HEALTH TECHNOLOGY ASSESSMENT

OVERVIEW OF COMMON ISSUES AND RESEARCH METHODS USED IN HTA

MODULE 3

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 **third** module of six in a series on Health Technology Assessment (HTA). The primary objective of this third module and workshop is to provide you with an overview of the main steps involved in conducting Health Technology Assessment, as well as the various research methods employed by studies involved in the collection of primary data. Furthermore, you will be able to identify the various ethical, socio-cultural, and legal issues related to conducting HTAs.

We hope that the fundamentals presented in this module will not only assist you in conducting your own Health Technology Assessment (HTA), but also provide you with the tools required to critically evaluate assessments prepared by others.

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 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 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

The main goals of this Health Technology Assessment (HTA) module are:

(1) To present an overview of the basic steps in the conduct of HTA,

(2) To present an overview of the research designs within experimental and non-experimental methodologies relevant to HTA evaluation, and

(3) To present the ethical, socio-cultural, and legal issues related to the way scientific evidence is gathered, synthesized, and subsequently utilized.

By the end of this HTA module, participants will be able to:

(1) Discuss a model of how to conduct a HTA,

(2) Understand the difference in the evidence collected from experimental (randomized) and non-experimental (non-randomized) research studies, and

(3) Identify ethical, socio-cultural, and legal issues relevant to research and HTA.

2.0 INTRODUCTION

According to Relman (1988), there have been two revolutions in medical care since the end of World War II. The first is described as the Era of Expansion and the second, the Era of Cost Containment. Having entered the third era, the Era of Assessment and Accountability, Relman (1988) argued we can no longer afford to provide health care, without knowing more about its successes and failures.

Keeping up with the latest advances in diagnosis and treatment is a challenge faced by all physicians and surgeons alike. Information is needed that is both valid (i.e., accurate and correct), and relevant to patients and practice (Flaherty, 2004). One of the tools viewed as critical in assisting patients, physicians, and policy-makers in making important health care related decisions is Health Technology Assessment (HTA). Technology assessment is said to begin from a desire for information, usually of a technical nature (Battista, 2000 – 2001). HTA has been conceptualized as a bridge between the world of research and the world of decision- or policy-making. According to Battista (2000 – 2001), HTA is similar to a tree "firmly rooted in scientific inquiry but with its foliage turned towards policy making." Based on the needs of policy makers for information, sunlight to the tree of technology assessment is provided.

Although HTA is not simply more research, it is important to have an understanding of the general research methodologies available and typically consulted in HTAs. The character and strength of HTA comes from integrating the efforts of multiple disciplines (Battista & Hodge, 1999). HTA accomplishes this integration by synthesizing information, examining databases, and at times, generating primary data. Choices among these methods are driven by the relevance of the results to improve decision-making (Battista & Hodge, 1999; MRC, 2003). Evidence can help close the gap between what we know and what we do in health care. While appropriately used medical technology can improve health and possibly reduce costs, the availability of evidence-based technology assessment is not enough to improve practice, reduce variation, and achieve better outcomes (Eisenberg, 1999). The findings of research need to be translated into information that is useful for making health care decisions.

In the following section we examine the steps involved in conducting a HTA. Subsequently, we direct our attention to two main categories of research methods of collecting primary data (experimental (randomized) and non-experimental (non-randomized)) as well as describing the main ethical, socio-cultural and legal issues that can be considered in HTAs.

3.0 BASIC STEPS IN HEALTH TECHNOLOGY ASSESSMENT (HTA)

Health Technology Assessment (HTA) applies rigorous, systematic methods of scientific inquiry to the evaluation and use of new or existing health care technologies. The main purpose of HTA is to inform decisions made at the individual or patient level, the level of the health care provider or institution, or at the regional, national and international levels (i.e., policy-making for technology in health care; http://www.nlm.nih.gov/nichsr/hta101/ta10104.html).

Vast amounts of practical experience have been accumulated around the world, and a large body of relevant literature (both on methods and on particular technologies) exists. The term **evidence-based medicine** refers to the use of current best evidence from scientific and medical research, and the application of clinical experience and observation, in making decisions about the care of individual patients (http://www.nlm.nih.gov/nichsr/hta101/ta10104.html).

Although there is great variation in the scope, selection of methods, and level of detail in the practice of HTA, most HTA activity involves some form of the following basic steps (http://www.nlm.nih.gov/nichsr/hta101/ta10104.html):

- Identify assessment topics
- Specify the assessment problem
- Determine locus of assessment
- Retrieve evidence
- Collect new primary data (as appropriate)
- Appraise/interpret evidence
- Integrate/synthesize evidence
- Formulate findings and recommendations
- Disseminate findings and recommendations
- Monitor impact

Not all assessment programs conduct these steps (refer to Appendix A to view an alternative framework), and they are not necessarily conducted in a linear manner. Many HTA programs rely largely on integrative methods of reviewing and synthesizing data from existing primary data studies (reported in journal articles or from epidemiological or administrative data sets),

and do not collect primary data. Some assessment efforts involve multiple cycles of retrieving/collecting, interpreting, and integrating evidence before completing an assessment (http://www.nlm.nih.gov/nichsr/hta101/ta10104.html). Depending upon the circumstances of an HTA, the dissemination of findings and recommendations and monitoring of impact may not be parts of the HTA itself, although they may be important responsibilities of the sponsoring program or parent organization.

For more detailed information regarding HTA, we strongly recommend that the readers refer to the HTA Initiative Series published by the Health Technology Assessment Unit of the Alberta Heritage Foundation for Medical Research. The Unit has published a series of HTA Initiatives that address in detail a variety of important HTA issues. The HTA Initiative publications can be downloaded from:
http://www.ahfmr.ab.ca/publications/index.php?dept=1&search=

Of particular interest, the AHFMR Series contains the following HTA Initiative documents:

#1 "Framework for Regional Health Authorities to Make Optimal Use of Health Technology Assessment"

#2 "Making Managerial Health Care Decisions in Complex High Velocity Environments"

#3 "Proceedings of the Conference on Evidence Based Decision Making: How to Keep Score"

#4 "AHFMR Screening Procedures for Use When Considering the Implementation of Health Technology" Released April 2001

#5 "Priority Setting in Health Care: From Research to Practice"

#6 "Screening Procedures for Use When Considering the Implementation of Health Technology" Released April 2002

#7 "Local Health Technology Assessment: A Guide for Health Authorities"

#9 "Elements of Effectiveness for Health Technology Assessment Programs"

