

Chapter 1

INTRODUCTION

“Science is all those things which are confirmed to such a degree that it would be unreasonable to withhold one’s provisional consent”

Stephen Jay Gould (Lecture on evolution, Cambridge 1984. Quote from MacKay, 1991)

1.1 RATIONALE AND ORIGINS OF THE RESEARCH

This thesis is about the question “what is evidence?” asked within the academic world of medicine and public health. The question was formed as a result of the phrase “evidence-based medicine” or EBM, which was coined in 1992 and led to unprecedented currency for the term “evidence”.

Within a few years, questions about the definition and implications of evidence-based medicine were receiving significant attention in medical and other health literature. This led to broader questions about the meaning of evidence, which became a popular and highly contested topic in medical editorials, opinion pieces, letters to the editor, on the Internet and among clinical and public health colleagues.

In the mid 1990s at the University of Sydney, Faculty of Medicine, discussions about evidence were further intensified by the introduction of EBM into the new medical degree (Graduate Medical Program, GMP). Curriculum development brought together traditionally independent disciplines, and discussions about the evidence-base for medicine involved much disagreement and significant controversy. Both the international and local debates indicated that many teachers of public health and medicine approached and used the concept of evidence differently, despite their shared scientific background.

In the Oxford English Dictionary on Historical Principles, evidence is defined as “that which serves to make clear, obvious, certain and conclusive an understanding or judgement”. Evidence also “proves or disproves conclusions”. These definitions suggest that concepts such as certainty or proof may be important, but dictionary definitions are inevitably generic and devoid of context. The aim of this study was to step aside from this broad definition and to develop theoretical propositions from empirical data about how evidence was conceptualised in medical and public health research and practice. The study focused on how senior academics and practitioners, who teach and influence others in medicine and public health, approach, think about and construct the notion of evidence.

1.2 STUDY OBJECTIVES AND RESEARCH QUESTIONS

The aim was that findings of this study could inform and facilitate communication about the topic of evidence in health care. The objectives were as follows: to identify principles that underpin concepts of evidence; to map key conceptual components of the term evidence and generate theoretical propositions about their relationships; and to consider their significance and meaning.

The research also concentrated on the relation between concepts of evidence and approaches to EBM. My purpose was to gain insight into the different perspectives on EBM and to explain or interpret some of the surrounding debates and controversies.

The strategy was to tap, describe and interpret the ideas of those who influence colleagues and students in medical and public health research and practice. The method of Grounded Theory was adopted. The research questions developed were:

- 1 What concepts does the term evidence encompass?

- 2 What are the relationships between the conceptual components of evidence?

- 3 What are the relationships between concepts of evidence and evidence-based medicine?

1.3 THESIS STRUCTURE

The thesis is divided into six chapters, including the Introduction, chapter 1.

Chapter 2 (Setting the Scene) is a selected review of literature and a description of the study's broader and historical context. This includes philosophies of science and social science, approaches to inquiry in health care and a background to evidence-based medicine. There is also a description of local developments in the Faculty of Medicine at the University of Sydney, where the core data on concepts of evidence were collected.

Chapter 3 (Methodology) is a summary of the theory and principles underpinning this research and my *a priori* perspectives on concepts of evidence in health care. It includes a review on the method of Grounded Theory, an outline of the theoretical position adopted in this research, and the approaches used to maximise rigour.

Chapter 4 (Study Methods) provides a detailed account of the research process. It includes my decisions about data sources and triangulation, the

steps of data collection, and a description of the process of data analysis and theory generation. Data collection and data analysis have been separated for the purpose of reporting only. In practice data collection and analysis were an iterative process.

Chapter 5 (Results and Reflections) presents the main findings and at the end of each section, my reflections on the concepts presented. The research was conducted using a method that is both analytical and interpretive (see chapter 3). The format in chapter five is intended to reflect the close relationship between analysis and interpretation of qualitative data.

Chapter 6 (Final Reflections and Conclusions) positions the study findings and presents an overall discussion of the results and their implications.

1.4 THESIS REPORTING STYLE

1.4.1 Contemporary perspectives

Interpretive or qualitative inquiry is notable for the diversity of its philosophical underpinnings and theoretical forms, with different traditions of reporting style (Denzin and Lincoln, 1994). The prevailing trend in reporting qualitative research is towards reflexive, author-centered reports that have “more realistic situated speakers, with our subjectivities engaged”

(Richardson, 1994). Thus contemporary norms include a visible researcher as the primary instrument of analysis and interpreter of text, and writing in the first person (Wolcott, 1990; Denzin, 1994; Van Maanen, 1995).

Writing style is also determined by the researcher's personal preferences and the usual conventions and publishing requirements within their discipline.

There have been some concerns that excessive self-consciousness and disclosure distracts from the voice of those being researched, and writers may purposefully exclude themselves from their work (Fujimura, 1991; Johnson and Scott, 1997). A reflexive text is not inherently better but rather one of several options that include a "realist" style characterised by a "studied neutrality" or apparent absence of the writer (Van Maanen, 1988).

1.4 2 Position adopted

I have adopted a dual reporting style for this research. The combination reflects my "middle-ground" methodological position on Grounded Theory (chapter 3) as an inherently interpretive method that is applied to real social phenomena. It also allowed me to adopt the writing style that felt the most natural in different sections.

The Introduction, Methodology and Methods are written in the first person. This makes the subjective nature of my adopted theoretical perspective and

analytical decisions explicit, and is intended to highlight the propositional nature of my findings.

Setting the Scene (chapter 2), much of the Results and Reflections (chapter 5) and Final Reflections and Conclusions (chapter 6) are written in the third person. This is not an attempt to claim pure objectivity but to focus attention on the topics and findings, rather than on myself. I also use the first person in some sections of these chapters to stress an interpretative decision, or where I discuss how the study modified my *a priori* perspectives on evidence.

Chapter 2

SETTING THE SCENE

“The intensity of the conviction that a hypothesis is true has no bearing on whether it is true or not.”

Peter Medawar (Advice to a Young Scientist, 1979)

2.1 RATIONALE FOR LITERATURE REVIEW

Prior to this thesis, there had been no empirical research reported on how medical academics, researchers and practitioners conceptualise evidence. The aim of this study was to generate theoretical propositions rather than to evaluate an existing theory, and the *a priori* position was purposefully open and flexible (described in chapter 3). It was also important however to be aware of prevailing norms and trends in medical and public health research and practice, and of potentially important aspects of the concept of evidence.

A literature review was conducted at the start of the study, with three objectives. These were to provide a historical context to the contemporary focus on evidence in health care; to select a research strategy; and to identify trends in research and practice in medicine and public health to facilitate

meaningful data collection from researchers and academic practitioners on the topic of evidence. Examination of the literature continued throughout the time of the study and particular topics became more relevant as data collection and analysis progressed. Selection of the literature for this chapter was finalised after data analysis had been completed, in order to construct an appropriate introduction to the concepts presented and discussed in chapters 5 and 6.

The rationale for the study primarily emerged from interactions about evidence within the Faculty of Medicine, University of Sydney (described in section 2.9) rather than from the literature. Hence the purpose of this chapter is to provide an overview of the literature that reflects the broad academic history and context in which the study was undertaken, rather than to identify gaps in that body of knowledge. The review (sections 2.2 to 2.8) includes a summary of science and social science epistemologies, an introduction to health care research and a review of the formation and development of evidence-based medicine. The health literature is critically examined in the light of the study findings in chapters 5 and 6.

2.2 PHILOSOPHIES OF SCIENCE

2.2 1 The modern worldview

In a comprehensive overview of the ideas and philosophies that shaped the Western world, Richard Tarnas (1991) drew attention to the significance of philosophy's "momentous transfer of allegiance from religion to science". This accompanied the scientific revolution of the 16th and 17th Centuries, which he described as "final expression of the Renaissance."

Philosophy of science can be divided into the epistemology of science i.e. why should we believe scientists' claims about knowledge, and the metaphysics of science i.e. what does it mean if the scientists are right (Papineau, 1996). It is the domain of epistemology that raises questions of justification, confirmation or proof and thus, potentially, questions of evidence. The ascendancy of a scientific worldview led to an ongoing quest to understand the epistemology of science, and to explain or challenge its superiority.

The scientific revolution laid the foundations for the "modern worldview". This was built on intellectual developments such as the Copernican heliocentric model of astronomy; Kepler's discovery of elliptical orbits; Galileo's formulation of scientific method; Bacon's inductive empiricism; Descartes' deductive rationalism; and Newton's synthesis to formulate

unifying laws of gravity and motion. The modern worldview encompassed the following characteristics (from Tarnas, p 284-290):

- The world was seen as the result of innate, mechanical regularities generated by nature without a higher purpose
- Mind and matter were separated into the subjective, human consciousness and an objective, impersonal world
- Science was the intellectual authority, which employed human reason and empirical observation to generate objective knowledge (as distinct from the personal, subjective and speculative domains of metaphysics and religion)
- The modern world was comprehensible by man's rational and empirical faculties alone, and the other aspects (emotional, aesthetic, ethical, relational, imaginative) distorted objective understanding
- Evolution was an amoral, random and brutal struggle for survival and thus modern man was independent intellectually, psychologically and spiritually from God or religion.

2.2 2 Empiricism and the problem of induction

Scientific thinking continued to blossom during the Enlightenment. The resulting success in industrial and technical achievements was accompanied by further philosophical developments, which Tarnas called the "self-critique

of the modern mind". In this phase he included the philosophies of empiricists Locke, Berkeley and Hume, British philosophers from the 17th and 18th Centuries, and Kant from 18th Century Germany. The empirical perspective on the nature and limits of knowledge was driven by the notion that only beliefs based on observation and experience were truly justified. This was in direct contrast to the rationalist's view that truths about the world emerged from innate human reason.

Locke's thesis was that it was necessary to rely on sensory experience and that only reflection on sensation could lead to sound conclusions. Thus if science focused on phenomena that were objectively measurable (weight, shape) it could generate reliable knowledge. Berkeley's position was that objectivity existed as a result of regularity in the mind's perceptions of sensory input. He also cautioned however, that conclusive inferences could not be made.

Hume also grounded knowledge in sense experience. He distinguished between sensory impressions of reality, and ideas, which were only weak copies. Hume's major contribution was to propose that although the causal relationship between sensory impressions and ideas was the basis of knowledge, it could never be directly substantiated. Causality was presented as merely the repeated association of past events, which could provide no logical guarantee of future events. His philosophy undermined the logic of

induction and thus challenged empirical science, which sought to predict generic laws of nature from finite observations.

Kant's solution was to find the middle ground. He rejected extreme rationalism (knowledge from pure intellect) and empiricism (knowledge from pure sensation) and proposed that only their conjunction could produce valid scientific observations.

“The understanding cannot see. The senses cannot think. By their union only can knowledge be produced.” (Kant, Critique of Practical Reason, 1788. Quote from Mackay, 1991)

Kant's solution to the logical problem of induction was to propose *a priori* mental frames through which cognition was channeled. These included notions such as cause and effect, quantity and relation. Hence experience was identified as a construction of the mind rather than having an external source, that was brought to, and imposed on, the input from the senses. Tarnas described the period of self-critique as causing a partial loss of faith in the human mind and its capability of judging external reality (Tarnas, p333-354). With Kant's philosophy of science, a vindication of the scientific method prevailed.

It has been proposed that the 17th Century notion of induction and inductive evidence was a precondition for the emergence of modern probability and

contemporary scientific thinking. That is, in order to construct probability as a relation between a hypothesis and the supporting evidence, scientists needed to use non-deductive reasoning to conceive of evidence as something that points beyond itself (Hacking, 1975).

Empiricism was again taken up in the 20th Century and formalised within “logical positivism”. The movement originated in Vienna in 1922 and its adherents became known as the Vienna Circle. This group of philosophers and scientists met regularly until 1938, and their influence remained strong until the 1960s. Logical positivism proposed a radical scientific approach to knowledge, which claimed that only statements verifiable by facts derived from experience and observation, or from logical mathematical truths, were meaningful. Everything else was classed as meaningless metaphysics. The logical positivists also embraced reductionism and rejected the need for a philosophical justification of science or scientific knowledge (Audi, 1995, p836).

