Effectiveness of Medical Ethics Education
– A Review of the Literature

By
Aishwarya Palwai

A Master's Paper submitted to the faculty of
The University of North Carolina at Chapel Hill
In partial fulfillment of the requirements for
The degree of Master of Public Health in
The Public Health Leadership Program

Chapel Hill
2002

[Signature]
Advisor – Vaughn M. Upshaw, DrPH
24 April 2002

[Signature]
Second Reader – Dean M. Harris, JD
25 April 2002
Table of Contents

Introduction ................................................................................... 1
Definitions .................................................................................... 2
History ......................................................................................... 3
Context......................................................................................... 4
Literature Search ............................................................................ 7
Literature Review ............................................................................ 9
Study Characteristics (Table 1) .......................................................... 10
Research Design ............................................................................ 14
Participants .................................................................................. 14
Length and Type of Training .............................................................. 14
Study Results (Table 2) ................................................................... 16
Outcome Measures ......................................................................... 21
Statistical Analysis ......................................................................... 23
Results ....................................................................................... 23
Conclusions ................................................................................. 25
Limitations .................................................................................. 26
Discussion ................................................................................... 27
Summary .................................................................................... 30
References ................................................................................... 32
Introduction

Medical ethics education is a required part of the curriculum in every medical school in the United States today. Courses in medical ethics have also been introduced in the curricula of all medical schools in Canada and in most school in United Kingdom and around the world. The World Federation for Medical Education in 1994 (at the World Summit of Medical Education) released a consensus statement that ethics should always receive full attention in the medical school, in all clinical encounters, and in the community.

There is a wide disparity between the perceived need for education in medical ethics and the actual implementation of such programs. It is possible that this is, in part, because formal medical ethics education is a relatively infant field, especially so at the graduate medical education level. An inherent paradox in this situation is that 'only after reaching consensus on what should be taught can medical ethics educators go on to study whether such teaching has resulted in positive outcomes'. Enough interest to ensure discussion and consensus about this cannot be achieved until the field is evaluated sufficiently to claim the validity required for universal implementation. On the flip side, it has been argued that it is not necessary to demonstrate an effect of ethics education for medical students or residents in order to justify teaching ethics, and that ethics should be part of the basic curriculum along with the other core medical subjects.

Evaluation of the effectiveness of various methods of training in medical ethics takes on a critical role in today's academic environment because of
the background considerations of cost and outcomes as measures of the value of the education. Such evaluations would ideally test improvement in ethical behaviors of the study subjects. But the constraints of reality dictate that less tangible and easily measurable aspects of the training such as knowledge base or confidence of the study subjects are usually tested. This is a major concern when trying to decide whether an education intervention is effective or not.

This paper reviews the literature on the evaluation of ethics training in medical schools, residency programs and among practicing physicians, in the USA and abroad. It distinguishes the variable definitions of the terms 'medical ethics education' and 'effectiveness', provides a history of ethics education, and provides a context for the need for medical ethics education. This is followed by a review and analysis of the relevant literature, a discussion of the findings and the ensuing conclusions.

**Definitions**

There are two categories of definitions of the term "medical ethics education". One category includes definitions that view medical ethics education as part of a larger body of knowledge that encompasses philosophy, the social sciences, and the liberal arts, as well as the moral and legal foundations of medicine. The other category includes definitions that focus only on the practice of clinical medicine; they identify patient cases as the focus of medical ethics education. Whatever the definition, the ability to use ethical reasoning skills
has been identified as a "professional skill" which needs to be developed during the course of a physicians' medical education.

**History of medical ethics education**

Medical ethics has been around for as long as the practice of medicine itself, and certainly as long as the history of general ethical thinking. The modern period of bioethics refers to the relatively recent recognition of a formal academic discipline of "biomedical ethics", "medical ethics" or "clinical ethics". As recently as forty years ago, there was no concept of formal clinical ethics in the field of medicine. There was a universally accepted paradigm of paternalism in which it was assumed that physicians would make the right treatment choices in the best interests of the patients. This view was challenged in the 1960s due to the social events of the time including exposure of unethical research practices and concomitant questioning of professional autonomy. In his book entitled *Morals and Medicine*, John Fletcher made a landmark contribution to the field by arguing for importance being given to the prerogatives of the patient as opposed to the traditional view that focused on the prerogatives of the physicians. The field of medical ethics education is currently still in a state of evolution and "as a component in American medical education, medical ethics faces the joys and tribulations of a budding adolescence."
**Context**

Whenever there is the possibility of interaction among persons, there are questions about how those persons ought to act. The study of ethics may be described as the study of how persons ought to act. Health care involves, to a great extent, interaction between persons, and hence, the practice of health care is inextricably linked to the ethical issues in any particular situation\(^\text{10}\).

As previously stated, ethics education is a part of the required curriculum in undergraduate medical education in the USA and has also been recommended for inclusion in the postgraduate medical curriculum. The Liaison Committee on Medical Education (LCME), the nationally recognized accrediting authority for medical education programs leading to the M.D. degree in U.S. and Canadian medical schools requires all medical schools in the USA to include in their curriculum contents, the sciences basic to medicine, a variety of clinical disciplines, and ethical, behavioral, and socioeconomic subjects pertinent to medicine and also requires that there should be presentation of material on medical ethics and human values\(^\text{11}\). There are no guidelines as to the specific curriculum contents, which have been left to the discretion of the individual institutions. The United States Medical Licensing Examination (USMLE) contains questions pertaining to medical ethics but there is no examination that exclusively tests knowledge of medical ethics.