#11 "Decision-Making for Health Care Systems: A Legal Perspective"

#12 "Review of Health Technology Assessment Skills Development Program"

#13 "Standard Quality Assessment Criteria for Evaluating Primary Research Papers from a Variety of Fields"

#14 "Workshop Summary Knowledge-Brokers: Linking Researchers and Policy Makers"

#15 "Quantitative Approaches to Patient Safety. Research in Risk Analysis and Risk Management as Applied to Radiotherapy"

#16 "An Exploratory Review of Evaluations of Health Technology Assessment Agencies"

4.0 RESEARCH METHODOLOGY TYPICALLY CONSULTED IN HTA

The role of methodology may easily be underestimated because the greatest controversies and more visible impact of research tend to arise naturally in relation to evidence of the effectiveness, or lack of effectiveness, of specific interventions (Lilford, Richardson, Stevens, Fitzpatrick, Edwards, Rock, & Hutton, 2001). Evidence from evaluations of specific interventions is primarily intended to inform judgments about their value, and, ultimately, the appropriate extent of public provision. Considerable professional, scientific, and media and public attention is therefore devoted to the substantive results of evaluative studies, especially when such evidence strongly indicates the value (or lack of value) of a drug, form of surgery, diagnostic technique or other form of service (Lilford, Richardson, Stevens, Fitzpatrick, Edwards, Rock, & Hutton, 2001).

Research methodology may seem removed, abstract and less policy-relevant to "real world" decision-making in healthcare systems. However, there are several reasons for arguing that methodology may have greater significance and therefore warrant more direct attention (Lilford, Richardson, Stevens, Fitzpatrick, Edwards, Rock, & Hutton, 2001). Indeed, health professionals do not readily or automatically accept and act upon evidence-based medicine, in contrast to other sources of information about practice (Lilford, Richardson, Stevens, Fitzpatrick, Edwards, Rock, & Hutton, 2001). The merits and role of non-randomized evidence for healthcare interventions are considered very modest within the methodological paradigm of evidence-based medicine, which gives greatest weight to well-conducted randomized trials and meta-analyses of such trials (Lilford, Richardson, Stevens, Fitzpatrick, Edwards, Rock, & Hutton, 2001). Debates about the merits of observational evidence, however, although expressed in technical and methodological terms, can often appear to reflect more basic conflicts. In short, the position one takes on the value of data from different health service research designs will affect the final interpretation of the technology in question (Lilford, Richardson, Stevens, Fitzpatrick, Edwards, Rock, & Hutton, 2001).

Research design can be thought of as the structure of research; it is the "glue" that holds all of the elements in a research project together. 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 turn our attention to a brief overview of information gathered from: (1) Experimental studies, and (2) Non-Experimental studies. Although by no means an exhaustive or all-encompassing discussion, the following sections are intended to help familiarize the reader with some basic research terminology that will facilitate an understanding of how to evaluate evidence derived from differing methodological research designs.

Please note that the main research terminology described here (experimental vs non-experimental) are less well known than the major clinical research terminology well defined as clinical trials (randomized or not) and epidemiological studies (e.g., observational (cohort or case-control study) and outcomes research, and clinical audits). Observational and outcomes research, and clinical audits are forms of non-experimental research. The main difference between clinical trials and epidemiological studies is that in clinical trials, the investigators manipulate the administration of a new intervention and measure the effect of that manipulation. In contrast, epidemiological studies only observe associations (correlations) between the treatments experienced by participants and their health status or diseases (http://en.wikipedia.org/wiki/Clinical_trial). All clinical trials may or may not be randomized. (See Module 4 and Module 5 for further details on Clinical Trials and on Observational and outcomes research and clinical audits)

4.1 EXPERIMENTAL (RANDOMIZED) STUDIES

Experimental studies have been utilized to evaluate the efficacy of drug therapy and other therapeutic modalities. These studies include the common study designs such as clinical trials, the type of study where most drug comparisons take place (rational pharmaceutical management plus program, 2001). the experimental study tests the factor of influence under investigation with respect to its effect on the target variable, together with a probability of error in determining an effect in the respective experiment. the probability of type 1 error (typically 5%, sometimes 1%) is defined by the researcher a priori (dannehl, 1997). If random assignment is used, we call the design a randomized experiment or true experiment.

Experimental design is a fairly complex subject and there are many variations that attempt to accomplish different things or solve different problems (e.g., clinical trials, intervention studies, randomized controlled trials). In the simplest type of experiment (i.e., **two-group experimental design)**, two equivalent groups are created (Trochim, 2002). One group is the program or treatment and receives the program, while the second group is the comparison or control group and does not receive the program or treatment. In all other respects the groups are treated the same. For instance, the groups include similar people, with similar backgrounds, living in similar contexts (Trochim, 2002). Therefore, if differences in outcomes between these two groups are observed, then one concludes that the differences are due to the only thing that differs between them, the one received the treatment and the other did not. Equivalent groups are created using the approach of random assignment and this is the key to the success of the experiment. People are randomly assigned from a common pool of people into the two groups. We rely on the idea of probability and assume that the two groups are probabilistically equivalent (Trochim, 2002). If we randomly assign people to two groups and have enough people in our study to achieve the desired probabilistic equivalence, then we may consider the experiment to be strong in internal validity and we probably can assess whether the program or treatment causes the observed outcome(s).

Randomized Controlled Trials (RCT) have become the standard technique for testing and evaluating new drugs and clinical procedures (Medical Research Council; MRC, 2003). The RCT is the most important type of experimental study used and provides the most reliable results. In RCT designs, similar subjects are assigned at random to a treatment or a control group to see if they develop the outcome of interest. Confounding factors are eliminated and powerful statistics can be used to show causality (Jones-Harris, 2003). It is considered the "gold standard" of experimental studies (Rational Pharmaceutical Management Plus Program, 2001). Accordingly, RCTs are considered the best type of study from which to draw conclusions on effectiveness of treatments, since they are the only studies that will control for bias and confounding variables, and ultimately provide the most accurate and reliable results (Rational Pharmaceutical Management Plus Program, 2001). Randomized experimental studies are generally considered the strongest of research designs when one's interest is in establishing a cause-and-effect relationship. However, as we will see in a subsequent

section, RCTs are expensive and time-consuming, and can raise ethical concerns about treatment strategies (Jones-Harris, 2003; MRC, 2003; Rational Pharmaceutical Management Plus Program, 2001).