Radical logical positivism was mostly rejected during the 1960s, but moderate elements (i.e. realism, reductionism and hypothesis testing) are retained within “post-positivist” philosophy. Post-positivism is a philosophy or worldview to which many contemporary scientists and social scientists subscribe, including a proportion of those with a naturalistic or qualitative orientation to research (Guba and Lincoln, 1994).

2.2 3 Falsification and Confirmation

A response to Hume's problem of induction and the uncertainty about causal inference came from Karl Popper. His proposal, which was highly regarded by many scientists, was that science did not need to rely on the logic of induction. Rather science was based on a process of conjecture and refutation (Popper, 1963). He rejected the positivist's notion of verification (through induction) and focused on falsification (by deduction).

According to Popper's philosophy, scientists develop a hypothesis or conjecture and then attempt to falsify it by observation and experiment. If they succeed and falsify the theory, then the conjecture is replaced. The logical conclusion of falsification was that "real" scientific theories must be constructed in such a way as to be falsifiable through observation and experiment. Thus those who adopted this philosophical position on science excluded from the scientific domain many psychological or social theories on the grounds that they were not falsifiable.

Popper's philosophy was founded on a notion that one could only refute, but never confirm a conjecture: "our knowledge can only be finite, while our ignorance must necessarily be infinite" (Popper, 1963). If a theory remained unfalsified, it was retained as potentially true but not proven. On this basis

even scientific theories considered as established “truths” (i.e. that planets orbit the sun) could not be accepted as beyond possible refutation. This led some to conclude that Popper’s philosophy had side stepped, rather than addressed, the epistemological problem of confirmation (Papineau, 1996, p294).

Some historians of science have claimed that Popper’s philosophy did not reflect the actual story of many successful scientific theories. They pointed out that some theories were not rejected when they appeared to have been refuted. That fortunately for science many classic theories were pursued despite seemingly strong evidence, at the time of their development, as to their falsity e.g. Newton’s gravitational theory, Maxwell’s kinetic theory and Bohr’s theory of the atom (Chalmers, 1999 (1976), p91). Others have drawn attention to important observations, experiments and discoveries that were made as a result of chance, coincidence or simply trying things to see what would happen. Thus formal conjecture or hypothesis testing were put into the perspective of reflecting only one dimension of scientific progress (Hacking, 1983).

A merger between the philosophy of science and the history of science supported theories about scientists and their work in relation to their historical and cultural context. Contemporary philosophies of science have included

considerations of the social and contextual factors that influence scientific process and scientific knowledge.

2.3 SCIENCE AS A SOCIAL ACTIVITY

2.3 1 Thought styles

Wartofsky (1978) described human knowledge as a “cultural and historical emergent, which bears the impress of a variety of needs, purposes and practices”. It followed that rather than “an a priori feature of human rationality”, science and medical knowledge were described as a socially determined program “whose norms have evolved with the historical growth of science and the development of a world community of scientists”.

An early proponent of a social perspective on science was a Polish doctor and scientist called Ludwik Fleck, whose book titled *Genesis and Development of a Scientific Fact* was published in German in 1935. He presented a radically different perspective to the then dominant logical positivism espoused by the Vienna Circle. His thesis was that “at least three-quarters, if not the entire content of science is conditioned by the history of ideas, psychology and the sociology of ideas and is thus explicable in these terms” (Fleck, p21).

Although Fleck's ideas had little impact at the time of publication, they re-emerged in the latter part of 20th Century. Wittich (1981) has identified four important themes from Fleck's original work:

- The notion of historically determined “thought styles” that are common to members of a scientific community
- That thought styles determine a scientific community's research activity
- That thought styles have a “tendency to persist”, and when they break down they are replaced by a new one (these are discernable as separate phases)
- That training for science requires training in the obligatory thought style.

Fleck also recognised the significance of dissent from group-endorsed thinking and described differences in “thought styles” as a source of conflict.

2.3 2 Paradigms

Fleck's ideas were decades ahead of their time and he remained relatively obscure. The most influential exponent of a historical and social perspective on scientific progress (and knowledge) was Thomas Kuhn, who first published his now famous *The Structure of Scientific Revolutions* in 1962.

Kuhn cited Fleck's work as "an essay that anticipates many of my own ideas" (p ix).

Kuhn's book was taken up as a serious challenge to traditional accounts of science as objective theory-neutral processes of induction or deduction. His thesis was that science was conducted within theoretical frameworks or paradigms. He proposed also that scientific progress was revolutionary because over time, the dominant frameworks were overturned and replaced.

Important components of Kuhn's propositions were as follows (Kuhn, 1996):

- Scientific communities share a paradigm: a world-view made up of theoretical assumptions, laws, standards, accepted techniques and tacit knowledge. These determine the puzzle solving, communication and progress of *normal science*.
- When a paradigm is no longer useful a crisis occurs, resulting in a *revolution* or paradigm shift where the majority adopt a new world-view.
- Paradigms are *incommensurable*: criteria belonging to one cannot be used to judge propositions developed within another
- There are no external or non-paradigm-dependent criteria to judge competing paradigms and thus no common or external yard stick for their comparison

Kuhn's ideas were sometimes interpreted as inherently relativistic because there was no objective means of judging one worldview over another. His ideas were also interpreted to mean that decisions about scientific theories were not based on objective evidence, but rather on the assumptions, a priori beliefs and interests of the community of scientists involved. In his 1969 postscript however, Kuhn rejected the charge that his work presented science as "subjective and irrational" (p191).

In retrospect, Kuhn presented a blend of ontological realism (the world that is examined is unchanged when a "scientific revolution" occurs) and epistemological relativism (derived from the paradigm-dependent interpretations and constructs that are formed about the world). His text has been widely accepted and cited as starting a trend that continues to have currency today: that "a more adequate account of science must proceed from an understanding of the theoretical frameworks in which scientific activity takes place" (Chalmers, 1999 (1976), p104).

The importance of theoretical frameworks provided a key rationale for me to undertake this study to examine, and form propositions about, concepts of evidence among teachers of medicine and public health.

2.3 3 Social construction of science

One of the most radical social interpretations of science was associated with the University of Edinburgh's "strong program". This proposed that scientists' decisions (i.e. which theories to pursue) were inevitably socially determined, and attempts to structure scientific programs to exclude social factors were pointless (Barnes and Bloor, 1982). The strong program was in stark contrast to logical positivists' presentation of science as a superior pursuit with objective and exclusive access to the truth through observation and experiment.

Barnes and Bloor promoted a wholly social interpretation of science and the social construction of scientific knowledge. This encompassed a relativism that accepted theories have different degrees of truth-value, but argued that external factors, rather than evidence, determine which ones are pursued and believed (Couvalis, 1997, p144). Hence their approach went beyond the epistemological proposition that interpretation of data was "theory laden". They concentrated on the sociology of knowledge and how allegiances to social, economic and political groups constructed entire research agendas.

The lack of a simple and direct relationship between the reality of the world and our knowledge of it has been conceived as "undetermination". This means the natural world does not determine observations, evidence does not

determine the theories adopted and scientific results are not independent of the society in which they were generated (McMullin, 1995). It could be argued however, that these social constructivist interpretations of science have had greater attention from philosophers and social scientists than from scientists, and had greater impact on social science epistemology and social theory, than on scientific thought and process. The next section examines epistemologies of social science.

2.4 KNOWLEDGE OF THE SOCIAL WORLD

2.4 1 Naturalists and anti-naturalists

For much of the 20th Century, social science approaches to knowledge were divided into two major camps: the naturalists and anti-naturalists (or dual methodologists). Each camp had several manifestations and there were also intermediate positions. The naturalists believed social scientists should seek general laws like those in natural science. They believed in analytical and exact methods to provide explanations that allow prediction and intervention in the social world as for the natural world.

The anti-naturalists distinguished the science of culture (or spirit) from the natural sciences, both ontologically and/or methodologically. This distinction

introduced by Wilhelm Dilthey (1833-1911) a German philosopher and historian. In the anti-naturalists approach to social science, investigation was based on understanding, sometimes referred to as the German term *Verstehen* (Schwandt, 1997), rather than attempts at causal explanations. The focus of the anti-naturalists was often on the unique character of phenomena and their relationship with culture and values (Outhwaite, 1996 p86).

2.4 2 Post-modern influences

Except in social anthropology, the naturalist philosophy and positivism were strong influences on social sciences up to the 1960s, when they gave some way to challenges from interpretive hermeneutics and critical theory. Further influences came from the post-structuralist critique of grand or “meta” theories, which had returned to favour in the 1970s. Foucault was particularly influential in challenging the structural analysis of society, and led the way in identifying a potential for the simultaneous existence of multiple truths. Scientific and medical knowledge was identified as a “regime of truth”, which although dominant, could be challenged (Foucault, 1980; Macey, 1993).

Most recently the epistemology of social science has been influenced by post-modern challenges to the concept of scientific rationality and progress, on both epistemological and ethical grounds (Lyotard, 1984; Smart, 1996). In the context of scientific research as evidence, an important aspect of post-

modernism was a rejection of the possibility of proving or disproving hypotheses on universal, rational and scientific grounds (Turner, 1992). Some authors claimed that the distinction between natural science and social science had been undermined (Knorr-Cetina, 1981).

Post-modern approaches to inquiry rejected concepts of objectivity and absolute truth and pointed to the fragmented and interpreted nature of all knowledge, and to the political and social antecedents and consequences of its constructions (Denzin, 1994; Richardson, 1991). In a post-modern world, academic or scientific knowledge could be as self-perpetuating and deluding as myth.

“A myth is a fixed way of looking at the world which cannot be destroyed because, looked at through the myth, all evidence supports that myth”.

(Definition of myth from Edward De Bono, Po: Beyond Yes and No, 1972)

Many post-modern and critical theories celebrated partisanship, promoted new creations of reality through re-framing and alternative forms of representation, and promoted emancipation from traditional distributions of power in society (Kincheloe and McLaren, 1994; Harvey Brown, 1995).

2.4 3 Pragmatism

Another strong influence on the epistemology of social science has been the various forms of pragmatism. This emerged at the end of the 19th Century as a philosophy where the meaning of a concept was to be determined by the results of its application (Haack, 1996). Pragmatism greatly influenced the “Chicago School” of sociology and Symbolic Interactionism; a social theory that underpinned the development of the Grounded Theory, the method adopted in this research (see chapter 2).

Pragmatism developed into many forms but, primarily, truth was not about correspondence with reality but “the result of a process of agreement by a community of investigators and based on explanatory and practical success” (Outhwaite, 1996; p92). In radical forms of pragmatism, truth is anything which can be defended and rationality is the respect for another’s opinion (Rorty, 1991).

2.4 4 Eclectic social inquiry

Contemporary reviewers of the philosophy of social science highlight the prevailing diversity and the cross-disciplinary, blended approaches to inquiry (Fujimura, 1991; Outhwaite, 1996; Hollis, 1996). In order to classify ontological, epistemological and methodological distinctions and to provide

an analytical structure, Hollis has proposed a two-by-two table resulting in “four keys to analysing social action”. This is reproduced below (Hollis, 1996; p359).

| | Explanation | Understanding |
|----------------------|-------------------------|----------------------------|
| Holism | <i>Systems</i> | <i>Cultures</i> |
| Individualism | <i>Rational choices</i> | <i>Subjective meanings</i> |

The columns represent the naturalist’s quest for explanation and the anti-naturalist’s priority for understanding. The rows indicate that it is possible to approach explanation and understanding from holistic or individual perspectives. The proposition is that social theories focus on one of the four keys to investigating social action: systems, cultures, rational choices or subjective meanings. These quadrants anticipate the formulation of questions, selection of methods, and the type of evidence that is sought about social action. Each have their strengths and limitations, and because none have emerged as the “master key” the conclusion is that a total picture of the social world probably needs to be assembled from several, if not all, perspectives.