The American Medical Association’s Accreditation Council for Graduate Medical Education (ACGME) specifically requires all residency
programs to include education in medical ethics in their specialty’s curriculum. Educational material pertaining to clinical ethics is included in the specialty board examinations but, just like the LCME guidelines, there are no specific curriculum guidelines, no consensus regarding the content and skills required, and no specific examinations for testing knowledge or ethical behavior.

Depending on how broadly or narrowly ethics educators define “medical ethics education”, the goals of medical ethics education vary. If a broad definition is used, these goals include the following aspects:

i) Legal duties to secure informed consent, tell the truth and protect confidentiality

ii) Objective competencies including an understanding of Do Not Resuscitate (DNR) regulations and surrogate decision-making procedures.

iii) Discursive moral skills such as moral sensitivity, reciprocity and moral development that combine into the capacity for moral dialogue and debate

iv) Behavioral goals that challenge moral education to nurture a more humane, sensitive and communicative physician.

In such a scenario, the ability to deal with practical issues would constitute only one aspect of the training. The definitive goal would be to aid in the development of an ethical human being.
On the other hand, if medical ethics education is focused only on practical issues that are dealt with in day-to-day practice, the goals would be stated in a slightly different manner\textsuperscript{12}:

i) Recognition of ethical issues as they arise in clinical care and identification of hidden values and unacknowledged conflicts

ii) Critical thought about ethical issues in ways that lead to ethically justifiable courses of action

iii) Application of the practical skills needed to implement an ethically justifiable action;

iv) Ability to recognize when the management of a clinical situation requires consultation with individuals or institutional bodies with additional expertise or authority.

This view focuses on ensuring that physicians are capable of doing the legally right thing, are able to demonstrate familiarity with the ethical issues in the field, and are able to act in a manner that conforms to the consensus opinion on these issues. These goals would serve the purpose of keeping physicians out of trouble with respect to the laws of the land and of providing them with the ability to not only think and act ethically but also to justify their actions in terms of concrete concepts in the sphere of ethical thought.
Standardized evaluation of competence in ethics is important for the following reasons:\(^{13}\):

1) For formal empirical study of medical ethics teaching and modes of delivering it.

2) To justify access to curriculum time by demonstrating that ethics is a teachable subject that can be formally evaluated.

3) Widespread employment of a core curriculum would make it easier to evaluate students across institutions.

4) Medical ethics can be defined in part by how it is evaluated. There is a two-way process between refining the core curriculum of knowledge and skills, and their evaluation.

The ultimate goal of evaluation of competence in ethical issues is to assess whether ethics education contributes to increased frequency of ethical practice of medicine.

**Literature Search**

A review of the literature was done to identify studies that evaluated the effectiveness of ethics training on moral reasoning and/or behaviors in the clinical setting. The objective was to determine the effectiveness, if any, of such programs, and to identify key components and the principles underlying effective programs. The search strategy for the identification of studies was as follows: Searches were carried out in MEDLINE through Pub Med, Academic
Universe and JME online (Journal of Medical Ethics). The search strategy for each database was based on using each or combinations of the following MeSH headings and Keywords:

1) Ethics, Medical
2) Education
3) Teaching
4) Training
5) Effectiveness
6) Outcomes
7) Evaluation

Some studies were also identified through the bibliographies of related, published studies.

The review process consisted of the following stages:

1) Pre-screening of titles and abstracts
2) Identification and retrieval of relevant studies
3) Data extraction from included studies; and
4) Qualitative review of the results of all the studies.

_Inclusion and Exclusion Criteria_

Studies were included based on criteria such as participant type, intervention and outcome measures. The review was limited to studies published after the year 1980, in English. Studies that involved nurses, nursing students,
dentists, dental students, and other allied health professionals as participants, were excluded. Most of the studies reviewed were performed in the USA. However, a couple of studies, one performed in the United Kingdom and the other in Denmark, were included, as these studies could also contribute towards testing the efficiency of medical ethics education. A total of 14 studies were identified which satisfied the inclusion criteria.

**Review of Literature**

Each study was reviewed as to the evaluation methods and the results. The study characteristics such as research design, study subjects, and type and length of training are listed in Table 1. The outcome measures, statistical analysis, results and conclusions of individual studies are presented in Table 2.

