahrq.gov/clinic/outfact.htm

Pai, M. (2006). Observational Study Designs. SUNMED.org. Accessible at: http://www.sunmed.org/Obser.html

Peipert, J.F. & Phipps, M.G. (1998). Observational studies. *Clinical Obstetrics and Gynecology, 41(2)*; 235 – 244.

Sheldon, T.A. (2001). Biostatistics and study design for evidence-based practice. *AACN Clinical Issues, 12(4)*; 546 – 559.

Siderowf, A.D. (2004). Evidence from clinical trials: Can we do better? *NeuroRx: The Journal of the American Society for Experimental NeuroTherapeutics, 1*; 363 – 371.

Smith, R. (1992). Audit and Research. *BMJ, 305;* 905-906.

Stevens, A.J., Raftery, J., & Roderick, P. (2005). Can health technologies be assessed using routine data? *International Journal of Technology Assessment in Health Care, 21(1)*; 96 – 103.

The Royal College of Physicians. Surgical Knowledge. Surgical Audit – Lecture. http://www.edu.resed.ac.uk/Lectures/Lt17.htm. Accessed 01/08/2005 at 9:46pm.

Torabinejad, M. & Bahjri, K. (2005). Essential elements of evidence-based endodontics: Steps involved in conducting clinical research. *Journal of Endodontics, 31(8)*; 563 – 569.

UBHT Clinical Audit Central Office. How to…. tell the difference between audit and research. http://www.ubht.nhs.uk/ClinicalAudit/docs/HowTo/CA%20&%20Research.pdf (accessed 26 January 2006).

Wagaarachchi, P.T., Asare, K., Ashley, D., Gordon, G., Graham, W.J., Hall, M.H., Henneh, R., McCaw-Binns, A., Penney, G.C., Antwi, K.Y., & Bullough, C.H.W. (2001). Improving the quality of obstetric care in developing countries through criterion-based clinical audit. An international collaborative project funded by the United Kingdom Department for International development. A practical field guide. The Dugald Baird Centre for research on women's health. http://www.abdn.ac.uk/dugaldbairdcentre/cbca/cbca.pdf

World Health Report (WHR). (2000). "Chapter 1: Why do health systems matter?" www.who.int/whr/2000/en/whr00_ch1_en.pdf

Writing @ CSU: Writing Guide. Steps and methods used in qualitative observational research. (http://writing.colostate.edu/references/research/observe/pop4a.cfm). Accessed 01/08/2005 9:19pm.

9.0 APPENDICES

9.1 APPENDIX A: CHECKLIST FOR REPORTS OF OBSERVATIONAL STUDIES

Introduction:

(1) Is the hypothesis clearly stated?

(2) Is the literature review complete and current?

(3) Is the research design clearly identified and appropriate for the research question?

Methods:

(4) Is there a clear description about how the study groups were selected, identified, approached, assessed for the exposure, and diagnosed with their outcome or disease?

(5) Does the report contain a clear description of data collection, training of interviewers, timing of the interviews, gradient of exposure, and stage of disease?

(6) Are important confounding and effect-modifying variables measured, described, and analyzed?

(7) Have all subjects been accounted for at follow-up, and were methods taken to avoid losses and biased follow-up?

(8) Was the study conducted in an ethical manner?

(9) Were sample size calculations performed and were adequate numbers of subjects recruited?

Results:

(10) Are the baseline characteristics of the study groups compared?

(11) Is the problem of multiple comparisons addressed in the text or from a statistical point of view?

(12) Are appropriate measures of association (that is, relative risks, odds ratios) and their confidence reported?

(13) Is there an analysis of potential confounding or effect-modifying variables?

(14) Are multivariable analyses performed, when appropriate?

Conclusions:

(15) Are the conclusions based on the results of the analysis?

(16) Does the discussion include a distinction between the a priori hypothesis and ex post facto hypotheses?

(17) Are competing explanations for the study findings discussed?

(18) Are the statistical associations distinguished from causal relationships?

(19) Is a distinction made between clinical and statistical significance?