2.5 THE *NORMAL SCIENCE* OF MEDICINE AND PUBLIC HEALTH

2.5 1 Dominant conventions and alternative paradigms

Kuhn's notion of *normal science* described periods of research and practice where communities share accepted conventions of inquiry and functional modes of problem solving. Health professions are social communities where dominant conventions are formalised into the training, socialisation and accreditation of their members (Freidson, 1988). It is possible to reflect on the *normal science* of contemporary medical and public health research and practice. This draws on science and social science disciplines to determine the forms of inquiry and conventions employed for collecting and interpreting empirical data.

The literature on health research commonly distinguishes between realist (post-positivist) and relativist (social constructivist) epistemologies. Contemporary social theory has been employed to interpret medical knowledge (Jordanova, 1995). Yet it is also accepted that within medicine and public health, the dominant approaches to research are those based on traditional scientific principles. Thus scientific research and scientific reasoning are expected to determine medical and public health knowledge, and this is expected to represent, rather than construct, reality (Osborn, 1996).

There has also been a traditional divide between qualitative and quantitative methods in health research, and differences in ontology and epistemology have been identified as the primary basis for the split (Buchanan, 1992; Baum, 1995). Anthropology and ethnography have provided social, cultural and historical accounts of medical science and medical knowledge (MacCormack, 1994; Young, 1980). Naturalistic and qualitative methods have been employed to examine patient perspectives on disease and treatment in clinical medicine, and in social and cultural studies in public health and health promotion (e.g. Jamerson et al, 1996; MacCormack, 1992). However, because dominant criteria for research have been based on scientific concepts such as objectivity, validity and reliability, social constructivist approaches to qualitative inquiry have been labeled an “alternative paradigm” (Guba, 1990).

2.5.2 A focus on diagnosis and explanation

During much of the 20th Century, the “normal science” of both medicine and public health employed the logic of causal inference to explain and predict disease e.g. pathogenesis or epidemiology (Renton, 1994). Causal theories have provided a basis for the explanation and diagnosis of disease at both the individual and population levels. Biological, behavioural and social theories derived from the causal pathways of disease have also been employed to plan medical and public health interventions.

In the first half of the 20th Century the focus of inquiry in medicine and public health was firmly on diagnosis rather than cure. Many interventions were only marginally effective and often a doctor could only provide prognosis and comfort. “Explanation was the real business of medicine. What the ill patient and his family wanted to know was the name of the illness, and then if possible, what had caused it, and finally, most important of all, how it was likely to turn out” (Thomas, 1984; p28).

As technologies improved, medicine was able to develop and offer serious interventions, such as antibiotics and surgery. Although medicine became increasingly scientific, medical practitioners distinguished themselves from scientists. This was based on a difference between the aim of science, which was an acquisition of knowledge and understanding of the world, and the aim of medicine, which was the prevention and treatment of disease (Munson, 1981). Thus medicine would now be judged by the success of its therapies, which gave greater prominence to intervention evaluations. In public health there were parallel moves to balance descriptive and aetiologic research on the magnitude and cause of community health problems, with determining the effectiveness of interventions.

2.6 EVALUATING HEALTH CARE

2.6.1 Numerical methods and clinical epidemiology

Some contemporary health care evaluation techniques, such as a use of control groups to assess medical interventions, go back to the 17th Century. It was not until the 18th and 19th Century, however, that medical statistics were formally developed and applied (Matthews, 1995). The French clinician PCA Louis (1787-1872) is widely regarded as the forefather of the “numerical method”, that is, collecting standardised and reproducible clinical observations so they could be tabulated and counted. He used this method to show that blood letting using leaches (a commonly used medical treatment in his time) actually reduced patients’ overall chances of survival (Morabia, 1996).

The use of summarised clinical observations from groups of patients paved the way for clinical epidemiology, a term introduced in the 1960s by Alvan Feinstein (Feinstein, 1968). Clinical epidemiology refers to the application of epidemiological principles and methods to problems encountered in clinical medicine, including the evaluation of medical interventions (Fletcher et al, 1996).

2.6 2 The rise of randomised controlled trials

The 1980s have been described as a time when health research had increasing influence on health policy in the US. This coincided with the ascendancy of an “economising” model for health research that was one of “regulating expenditures for health services and assessing their results”. It was fed by the research disciplines of economics, biostatistics and epidemiology, but driven by social and political ideas “about individuals and social classes, the public and private sectors, efficiency, and the promise of medical science” (Fox, 1990). The randomised controlled trial in particular gained recognition and influence in health policy settings.

The use of controls was well established in medical trials by the 1930s, but patient allocation had been implemented by the treating clinician, who also decided on the patient’s eligibility for the trial. The decision about eligibility was a primary source of bias as it was highly influenced by the clinician’s knowledge of the treatment option next in line (Doll, 1998).

The technique of randomly allocating subjects to intervention and control groups was developed by Fisher for the purpose of minimising the effects of bias (systematic error) and chance (random error) in agricultural experiments (Fisher, 1926; 1935). Bradford Hill identified the largely untapped potential for employing statistics in clinical evaluative research (Hill, 1937). He was

involved in the design of the first randomised controlled clinical trial (RCT) to be reported in the medical literature (Medical Research Council, 1948; Doll, 1998).

The underlying principle of randomisation was to obtain truly comparable groups, in terms of prognosis and responsiveness to treatment, so that observed effects could be attributed solely to the intervention (Newell, 1992; Kleijnen et al, 1997). Bradford Hill's landmark trial of streptomycin with the British Medical Research Council included the following standards: a careful definition of eligible cases prior to allocation; randomisation to treatment and control groups in sealed envelopes; and blinded evaluation of the outcomes (Yoshioka, 1998). It heralded the rise of a new partnership between clinical expertise and statistical inference and "the confluence of these two separate traditions constituted the sine qua non for the emergence of the probabilistically informed clinical trial" (Matthews, 1995; p130).

By the 1960s the RCT was the primary research design for determining a causal relationship between interventions and their effects (Campbell and Stanley, 1963). Matthews (1995) however, attributes the ascendancy of RCTs in medicine to wider social and political pressures, than to internal consensus in the medical profession. The disastrous use of thalidomide and resulting birth defects resulted in a public outcry and calls for greater regulation of the medical professional. The evidence produced in RCTs would not have

detected the unanticipated effects, but the public outcry was a catalyst for change. The result was a series of 1960s government regulations in the US, Europe and Australia that had a “profound effect on the prestige of the clinical trial” (p141). The RCT continues to be the gold standard test for the efficacy of all pharmaceuticals.

Traditional public health epidemiology was observational rather than experimental, and criteria formalised by Bradford Hill (Hill, 1991(1965)) were an accepted standard for assessing a causal relationship. Although these criteria remained influential in evaluating interventions, over the next decades quantitative experimental and quasi-experimental designs became the gold standard for evaluating the effects of most interventions including those in a “field setting” (Cook and Campbell, 1979). These methods were often applied to evaluating public health interventions.

Determining appropriate methods and standards for evaluating the effectiveness of health promotion, or long term prevention of chronic disease, has proved to be an ongoing challenge. Community based evaluations conducted as randomised controlled experiments have been expensive and many have failed to demonstrate intervention success. It is often unclear whether this was due to inadequate evaluation methods or because the interventions failed (Ebrahim and Smith, 1998; Harvey, 1998; Hawe, 2000). Debates on how to interpret the evidence highlight a need to improve

intervention evaluation methods in order to generate convincing and useful results (Susser, 1995; Winkleby, 1994; Rychetnik and Frommer, 2000).

2.6 3 The research-practice “gap”

Differences between what occurs in policy or practice, and what should occur based on the findings of research, are often described as the research-practice “gap”. This gap has been discussed in relation to many aspects of health care (Buchanan, 1994; Haynes et al, 1995; Davis and Howden-Chapman, 1996; Nutbeam, 1996) and has been a cause for concern among some analysts for several decades. Forty years ago Lindholm (1959) described policy making as a complex process of “muddling through” to counter the overly rational models portrayed to policy students. He concluded that despite sincere attempts at rationality “policies will continue to be as foolish as they are wise”. A series of subsequent trends and developments have attempted to correct this problem, and in the 1990s they culminated (within health) in the “evidence-based” movement.

“Utilization focused evaluation” (UFE) is an approach that focuses on the utility of evaluation findings (Patton, 1997). This is in tune with the philosophy of Pragmatism described in section 2.4 3. In his 1997 publication, Patton placed UFE into a historical context by describing the 1960s enthusiasm, within Governments, for rational research-based decision making.

This was followed by a growing sense that the “utopian hopes for a scientific and rational society had somehow failed to be realised” (p7). The impact of evaluation research on policy had been too slow or non-existent and incorporation of important findings into practice was highly variable.

In clinical medicine, the divide between evaluative research and mainstream practice was sharply brought to the attention of the medical profession by Archie Cochrane (1972) in his monograph *Effectiveness and Efficiency*. This publication collated examples of ongoing use of ineffective, unproven and harmful treatments and decried the lack of systematically available information on health care output. He promoted quantification of health service input and outcomes, and measurement of that relationship using RCTs. Cochrane presented in a “single volume...an articulate challenge to the modern medical establishment” (Thomas, 1997).

Recognition of controlled trials as an arbiter of treatment effectiveness meant that scientific theories and uncontrolled experience were no longer adequate for rational clinical reasoning (Wulff, 1976 and 1986). This principle was contrasted with the realities of practice, where it was found that the key determinant of health care was rarely scientific evidence on effectiveness and efficiency. Rather determinants were undergraduate training, the expertise of senior clinicians, and regional or personal incentives (Sackett et al, 1977; Becker, 1995; McPherson, 1990; Mooney, 1994).

2.6 4 The Cochrane collaboration

Despite its UK origin, initial enthusiasm for Cochrane's publication was primarily found in the US and Canada. Later this had a "reactive effect" in Britain (McLachlan, 1997). Tudor-Hart (1997) takes a sociological perspective on the uptake of Cochrane's ideas: "by the 1970s, medical care the world over was becoming too effective, too expensive and too potentially profitable for investment to be left to an unaccountable profession". By the 1990s "Cochrane's book provided just what the new managers thought they needed".

Within the medical profession the focus was on effectiveness rather than efficiency. Cochrane had especially criticised obstetrics for not seeking good evidence as a basis for practice. In the 1980s clinicians and researchers in the UK reclaimed his legacy by producing the first collection of systematic reviews of RCTs regarding pregnancy and childbirth (Chalmers et al, 1989, 1997). Their work was later consolidated by the establishment of the now international Cochrane Collaboration, the primary role of which has been "preparing, maintaining and disseminating" systematic reviews (Chalmers and Haynes, 1994; Bero and Rennie, 1995).

The Cochrane Collaboration was set up to deal with the overwhelming volume of clinical evaluative research that had become unmanageable, and was inaccessible and underused by the medical profession. Also thousands of RCTs were conducted each year and reported inconclusive or contradictory results. The aim was to identify the potential clinical importance of the findings through systematic review of published and unpublished trials, and quantitative pooling or meta-analysis of the data (L'Abbe et al, 1987; Sackett, 1995).

2.6 5 Calls for methodological flexibility

The prestige of experimental evaluative research in both clinical and community settings has led to concerns about the impact of its dominance. An exclusive focus on RCTs and reliance on quantitative pooled results have been challenged as too narrow and limiting (Black, 1996; Mackenbach, 1995; Feinstein and Horwitz, 1997; Rychetnik and Frommer, 2000). Concerns highlighted the value of observational evidence in addressing the important questions that could not be addressed by experiments.

For decades those with an interpretive or qualitative perspective have reminded researchers that questions should determine the selection of methods rather than allowing methods to drive research agendas (Hammersley and Atkinson, 1983). Calls continue for “methodological flexibility” in the

evaluation of health care, to ensure that important questions are not ignored or forgotten (Daly et al, 1992; Pope and Mays, 1995; McDonald and Daly, 2000).