There are limitations with this literature review process: Because of the inclusion criteria mentioned, this review does not include unpublished data, or studies published in other languages. It also does not address the effectiveness of ethics education on other health professionals like dentists, nurses, physical therapists and occupational therapists who play a significant role in patient care.
<table>
<thead>
<tr>
<th>Author; Location of Study, Year</th>
<th>Research Design</th>
<th>Study subjects</th>
<th>Type of Training</th>
<th>Length of training</th>
</tr>
</thead>
<tbody>
<tr>
<td>Siegler M., University of Chicago, USA, 1982</td>
<td>Pretest and post test study group with a control group</td>
<td>Third year medical students</td>
<td>Medical ethics program - case study method</td>
<td>One month; 12 to 14 one-hour sessions</td>
</tr>
<tr>
<td>Self D.J., Texas A&amp;M University, USA, 1989</td>
<td>Non-randomized natural experimental design; pre-test and post-test study group with a control group</td>
<td>First-year medical students and veterinary medicine students</td>
<td>Two types of teaching methods compared: a) Lecture format instruction in medical ethics b) Half of the instruction in lecture format and half in a small-group, case-study format</td>
<td>2 hours each week for two quarters - a total of 44 contact hours</td>
</tr>
<tr>
<td>Self D.J., Texas A&amp;M University, USA, 1992</td>
<td>Pre- and post-test study group with a control group</td>
<td>First-year medical students</td>
<td>Course in medical ethics - half the instruction in large group lecture format and half in a small-group, case-study format</td>
<td>2 hours each week for two quarters - a total of 44 contact hours</td>
</tr>
<tr>
<td>Shorr A.F., University of Virginia, USA, 1994</td>
<td>Prospective design; pre- and post-test study group</td>
<td>First-year medical students</td>
<td>Formal classroom course - Introduction to Clinical Ethics</td>
<td>Twelve 2 hour sessions that include both lectures and small group seminars</td>
</tr>
<tr>
<td>Author; Location of Study, Year</td>
<td>Research Design</td>
<td>Study subjects</td>
<td>Type of Training</td>
<td>Length of training</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>-----------------</td>
<td>----------------</td>
<td>-----------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Holm S., University of Copenhagen, Denmark, 1995</td>
<td>Pre- and post-test study group without a control group</td>
<td>Medical students</td>
<td>Course on Philosophy of Medicine and medical ethics</td>
<td>80 hour course divided into two parts - part 1 in 3rd semester, part 2 in 7th &amp; 8th semesters</td>
</tr>
<tr>
<td>Goldie J., University of Glasgow Medical School, UK, 2001</td>
<td>Quasi-experimental, pre- and post-test, non-equivalent control group design</td>
<td>First-year medical students</td>
<td>Small-group ethics teaching complemented by plenary seminars</td>
<td>One year (Eight 3 hour sessions)</td>
</tr>
<tr>
<td>Sulmasy D.P., The Johns Hopkins Hospital, USA, 1993</td>
<td>Randomized controlled trial</td>
<td>Medical houseofficers in internal medicine residency program</td>
<td>Limited intervention (LI) - lecture series or Extensive intervention (EI) - lectures and case-conferences with an ethicist in attendance</td>
<td>Over a seven month period: LI - A series of six lectures EI - Six ethics lectures plus six case conferences</td>
</tr>
<tr>
<td>Sulmasy D.P., The Johns Hopkins Hospital, USA, 1994</td>
<td>Prospective study, with an initial randomized phase (see above)</td>
<td>Medical houseofficers in internal medicine residency program</td>
<td>A monthly 30-minute lecture, monthly conference with case discussions</td>
<td>Two year curriculum with lectures and case conferences alternating monthly</td>
</tr>
<tr>
<td>Author; Location of Study, Year</td>
<td>Research Design</td>
<td>Study subjects</td>
<td>Type of Training</td>
<td>Length of training</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>----------------</td>
<td>----------------</td>
<td>-----------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Sulmasy D.P., Georgetown University Medical Center, Washington DC, USA, 1997</td>
<td>Long-term cohort study</td>
<td>Medical houseofficers in internal medicine residency program</td>
<td>Ethics morning report; didactic conferences</td>
<td>Two year curriculum - alternate monthly sessions plus occasional teaching through the ethics consultation service, occasional medical grand rounds on ethics</td>
</tr>
<tr>
<td>Berseth C.L., Mayo Clinic, USA, 1990</td>
<td>Pre- and post-test study group with a control group</td>
<td>Residents in Pediatrics, Family Medicine and Anesthesiology</td>
<td>Human Values Seminar series with didactic material presentation followed by small-group discussion</td>
<td>3 hour seminars bimonthly for a year</td>
</tr>
<tr>
<td>Holloran S.D., Harvard Medical School, USA, 1995</td>
<td>Retrospective study - Review of medical records and LOS trends</td>
<td>Surgical residents</td>
<td>Case study method</td>
<td>Four 60-minute discussion groups</td>
</tr>
<tr>
<td>Angelos P., Northwestern University, USA, 1999</td>
<td>Pre- and post-test study group</td>
<td>Surgical residents</td>
<td>Focus on 5 topical areas with each module consisting of session objectives, teaching plans, handouts, and audiovisual aids</td>
<td>Four 90-minute interactive teaching sessions over a six-month period</td>
</tr>
<tr>
<td>Author; Location of Study, Year</td>
<td>Research Design</td>
<td>Study subjects</td>
<td>Type of Training</td>
<td>Length of training</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>----------------</td>
<td>----------------</td>
<td>------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>White B.D., St. Thomas Hospital, USA, 1993</td>
<td>Prospective study</td>
<td>Attending physicians in a hospital</td>
<td>Discussion of articles, landmark cases, literature review, topical study</td>
<td>Four modules of 6 sessions each - each session 3 hours long; Overall, 18 hours of clinical ethics instruction and discussion per module</td>
</tr>
<tr>
<td>Malek J.I., Georgetown University, Washington DC, 2000</td>
<td>Pre- and post-test study group with randomized cross-over design</td>
<td>Practising health care professionals</td>
<td>Series of lectures and small group discussions on bioethical principles and theoretical approaches as well as specific topics in medical ethics</td>
<td>Six day intensive bioethics course</td>
</tr>
</tbody>
</table>
Research Design

Twelve studies used a non-randomized natural experimental design with pre-test and post-test study groups with (five studies) or without (seven studies) control groups. There was only one randomized control trial and, a prospective study was later done based on this randomized trial \(^{20,21}\). Both of these studies are included in this review. One retrospective study was conducted by Holloran et al, in which medical records review and Length of Stay (LOS) figures were used to assess the improvement in the subjects’ ethical skills\(^{24}\).