Since 1992, there have been many developments in the methodology of RCTs and in other methods of HTA (MRC, 2003). Although there is still a pressing need for more surgeons to be effectively trained in trials methodology, surgical procedures have increasingly been validated in rigorous comparative studies (MRC, 2003).

For instance, as an alternative or extension to controlled trials, where it is not pragmatic or ethical to conduct a trial, or where questions remain unanswered following a trial, research audits may be the method of choice. A national research **audit** aims to answer specific research questions, such as the long-term safety and effectiveness of a procedure, by the use of routine data collection. An audit may show whether the benefits expected from the evidence provided by a randomized controlled trial were achieved when a procedure has been adopted into the wider framework of clinical practice. ASERNIP-S believes that surgeons, government and consumers stand to benefit from this type of comprehensive national research audit because they will help improve Australian health outcomes (AERNIP, 2003). A clinical audit is a "quality improvement process that seeks to improve patient care and outcomes through systematic review of care against explicit criteria and the implementation of change" (AERNIP, 2003).

[For a more comprehensive look at how to assess clinical trials in surgery, consult Module 4 of our series of Health Technology Assessment Modules and Workshops. For more information on research audits, refer to Module 5 of our series of Health Technology Assessment Modules and Workshops.]

4.2 NON-EXPERIMENTAL (NON-RANDOMIZED) STUDIES

Most methodologists agree that the experimental study is the best method for medical research. However, often one has to be satisfied with non-experimental designs for gaining knowledge. This is due to organizational and economic, as well as legal and ethical limits that are often encountered when conducting experiments with humans. Consequently, some have described the non-experimental (non-randomized) study as a methodological compromise, with the experimental study serving as the guiding method (Dannehl, 1997). If random assignment is not used, and multiple groups and multiple waves of measurement are not employed, then we are dealing with a **non-experimental design**.

Much research is not experimental (i.e., no variable is manipulated), but rather "observational" (and some use the term "correlational"). A non-experimental study is

considered the weakest type of design with respect to internal validity and causal assessment. The non-experimental study can only, if at all, test differences between compared groups with respect to the target variable. The probability of type 1 error (e.g., typically 5% or in some cases 1%) in determining a difference in the respective target vairable is defined a priori by the researcher. Upon obtaining a statistically significant difference in a non-experimental study, it remains unknown whether the difference detected stems from the target factor under investigation, or whether it is due to other confounding and intervening factors (Dannehl, 1997). However, for some research questions, especially descriptive ones, the non-experimental design is clearly the best choice.

Although we will maintain a select focus on observational studies as forms of nonexperimental methods in collecting empirical data for health technology assessments, it is important to recognize the various classifications of non-randomized studies. In the following section we will briefly introduce various types of non-randomized studies: **(1) quasi-experimental studies and (2) observational studies**. Despite all the differences between these types of non-randomized studies, the differences are only gradual in nature. In reality, the greatest leap of quality lies between "experimental" (randomized) and "non-experimental" (non-randomized) studies.

4.2.1 QUASI-EXPERIMENTAL STUDIES

Sometimes quasi-experimental studies are classified as an independent design, falling between the experimental and non-experimental design spectrum. A quasi-experimental design is one that looks similar to an experimental design, but lacks the key ingredient of random assignment. However, it does use either multiple groups or multiple waves of measurement. With respect to internal validity (i.e., approximate truth about inferences regarding cause-effect relationships), quasi-experimental studies often appear inferior to randomized experiments (Trochim, 2002). One of the intended purposes for employing quasi-experimental research is to capture longer time periods and a sufficient number of different events to control for various threats to validity and reliability. Typically, the word "trend" is used instead of "cause" as finding the one true trend is an important goal of quasi-experimental designs.

Among the assortment of quasi-experimental designs that have specific applicability and noteworthy features are the following:

(1) Time Series Designs

A time series is the most common type of longitudinal (over time) research, and it can be interrupted or non-interrupted. Both types examine changes in the dependent variable over time, with only an interrupted time series involving before and after measurement. This kind of research is sometimes

referred to as impact analysis or policy analysis. A single group of participants is tested repeatedly both before and after a manipulation or a natural event. The multiple measures permit the detection of confounding variables. With interrupted time series designs, multiple observations are made over time and are "interrupted", usually by an intervention or treatment (http://www.epoc.uottawa.ca/inttime.pdf). The investigators must indicate a specific point in time when the intervention occurred. A control group may or may not be present.

(2) Proxy Pretest Design

Resembles a standard pre-post design, but the pretest is collected after the program or treatment is administered. A "proxy" variable is used to estimate where the groups would have been on the pretest. This design is not one that should be selected by choice (Trochim, 2002). However, it is useful in situations where a program or treatment needs to be evaluated, but has already started or been administered.

(3) Separate Pre-Post Samples Design

In this design, the people used for the pretest are not the same as the people used for the posttest. Of the four groups, two come from a single, nonequivalent group and the other two also come from a single nonequivalent group (Trochim, 2002). This is not a strong design because individual participant responses cannot be matched from pre to post, and changes can only be examined across the "average" person.

(4) Double Pretest Design

A very strong quasi-experimental design with respect to internal validity. It includes two measures prior to the program. If the program and comparison group are maturing at different rates, this will be a detected change from pretest 1 to pretest 2. Thus, this design explicitly controls for selection-maturation threats (Trochim, 2002).

(5) Switching Replications Design

This is also a very strong design with respect to internal validity, as well as external validity or generalizability because it allows for two independent implementations of the program (Trochim, 2002). The design has two groups and three waves of measurement. In the first phase of the design both groups are pretests, one is given the program and both are posttested. In the second phase of the design the original comparison group is given the

program while the original program group serves as the "control". This design is identical in structure to the randomized experimental version, but lacks random assignment to groups. In addition, it assures all participants eventually receive the program or treatment, and thus is one of the most ethically feasible quasi-experiments (Trochim, 2002).

(6) Nonequivalent Dependent Variables Design

In its simple form, it is an extremely weak design with respect to internal validity, but in its pattern matching variations, it opens the door to an entirely different approach to casual assessment that is extremely powerful (Trochim, 2002). The idea with this design is that a program is designed to change a specific outcome. The "control" variable must be similar enough to the "target" variable to be affected in the same way by history, maturation, and other single group threats to internal validity, but not so similar that it is affected by the program (Trochim, 2002).

(7) Pattern Matching Design

This design can be quite strong with respect to internal validity, especially if there is a larger set of variables and the expectation pattern matches well with the observed results. It requires that the researcher specify expectations prior to the institution of the program, as well as a large set of outcome variables and a detailed sense of how they are related to each other.