2.7 GRADING EVIDENCE

2.7 1 Levels of evidence and recommendations for practice

Serious debate about what should count as adequate evidence for health care decisions was taken up in the medical literature following the introduction of the phrase “evidence-based medicine” or EBM (EBMWG, 1992). The principles espoused in the new evidence-based thinking had however, been developed over two decades, building on Cochrane’s broad ideals of *Effectiveness and Efficiency*. Hence the foundations of EBM were the probabilistic approach to clinical medicine known as clinical epidemiology (Sox et al, 1988; Sackett et al, 1991; Fletcher et al, 1996) and systematic approaches to critically appraising evaluative research.

One of the early examples of this approach was a report from the Canadian Task Force on the periodic health examination (CTFPHE, 1979). This was established in 1976 to formulate recommendations on 79 conditions that were to be included in periodic health examinations. Three aspects of these conditions were considered: the burden of disease; the validity and

acceptability of the detection maneuver; and the effectiveness of any subsequent intervention. The recommendations for action for each condition were classified according to the strength and quality of evidence that was available on all treatment or preventive options. See [Table 1](#) on grades of evidence and [Table 2](#) on classifying recommendations for action.

Table 1: Grades of evidence on intervention effectiveness (CTFPHE, 1979)

| Grade of evidence | Criteria |
|-------------------|---|
| I | Evidence from at least one good randomised control trial |
| II-1 | Evidence from well designed cohort or case-control analytic study |
| II-2 | Evidence from comparisons between times or places with or without intervention |
| III | Opinions of respected authorities, clinical experience, descriptive studies or expert reports |

Table 2: Classification of recommendations (CTFPHE, 1979)

| Class of recommendation for action | Criteria |
|------------------------------------|---|
| A | Good evidence to support the recommendation that condition be considered in a periodic health exam |
| B | Fair evidence to support the recommendation that condition be included |
| C | Poor evidence regarding the inclusion of the condition in a periodic health exam |
| D | Fair evidence to support recommendation that condition be excluded from consideration in a periodic health exam |
| E | Good evidence to support recommendation that condition be excluded |

The Canadian report demonstrated the paucity of evidence on the effectiveness of interventions in the area of preventive care, as very few conditions were supported with Grade I or Grade II evidence (Fletcher and Spitzer, 1980). Yet the analytic principles and the focus on demonstrated effectiveness from outcomes-based evaluation research, represented a key development in the notion of evidence-based decision making (Woolf et al, 1990). The study design hierarchy was similar to Campbell and Stanley's from 1963, but its application in "rules of evidence" and the explicit and direct link to formulating recommendations or guidelines for practice were innovative.

The Canadian recommendations on periodic health examinations were regularly updated in the 1980s and 1990s e.g. Douketis et al (1999). In 1989, the US Preventive Services Task Force published its own guide to clinical preventive services, with a second edition published seven years later (USPSTF, 1989, 1996). They adopted a similar approach to formulating graded recommendations for practice. This included identifying the outcome of interest, systematically locating and appraising individual studies and linking recommendations to the strength of available evidence (determined by the number and quality of studies, consistency of results and generalisability of findings) (Woolf et al, 1996).

The Americans also ranked evidence by study design: (from highest to lowest) randomised controlled trials; cohort studies; case-control studies; uncontrolled experiments; descriptive studies; and expert opinion. This hierarchy was labeled “levels of evidence” and this term has been used in subsequent publications on how to review intervention effectiveness (Guyatt et al, 1993; Sackett et al, 1997; NHMRC, 1999 and 2000).

Levels of evidence became a key focus of debate and contention in relation to evaluating evidence in health care, and led to debates about giving more value to some forms of evidence over others. Concerns were raised about the many forms of evidence that were essential ingredients to clinical judgement, but which were seemingly undervalued in narrow definitions of “best available” evidence (Feinstein and Horwitz, 1997). Evidence hierarchies and the valuation of evidence are discussed in chapters 5 and 6.

2.8 EVIDENCE-BASED MEDICINE

2.8 1 Coining a phrase

The group who coined the phrase evidence-based medicine (EBM) and introduced it as “a new paradigm for medical practice” was led by Gordon Guyatt from McMaster University (EBMWG, 1992). The evidence-based medicine working group outlined the difference between traditional practice

and the “way of the future” which “de-emphasises intuition, unsystematic clinical experience and pathophysiologic rationale as sufficient grounds for clinical decision-making and stresses the examination of evidence from clinical research.” Hence EBM was presented as a new, more explicit and scientific approach to clinical practice. In this context, evidence was defined as clinically relevant, quantitative research that informed a probabilistic approach to clinical decisions.

One of the key requirements for practising EBM was to search and critically appraise the literature by applying “formal rules of evidence” (EBMWG, 1992). The purpose was to close the gap between scientifically reviewed clinical research and medical practice (Oxman and Guyatt, 1993). The importance of scientific critical appraisal was reinforced by editorials highlighting the “poverty of medical evidence” (Smith, 1991) and the “scandal of poor medical research” (Altman, 1994).

Critical appraisal skills for assessing clinical research had also been cultivated at McMaster University, where they had been part of the medical curriculum since the 1970s. This expertise was incorporated into a timely (and ongoing) series of “User’s Guides to the Medical Literature”, first published in the *Canadian Medical Association Journal* in the 1970s, and republished in revised form in *JAMA* in 1993 by members of the EBM working group (Oxman et al, 1993). These guides identified steps on how to determine the

validity and applicability of clinical research. Also new journals such as *ACP Journal Club* and *Evidence-Based Medicine* were promoted for their efficiency because they published summaries of the “best” (screened and appraised) evidence (Sackett and Rosenberg, 1995).

2.8 2 Exponential growth of EBM

Whether it was supported or criticised, EBM was a catchy phrase. For the year it was introduced (1992) 26 articles with the keyword “evidence based” in the title or abstract were stored on Medline. For 1996 (the start of this study) there were 250 hits for the “evidence based” as a keyword. By 1999 (when the study was completed) this had ballooned to 908 hits for the year. By that time “evidence-based” was not only a prefix to medicine, but also to nursing, allied health, dentistry, public health, health promotion and health policy, health care etc.

In both concept and name, the proponents issued a provocative challenge by staking the claim of “evidence”. As a result, EBM rationale, process and philosophy were challenged, clarified and argued in many subsequent publications (Graham-Smith, 1995; Lancet editors, 1995a; Carr-Hill, 1995; Sackett et al, 1996; Feinstein and Horwitz, 1997; Charlton, 1997). Some of these opinions are reflected and discussed in chapters 5 and 6.

2.9 THE STUDY CONTEXT

The core data for this research were collected in the Medical Faculty at the University of Sydney between 1996 and 1999 (data collection is described in chapter 4). At the start of the study the Faculty was in the last year of preparing the curriculum for its new Graduate Medical Program (GMP), with the first intake of students due in 1997. This curriculum was structured around four themes: Basic and Clinical Science; Community and Doctor; Patient and Doctor; and Personal and Professional Development. EBM was a new and significant component of the fourth theme.

The Faculty's shift from a traditional six year, undergraduate degree to a four year GMP required major restructuring of teaching style and content. Traditional discipline-based lectures were largely replaced by cross-disciplinary problem-based learning (PBL) sessions. In the same period, the Master of Public Health (MPH) taught in the Department of Public Health and Community Medicine underwent several internal and external reviews (Douglas et al, 1995). Subsequent developments included integration of previously distinct subjects into a core problem-based component of the MPH. These curriculum changes, particularly the GMP, led to an unprecedented level of regular interaction between senior teachers from previously distinct disciplines.

These developments resulted in a time of increased discourse and mutual influence between medical and public health teachers. All Faculty members who were to participate in PBL tutoring in the medical degree attended Faculty development sessions about the new GMP, which included a session on EBM. The early advocates for EBM, who also did most of the EBM curriculum development and teaching, were based in the Department of Public Health and Community Medicine. In addition all staff in the Department of Public Health and Community Medicine, as members of the Faculty, were expected to participate in PBL tutoring or other GMP teaching. As a result they interacted with other members of the Faculty in tutor meetings and discussed aspects of the basic science or clinical medicine curriculum.

Also in the 1990s, the Department of Public Health and Community Medicine expanded its program of short courses in Clinical Epidemiology, and the Masters of Clinical Epidemiology had a growing intake of clinicians from local teaching hospitals. Many clinicians involved with promoting and teaching EBM in the clinical schools had either been students in, or had formal teaching links to, the Department.

The radical changes to the medical curriculum generated heated debate within the Faculty. An important related development was that the Faculty funding formula was being restructured, so that Faculty resource allocation was linked to Departments' contributions to the new curriculum. This gave additional

importance to negotiations between Departments and disciplines about what was to be included and excluded from the new medical program.

Particularly challenging aspects of the curriculum changes were the cuts in traditional teaching time and to core content (i.e. fewer lectures, particularly in basic science, to allow for more self-directed learning). This required difficult decisions on what was the *essential core* of knowledge and the evidentiary basis for various aspects of medicine. The introduction of EBM was often controversial, not only because EBM was new and because some did not agree with its premise, but also because it was perceived to be taking valuable teaching time from traditional subjects such as physiology or anatomy.

As GMP teaching materials and tutor guides were developed, Faculty members discussed and disputed their substantiation and the type of references that were to be provided to students. Each week the teaching materials were related to a particular problem case or topic. Staff and students needed to decide what was the most important information for understanding the case and determining a diagnosis and management. Thus collaboration or executive decisions were required to determine the most pertinent form of evidence that students should seek and present as a basis for their learning.

The Faculty discussions about EBM in the GMP curriculum, and the exponential growth in literature on EBM, led to the topic of EBM forming a

prominent backdrop against which interaction about evidence took place. Equally, discussions about what was and what wasn't evidence-based led many to focus on the broader and more philosophical question about the meaning of evidence. It was in this primed and often highly charged context that I conducted my research.

Chapter 3

METHODOLOGY

“People who write about methodology often forget that it is a matter of strategy, not of morals”

George Homan ([1949] Quoted in Miles and Huberman, Qualitative Data Analysis, 1994)

3.1 INTRODUCTION

The traditional focus in medical and public health research is on rigour in research method as determining the quality and value of findings.

Contemporary approaches that highlight the partial, local and historical nature of knowledge also focus on making explicit the researcher’s position and assumptions, and the rationale and theoretical basis of the strategies selected (Denzin and Lincoln, 1994). The method used in this study was Grounded Theory. It is particularly relevant to make my rationale and interpretation of the theoretical base of Grounded Theory explicit because the method has been interpreted in multiple ways and continues to evolve.

3.2 RATIONALE FOR USING GROUNDED THEORY

The examination of the concept of evidence as reported in this thesis was new in medicine and public health. Hence the research questions identified required an exploratory and qualitative approach. I decided a research strategy was required that would support the proposition that a premature classification of new ideas can stunt their growth (de Bono, 1967).¹

Grounded Theory is a “general methodology for developing theory that is grounded in data systematically gathered and analysed” (Strauss and Corbin, 1994). The aim is to produce a “micro-theory” that is developed from a particular context (called substantive theory by Glasser and Strauss) which would be potentially be more relevant for making interpretations about events in that context than existing meta-theories (called formal theory) or non-empirical deductions.

Thus Grounded Theory was selected as the appropriate method because it is exploratory and highly inductive, particularly at the early stages of a study. Another strength attributed to Grounded Theory is that it promotes in readers an understanding of the topic at a conceptual and explanatory level, because

¹ Edward de Bono (1967) has examined the difference between vertical, logical thinking (high probability) and more open and free floating (low probability) thinking that generates new ideas. He illustrated this with an analogy of digging holes: vertical thinking is excellent for digging better and bigger holes, but if we are digging in the wrong place, lateral thinking is required to find a new spot.

“facts are easily forgotten, ideas are not” (Glaser, 1991, p 12). Grounded Theory was also selected because the steps for analysing empirical data are systematic and explicit.