Participants

Of the 14 studies reviewed, six studies evaluated the effects of specific training programs on undergraduate medical students, mostly first year students\(^ {14-19}\). Three studies involved medical house officers and three others involved residents\(^ {20-25}\). Of the remaining two studies, one looked at the effect of ethics training intervention on attending physicians in a hospital setting\(^ {26}\). One study mentioned that it involved health care professionals but the authors did not elaborate on the academic standing or the practice settings of the study subjects\(^ {27}\).

Length and Type of Training

There was a wide variation in the length and type of training that was provided in different settings. They ranged from a six-day intensive bioethics course for practicing professionals to a two-year integrated curriculum for
undergraduate medical students. The most frequent type of training involved case studies, either real life cases in the hospital or simulated cases with inherent ethical issues. Many studies evaluated the effects of a combination of teaching methods, which included didactic lectures, small group discussions, and other educational tools including audio-visual aids and published articles.

The length of training did not seem to be related to the previous training experiences of the study subjects in medical ethics. For instance, one of the brief courses lasting a month (total of 14 contact hours), was targeted at third year medical students\textsuperscript{14}, while two groups of house officers were enrolled in a two year curriculum\textsuperscript{21,22}.

All of the studies included in this review evaluated the change in ethical attitudes and/or behaviors on a dichotomous scale but most did not use a continuous scale to detect the correlation if any, between length, type or intensity of training and extent of change on a continuous scale. Two studies used continuous scales to evaluate the outcome measures. One was the randomized trial conducted by Sulmasy et al in 1993, which looked at the impact of an extensive intervention versus limited intervention training program on the knowledge and behaviors of medical house officers\textsuperscript{20}. Self et al used a continuous scale to evaluate and compare the effectiveness of two types of teaching methods – lecture format instruction vs. small-group case study\textsuperscript{15}. 
<table>
<thead>
<tr>
<th>Author; Location</th>
<th>Outcome Measure</th>
<th>Statistical Analysis</th>
<th>Results</th>
<th>Conclusions</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Siegler M., University of Chicago, USA, 1982</td>
<td>Score on Simulated clinical-ethical cases</td>
<td>t-test for independent samples used to compare the mean score of the experimental group with the mean score of control group</td>
<td>Mean score of experimental group significantly higher than that of control group</td>
<td>Medical students who take a clinically oriented course in medical ethics show increased reflectiveness regarding ethical decisions</td>
<td>Relatively arbitrary weighting of categories; Not enough deliberation and validation of testing instrument</td>
</tr>
<tr>
<td>Self D.J., Texas A&amp;M University, USA, 1989</td>
<td>Sociomoral Reflection Measure (SRM)</td>
<td>Scheffe paired comparison tests done between pretest and post test scores, and between both the medical school groups and the control group</td>
<td>The hypothesis that teaching medical ethics produces significant development of moral reasoning could not be rejected. The hypothesis that small-group case-study format was more effective than the lecture format could be rejected if rigorous statistical tests were used</td>
<td>Teaching medical ethics makes a significant difference in the development of moral reasoning in medical students; greater gains using small-group case-study format</td>
<td>Assessment instrument not sensitive enough to detect subtle changes over small time periods; sample size too small to obtain statistical significance for one hypothesis.</td>
</tr>
<tr>
<td>Self D.J., Texas A&amp;M University, USA, 1992</td>
<td>Defining Issues Test (DIT)</td>
<td>Statistical analysis done but not described.</td>
<td>The hypothesis that formal teaching of medical ethics promotes a significant increase in the growth and development of moral reasoning could not be rejected. Highly significant difference between the DIT scores of experimental and control groups</td>
<td>Sound empiric research can be carried out on the effectiveness of teaching medical ethics. Further research needed on the evaluation of accountability for teaching medical ethics</td>
<td>Assessment instrument not sensitive enough to detect subtle changes over small time periods</td>
</tr>
<tr>
<td>Author; Location; Year</td>
<td>Outcome Measure</td>
<td>Statistical Analysis</td>
<td>Results</td>
<td>Conclusions</td>
<td>Limitations</td>
</tr>
<tr>
<td>------------------------</td>
<td>-----------------</td>
<td>----------------------</td>
<td>---------</td>
<td>-------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Shorr A.F., University of Virginia, USA, 1994</td>
<td>New test instrument that employed clinical vignettes, multiple-choice, true-false, Likert scale questions</td>
<td>Wilcoxon matched-pairs signed-ranks test (clinical vignettes); Chi-square test for association (factual knowledge questions)</td>
<td>No statistically significant change for clinical vignettes; statistically significant changes in level of factual knowledge and specific opinions of study participants</td>
<td>Formal course work in medical ethics has a limited influence on medical students.</td>
<td>Students were homogenous in their beliefs prior to enrolling in the class; nearly a third of the students had completed a prior course in medical ethics.</td>
</tr>
<tr>
<td>Holm S., University of Copenhagen, Denmark, 1995</td>
<td>Moral reasoning scores measured by Defining Issues Test (DIT) questionnaire</td>
<td>Pre- and post-test scores compared using the Mann-Whitney U-test or the Wilcoxon matched-pairs test.</td>
<td>Analysis of paired data showed significant increase in the D-score and the P-score, and a significant decrease in the M-score. No significant change in the A- or U-scores* Positive change in DIT scores</td>
<td>Improvement in the ability to reason with concepts like justice and impartiality</td>
<td>Positive change in DIT score may be considered improvement or deterioration depending on the ethical viewpoint (care vs justice).</td>
</tr>
<tr>
<td>Goldie J., University of Glasgow Medical School, UK, 2001</td>
<td>Ethics and Health Care Survey Instrument (12 case vignettes)</td>
<td>McNemar's test for each question to find movement towards or away from consensus.</td>
<td>Significantly higher average number of consensus answers post-test in the experimental group especially in the areas of autonomy, confidentiality and consent.</td>
<td>Small-group teaching had a positive impact on first year students' potential ethical behavior and is more effective than a lecture and a large-group seminar-based course</td>
<td>Inherent selection bias; students' religious views could potentially influence their responses.</td>
</tr>
<tr>
<td>Author; Location</td>
<td>Outcome Measure</td>