(8) Regression Point Displacement Design

The regression point displacement design attempts to enhance a single program unit situation by comparing the performance on that that single unit with the performance of a large set of comparison units. For example, in community research, the pre-post results for the intervention community would be compared with a large set of other communities (Trochim, 2002). The advantage of doing this is that we do not rely on a single nonequivalent community, but rather use results from a heterogeneous set of nonequivalent communities to model the comparison condition, and then compare the single site to this model (Trochim, 2002). This design is especially applicable in situations where a treatment or program is applied in a single geographical unit (e.g., city, hospital, hospital unit) instead of an individual, where there are lots of other units available as control cases and where there is routine measurement of relevant outcome variables (Trochim, 2002).

(9) non-equivalent groups design

The most commonly used quasi-experimental design is the **non-equivalent groups design**. In its simplest form, it requires a pretest and posttest for a treated and comparison group, but the groups are not created through random assignment (Trochim, 2002). The researcher uses intact groups that are thought to be similar as the treatment and control groups. Since the researcher does not control the assignment to groups through the mechanism of random assignment, the groups may be different prior to the study and these differences may affect the study outcome.

(10) regression-discontinuity design

The second common quasi-experimental design is the **regression-discontinuity design**. The distinguishing characteristic with this design is assignment to treatment using a cutoff score on a pretreatment variable (Trochim, 2002). This design is useful since it allows the assignment of subjects to the program who need or deserve it most. The regression-discontinuity design does not require the assignment of potentially needy individuals to a no-program comparison in order to evaluate the effectiveness of a program. Also known as causal comparative, as such studies describe existing differences and try to identify cause (Trochim, 2002). From a methodological point of view, inferences drawn from a well-implemented regression-discontinuity design are comparable in internal validity to conclusions from randomized experiments. From an ethical perspective, such a design is compatible with the goal of getting the program or treatment to those in most need (Trochim, 2002).

4.2.2 OBSERVATIONAL STUDIES

In biomedical research, observational studies are used for detecting potential causes of health care problems and to examine how exposure to risk factors that 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 issue with the use of observation is that it does not establish causal links between variables. There is always a possibility that some unknown factor is exerting an influence over the dependent variable. Moreover, 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 designs range from relatively weak studies like descriptive and ecological studies to strong designs like case control and cohort studies (Pai, 2006). The following discussion will provide a general overview of the various observational designs. [For a more comprehensive discussion and information on how to assess observational, outcome, and audit research in surgery, consult Module 5 in the Health Technology Assessment Module and Workshop Series.]

(1) Correlational or Descriptive Studies (Case-Report/ Case-Series)

- Considered the weakest epidemiological design, the investigators merely 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 series is an example of a descriptive study.
- 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 further hypothesis testing. No matter how convincing data from descriptive and correlational studies may appear, because they have less control over the variables and environments that they study, these non-experimental designs cannot establish cause and effect relationships.

(2) Ecological Studies

- While also considered a weak design, ecological studies can be useful in generating hypotheses. Like the descriptive study design, while apparent association between variables may be observed, no causal inference can be drawn. An apparent ecological link may not be a true link, as it could be confounded by several other factors (Pai, 2006). The units of study are populations rather than individuals.

(3) Cross-Sectional Studies

- Widely used to estimate the prevalence of disease or the prevalence of exposure to risk factors or both. Estimating the prevalence of conditions can be useful in predicting the need for health service use.
- These types of studies are easy to do and tend to be economical since repeated data collection is not done (Pai, 2006). Measurements are made on a population at one point in time. Cross-sectional designs measure the prevalence of disease and thus are also called prevalence studies. Since there is no longitudinal component, cross-sectional surveys cannot possibly

measure incidence of any disease. The main problem 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).

(4) Case Control Studies

- 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).
- 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). Other disadvantages are that case control studies rely on records to determine exposure, it is often difficult to select control groups, and it is difficult to eliminate confounding variables. According to Greenhalgh (1997), 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.

(5) Cohort Studies

- 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.
- Cohort studies are usually prospective (i.e., forward looking). They 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).

- Cohort studies are 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). In addition, 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).Nonetheless, the clear temporal (time) sequencing is extremely important while making causal inference.

5.0 ETHICAL, SOCIO-CULTURAL, AND LEGAL ISSUES

Similar to clinical practice, research involving human subjects can raise difficult and important ethical and legal questions. The field of research ethics is devoted to the systematic analysis of such questions to ensure that study participants are protected and, ultimately, that clinical research is conducted in a way that serves the needs of participants and society (Weijer, Dickens, & Meslin, 1997). To meet these ends and ensure that clinical research is conducted with the highest scientific and ethical standards, various ethical principles, legal requirements and policy statements have been formulated. In the proceeding sections, we briefly present some of the more salient issues within each of these domains.

5.1 ETHICS

In a document entitled the Belmont Report, the United States National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research set out a predominant ethical framework for human experimentation (Weijer, Dickens, & Meslin, 1997). The report articulated three guiding principles for research.

(1) **Respect** for persons requires that the choices of autonomous individuals be respected and that people who are incapable of making their own choices be protected. This principle underlies the requirement to obtain informed consent from study participants and to maintain confidentiality on their behalf (Weijer, Dickens, & Meslin, 1997).

(2) The principle of **beneficence** requires that participation in research be associated with a favorable balance of potential benefits and harms (Weijer, Dickens, & Meslin, 1997). When effective standard treatment exists for a disease, it is unethical to expose patients to the risk of "treatment" with placebo alone, since placebo is an inferior treatment (Weijer, Dickens, & Meslin, 1997).

(3) The principle of **justice** entails an equitable distribution of the burdens and benefits of research. Researchers must not exploit vulnerable people or exclude without good reason eligible candidates who may benefit from participation in a study (Weijer, Dickens, & Meslin, 1997).

In addition to these main principles, other ethical requirements have been identified as important for researchers and assessors to observe. Weijer, Dickens, and Meslin (1997) cited the following:

- A study must employ a scientifically valid design to answer the research question.

- A study must address a question of sufficient value to justify the risk posed to participants. Exposing subjects even to low risk to answer a trivial question is unacceptable.

- A study must be conducted honestly. It should be carried out as stated in the approved protocol, and research ethics boards have an obligation to ensure that this is the case.