In summary, the method allowed me to:

- Derive the main components underpinning the complex topic of evidence from empirical data
- Identify recurring patterns and relationships between the main conceptual components in an explicit and systematic way
- Improve understanding of the topic at a conceptual and theoretical level

3.3 DEVELOPMENTS IN GROUNDED THEORY

3.3 1 Core techniques

Grounded Theory is a highly documented qualitative research method (Glaser and Strauss, 1967; Glaser, 1978; Strauss and Corbin, 1990; Glaser, 1994; Morse, 1994; Llewellyn, 1997). The core techniques have remained constant throughout the alternative philosophical interpretations (Annells, 1997) and employ both induction (data giving rise to concepts) and deduction (testing on

the data concepts derived from experience or other sources). The key features of the core techniques are identified below:

Developing theoretical sensitivity: Theoretical sensitivity is about the researcher gaining familiarity with the topic to be investigated, and the context in which it is to be studied. The aim is to have adequate knowledge and local awareness to recognise and name the most important concepts in the data.

Theoretical sampling: Sampling in Grounded Theory is determined by the stage of analysis. It is also driven by the information required to explore fully the concepts identified as important, and to develop theoretical propositions about those concepts. Thus, theoretical sampling is conducted to gather data that are required to fill identified gaps in knowledge and gaps in the developing theory.

Conceptual and theoretical coding: Coding defines the classification of segments of data. Conceptual codes are derived from the concepts that the researcher attributes to particular segments of data. Theoretical codes are based on their propositions about the relations among key concepts.

Constant comparison: This is a systematic process of data analysis where consecutive segments of data are compared and contrasted with each other to

identify and distinguish key concepts. For example, each new segment of data is compared or contrasted to those examined earlier.

Memo writing: Memos are prepared about the codes developed, with details of their properties, dimensions and relations to each other.

Identifying a core category: In the latter stages of data analysis a concept is selected as a central idea or activity related to the research topic. It is regarded as a core category in that all other concepts are examined in relation to that category.

Working toward theoretical saturation: Theoretical saturation is the ideal endpoint of data collection and analysis. It means that data collection and analysis continue until all significant variations of a concept (particularly the core category) have been explored and the main theoretical links hypothesized.

3.3 2 Origins of Grounded Theory

The originators and primary writers on Grounded Theory were two sociologists, Anselm Strauss and Barney Glaser. In 1965 they collaborated in a study of death and dying in hospitals (Glaser and Strauss, 1965). The steps

and procedures they adopted for analysing data were compiled and identified as Grounded Theory (Glaser and Strauss, 1967).

Strauss was an established sociologist from the “Chicago school”, which placed great emphasis on empirical fieldwork and qualitative research. In an intellectual biography of Strauss, Juliet Corbin identified his early theoretical perspective as “pragmatist interactionist”, highlighting the influences of George Herbert Mead, Herbert Blumer and John Dewey. Significant features of this perspective include: regarding human socialisation as a process of ongoing change; the importance of interactions for individual and collective development; focus on the particular rather than the universal; a special interest in action and change; and anti-essentialist notions of truth and knowledge. Strauss combined these symbolic interactionist influences with his strong sense that social-structural contexts determined the conditions and consequences of action (Corbin, 1991).

Glaser’s background was in quantitative, theory-testing research. He had come “straight out of the world of conjectural sociology” and identified his contribution as building on Strauss’ empirically-derived conceptualisation and adding the systematic process of “integration and formulation of theory” (Glaser, 1991). Glaser’s methodic approach, and strong belief in strict adherence to the constant comparison process that is central to Grounded Theory, is spelt out in his later publications (1978, 1992 and 1994).

3.3.3 Diverging methodological perspectives

A notable development in Grounded Theory was the public split between its originators. It occurred following Strauss' collaboration with Corbin to publish a detailed elaboration of the method's procedures and techniques (1990). They built on the broader rationale proposed in 1967, but added "the conditional matrix" as a significant new dimension that presented a strong focus on the context, conditions and consequences of action. It also included alternative suggestions for "opening up" the data; that is using external sources and creative processes for forming ideas, which were "tested" against the local data. Strauss and Corbin's text has been very popular and influential in qualitative health research. This is possibly because it provides a clear and elaborate formula for examining qualitative data, which is also amenable to formulating propositions about explanatory or potentially causal relationships.

Strauss and Corbin's reformulation of Grounded Theory infuriated Glaser. He condemned it for betraying the "true emergence" spirit by "forcing" data into preconceived frameworks so that "the true nature of the data is lost forever" (Glaser, 1992). In his subsequent publications with Sociology Press (a lesser known publisher than Strauss and Corbin's Sage Publications) Glaser continued to stress the importance of the method's core procedures as the only reliable means to obtain a "truly" grounded theory (1992, 1994). In 1992

Glaser condemned Strauss and Corbin's suggestions for using previous experience in theory building as "ultra-pretentious preconception", where the "rigour of sticking to the data is gone" (Glaser, 1992 p53). He also criticised them for describing Grounded Theory as a process of theory building and "verification". Glaser was quite adamant that verification of theory can only take place through follow-up quantitative research and that Grounded Theory can yield only hypotheses.

3.3 4 Contemporary interpretations of Grounded Theory

There has been much discussion of the differing philosophical orientations that underlie the "Glaserian" or "Strausserian" approach to Grounded Theory (Stern, 1994; Melia, 1996; Locke, 1996; Annells, 1996, 1997). Guba and Lincoln's (1994) classification of qualitative research paradigms and Denzin and Lincoln's (1994) review of the history and developments in qualitative research (the five "moments") have been used as yardsticks. There is some consensus that Glaser's methodology is firmly based in the post-positivist tradition, because it assumes a reality in the data that will naturally "emerge" if the methods are systematically applied. In this formulation the method has precedence and researchers are potentially inter-changeable, as long as they apply the procedures correctly.

Strauss and Corbin's approach has also been classified as post-positivist due to their formulaic procedures and definition of Grounded Theory as a "scientific method". However, there is also a social constructivist interpretation of their work. This is based on their emphasis on the researcher's creative theory "development" role, and on the incorporation of social-structural contexts in their potential theories (Locke, 1996; Annells, 1997).

Grounded Theory had its founding roots in Symbolic Interactionism (SI). This is a sociological theory that highlights the significance of human interaction as the source of the meaning that is attributed to words and other symbols. The important components of SI (which also underpinned this study) are as follows: that meanings and interpretations can be fluid, evolving and varied; they emerge from the process of interaction and communication between people; and that these meanings can be examined through the study of interactions. The social constructivist interpretation of Grounded Theory is in line with post-modern developments in SI, which incorporate propositions about macro-sociological factors, and thus counteract criticism of it being "astructural, apolitical, ahistorical" (Plummer, 1996 p238).

3.4 MY METHODOLOGICAL POSITION

Following is a summary of the epistemological and methodological position adopted in this research. It is based on my interpretation of the theories and assumptions underpinning Grounded Theory, and on my approach to issues of rigour that impact on the type of evidence that can be generated.

Primarily I ascribe to a position that Miles and Huberman describe as “transcendental realism” (1994, p4). This is a combination of realist and constructivist worldviews i.e. social phenomena exist in the objective world (not just in our minds) and some form relationships that are stable enough to be amenable to investigation. Our knowledge of these constructs however, can only be indirect and partial, and is bounded by the social and historical context within which it is created.

Procedurally, I accepted Glaser’s emphasis on the constant comparative method as the primary source of theoretical formulation. Although I did not set out to develop Strauss and Corbin’s “conditional matrix”, I followed their direction to consider and include different levels of context.

The findings in chapter 5 are not presented as verified theory, although Strauss and Corbin (1990) use the term verification to highlight Grounded Theory’s empirical base, rather than to mean it was “proven”. My approach

towards verification is relatively conservative, that is, influenced by previous quantitative training and my participation in the evidence-based medicine curriculum in our Faculty. Hence I associate theory verification with greater degrees of certainty than are achievable through the procedures adopted in Grounded Theory. This position is more consistent with Glaser's, and in chapters 5 and 6, I have presented "theoretical propositions on evidence".

My realist view of the natural world is combined with a constructivist view of the method of Grounded Theory. The analytical procedures are not researcher-independent and I class the concepts and theoretical propositions as having been "developed" from an empirical base, rather than naturally "emerging" from data. Harry Wolcott described three levels of qualitative data "transformation": description (observations made by researcher); analysis (identification and sorting of essential features and systematic examination of their interrelationships); and interpretation (what is to be made of it all) (Wolcott, 1994). While much qualitative research can be classified into these stages, it is difficult within Grounded Theory where the procedures are simultaneously descriptive, analytical and interpretive.

My analysis included descriptive summaries of the study context, participant observations and interview data. It also included coding of interview transcripts, which was intended to lift the data beyond ethnographic description by classifying it into conceptual categories. In this analysis, early

concepts represent what is “going on” in each segment of data. As analysis progresses, these are grouped into increasingly abstract categories.

Hypotheses are then developed about potential relationships.

The focus on analytical coding and theoretical formulation encourages data dissection and re-construction into the researcher’s translation and interpretation of the phenomena. Hence, although it is important that all categories are derived from and checked against the data, the “higher level” conceptual formulations quite clearly represent my interpretations, rather than those of the participants. I “must take responsibility and hold myself accountable for the final perspective” (Fujimura, 1991).

3.5 ISSUES OF RIGOUR

Rigour is an overarching concept for determining worth and merit in qualitative research. It is an attempt to address the standard epistemological “problem of the criterion” i.e. how can one determine the extent of knowledge, which resurfaced with the critique of traditional empiricism and knowledge based on induction. As this is a thesis on “evidence”, it is pertinent to provide a summary of recent developments on issues of rigour in qualitative research, and how they apply to this study.

Theoretical perspectives impact on definitions of rigour and influence the criteria by which it is judged (Guba, 1990, p167-201, Guba and Lincoln, 1994). From a realist perspective one would aim to produce descriptions that were a close approximation or representation of the “truth” about the world. From a more constructivist perspective, the aim is to produce clear and understandable interpretations. Within health, the rigour of qualitative research has often been discussed in relation to, or as a reformulation of, the traditional post-positivist approaches to ensure validity and reliability (Sandelowski, 1986; Mays and Pope, 1995; Seale and Silverman, 1997).

In this study, like other research which is a “cohabitation of realist and interpretivist canons” (Miles and Huberman, 1994, p437), my aim is to enhance the confirmability, auditability and credibility of findings (p277-279). This requires that I am “accountable for the rationality and trustworthiness” of the methods (Miles and Huberman, 1994, p5) and leave a clear audit trail so others can reconstruct the process and make informed decisions about the research findings (Morse, 1994b).

Procedural rigour i.e. relying on detailed and systematic process, was especially championed by Glaser (1978, 1992, 1994). Although there is not one “right” way to do Grounded Theory, I have attempted to be explicit about steps taken and decisions made (addressed in chapter 4). I also identify how I

came to be interested in this research and my *a priori* “bias”, as required by the constructivist tradition.

The most recent developments in determining rigour of health-related qualitative research has come from the evidence-based medicine movement itself. This has taken the form of “critical appraisal” to assess publications of qualitative research (Greenhalgh, 1997; Popay et al, 1998; Giacomini and Cook, 2000). These critical appraisal guides present concrete and pragmatic criteria for assessing qualitative research as “evidence”. Mostly they appear to support a post-positivist perspective or one similar to Miles and Huberman’s transcendental realism. They are a useful checklist for what to make explicit in research reports and provided a guide for writing the methods chapter of this thesis.

The highly interpretive nature of Grounded Theory also calls for the recognition of “fuzzy” thinking, intuition and tacit knowledge (Altheide and Johnson, 1994, p492; May, 1994). Although these implicit pathways are difficult to assess, they can be rationally enhanced through developing “theoretical sensitivity” (Glaser, 1978) as outlined at the end of this chapter. They are also promoted by adopting an experienced Grounded Theory mentor (Stern, 1994), such as my associate supervisor G Llewellyn.

I have also attempted to take note of another constructivist argument that rigorous methods are not the only criteria for research (Smith and Biley, 1997). This stems from the position that research will reveal what is already implicit in the methodology and inherent in the researcher: "...each method implies a different line of action towards reality – and hence each will reveal different aspects of it" (Denzin, 1989, p235). From this perspective, the aim of research is understanding and reconstruction, rather than representing an independent reality. "I do not go about trying to discover a ready-made world; rather, I seek to understand a social world we are continuously in the process of constructing" (Wolcott, 1994, p 368).