<td>Statistical Analysis</td>
<td>Results</td>
<td>Conclusions</td>
<td>Limitations</td>
</tr>
<tr>
<td>------------------</td>
<td>-----------------</td>
<td>----------------------</td>
<td>---------</td>
<td>-------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Sulmasy D.P., The Johns Hopkins Hospital, USA, 1993</td>
<td>Confidential self-administered questionnaire containing MCQs and simulated case</td>
<td>Dichotomous outcomes-Chi-square/Fisher's exact test; Knowledge scores-Student's t-test; Confidence scales-Wilcoxon rank-sum test</td>
<td>% correct on knowledge test-uniformly low for all groups; Intervention associated with significantly higher confidence for EI group vs controls (only for those reporting poorer preparation in medical school)</td>
<td>The interventions had no significant impact upon house officers' knowledge but had a significant impact on their confidence and responses to a simulated case</td>
<td>Relative effects of the various components of the intervention could not be studied; Small number of house officers limits ability to detect small differences</td>
</tr>
<tr>
<td>Sulmasy D.P., The Johns Hopkins Hospital, USA, 1994</td>
<td>Chart review to note documentation of DNR orders, and attention to CCCs</td>
<td>Dichotomous variables-chi-square/Fisher's exact test; Normally distributed data-Student's test; non-parametric data-Wilcoxon rank sum test; correlations-Pearson r.</td>
<td>Increase in proportion of DNR orders written by house officers; increased documentation of attention to CCCs</td>
<td>Ethics education can improve certain aspects of the quality of care rendered to patients who have DNR orders</td>
<td>Impact of full curriculum less than EI used during the pilot study - need to study which components of program were most effective; Focus on documentation, not behaviors</td>
</tr>
<tr>
<td>Sulmasy D.P., Georgetown University Medical Center, Washington DC, USA, 1997</td>
<td>Anonymous questionnaire-a previously validated instrument to measure knowledge, confidence and attitudes</td>
<td>Categorical variables analyzed using the chi-square test, continuous variables analyzed using Student's t-test; correlations reported as Pearson's r</td>
<td>Knowledge score improved 14%, confidence rose from 3.3 to 3.8 on a five-point scale, positive correlation between knowledge and confidence; attitudes changed only on one of the test items</td>
<td>Curriculum practical, popular, and effective; Knowledge and confidence improved significantly, attitudes remained essentially unchanged</td>
<td>Uncontrolled study; Cannot rule out probable influence of site-specific factors</td>
</tr>
<tr>
<td>Author: Location Year</td>
<td>Outcome Measure</td>
<td>Statistical Analysis</td>
<td>Results</td>
<td>Conclusions</td>
<td>Limitations</td>
</tr>
<tr>
<td>-----------------------</td>
<td>-----------------</td>
<td>----------------------</td>
<td>---------</td>
<td>-------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Berseth C.L., Mayo Clinic, USA, 1990</td>
<td>Written questionnaire - MCQs &amp; open-ended; six scales of attitude changes</td>
<td>Analysis of variance to compare attitude changes pretest to posttest; Pearson's correlations between scale of attitude change and variables hypothesized to influence medical decisions</td>
<td>Changes in attitude occurred in both the study and control groups with no difference between the two groups for 5 of 6 scales; formal training altered only one attitude scale significantly</td>
<td>The process of residency training profoundly affects ethical attitudes; The ability of formal ethics training to influence ethical attitudes may be limited</td>
<td>Personal issues such as societal needs and cost concerns may have been confounding factors that influenced residents' attitudes.</td>
</tr>
<tr>
<td>Holloran S.D., Harvard Medical School, USA, 1995</td>
<td>Medical record review; Length of stay assessment for patients who died in the SICU during 1990 through 1993</td>
<td>Categorical variables analyzed using the chi-square test</td>
<td>Significant decrease in LOS in SICU for dying patients; significant increase in documentation of advance directives in patient records</td>
<td>Increased knowledge &amp; skill of residents in addressing and integrating ethical issues into practice; improved patient care with reduced SICU LOS and reduced cost</td>
<td>Confounding factors (managed care concern for reducing costs) may have caused the decrease in LOS</td>
</tr>
<tr>
<td>Angelos P., Northwestern University, USA, 1999</td>
<td>Modified published survey instrument - close ended surveys; each session assessed on a Likert-type scale of 1 to 5</td>
<td>Paired t test-change in confidence levels pretest to posttest; Descriptive statistics to summarize end of session evaluation forms; Pearson correlation coefficients-relationship between PGY &amp; responses to test q</td>
<td>High scores given for value of topic &amp; quality of presentation; Statistically significant improvement in resident confidence in addressing ethical issues (13 of 22 items); No significant improvement in communication skills</td>
<td>A focused curriculum can be developed that has a measurable impact on residents' confidence in addressing ethical issues</td>
<td>Performance based evaluation system is a better way to assess competence.</td>
</tr>
<tr>
<td>Author; Location Year</td>
<td>Outcome Measure</td>
<td>Statistical Analysis</td>
<td>Results</td>
<td>Conclusions</td>
<td>Limitations</td>
</tr>
<tr>
<td>-----------------------</td>
<td>-----------------</td>
<td>----------------------</td>
<td>---------</td>
<td>-------------</td>
<td>-------------</td>
</tr>
<tr>
<td>White B.D., St. Thomas Hospital, USA, 1993</td>
<td>Self assessment by participants through weekly surveys, CME evaluation, and open-ended questionnaires</td>
<td>Qualitative analysis of responses; Statistical analysis not done</td>
<td>Greater awareness and interest in discussed issues; greater willingness to share concerns with other practitioners; discussions more open to patient and family concerns; specific cases directly influenced by the program</td>
<td>Providing physicians with clinical ethics perspectives and skills has a positive effect both on the individual physician's practice and on the institutional context of patient care.</td>
<td>No attempt at identification or measurement of the influence of all the factors and variables that might have contributed to the changes in physicians' attitudes</td>
</tr>
<tr>
<td>Malek J.l., Georgetown University, Washington DC, 2000</td>
<td>Coding sheet developed to quantify open-ended responses and qualitative analysis done with the coded responses</td>
<td>Statistical analysis done but not described</td>
<td>Clear differences in pre-test and post-test responses in three qualitative domains: justification, ranking of elements, and recognition of conflicting elements; trend within each vignette towards explicit use of ethical principles compared to implicit use</td>
<td>The course imparted to the participants an enhanced capacity to better support their proposed resolutions and an improved ability to articulate their thoughts on relevant issues.</td>
<td>Inherent self selection bias and response bias</td>
</tr>
</tbody>
</table>
**Outcome Measures**