- Study findings must be reported accurately and promptly. Methods, results and conclusions must be reported completely and without exaggeration to allow practicing clinicians to draw reasonable conclusions. Whenever possible, study results should be reported quickly to allow physicians timely access to potentially important clinical information.

A very useful peer-reviewed journal devoted exclusively to research ethics is *IRB: A Review of Human Subjects Research* (Weijer, Dickens, & Meslin, 1997). Another valuable resource is the *Canadian Medical Association Journal's* series on bioethics topics for clinicians. Please refer to the appendices at the end of the module for further information on useful resources on Canadian ethics and guidelines (e.g., http://sprojects.mmi.mcgill.ca/ethics/X/topics/research/researchmain.htm; http://www.wma.net/e/policy/b3.htm).

5.2 LAW

Many of the ethical principles delineated above are also established in law. For instance, the legal doctrine often described as "informed consent" is better understood as "**informed choice**", since it is a physician's legal duty to inform the patient so (s)he may exercise choice, which does not necessarily always result in consent (Weijer, Dickens, & Meslin, 1997). Failure by the physician to disclose information relevant to the choice a patient is asked to make falls under the **law of negligence** within the civil law (Weijer, Dickens, & Meslin, 1997).

Similarly, patients who are invited to enter a study must be informed of the nature and extent of the known risks of participation, the possibility that participation may present unknown

risks, and the intended benefit of the study to participants and others (Weijer, Dickens, & Meslin, 1997). Providing treatment without the patient's consent may be grounds for legal action on the basis of "**unauthorized touching**" which in criminal law is **assault** and in civil law is **battery** (Weijer, Dickens, & Meslin, 1997).

Fiduciary duty is the highest standard of duty implied by the law and requires that physicians disclose information about a patient only in the patient's best interests and that they avoid any conflict of interest in the disclosure of patient information, even if that information is contained in the records physicians lawfully hold (Weijer, Dickens, & Meslin, 1997). **Unauthorized disclosure** is actionable as a breach of fiduciary duty. It may also violate a duty of confidentiality enacted in provincial legislation.

5.3 POLICY

A number of international policies guide the conduct of research. Some of these include the:

- Nuremberg Code
- International Covenant on Civil and Political Rights
- World Medical Association's Declaration of Helsinki

The Declaration of Helsinki is probably the most influential document governing research worldwide (Weijer, Dickens, & Meslin, 1997). The Declaration emphasizes that patients' participation in research should not put them at a disadvantage with respect to medical care (Weijer, Dickens, & Meslin, 1997). For those conducting studies funded by different countries, it is expected that they conduct their research in accordance with the regulations of that country. For specific information, it is recommended that researchers and assessors consult the guidelines of the Council for International Organizations of Medical Sciences (Weijer, Dickens, & Meslin, 1997).

Medical research in Canada is governed by guidelines of the Medical Research Council (MRC) of Canada (Weijer, Dickens, & Meslin, 1997). Proposals for research involving human subjects must be submitted to a local research ethics board for review. However, research that will not generate generalizable knowledge (e.g., quality assurance research for internal use and not intended for publication) is generally considered exempt from such review (Weijer, Dickens, & Meslin, 1997).

5.4 THE IMPORTANCE OF ETHICS IN ASSESSMENT

Empirical studies have much to contribute to our understanding of informed consent and the risks and benefits of participation in research (Weijer, Dickens, & Meslin, 1997). If the principle of respect for persons is to be upheld, it follows that research subjects must be **informed** of the purpose, nature, risks, benefits and alternatives associated with their participation, as well as **understand** this information (Weijer, Dickens, & Meslin, 1997). Physicians, assessors, and researchers alike must establish and maintain effective strategies to ensure that research subjects comprehend the information they are given during the consent process. Ethical issues in research must not be addressed as an afterthought. Ethical issues permeate research and must guide research design (Weijer, Dickens, & Meslin, 1997).

The development of the contemporary field of medical ethics owes much to the growth and increasing sophistication of biomedical technology. Technology has affected both the ethical questions that we ask and the ways in which we ask them. Although the basic focus of medical ethics is the moral nature of human interaction in the realm of health care, the ways in which technology affects that interaction and how medical technology can be used to improve the human condition are essential questions for the discipline (Heitman, 1998; Macklin, 1999). According to Singer (2000), the revolution in information technology will continue to dramatically change medical practice. This subject raises many ethical issues, including confidentiality of electronic medical records, and the relation of clinical records to research and management of health systems (Singer, 2000).

Surgical procedures are frequently introduced into general practice on the basis of uncontrolled studies that are less rigorous than those required for the approval of medical interventions (Freeman, Vawter, Leaverton, Godbold, Hauser, Goetz, & Olanow, 1999). The standard for the evaluation of surgical therapy is lower because of the complexity of designing and conducting scientifically valid and ethically acceptable clinical trials of surgical procedures (Freeman, Vawter, Leaverton, Godbold, Hauser, Goetz, & Olanow, 1999). As a result, many surgical trials fail to control for investigator bias or placebo effects (Freeman, Vawter, Leaverton, Godbold, Hauser, Goetz, & Olanow, 1999). Ultimately, ethically responsible and scientifically sound human-subjects research should seek to advance the state of medical knowledge and clinical practice without knowingly compromising the welfare and integrity of individual trial participants in the process (London & Kadane, 2002).

5.5 HEALTH TECHNOLOGY ASSESSMENT AND ETHICAL CONSIDERATIONS

The field of health care technology assessment has developed concurrently with the discipline of medical ethics. Health care technology assessment is focused evaluation of the nature, purposes, use, and consequences of technology used in the pursuit of health and improved

quality of life (Heitman, 1988). Although not all of the specific applications of technology assessment have ethically-oriented goals, there are ethical dimensions in all of the varied techniques and functions of technology assessment (Heitman, 1998). Many leaders in technology assessment claim that its ultimate ethical purpose is to change health care policy and practice to improve the health of individual patients and/or society at large (Heitman, 1998).

The general ethics of technology assessment rests on the fundamental presupposition that health care technologies should help those to whom they are applied (Heitman, 1998). The measure of ethical technology assessment must be the integrity of the specific project and its practitioners, from a critical appraisal of the goals of the project and its place in the overall mission of its sponsors to the intended use for the conclusions and the areas in which other interests might compromise the project. The internal consistency of the project's goals, procedures, and effects, and evaluators' open and honest acknowledgment of their purposes, are minimum standards for ethically and methodologically sound assessments (Heitman, 1998).