These perspectives on rigour and quality highlight communication and the impact of the product on the reader i.e. its ability to convey meaning and tell a story. Quality lies in "the power of its language to display a picture of the world" (Buchanan, 1992). I have presented my research and obtained feedback at multiple conferences and seminars throughout the project. This has been an integral part of the research process and my role as a participant observer of debates on evidence. It has also been a means of improving the communication and potential impact of my work. A list of conferences and seminars where I presented papers is in section 4.3 1 of the methods chapter.

3.6 *A PRIORI* POSITION

In using Grounded Theory researchers aim to minimise *a priori* expectations of results and need to be willing to modify their existing assumptions. Glaser (1992) was emphatic about the primary role of induction rather than deduction. My approach was to identify key prior assumptions and document in a research log how these changed over time. This journal was maintained throughout the study and used as a means of reflection or to record many aspects of the research process. Entries were made on conceptualisations of the structure and aims of the project, developing research questions, clarifying my position on the methodology, and on interactions between myself and the research context.

At the start of this study in 1996, I was completing a Master in Public Health (MPH) degree and was employed as an associate lecturer in biostatistics and epidemiology in the Faculty of Medicine at the University of Sydney. I was also employed to assist with the integration of EBM into the new Graduate Medical Program curriculum. As a member of the Faculty and participant in curriculum development and teaching I was suitably placed to undertake research into the interpretation of evidence by researchers and academic practitioners.

My *a priori* perspectives on the concept of evidence were determined by my MPH studies and employment experience in the Faculty. I had observed discrepancies between disciplines in standards of proof and approaches to evaluation while doing the MPH coursework in 1994. My interest in these issues increased when multiple perspectives were revealed in the often heated debates about what was meant by the term “evidence-based”.

Three significant features of my *a priori* perspective, which remained constant throughout the research, were:

- There are differences in how the concept of evidence is understood and interpreted, even among people with a common background in science or health
- Evidence is a fundamental and emotive topic among researchers and health professionals
- There is potential benefit in making explicit the differences and commonalities in perspectives on evidence among those who teach medicine and public health

Other early ideas required modification as the project developed. These were as follows:

- Firstly, I started with a concept of evidence as a relatively stable phenomenon, almost ready and waiting to be discovered. I anticipated individuals to have existing positions on evidence, which may conflict with those outside their discipline, but would be established within their own. These assumptions were modified when I often found myself participating in and observing active constructions of evidential frameworks. This occurred both within the interviews and in the broader context. My developing theoretical understanding of this phenomenon was informed by literature on grounded theory methodology, symbolic interactionism and constructivist approaches to research, which explained the fluidity and social processes which became apparent in my data.
- Secondly, I anticipated interviews and discussions on evidence would be informative, but dry and based on fact and theory. I was surprised by the humanistic and emotional component of the topic. The interviews were often “fun” - creative, imaginative streams of consciousness, which revealed a strong sense of the individuals’ values and passions about their work.

- Finally, I had a relatively dichotomous (right or wrong) approach to research methods. I expected to find more agreement than I did, on what is “correct” Grounded Theory and research in general. Interactions with researchers from various backgrounds, plus tackling the philosophical aspects of evidence, have led to my developing a view of the research endeavor as more value-laden and paradigm dependent. Some of these developments are addressed in more detail in the discussion.

3.7 THEORETICAL SENSITIVITY

Theoretical sensitivity is an essential first step in Grounded Theory, and requires familiarisation with the territory (content and context) in order to recognise and label key concepts (Glaser, 1978; Strauss and Corbin, 1990).

My active involvement within the Faculty has been a key feature of the research. It allowed me to conduct participant observation and gain theoretical sensitivity on evidence in relation to science, medicine and public health. I maintained my position in the EBM curriculum development from March 1995 to March 1998, and continued to take part in Faculty activities such as teaching, research, seminars and conferences throughout data collection and analysis (1996-1999) and writing up (2000). Theoretical

sensitivity was also obtained through reviewing health-related literature and recommended works from other disciplines that formed the basis of chapter 2.

To facilitate analytical thinking, I attended an undergraduate course on history and philosophy of science at the University of Sydney. I also attended a course on qualitative health research at LaTrobe University, Melbourne in order to identify potential strategies of inquiry. Both courses provided a starting point for orientation within this broad and unfamiliar territory. In March 1997 I formed an ongoing, monthly qualitative research study group within the Faculty of Medicine, which continues to run in 2000. Throughout the study, this group was an invaluable source of feedback on my research and opportunity to develop sensitivity to methodological issues in qualitative inquiry.

Chapter 4

STUDY METHODS

4.1 DATA COLLECTION

4.1 1 Principles of data collection

The aim of collecting empirical data in this study was to tap into influential ideas about evidence within the context of academic medicine and public health. Three main tenets underpinned decisions about data collection:

- Interactions between people provide data useful for exploring concepts and their meaning.

- Data collection and analysis can be efficient by purposefully selecting and sampling data sources and collecting rich or dense data.

- Rigour is improved by adopting the principles of triangulation: examining the research subject from various vantage points (Schwandt, 1997).

4.1 2 Data sources

Grounded Theory promotes local and highly contextualised research. My topic was distinguished by a strong interaction between various levels of study context due to a frequent blurring of boundaries between local, national and international forums where evidence was discussed e.g. Faculty, EBH Email list, conferences, and the medical literature.

I selected sources of data to achieve in-depth investigation from a rich core in the form of interviews within the Faculty of Medicine at the University of Sydney. To this core I added breadth by collecting data from sources that reflected the local, national and international contexts where evidence-based health care was a topic of debate.

Other local data were collected as field notes from participant observation in Faculty curriculum development, teaching, seminars and other activities. Many members of the Faculty interacted with developments in the broader evidence-based medicine / health care movement. This is a characteristic of the study participants i.e. senior academics and practitioners who participate in a global academic discourse, enabled by advances in computer technology and electronic communication. Thus debates about evidence on the Internet, in the health literature and at national and international conferences were important sources of data to supplement the core interviews.

Selected data sources:

- In-depth, unstructured, one-to-one interviews with those involved in teaching or research in basic medical science, clinical medicine and public health at the Faculty of Medicine, University of Sydney (collected 1997-1999)
- Participant observations of local curriculum-related activities and local, national and international conferences related to evidence in health care (collected 1996-1999)
- Interactions on the international “evidence-based-health” Email discussion list (collected 1995-1999)
- Editorials, letters to editor and opinion pieces in medical and public health literature (collected 1995-1999)

4.1 3 Triangulation

The notion of triangulation was used to examine the concept of evidence from multiple perspectives and thus to overcome the inherent limitations of a single perspective or measurement. Miles and Huberman (1994, p267) describe triangulation as “not so much a tactic as a way of life” that self-consciously

sets out to “double check findings, using multiple sources and modes of evidence”. Different forms of triangulation were used to achieve and assess the completeness of findings.

In the literature on triangulation, confirmation is used to reduce bias or “validate” findings in situations where agreement is feasible and required. A more constructivist approach would adopt the notion of completeness, where “validity is not the prime issue, but ensuring breadth, depth, richness and wholeness of data” (Begley, 1996). In the latter, consensus is not required or even expected (Hammersley and Atkinson, 1995).

In this study I employed three forms of triangulation (Denzin 1989, p236):

- Data triangulation: time (data collection at different moments); space (multiple sites); person (different individuals and groups)
- Within-method triangulation (data collection strategies; observation, interview, collection of published materials)
- Respondent confirmation (checking findings with participants) (Hammersley and Atkinson, 1995).

Figure 1 overleaf displays three levels of data triangulation in this study (starting from bottom of page).

At the early stages of the study, the participant observations, literature, EBH Email list discussions, and the ideas I developed when attending the history and philosophy of science course were triangulated to develop maximum theoretical sensitivity. They also informed the first draft of my interview questions, which were tested in two pilot interviews.

In-depth interviews formed the core source of data for detailed coding and analysis to build theoretical propositions. A form of respondent confirmation was applied to the interview process. The aim was to check my understanding and portrayal of participants' views by preparing summaries of first interviews and discussing these in follow-up interviews. This confirmation of key issues also helped in developing descriptive and primary conceptual categories. The interview process is outline in section 4.2 and data analysis is described in section 4.4.

Throughout the study I supplemented the core interviews by sampling the other data sources: participant observation locally and at conferences; Email discussion list; and discourse in the literature. Sampling was determined by my theoretical propositions (i.e. theoretical sampling, Glaser, 1992; chapter 3). The purpose was to enrich and complete the developing theory by examining a range of interactions, which occurred at local, national and international levels of context. These interactions were used to interpret new conceptual categories, identify their position in relation to each other, and to

modify my theoretical propositions. Theory development is described in the analysis section 4.4.

Conferences and seminars were also a means of presenting and obtaining feedback on my developing findings. These provided opportunity for a wider confirmation than with individual interview participants. In the last 12 months of the study, data collected from conferences, the Email discussion list and literature were also used to test the theoretical generalisability of my findings. This involved an ongoing assessment of my propositions to see if they “fit” new interactions about evidence, whether they assisted in explaining developments in the evidence-based health care movement.

4.1 4 Ethical considerations

My part-time positions at the Faculty of Medicine allowed me to examine from within, as a participant observer, how evidence was being discussed and negotiated (Bogdewic, 1992). The ethical considerations related to this role and to the in-depth interviews are discussed below. The other data were sampled from public information such as the literature, the Internet and presented papers and discussions at conferences. I have used direct quotes from the published literature (cited) and a few from the Internet (writers not identified) in presenting my findings. The other information has been incorporated as general description.

The National Health and Medical Research Council's guide on ethical aspects of qualitative methods was used in preparing and conducting this research (NHMRC, 1995). The research proposal, participant information sheet outlining the aim of the research, and an interview consent form were approved by the University of Sydney human research ethics committee: the participant information and consent form are attached in Appendix 1.

Two major considerations in relation to the interviews were: the sensitive and political nature of evidence within the changing local environment, and that my main supervisor (SR Leeder) was appointed as Dean of the Faculty of Medicine in 1997. I wanted to ensure that people did not feel compelled to participate in interviews, and made it clear that my supervisors would not be informed who was approached or if anyone declined. I also wanted to ensure that participants would feel safe to speak openly and freely. Individual interviews were confidential. Information available to others about who took part, and what was discussed, was determined by participants themselves; some decided to remain anonymous while others openly discussed their involvement with colleagues and my supervisors.

All interviews were taped and transcribed. Participants were informed that apart from myself, only a professional transcriber (independent of the University) would listen to the tapes, identified by date and number.

Interview transcripts were de-identified (by removing references that identify the speaker) before being shared with my supervisors. The quotes used throughout this report and subsequent publications are not identified.

There were many advantages to having an established role within the Faculty. It also raised questions however, about overlap between my role as an employee and as a researcher. The participant observation literature considers issues of gaining access and “leaving the field”, but does not provide information about conducting research in a context where the researcher is an ongoing participant before and after the research.

An important recommendation for participant observation is to be open and honest about the researcher’s role and the purpose of the research (Bernard, 1988; Bogdewic, 1992). My research was approved by those overseeing EBM teaching and in charge of the GMP curriculum activities. I did not formally introduce my research role at all Faculty meetings, as my focus in observing interactions was to elaborate on my conceptual categories, rather than to study members of the Faculty. However, many people with whom I had contact were aware of my research topic and keen to discuss it with me. My logs and theoretical memos were always written afterwards, contained no identifying details, and were kept at a conceptual level.

On contacting interview participants I identified my involvement in the EBM curriculum, but distinguished this from my research role as a PhD student interested in the broader topic of evidence. I wanted especially to avoid any perception that they were being assessed on their knowledge or support for EBM curriculum activities. The establishment of these “ground rules” was standard for all interviews, aimed at building trust and ensuring informed consent.