The instruments used for pretest and posttest measurements varied widely from known instruments like the Moral Judgment Interview of Kohlberg (MJI) and the Defining Issues Test of Rest (DIT) to written essay examinations, to self-assessment surveys. In six cases, this entailed the development and testing of new instruments\(^{14,17,20,23,27}\) or modification of existing instruments\(^{25}\).

Siegler et al, in 1982, evaluated medical ethics teaching by measuring ethical reasoning using simulated clinical-ethical cases\(^{14}\). Other studies that utilized a clinical-vignette based testing instrument were conducted by Shorr et al in 1994, and Goldie et al in 2001\(^{17,19}\). Siegler and Shorr developed their own evaluation instruments while Goldie et al used an adapted version of the ‘Ethics and Health Care Survey Instrument’, which had been previously developed and tested by Kipnis and Gerhard at the University of Hawaii. Self et al, 1989, used the Socio-moral reflection measure (SRM) of Gibbs, a paper-and-pencil version of the original Moral Judgment Interview (MJI) of Kohlberg\(^{15}\). The authors felt that it was not feasible to use the MJI because of the time and expense involved. The SRM was chosen because of the low cost and ease of administration associated with it, and the authors’ belief that, like the MJI, it assessed the spontaneous generation of moral reasoning and justification. The DIT is a paper-and-pencil written test based on the cognitive moral development theory, which also forms the basis for the MJI. The DIT was used by two of the fourteen studies reviewed here\(^{16,18}\). The benefit of using the DIT is that it is computer scorable, group
administrable, relatively inexpensive, and provides rapid results. In their first study in 1989, Self et al used the SRM scale to compare two types of training, and in 1992 they evaluated the more effective training method using the Defining Issues Test (DIT) of Rest (1979)\textsuperscript{15,16}. The study done in 1992 by Self et al confirmed the findings of their previous study using a different instrument of assessment and strengthened the authors' conclusion about the effectiveness of the teaching methods tested\textsuperscript{16}. Six studies used questionnaires or survey instruments that contained either multiple choice, open-ended or closed-ended questions or a combination of these\textsuperscript{20,22,23,25-27}. Two studies employed medical records reviews to evaluate behavioral and practice changes in the study subjects\textsuperscript{21,24}.

Outcome measures in most studies were not simple instruments measuring just one facet of the subjects' 'ethical development' but rather a combination of measures were used to detect the various subtle changes in study subjects. The instruments in many cases were designed to detect separately, changes in knowledge, attitudes and behaviors of the subjects. Since there is no consensus on evaluation methods and testing instruments in the field of medical ethics education, a number of studies simultaneously evaluated the testing instrument itself. Of these, pilot testing of the new testing instrument was done only in the study by Shorr et al\textsuperscript{17}. Evaluation of test instrument by independent coders/ scorers was performed by Siegler et al, and by Malek et al\textsuperscript{14,27}.
**Statistical Analysis**

Statistical analyses were performed by thirteen studies and reported by eleven studies. Two studies did not specify the type of statistical analysis done\(^\text{16, 27}\). One study performed a qualitative analysis of physicians’ self-assessment of the training program\(^\text{26}\).

A variety of statistical tests were used to compare pre and post-test scores: Scheffe paired comparison tests\(^\text{15}\), Wilcoxon matched-pairs test\(^\text{17-21}\), Mann-Whitney \(U\) -test\(^\text{18}\), the paired t-test\(^\text{25}\) or analysis of variance\(^\text{23}\). Categorical variables were analyzed using Chi-square test or Fisher’s exact test\(^\text{17, 20-22, 24}\). Continuous variables were analyzed with Student’s t-test\(^\text{14, 20-22}\). Correlations were reported using Pearson’s \(r\)\(^\text{21-23, 25}\). In addition, descriptive statistics were used to summarize changes at the end of the training program\(^\text{25, 26}\).