Health technology assessment (HTA) is a form of policy research that systematically examines short-and long-term consequences in terms of health and resource use, of the application of a health technology, a set of related technologies, or a technology-related issue (Henshall, Oortwijn, Stevens, Granados, & Banta, 1997). The goal of HTA is to provide input to decision making in policy and practice (Henshall, Oortwijn, Stevens, Granados, & Banta, 1997). The most frequent activity in HTA is a synthesis or systematic review of available information, especially on efficacy and cost-effectiveness, to assist different types of policy decisions (Henshall, Oortwijn, Stevens, Granados, & Banta, 1997). Despite its policy goal, HTA is and must be firmly rooted in science and the scientific method. The process of technology assessment must be carried out with integrity and the results must be valid (Henshall, Oortwijn, Stevens, Granados, & Banta, 1997).

In the sections to follow, each of the steps of conducting HTA is presented in relation to their relevant ethical considerations.

5.5.1 IDENTIFICATION OF TECHNOLOGIES FOR ASSESSMENT

The identification of subjects for assessment should depend on criteria explicitly stated in advance, consistent with the mission of the group or organization sponsoring the project and its identified motives for conducting the technology assessment generally (Heitman, 1998). Such criteria should conform to any formally stated ethical ideals or goals in order to prevent internal conflicts of interest (Heitman, 1998). Technologies that are controversial are often obvious candidates for assessment. An assessment of the routine

use of a common technology may be more useful than an evaluation of a less frequently used, high-profile intervention (Heitman, 1998).

5.5.2 SPECIFICATION OF THE ASSESSMENT PROBLEM

How a problem is defined inherently influences its evaluation. An assessment should describe its central variables carefully and justify their inclusion in light of the evaluators' own purposes for conducting the assessment and the projected use by its target audience (Heitman, 1998). Assessments may be applied in ways never intended by the evaluators, at times even in conflict with their intended purpose (Heitman, 1998).

5.5.3 RETRIEVAL OF THE EVIDENCE

The quality of evidence in health care is typically determined by the process of peer review; authoritative information is that which has been published after scrutiny and endorsement by experts in the field (Heitman, 1998). This standard also applies in health care technology assessment. Before considering data for selection, evaluators should establish the criteria by which they will define reliable and relevant evidence. These criteria may need to be reassessed as evaluators discover the extent and quality of available evidence (Heitman, 1998).

5.5.4 COLLECTION OF NEW PRIMARY DATA

Because the careful collection of primary data on new technologies can be as logistically complex, time consuming, and expensive as the rest of the assessment combined, few organizations are willing to evaluate cutting-edge technologies early in their life cycle when the potential return on their investment is unknown. Original evaluation taxes institutional resources to such an extent that priority is typically given to the assessment of technologies for which substantial data already exist (Heitman, 1998). This constraint creates an ironic cycle in which technologies remain unevaluated until they are widely accepted into practice, at the risk of harmful physical consequences for patients and financial consequences for institutions and society (Heitman, 1998). When there is sufficient experience with a technology to evaluate its use meaningfully, there may still be areas in which new data must be gathered. This is particularly an issue with outcomes research, upon which practice guidelines claims to rely (Heitman, 1998).

The key ethical questions in this context are: (1) the extent to which the assessment group and its target audience are willing to tolerate uncertainty about the technology's use and effects, and (2) whether an evaluation based on incomplete data is better than no evaluation at all (Heitman, 1998).

5.5.5 INTERPRETATION OF THE EVIDENCE

The initial step in interpreting evidence is the establishment of criteria for its inclusion and role in the review; not all data available on a given technology may be suitable or equally useful for the purposes of a specific assessment (Heitman, 1998). The methods and presuppositions of both peer-reviewed publications and other material should be scrutinized on several grounds. Using formal criteria of methodological rigor and clarity is essential for grading the data and their applicability to the assessment (Heitman, 1998).

5.5.6 SYNTHESIS OF THE DATA

The use of good data in policy making cannot, in itself, ensure the quality of policy. The synthesis of evidence from different studies and different aspects of the same issue require both a clear understanding of the questions that each is capable of addressing and a considered choice of methods (Heitman, 1998). Of the techniques commonly used to synthesize data, meta-analysis, cost analysis, and decision analysis are often presumed to be objective and verifiable because they employ standardized mathematical methods to reach their conclusions (Heitman, 1998). Because each of these methods is influenced by subjective interpretation of facts and values and may vary with the purpose of specific assessments, procedural and conceptual integrity in the synthesis of evidence is a more appropriate goal than absolute objectivity (Heitman, 1998).

5.5.7 FORMULATION OF FINDINGS AND RECOMMENDATIONS

A project's findings and recommendations are the central elements of interest for most readers. The primary ethical concern for the construction of the final report is that it be as intelligible as possible to its intended audience (Heitman, 1998). Findings and recommendations should be phrased in a format parallel to that of the statement of the original questions; where conclusions cannot be reached from the evidence considered, some commentary is needed on why certain questions cannot be answered (Heitman, 1998). Limitations should be clearly acknowledged and described.

5.5.8 DISSEMINATION OF FINDINGS AND RECOMMENDATIONS

One of the fundamental ethical precepts in biomedical research is that results should be available to others interested in the problem through the published literature and informal collegial communication (Heitman, 1998). Technology assessments that appear inconclusive should not be withheld from distribution. If they are methodologically sound, such assessments may be both useful to policy makers and clinicians and also serve as a point of departure for future research (Heitman, 1998). While the standard of openness is well accepted by academic researchers and governmental agencies, researchers and administrators in the drug and device industry, insurance companies, and some health care institutions may be reluctant to share assessments that they perceive to be

proprietary information (Heitman, 1998). In a competitive health system, it is hard to argue for public access to privately conducted privately funded assessments intended to enhance an organization's profitability (Heitman, 1998). However, researchers and companies that engage in technology assessment would be wise to make the process open to external review and criticism to avoid the dangerous consequences of self-deception and unchecked self-interest (Heitman, 1998).

5.5.9 MONITORING THE EVALUATION'S IMPACT

Like technology itself, technology assessment can have both intended and unexpected consequences. Monitoring the impact of an evaluation is essential to maximizing its intended effects and preventing the harmful repercussions of misinterpretation or misapplication (Heitman, 1998). An assessment project should include a plan for the follow-up evaluation of its report as well as strategies for responding to any questions and criticisms (Heitman, 1998). Because technology assessment is an iterative process, new information or changes in the technology may require the re-evaluation of the project's original conclusions (Heitman, 1998).