4.2 INTERVIEWS

4.2.1 Introduction

Unstructured, in-depth interviews (Minichiello, 1990) were conducted with academics, clinicians and researchers involved in teaching medicine and public health within the Faculty of Medicine at the University of Sydney. They provided intense and concentrated interactions where the ideas and reasoning underpinning concepts of evidence could be explored in depth. Holstein and Gubrium (1995) identify the interview as an active process of “meaning making”, where the researcher “may provoke interpretive developments that might emerge too rarely to be effectively captured in their “natural habitat...” (p18). The interviews were conducted locally to allow me

to interpret them within the bounds of an accessible and familiar context (known from my participant observation and employment).

First interviews lasted 1 to 1.5 hours. Second interviews took place between 3 and 12 months after the first, and usually lasted between 45 minutes and 1.5 hours. They were conducted as a means of respondent validation and to explore how people's ideas had developed. I prepared one or two page summaries of first interviews, which were provided before the second interview to those who participated in follow-ups. Many people said they did not want to see the full transcripts, not only because of their length, but joking that they would be "horrified" at the lack of coherence or "too embarrassed, I think I'd want to correct it". The interview summaries provided a brief and accessible substitute and discussing them at the start of the second interview was a useful warm-up activity to remind us where we had left off. They were also an opportunity to check if I had accurately understood and portrayed their views and if they had any corrections or modifications.

Participants all approved the summaries, although they would often expand and elaborate on an idea or add a new perspective based on events from the intervening period. I had planned to conduct the interviews a few months apart, but found it was more useful to wait even up to 12 months later.

Participants were better able to discuss developments in their own ideas and in the wider context of evidence-based health care. The longer time-gap did not

appear to impact on the usefulness of the summaries as a reminder of the topics already covered.

4.2 2 Sampling

Interviews were held mainly from 1997 to 1998 (with one in 1999) and in total 35 formal interviews were conducted with 22 participants (13 people were interviewed twice and 9 people were interviewed once). Single interviews included the 2 pilot interviews with visiting scholars and one “contrast interview” from a person in the Faculty of Arts, the latter being used to “open up my data” (Strauss and Corbin, 1990; p84) by providing a trigger for comparisons. Data from all interviews contributed to analysis.

Sample size was determined by progress in the analysis. I stopped recruiting new participants or conducting follow-up interviews when I achieved “theoretical saturation” of categories and their relations (Strauss and Corbin, 1990; Glaser, 1992). In addition to the 35 taped interviews, many hundreds of informal conversations on evidence contributed to my research. These were not recorded verbatim, but reports were included in my logs. After I started developing codes these conversations regularly informed theoretical memos on categories and their properties and memos on development of theoretical propositions.

My purpose when selecting participants was to reach respected and articulate thinkers and to obtain a diverse range of perspectives. I used a modified form of snowball sampling (Patton, 1990) which I have called “split snowball” sampling. Each participant was asked to nominate two others whom they regarded as clear conceptual thinkers, but who they believed to have quite different views on evidence to each other. I also asked that at least one nominee would be someone with views significantly different to their own, although the second nominee could be someone to enrich and build on the nominator’s perspective.

In summary, the aim of the split snowball sampling technique was to facilitate:

- The selection of valuable interviewees (i.e. rich sources of data)
- Maximum variation in the perspectives presented
- Efficient coverage and saturation of developing categories
- Generation of interest and willingness to participate by senior academics through introduction and recommendation by respected colleagues.

Participants were mostly happy to be identified to those whom they nominated and often established contact on my behalf. I also presented the option for them to act as an anonymous referrer, in which case their nominees were informed that “a previous participant” had recommended them. The snowball sampling proved to be a useful means of introduction and participants often appeared pleased to have been esteemed by their colleagues, even when they remained anonymous.

Apart from one non-respondent with whom I was not able to establish contact, all those who were approached agreed to be interviewed, which dispelled the notion of “elites” as often inaccessible research subjects (Hertz, 1995). This was likely to be a reflection of a genuine interest in the topic of evidence, the respect afforded to my supervisors, that the project was conducted from within the Faculty and the personal contact established via the snowball sampling technique.

4.2.3 Interview process

I adopted Holstein and Gubrium’s position that both the researcher and participant have an “active” role within an interview (1995). They describe the participant’s interpretive capabilities as “activated, stimulated and cultivated” (p 17) by the interviewer, rather than a “vessel-of-answers” to be tapped. I selected an open and unstructured format in order to examine how

people approach the topic of evidence, and their progress in thinking through the issues.

In setting up and starting the interviews I wanted to allow participants to determine the direction as much as possible and tried to limit my introduction of new concepts to the second half. My role was to set the scene to encourage freedom of expression and then to prompt for explanation and elaboration. As interviews progressed I became more proactive in relation to exploring themes identified from the other data sources and recurring in previous interviews. After further data collection and analysis, I included direct questions based on my developing theory.

The introduction for each interview highlighted the exploratory nature of the research: that there were no “correct” answers and that I was seeking a theoretical understanding of evidence, not evaluating my participants. The first question was always the same:

“If we start broadly, what does the term or the concept of evidence mean to you?”

This very open beginning was challenging and often required a conceptual step back from daily activities. Participants would introduce humour, a little clarification or even ask for another prompt.

“Gee [laugh]...well since I have absolutely no ideas at all I think you could just finish there Lucie [laugh].”

“That's a really good question, do they get easier? (Both laugh)”

”Okay, well, I guess it's information. Information which is em... this is hard, I need a dictionary [laugh].”

The question may have appeared initially daunting, but it did have the desired effect of allowing participants to select the starting point and progress in almost any direction. My pilot interviews demonstrated that it was acceptable to “jump in at the deep end’ given the nature of the participants. They were all senior academics/clinicians (professors or associate professors and two senior lecturers) who are familiar with thinking things through, and appraising and verbalising ideas. These were skills for which they had been selected, as identified by their peers.

Once I included the reassurance that I did not expect established and thought out positions on evidence, and that it was quite usual and acceptable to work through the issues within the context of our conversation, all participants took up the challenge with great skill and often with enthusiasm and relish.

“I’m always keen on people asking me about these kinds of things because in many ways it does actually crystallise my own ideas and I always like that.”

“I mean just to be able to expound on something for an hour I think is a real, a great luxury..”

“the process of interview I just view almost as if you are a sounding board for my ideas”

An unstructured interview can promote reflection and free association of ideas. The spontaneity and creativity of the process in which participants engaged became quickly apparent. My most common role was to encourage and follow their “stream of consciousness”.

“what you’ve done is provoke some very tantalising ideas, which as you can see I’m playing around with”

“I would not have been able to predict at all that we would have got to this point in the conversation”

In addition to the very open introduction, I had a list of back-up topics. As interview data collection proceeded, this list was developed and expanded.

An example of the checklist (from December 1997 when many first interviews had been completed) is included in Appendix 2.

The topics acted as a checklist for common themes, the dimensions of which were explored in subsequent interviews. Mostly I only referred to these prompts at the end of an interview, in case pertinent topics had been missed. Although each interview initially appeared to be different, similar issues were often raised by the participants or emerged from our conversation and did not require prompting or direct questions.

The active interview process required time for pause and reflection: for participants to go back and forth over a topic and develop their ideas and for me to absorb and digest their ideas and formulate a response. My contributions fluctuated between holding back, gentle prompts and direct questions. Occasionally I persisted with a particular topic and pushed for clarification or explanation if something appeared incongruent or missing. This level of interaction was usually accepted with some humour but also a genuine willingness to work through the issues.

“Yeah I know that's what you asked me and I thought I'd escaped [laugh]”

“You are going to tie me down now aren't you [laugh]...”

At the end of each interview I asked participants to reflect on the process. Based on their feedback and my own observations, I concluded that overall participants found it challenging, but stimulating and rewarding in terms of developing their own thinking. The unprepared and unstructured approach was fairly demanding on both sides but stimulated a valuable spontaneity and creativity, which captured a more “natural” and honest response.

“if I had time to think about the topics, it would be more stilted and I'd be more nervous probably”

“I was looking to see if you were going to push me [laugh] and no you weren't, so I bumbled about. I don't know what you made of all of that, but it actually meant that I had to do more thinking than if you had sort of bowled the thing along”

A matter raised by several participants, both implicitly and explicitly, was the commonly felt pressure in scientific academia to be sure and “to get it right”. It is captured by one person’s comment in response to my question on the interview process (below). This also touches on concepts highly pertinent to the topic of evidence, such as knowing the truth, justification and degrees of certainty, which are addressed in detail in the results.

“Terrifying, (did you?) oh yeah, terrible [laugh]... you do put people at ease and everything like that, but they're fairly high level intellectual issues aren't they, and I think in academic life the reason why we proceed so slowly is we all like to really make sure that we've got it all right and not going to say anything stupid that's going to embarrass...it's always so difficult to unburden yourself on something you know absolutely nothing about [laugh]”

In summary, the interviews were unstructured and used a process of reflection and active construction. Participants engaged in intellectual and philosophical reflection with a high degree of free association and sometimes found themselves exploring unanticipated dimensions of evidence. This dynamic aspect in my data is addressed in section 5.2 in the results.

4.3 OTHER DATA SOURCES

4.3.1 Participant observation

Participant observation was undertaken from mid-1995 to the end of 1999 while I was active in curriculum development, teaching and research. Hence my exploration of the topic of evidence was as an active participant from within academia rather than as an external or passive observer.

I considered all interactions on evidence to be a useful source of data, whether the topic was merely introduced or teased out and debated. I kept a log of concepts, exchanged ideas and observations on developments in thinking within the evidence-based health care movement. I also kept a personal log of my thoughts about the topic and my conceptualisation of the project. Once I started to develop codes I mostly incorporated my observations directly into analytical memos linked to the developing concepts and categories.

A major source of participant observation data was in the Faculty of Medicine, including:

- Evidence-based medicine curriculum development meetings
- Evidence-based medicine Faculty development and training sessions
- Graduate Medical Program problem-based learning (PBL) tutors meetings
- PBL teaching sessions with students (who reflected on Faculty debates and disagreements about evidence)
- Department of Public Health and Community Medicine seminars and follow-up discussions by staff and students
- Informal discussions within the Department about evidence; one-to-one, in groups and on Email
- Research and development projects on evidence and effective health care

Other data sources were local and international conferences and seminars, which are listed below.

- Cochrane Collaboration Colloquiums (1996 Adelaide, 1997 Amsterdam)
- University of Sydney, Faculty of Medicine curriculum conference (1997, Wollongong)
- University of Sydney, College of Health Sciences research conference (1998, Sydney) *
- National Health and Medical Research Council and Menzies Foundation Workshop on Evidence-Based Health Advice (1998, Melbourne) *
- Second and Third International Conference on Scientific Basis for Health Services (1997, Amsterdam; 1999, Toronto) * *
- Monash University Centre for Clinical Effectiveness conference on Evidence-Based Medicine; has it fulfilled its early promise? (1999, Melbourne)
- Centre for Values, Ethics and Law in Medicine Seminar on “Evidences in Medicine; arguments for a pluralist world” (1999, Sydney) *
- Department of Public Health and Community Medicine seminar, University of Calgary (1999, Canada) *

An important component of the process of data collection and analysis was to present my findings and developing ideas at conferences and seminars to academics, clinicians and scientists and reflect on the ensuing discussion. The sessions marked (*) are where I presented papers on my research.

4.3 2 Evidence-based health discussion list

Another important source of data was an international “evidence-based health” (EBH) Email discussion list. The host server is Mailbase (in the UK), and the list was established in 1994 by a medical statistician from the London Hospital Medical College, in collaboration with leading EBM practitioners from Oxford and Cambridge. From 1995 to 1999, the Email discussion list was run by researchers from the NHS R and D Centre for Evidence-Based Medicine (CEBM) in Oxford. I started to collect data from this discussion list from the announcement of the official opening of CEBM in March 1995 and stopped data collection mid-1999, although I continued to follow the postings to the list.

The list has been a forum for some of the world’s leading evidence-based health care experts, originators and proponents, as well as observers and students of the movement. It was a source of data that represented mostly constructive and pro-EBM perspectives, where ideas were formed, developed

and shared about how to deal with controversies, overcome obstacles and teach EBM skills to colleagues.