(20) Are sample size and power addressed in the discussion?

(21) Are the findings placed into context of existing data? Are the potential reasons for the differences addressed?

(22) Are policy implications and areas for future research discussed?

Taken from: Peipert, J.F. & Phipps, M.G. (1998). Observational studies. *Clinical Obstetrics and Gynecology, 41(2)*; 235 – 244.

Summary of Questions to Assess the Validity of Studies of Causation

Chance

- Were there clearly stated, justified, a priori hypotheses?
- Was there evidence to reject the null hypothesis (presentation of p values and 95% confidence intervals)?
- Was there evidence that the study was sufficiently powered?

Bias

- What was the study design? Cross-sectional study, case-control study, or cohort study?
- In a case-control study, how was the sample selected? Is this likely to be associated with the exposure?
- In a cohort study, how were the participants allocated to exposure status? Is this likely to be associated with the outcome?
- Where possible, were objective or valid and reliable measurements used?
- In a case-control study, were the assessors of exposure blinded to outcome status?
- In a cohort study, were the outcome assessors blinded to exposure status?
- In a cohort study, was there substantial loss to follow-up? What were the characteristics of those who left the study?

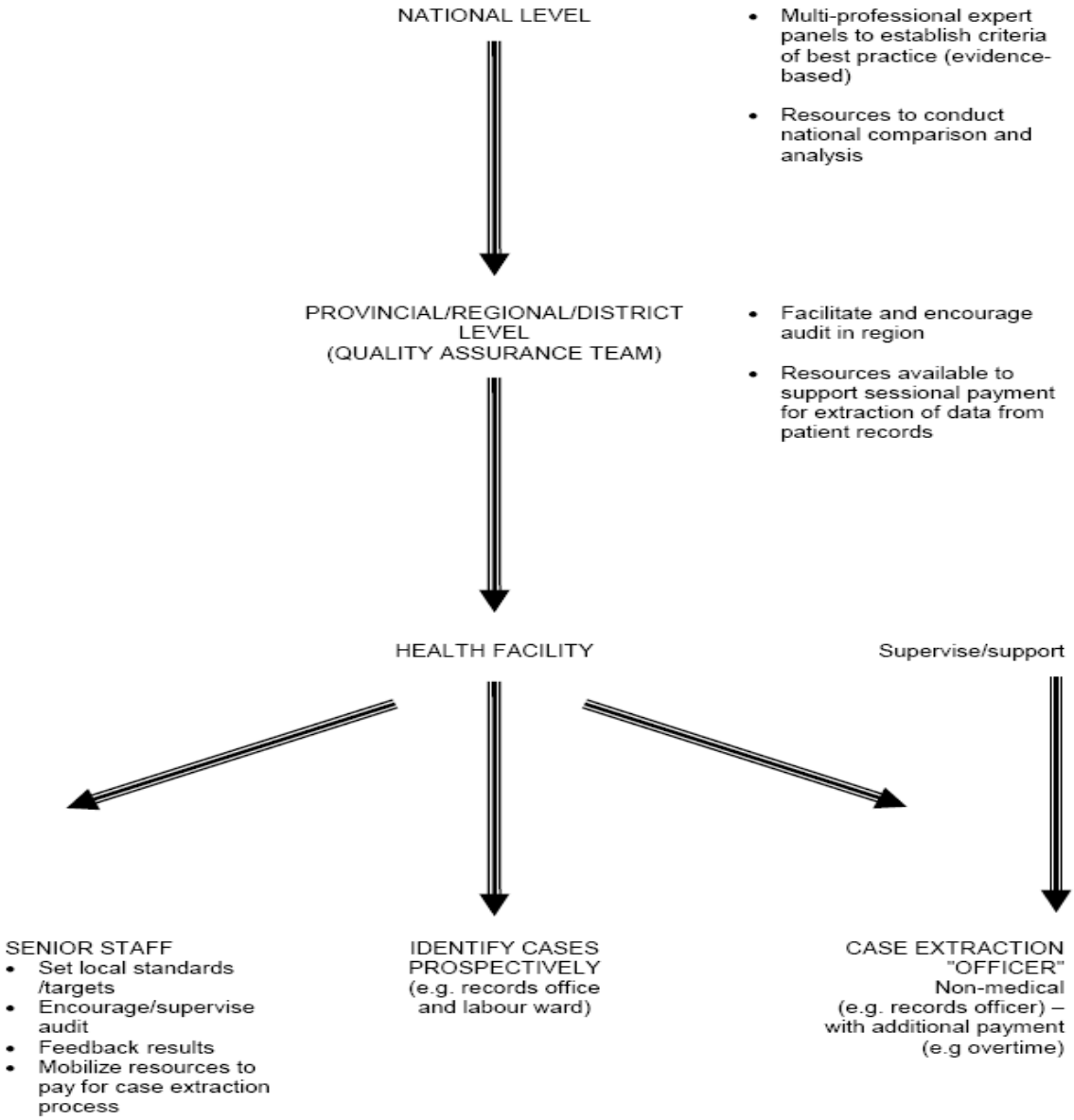
Confounding

- Did the authors consider possible confounding factors? Were they accurately measured?
- Did the authors analyse the data to take into consideration the effects of these potential confounders (restriction, stratification, or statistical analysis?)
- If matching was used, did the authors perform a matched analysis?

Taken from: Adamson, J. (2004). Evaluation of studies of causation (aetiology). *Evidence Based Nursing, 7*; 36 – 40. Downloaded from ebn.bmjjournals.com on 1 April 2006.

9.2 APPENDIX B: THE CLINICAL AUDIT FRAMEWORK

Figure 1 - Audit Framework



Taken from: Wagaarachchi, P.T., Asare, K., Ashley, D., Gordon, G., Graham, W.J., Hall, M.H., Henneh, R., McCaw-Binns, A., Penney, G.C., Antwi, K.Y., & Bullough, C.H.W. (2001). Improving the quality of obstetric care in developing countries through criterion-based clinical audit. An international collaborative project funded by the United Kingdom Department for International development. A practical field guide. The Dugald Baird Centre for research on women's health.
http://www.abdn.ac.uk/dugaldbairdcentre/cbca/cbca.pdf