**Results**

The results reported by various studies reflected the variation in the outcome measures evaluated in the studies. Statistically significant difference was found in at least one or more or all of the outcome measures between the scores of pretest and posttest groups or experimental and control groups in all fourteen studies reviewed.

Six studies had equivocal findings with respect to the various outcome measures. Self et al, 1989, found that the hypothesis that small-group case-study format was more effective than the lecture format could be rejected if
rigorous statistical tests were used. Shorr et al, 1994, detected a statistically significant change in levels of factual knowledge and attitudes of study subjects, but no significant change in responses to simulated cases presented as clinical vignettes. Holm et al, 1995, using the DIT questionnaire, found a significant increase in the D and P-scores indicating a shift in the moral reasoning of students towards a more justice based approach, and a decrease in M-score, indicating an increased ability to recognize nonsense items on the test questionnaires. There was, however, no significant change in the A-score (anti-authority attitudes) or the U-score (relations between the items chosen as important and the choice of action in the hypothetical situations presented in the questionnaires). Sulmasy et al, 1993, observed a significant change in confidence levels after the education intervention, but no difference in the knowledge test scores between the experimental and control groups. Sulmasy et al, 1997, documented an increase in knowledge and confidence scores with very small change in the attitude scale. Similarly, Berseth et al, 1990, observed that formal training in ethics altered only one of six attitude scales significantly.

The remaining eight studies acknowledged that there was a clear benefit from ethics education in terms of improvement in one or more of the following aspects: knowledge, confidence and behavior.
Conclusions

Four studies concluded that medical ethics education significantly improves the moral/ethical reasoning of the study subjects\textsuperscript{14, 15,16,18,27}. This improvement is seen as increased reflectiveness regarding ethical decisions, an improved ability to reason with concepts like justice and impartiality, an enhanced capacity to support proposed resolutions and an improved ability to articulate thoughts on relevant issues. Of the five studies that measured change in factual knowledge, three studies noted a positive change\textsuperscript{17, 22, 24} while two did not note any significant improvement\textsuperscript{20, 25}. Improvement in confidence levels was demonstrated by all three studies that assessed this aspect\textsuperscript{20, 22, 25}. Five studies measured change in attitudes of the study subjects while dealing with ethical issues; significant changes were noted in three studies\textsuperscript{17, 19, 26}. Two studies concluded that attitudes were unchanged following educational intervention\textsuperscript{22, 23}. Ethical behaviors were assessed in three studies and improvement was noted in all three, as far as the outcome measures for each study were concerned\textsuperscript{21, 24,26}. Ethics education was found to improve certain aspects of quality of care rendered to patients and helped integrate the ability to address ethical issues into practice.

The hypothesis of Self et al (1989) that a small-group case-study format was more effective than a lecture format could be rejected if rigorous statistical tests were used. Though the findings were not statistically significant, the authors concluded that students made greater gains in moral reasoning when taught in the small group format\textsuperscript{15}. The study by Sulmasy et al in 1993 compared
extensive intervention to limited intervention training; the authors concluded that extensive intervention (which included ethics lectures, a series of case-conferences and clinical rounds with an ethicist in attendance) was more effective than limited intervention (consisting only of ethics lectures)20. Goldie et al, in 2001, compared small group teaching in the present study to previous experiences with teaching ethics in the university, and concluded that it was more effective than the previous methods19.

Limitations

Many of the studies' conclusions were limited by the absence of valid testing instruments. Siegler et al, 1982 admitted that they had used relatively arbitrary weighting of scoring categories14. Self et al recognized, in both their studies, that the instrument they had used, the SRM and the DIT scales, and the original MJI scale on which these are based, had actually been designed to test changes in moral reasoning that took place over years, and not over months, as in their study. Hence, subtle changes could have been missed due to inadequate sensitivity of the instrument 15, 16. Six studies acknowledged that confounding factors could have contributed to the changes observed and such factors had not been sufficiently accounted for in their analyses17, 19, 22, 24, 26. Such factors included, among others, prior ethics education, religious views, site-specific factors, societal needs, and cost concerns. Two studies noted that the small sample sizes in their studies had resulted in failure to obtain statistical significance for results15, 20.
The possible influence of selection bias on the findings was noted in two studies\textsuperscript{19, 27}. Two studies recognized the need for further research to identify the relative effects of the various components of the intervention measures\textsuperscript{20, 21}. Two studies acknowledged that an ideal evaluation system would measure actual behaviors and not factors like factual knowledge, attitudes or documentation of actions\textsuperscript{21, 25}.

**Discussion**

All the studies that tested the relationship between training and ethical reasoning and/or practice defined these terms differently, and measured them using different scales. Furthermore, the study subjects were taken from different programs and levels of study or practice. Hence, it is difficult to integrate and summarize the findings of all the studies.

There are very good reasons to start medical ethics education early on in medical school and build on this foundation in later years. Medical school ethics training cannot create professionals who are fully competent to deal with practical ethical issues. However, it does present the general core of basic knowledge, attitudes and skills that residents and practicing physicians can use to achieve proficiency in their field of practice\textsuperscript{12}.

One would have expected that undergraduate students just entering medical school would be enrolled in longer or more intensive training programs while those with prior training would be involved in refresher courses so as to reacquaint them with previously acquired knowledge and to expose them to newer
developments and schools of thought in the field. However, this review noted that this was not the case, and there was no correlation between prior ethics education and length or intensity of the present training intervention.