6.0 CONCLUDING SUMMARY

Many HTA programs have described a number of basic steps involved in HTA, even though not all of these steps may be utilized in any one HTA program. Health Technology Assessment relies primarily on the use of integrative methods of reviewing and synthesizing data from existing primary research studies. In some cases, however, the process of doing effective HTA can also include the process of collecting primary data through local research initiatives. As we have presented in this modules, HTA needs to incorporate information collected from a wide variety of research based sources of evidence.

Some have argued that in the "known" absence of experimental studies, one does what one can with what is available. While this is common sense, it is important for decision- and policy-makers to have a basic understanding of the basic research methods and ethical, socio-cultural and legal issues that may be taken into consideration in a health technology assessment.

There are many areas where neither experimental nor quasi-experimental designs are technically or ethically possible (Macdonald, 2003). There is little argument that RCTs are an accepted high standard for testing effectiveness under ideal circumstances, but they may not be the best way to evaluate all the interventions and technologies that decision makers are considering (Eisenberg, 1999). For example, when the question is "Which service or technology has more adverse events associated with it?" a case-control study using a large database may be better than an analysis of the number of patients who dropped out of a trial because they could not tolerate an intervention (Eisenberg, 1999). Observational studies with analyses that consider potential bias offer an opportunity to capture data from community practices costing less than randomized trials.

As we have presented in this module, the randomized trial is unlikely to be replaced, but it should be complemented by other designs that address questions about technology from different perspectives (Eisenberg, 1999). Researchers need to develop and test new ways of evaluating technologies that can be accomplished quickly and can take advantage of emerging databases and information systems (Eisenberg, 1999). According to Eisenberg (1999), those who conduct technology assessments should be as innovative in their evaluations as the technologies themselves. Physicians should be active users of technology and technology assessments and should seek out evidence and use it every day. Clinicians must do their part to build the evidence base for health care: conduct research, become involved in clinical trials, and ask: "What is the evidence?" Managers and decision makers need to be more informed about basic research designs and develop a better understanding of the implications this research has on their own departments and organizations. Continued development of medical technologies has brought enormous benefits to patients, but also a collective responsibility to ensure that these technologies are deployed appropriately (Eisenberg, 1999).

6.1 REVIEW OF MODULE OBJECTIVES

By the end of this third HTA module, participants should be able to:

(1) Discuss a model of how to conduct a HTA,

(2) Explain the difference between primary data collection procedures, specifically experimental (randomized) and non-experimental (non-randomized) studies, and

(3) Identify the ethical, socio-cultural, and legal issues relevant to conducting research and HTA.

7.0 REFERENCES

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.

Alberta Heritage Foundation for Medical Research; AHFMR. *Health technology assessment Initiative Series at :* http://www.ahfmr.ab.ca/publications/index.php?dept=1&search=

American Medical Association. (1992). Evidence-based medicine: A new approach to teaching the practice of medicine. Based on the *Users' Guides to Evidence-based Medicine and reproduced with permission from JAMA. 268 (17)*: 2420-5.

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

Battista, R.N. (2000-2001). Unit 2: Towards a paradigm for technology assessment. In 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

Battista, R.N. & Hodge, M.J. (1999). The evolving paradigm of health technology assessment: reflections for the millennium. *JAMC, 160(10)*; 1464 – 1467.

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.

Dannehl, K. (1997). Nonrandomized comparative clinical studies. *Proceedings of the International Conference on Nonrandomized Comparative Clinical Studies in Heidelberg*. April 10 –11. http://www.symposion.com/nrccs/dannehl.htm

Eisenberg, J.M. (1999). Ten lessons for evidence-based technology assessment. *JAMA, 282(19)*; 1865 – 1869.

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

Flaherty, R.J. (2004). A simple method for evaluating clinical literature. *Family Practice Management*. http: www.aafp.org/fpm/20040500/47asim.html

Freeman, T.B., Vawter, D.E., Leaverton, P.E., Godbold, J.H., Hauser, R.A., Goetz, C.G., & Olanow, C.W. (1999). Use of placebo surgery in controlled trials of a cellular-based therapy for parkingson's disease. *New England Journal of Medicine,* 341; 988 – 992.

Greenhalgh, T. (1997). Assessing the methodological quality of published papers. *BMJ, 315*; 305-308.

Heitman, E. (1998). Ethical issues in technology assessment. Conceptual categories and procedural considerations. *International Journal of Technology Assessment in Health Care, 14:3*; 544-566.

Henshall, C., Oortwijn, W., Stevens, A., Granados, A., & Banta, D. (1997). Priority setting for health technology assessment. *International Journal of Technology Assessment in Health Care, 13(2)*; 144-185.

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

Lilford, R.J., Richardson, A., Stevens, A., Fitzpatrick, R., Edwards, S., Rock, F., & Hutton, J.L. (2001). Issues in methodological research: perspectives from researchers and commissioners. *Health Technology Assessment, 5(8)*; 1 – 62.

London, A.J. & Kadane, J.B. (2002). Placebos that harm: sham surgery controls in clinical trials. *Statistical Methods in Medical Research, 11;* 413 – 427.

Macdonald, G. (2003). Using systematic reviews to improve social care. Social Care Institute for Excellence. SCIE Reports: No.4

Macklin, R. (1999). The ethical problems with sham surgery in clinical research. *New England Journal of Medicine, 341*; 992-996.

Medical Research Council (MRC). (2003). Health technology assessment in surgery. www.mrc.ac.uk Accessible at: http://www.mrc.ac.uk/pdf-health_tech_assessment.pdf

Moher, D., Schulz, K.F., & Altman, D.G. (2001). The CONSORT statement: Revised recommendations for improving the quality of reports of parallel-group randomized trials. *Annals of Internal Medicine, 134(8)*; 657-662.

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

Rational Pharmaceutical Management Plus Program. (2001). Drug and therapeutics Committee Training Course. Session 3: Assessing Drug Efficacy. Management Sciences for Health. E-mail: rpmplus@msh.org

Relman, A.S. (1988). Assessment and accountability. The third revolution in medical care. *The New England Journal of Medicine, 319(18)*; 1220-1222.

Singer, P.A. (2000). Recent advances. Medical ethics. *British Medical Journal, 321*; 282 – 285.

Trochim, M.K. (2002). Research methods Knowledge Base. Accessed on: www.socialresearchmethods.net/kb/quasiexp.htm

Weijer, C., Dickens, B., & Meslin, E.M. (1997). Bioethics for clinicians: 10. Research ethics. *Can Med Assoc J; 156(8)*; 1153-1157.