The Clinical Audit Cycle

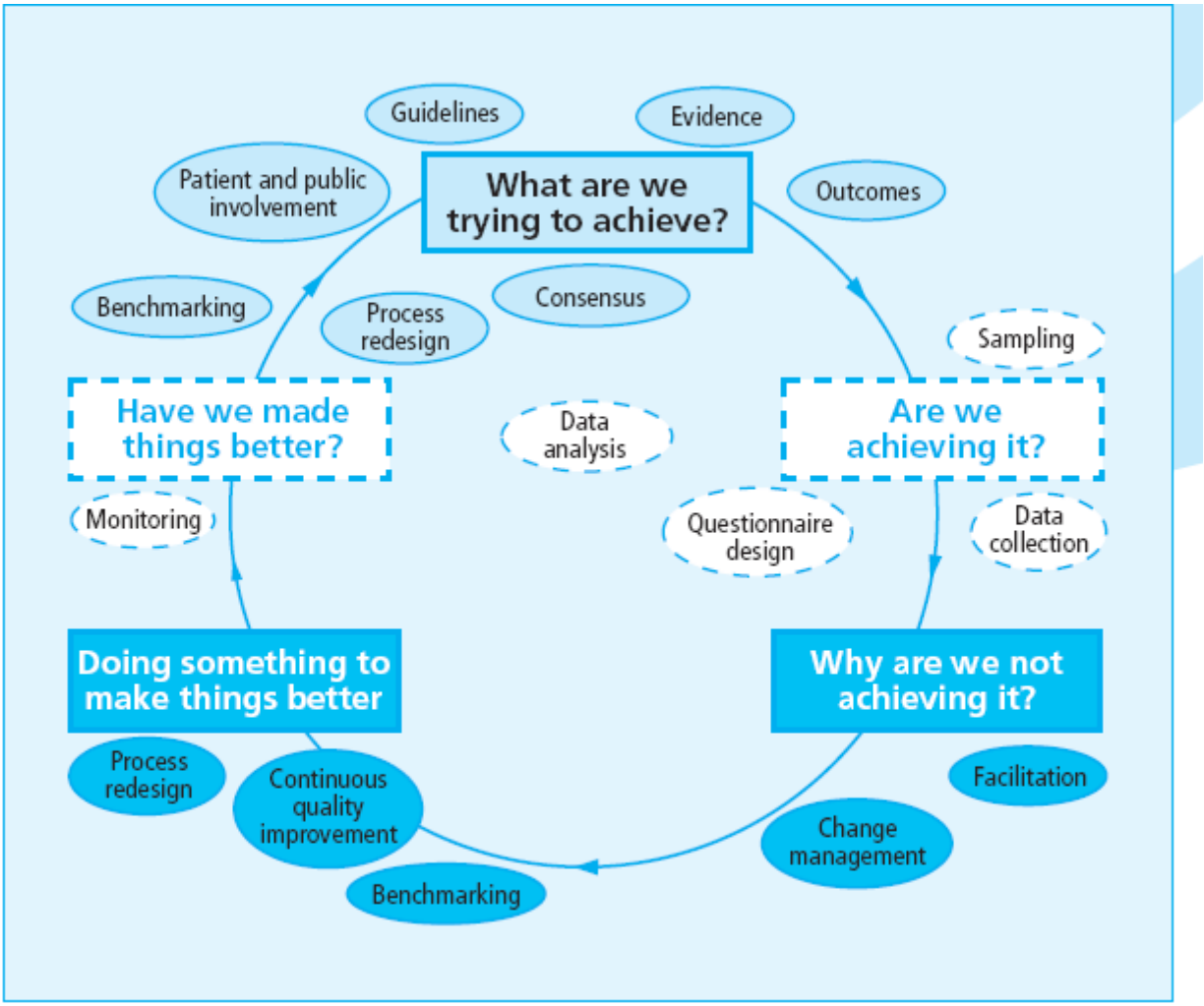
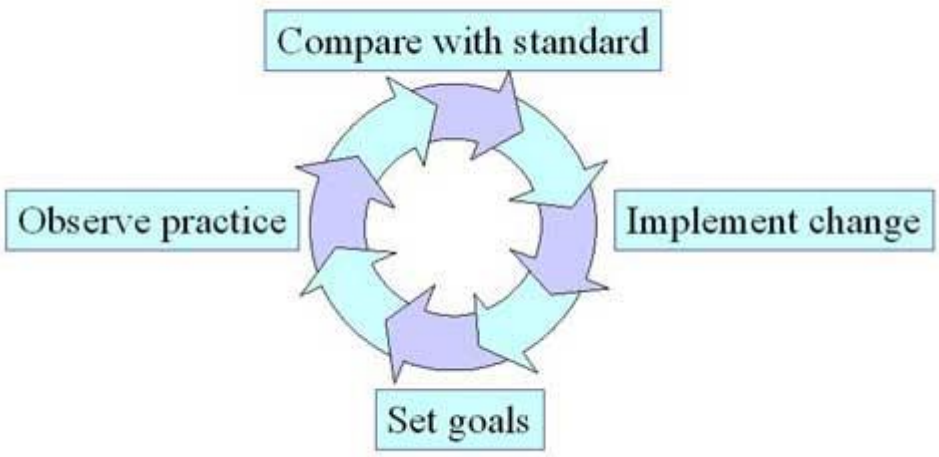


Figure 1. The clinical audit cycle[11]

Taken from: Jones, T. & Cawthorn, S. (2002). What is clinical audit? *"What is …", Volume 4(1)*; 1 – 8. www.evidence-based-medicine.co.uk

The Clinical Audit Cycle



Taken from: The Royal College of Surgeons of Edinburgh. Surgical Knowledge and Skills Website. http://www.edu.rcsed.ac.uk/Lectures/Lt17.htm