As stated earlier, a number of studies simultaneously evaluated the testing instrument but pilot testing of instrument was done in only one case. This approach leads one to question the validity of the findings of the studies.

The studies varied in their categorization of what exactly constituted ‘effectiveness’ in medical ethics training. Consequently, although all the studies measured effectiveness of clinical ethics training, their definitions of effectiveness extended across a wide spectrum and hence they looked for different end results in order to determine whether or not a program was effective.

Ethics training programs may simply assess the subjects' knowledge or may assess their moral reasoning or may go one step further and assess the subjects' behaviors in a clinical setting. Acquisition of clinical ethical competence has been described in terms of a learning pyramid with the cognitive or knowledge elements forming the base, the ability to apply this knowledge in the middle, and performance of actual behavior in the clinical context at the top. This concept has also been described as the Know-Can-Do model\textsuperscript{28}. All the reviewed studies assessed one or more of the following aspects of clinical ethics competence: 1) factual knowledge 2) moral /ethical reasoning 3) confidence or attitudes when dealt with ethical issues and 4) ethical behaviors/skills. Five studies assessed the ‘Know’ component of the Know-Can-Do model; they
measured changes in factual knowledge regarding ethical issues following the intervention. Eleven of the fourteen studies assessed ‘moral reasoning’, confidence, ‘ethical attitudes’ or ‘potential behaviors’ (which reflected the subjects’ attitudes) of the subjects, i.e. the ‘Can’ component of the model. Actual behavior of the subjects in situations dealing with ethical issues, the ‘Do’ component, was assessed in three studies. In two of these, actual behavior was assessed using medical case report reviews. The third study assessed participants’ behaviors in clinical settings using self-assessment questionnaires.

Eleven of the studies concluded that ethics education had a significant impact on the outcome being measured whether it was knowledge, attitudes or behaviors. One study concluded that formal course work in medical ethics had a limited influence on the study subjects. The two studies by Sulmasy et al in 1993 and 1997 came to diametrically opposite conclusions in that while the earlier study in 1993 detected a significant impact on house officers’ confidence levels and responses to a simulated case with no impact on their knowledge base, the latter study in 1997 showed improvement in knowledge and confidence while attitudes remained essentially the same.

The three studies that measured changes in ethical behaviors as outcome measures had areas of incongruity that considerably weakened the validity of their findings. The study by Sulmasy et al focused on documentation of Do Not Resuscitate (DNR) orders and documented attention to concurrent care concerns (CCC); they did not measure overall ethical behaviors of physicians.
while dealing with the patients. The study by Holloran et al measured the increase in attention to advance directives of dying patients in the Surgical Intensive Care Unit (SICU) that translated into lower length of stay in the SICU. This outcome could have been confounded by the effect of managed care cost constraints and this factor was not included in the statistical analyses. The third study that measured ethical behaviors, by White et al, had an extremely debatable study design in that it evaluated improvement in clinical ethics activities (before and after a clinical ethics program) by means of participants’ personal evaluation of clinical ethics activities; this invariably would lead to bias in the responses and would affect findings.

**Summary**

Very little research has been done in the field of medical ethics to evaluate the effectiveness of ethics education. Oft cited arguments include the claim that medical ethics deals with intangible concepts like values and attitudes and hence that the teaching of ethics cannot be objectively assessed, measured or reported. The studies that have been reviewed here establish that the effectiveness of different teaching methods can be proved or disproved one way or the other by demonstrating changes in the specific outcome measures.

Even if it is assumed that these studies were done with good study designs, one would have to bear in mind that a demonstrated change in the knowledge base of the participants or their responses to simulated cases cannot be interpreted as changes in actual behaviors.
While all the three studies that measured changes in behaviors showed significant changes following the training intervention, this merely reflects changes in the specific aspects of ethical behaviors that were measured and does not imply an overall improvement in ethical behavior of the physicians.

There is a need for further research on the evaluation of and accountability for teaching medical ethics. This would entail reaching a consensus on the definitions and techniques of measurement of ‘ethical reasoning’ and ethical behaviors. This would have to be followed by the development of standardized curricula, standardized testing instruments, better study designs and greater number of study subjects at different levels of practice so that accurate and statistically significant conclusions can be made about the effectiveness of education for different study populations.

Further research is required to investigate predictors for ethical practice other than education and moral reasoning. Work environmental factors such as institutional policies and procedures, inter-professional relationships etc. may influence ethical practice. These factors need to be identified and utilized to derive the greatest benefit from ethics education.

Small-group case study format and combinations of other training methods such as case conferences, seminars, and clinical rounds with ethicists have all been identified as effective teaching methods for medical ethics education. More research needs to be done in this field to identify, evaluate, compare and implement such effective methods.
References


5. Iserson K.V. et al., Requirements for Ethics, Socioeconomic, and Legal Education in Postgraduate Programs. The Journal of Clinical Ethics, Fall 1993; 4, 3: 225-228.


15. Self D.J. et al., The Effect of Teaching Medical Ethics on Medical Students' Moral reasoning. Academic Medicine, December 1989, 64: 755-759


17. Shorr A.F. et al., The Effect of a Class in Medical Ethics on First-year Medical Students. Academic Medicine, 1994; 69: 998-1000


26. White B.D. et al., Clinical Ethics Training for Staff Physicians: Designing and Evaluating a Model Program. The Journal Of Clinical Ethics, Fall 1993; 4-3: 229-235


30. ACGME website: http://www.acgme.org/