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

8.0 APPENDICES

APPENDIX A: FRAMEWORK TO CONDUCTING A HEALTH TECHNOLOGY ASSESSMENT (HTA)

An alternative framework for HTA offered by the European Collaboration for Health Technology Assessment includes the following steps:

- Submission of an assessment request/identification of an assessment need
- Prioritization
- Commissioning
- Conducting the assessment
- Definition of policy question(s)
- Elaboration of HTA protocol
- Collecting background information/determination of the status of the technology
- Definition of the research questions
- Sources of data, appraisal of evidence, and synthesis of evidence for each of:
    - Safety
    - Efficacy/effectiveness
    - Psychological, social, ethical
    - Organizational, professional
    - Economic
- Draft elaboration of discussion, conclusions, and recommendations
- External review
- Publishing of final HTA report and summary report
- Dissemination
- Use of HTA
- Update of the HTA

Adapted from: http://www.nlm.nih.gov/nichsr/hta101/ta10104.html

APPENDIX B: The Nuremberg Code (1947)

BMJ 1996;313:1448 (7 December)

NUREMBERG DOCTORS' TRIAL

The judgment by the war crimes tribunal at Nuremberg laid down 10 standards to which physicians must conform when carrying out experiments on human subjects.

PERMISSIBLE MEDICAL EXPERIMENTS

The great weight of the evidence before us to effect that certain types of medical experiments on human beings, when kept within reasonably well-defined bounds, conform to the ethics of the medical profession generally. The protagonists of the practice of human experimentation justify their views on the basis that such experiments yield results for the good of society that are unprocurable by other methods or means of study. All agree, however, that certain basic principles must be observed in order to satisfy moral, ethical and legal concepts:

1. The voluntary consent of the human subject is absolutely essential. This means that the person involved should have legal capacity to give consent; should be so situated as to be able to exercise free power of choice, without the intervention of any element of force, fraud, deceit, duress, overreaching, or other ulterior form of constraint or coercion; and should have sufficient knowledge and comprehension of the elements of the subject matter involved as to enable him to make an understanding and enlightened decision. This latter element requires that before the acceptance of an affirmative decision by the experimental subject there should be made known to him the nature, duration, and purpose of the experiment; the method and means by which it is to be conducted; all inconveniences and hazards reasonably to be expected; and the effects upon his health or person which may possibly come from his participation in the experiment. The duty and responsibility for ascertaining the quality of the consent rests upon each individual who initiates, directs, or engages in the experiment. It is a personal duty and responsibility which may not be delegated to another with impunity.

2. The experiment should be such as to yield fruitful results for the good of society, unprocurable by other methods or means of study, and not random and unnecessary in nature.

3. The experiment should be so designed and based on the results of animal experimentation and a knowledge of the natural history of the disease or other problem under study that the anticipated results justify the performance of the experiment.

4. The experiment should be so conducted as to avoid all unnecessary physical and mental suffering and injury.

5. No experiment should be conducted where there is an a priori reason to believe that death or disabling injury will occur; except, perhaps, in those experiments where the experimental physicians also serve as subjects.

6. The degree of risk to be taken should never exceed that determined by the humanitarian importance of the problem to be solved by the experiment.

7. Proper preparations should be made and adequate facilities provided to protect the experimental subject against even remote possibilities of injury, disability or death.

8. The experiment should be conducted only by scientifically qualified persons. The highest degree of skill and care should be required through all stages of the experiment of those who conduct or engage in the experiment.

9. During the course of the experiment the human subject should be at liberty to bring the experiment to an end if he has reached the physical or mental state where continuation of the experiment seems to him to be impossible.

10. During the course of the experiment the scientist in charge must be prepared to terminate the experiment at any stage, if he has probable cause to believe, in the exercise of the good faith, superior skill and careful judgment required of him, that a continuation of the experiment is likely to result in injury, disability, or death to the experimental subject.

Accessed from:
http://bmj.bmjjournals.com/cgi/content/full/313/7070/1448?ijkey=ETEcu2J5M00pM%2520

Appendix C: Tri-Council Policy Statement

Ethical Conduct for Research Involving Humans

Canadian Institutes of Health Research, Natural Sciences and Engineering Research Council of Canada, Social Sciences and Humanities Research Council of Canada, *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans*. 1998 (with 2000, 2002, and 2005 amendments).

Context of an Ethics Framework

1) The need for research
2) A moral imperative: Respect for human dignity
3) Guiding ethical principles
4) A subject-centered perspective
5) Academic freedoms and responsibilities
6) Ethics and law
7) Putting principles in practice

Section 1: Ethics Review

a) Research requiring ethics review
b) Research ethics boards (REBs)
c) Analysis, balance and distribution of harms and benefits
d) Review procedures
e) Conflicts of interest
f) Review procedures for ongoing research
g) Review of multicentred research
h) Review of research in other jurisdictions or countries

Section 2: Free and Informed Consent

a) Requirement for free and informed consent
b) Voluntariness
c) Naturalistic observation
d) Informing potential subjects
e) Competence
f) Research in emergency health situations

Section 3: Privacy and Confidentiality

a) Accessing private information – personal interviews
b) Accessing private information – surveys, questionnaires and the collection of data
c) Secondary use of data
d) Data Linkage

Section 4: Conflict of Interest

a) Conflicts of interest involving researchers
b) Conflicts of interest by REB members

c) Institutional conflicts of interest

Section 5: Inclusion in Research

a) Introduction
b) Research involving women
c) Research involving those who are incompetent to consent for themselves

Section 6: Research Involving Aboriginal Peoples

a) Introduction
b) Good practices

Section 7: Clinical Trials

a) Clinical equipoise
b) Phases of pharmaceutical research
c) Multicentre clinical trials
d) Placebo-controlled studies
e) Analysis and dissemination of the results of clinical trials

Section 8: Human Genetic Research

a) The individual, families, and biological relatives
b) Privacy, confidentiality, loss of benefits and other harms
c) Genetic counseling
d) Gene alteration
e) Eugenic concerns
f) Banking of genetic material
g) Commercial use of genetic data

Section 9: Research Involving Human Gametes, Embryos, or Foetuses

a) Research involving human gametes
b) Research involving human embryos
c) Research involving fetuses
d) Research involving foetal tissue

Section 10: Human Tissue

a) Privacy and confidentiality
b) Free and informed consent
c) Previously collected tissue