Others used the list as an opportunity to air and discuss the emerging critiques, particularly between 1995 and 1997 when there was often heated debate and controversy about EBM in the health literature. Local Faculty members involved in developing the Sydney EBM curriculum (and several of my interview participants) subscribed to the EBH list for various periods between 1995 and 1999. Thus there was significant interaction of ideas about evidence between the data collected from interviews, participant observation and the EBH list.

I followed all discussions on the list and used it to keep up with broad developments in the evidence-based health care movement. I regularly incorporated broad themes from the list into my analytical memos, but for subsequent and more detailed analysis I sampled discussions that particularly addressed my research topic. Those discussions were concerned with the following:

- The meaning and definitions of evidence and evidence-based medicine
- Critiques and controversies about EBM
- Changes and developments in the evidence-based movement

- Recommendations for evidence-based practice and reactions from colleagues
- Philosophical discussions about truth, proof and the role of science and evidence in health care
- Different types, forms and levels of evidence, their strengths, weaknesses and relative value.

Discussion on these topics tended to come in bursts, ranging from one or two postings to interactions lasting several weeks and comprising dozens of Emails. I printed out hard copies of each message posted on these issues and sampled them for my analysis to either draw out new dimensions of a concept or explore the relevance of local Faculty issues in the international context. For example, reactions from clinicians overseas to an incorporation of EBM concepts into medical teaching; or discussions about the difference between epidemiological evidence and evidence from clinical experience.

4.3 3 Medical and public health literature

Discourse and interactions in the medical and public health literature were also used as data to explore concepts of evidence and follow developments in evidence-based thinking. Editorials, letters to the editor, opinion articles and reviews of the evidence-based movement from 1995 – 1999 were theoretically sampled as data to generate themes, inform development of key analytical

concepts and explore their broader applicability. (The analytical process is outlined in the next section).

Medical and public health literature was also used to gain insight into international academic and scientific forums. This formed the broadest context within which local discussion of evidence-based medicine took place and with which all other data sources interacted. It was quite common for participant observation, interviews or the EBH list to highlight a recent or topical publication, which I was able to follow up and incorporate into my analysis.

The literature was also used in the traditional way for setting the scene and the literature review in chapter 2 and to discuss my findings in chapters 5 and 6.

4.4 DATA ANALYSIS

4.4 1 Introduction

The process of data analysis ran along side, and informed, data collection. The first step was to identify, name and explore the descriptive and conceptual components in my data. After generating key analytical categories, I concentrated on exploring and developing their relationships. Ultimately my

aim was to construct theoretical propositions about approaches to evidence within the context of academia and the evidence-based health care movement. My conceptual and analytic process is outlined here and the results and discussions are presented in chapters 5 and 6. I wanted the findings to be readable and accessible for those not familiar with Grounded Theory methods. Thus although I have used Grounded Theory terminology in the section on data analysis, but I have minimised the use of this language in chapter 5.

4.4 2 Summary of analytic strategies

I combined several strategies to maintain a balance between detailed coding of segments of text, and broader more holistic perspectives on the data. The analytical strategies used are outlined below and represented in the data analysis flow chart in [Figure 2](#) overleaf:

- Formulating summaries of individual interviews
- Developing descriptive themes from interviews, supplemented with theoretical sampling from other data sources
- Detailed analytical coding of interview transcripts

- Generating broad conceptual categories by merging individual single analytical codes and incorporating the descriptive themes
- Examining the dimensions of key categories and exploring their generalisability and “fit” across interviews and with other data sources
- Exploring and mapping theoretical links between categories and preparing regular summaries of “the story so far” to develop theoretical propositions

Figure 2

4.4 3 Descriptive themes

I started data analysis by listing some key themes. These were identified and developed from examination of the pilot interviews, supplemented by the medical literature, EBH list and local discussions. Over time and with each subsequent interview the themes were modified and expanded. The aim was to generate a list of some of the main and recurring broad themes in my data and to interpret these using multi-level perspectives from other data sources.

The construction of descriptive themes proceeded in parallel with detailed analytical coding of interview transcripts. The developing themes and the codes that emerged from analytical coding often overlapped. Eventually they were combined to form the key conceptual categories, which are described in the next sections. The early themes were also used to inform questions for new interviews.

Themes were separated into those related to evidence as a broad concept (e.g. what counts as evidence; its valuation; its use in different contexts; and its abuse) and specific issues emerging around EBM (e.g. critique and controversy, strengths of EBM, religious metaphors).

4.4.4 Interview summaries

A second strategy was to distill the 30 or 40 page interview transcripts into representative summaries of one or two pages. My aim was to draw out the points the participants had wanted to make so that they could be checked with the respondents, to gain an overall sense of each interview, and to use the summaries as a basis for follow-up interviews. They were also useful when developing the descriptive themes.

Below are extracts from three of my interview summaries. Such segments of one or two summary paragraphs represent 15 minutes to half an hour of real-time interview discussion:

First extract from interview summary:

'This person was reflecting on evidence within the context of making decisions, therefore in its applied context...

Believability of information is what constitutes evidence, the criteria for which come from the theoretical constructs of science (scientific religion) e.g. RCTs. These determine what we see as the most appropriate means of getting as close as possible to the truth. Believability is not just scientific rules of evidence as it also depends on what you want to apply it to (utility). The evidence may be valid but it may not be enough information for a particular

decision if other information is missing, such as quality of life outcomes of an intervention.

The scientific rules of evidence require a certain amount of faith, in that we cannot test that they are getting us closer to the truth other than using the same rules themselves. This is a cyclic problem.”

Extract from a different interview summary:

“Evidence provides the underpinning or support for a belief or point of view. In a medical framework it provides explanation of what is going on and evidence is associated with facts that are demonstrably true. Often we work with fairly linear, logical, concrete reasoning processes that allow another person to see why you have come to a particular conclusion.

Evidence related to other types of beliefs such as political ones, are often more nebulous and come from a variety of different sources, some rational and linear and some irrational and diffuse. ... Evidence should be reproducible and subject to scientific inquiry, which is different to the often unique factors related to the context in which it is applied.”

Third extract from interview summary:

“Evidence is about the probability of having the right answer. Experiments need to be done several and sometimes many times, depending on how

variable they are and how important it is and how much certainty you need. Prefers to do experiments with open and shut answers (deterministic) rather than those that require the application of statistical tests in order to show a significant effect (probabilistic). It's better if it works every time, under all conditions. If it is confirmed in thousands of experiments then you can't get any harder evidence, other than an axiom, which is absolute without evidence. If an experiment works 80% of the time, then 20% of the time you have not got the conditions right, in the sense that you have not explained and accounted for all the variations that can take place. All evidence is conditional, there will always be provisos and conditions under which it will or will not apply. Evidence is never absolute, so it will always be a probability in the broader sense but it's just a matter of what you know and what you don't."

The interview summaries were a holistic record of the stream of thought followed by each participant. As such they provided easily accessible overviews of their particular perspective on evidence and complemented the analytical coding of small segments of data. In advanced stages of analysis, I returned to the interview summaries to compare the perspectives of individual participants with my theoretical propositions about concepts of evidence.

4.4 5 Generating analytical codes and categories

Analytical codes were primarily developed from the interview transcripts. They were supplemented by theoretically sampling the other data sources to enhance exploration of their properties and dimensions. Over time, analytical codes were grouped and combined, together with the descriptive themes, to develop broader conceptual categories.

The systematic Grounded Theory method of constant comparison was applied. Each consecutive segment of data from each interview was compared and contrasted to the previous segments to determine how they could be distinguished and labeled. When a new code was developed, I attached a memo documenting its definition and characteristics. Each additional piece of data contributed to the coding schedule as it was checked against existing codes and their properties. This process has been referred to as “open coding” (Glaser, 1978, Strauss and Corbin, 1990) and is driven by the following analytical questions:

- What is this segment of data about or what is the main theme?
- How does this segment differ to the one before?
- How are these codes different to each other?
- Do I need a new code to classify this data?
- How are these codes alike, and can they be grouped into a category?

Rather than line by line coding, I varied the size of data segments coded. This allowed me to use phrases, sentences or paragraphs, according to the richness of each piece of data and the number of codes it appeared to represent.

I used the NUDIST (Non-numerical Unstructured Data Indexing Searching and Theorizing) package from Qualitative Solutions and Research Pty Ltd for data management, to store transcripts, attribute codes and develop memos.

Most of the coding was done on-screen and NUDIST facilitated efficient collation and retrieval of data segments and memos related to each code. This is conceptually represented in [Figure 3](#) below. Memos are linked to the transcripts and to codes developed from these data. Data retrieval consisted of collecting all data segments and memos linked to particular codes and categories.

Figure 3

The linked data-code-memo facility allowed me to restructure coding schedules as one function, and to keep track of the changes, by moving or merging codes with the relevant memos and data segments attached.

The analysis process was iterative - moving back and forth between data collection, modifying codes, exploring their relationships and collecting more data. In addition to on-screen coding I read transcripts away from the computer, making notes in the margins that were later transferred to analytical memos. I found it necessary to combine on-screen and off-screen reading of the transcripts as each process led to different perspectives on the same data. The off-screen reading stimulated broader and more lateral thinking, while on-screen coding tended to focus on detail and developing particular codes.

My early codes were predominantly descriptive; classifications depicting content of discussions and interviews, such as the characteristics attributed to evidence, the types of metaphor used and criteria determining the quality of evidence. As I gained confidence in the method of comparing segments of data my codes became more analytical. I began to group descriptive codes (and combine them with the developing themes) to formulate key categories representing conceptual processes e.g. evaluating evidence, seeking truth, misuse of evidence, socialising or attributing social characteristics to evidence.

The key categories from this stage of analysis (March 1999) are described below. Names given to key categories and important sub-categories are identified in *italics*. Appendix 3 contains classification schemata from three preceding time periods to illustrate their development (May 1997, December 1997 and March 1998). Identifying the key categories represented a turning point in the analysis. I shifted focus from creating new categories to that of achieving their “saturation” (by developing their properties and dimensions) and exploring their potential relationships. This process is described in the next two sections (4.4 6 and 4.4 7).

Key categories:

Accessing truth

Evidence was often discussed in relation to ideas about truth and reality, and the process of seeking or knowing the “truth” about the natural world.

Constructing frameworks

The concept of evidence was a composite, assembled in multiple ways and comprised of several dimensions. The elaboration of the components of the topic of evidence was a dynamic and active process that required interaction and reflection, rather than being presented “ready made”. Frameworks for evidence included its purpose, criteria for evaluation, application context and

individual value systems. These concepts are considered as individual categories below.

Definitions of evidence

There were varied approaches or “points of entry” taken to thinking about and defining evidence.

Qualities of evidence

Definitions of evidence were linked to descriptions of its attributes and the implications of these qualities.

Evaluating evidence

The evaluation of evidence was closely related and sometimes overlapped with definitions of evidence, i.e. evidence was defined as “valid” or “reliable” information, or research that met particular standards and criteria. Evaluation of evidence was conducted along different dimensions, and often against specific criteria.

Relating evidence to knowledge

Evidence was discussed in relation to epistemology: what is known and what can be known. Relating evidence to knowledge led participants to elaborate on their *epistemic positions* on notions of proof, certainty and uncertainty. Relying on evidence was often a means of managing inevitable uncertainty.

Relating evidence and action

Evidence was related to participants' professional roles: their research judgements, decisions about practice and other social agendas. These were elaborated on in the context of discussions about the *role and functions* of evidence; how evidence is used, applied and interpreted. The relationship between evidence and action was also linked to a notion of *pragmatic compromise*, which was often perceived as inevitable in relation to applying evidence to make decisions about practice.

Misusing evidence

The notion that the concept of evidence can be misused, either through poor judgement or with a hidden agenda.

Relating evidence and values

The professional values of researchers and practitioners were sometimes considered in terms of their impact on the generation, evaluation and application of evidence. These values were highlighted as determining *definitions* and *criteria* for evidence. There were questions about how well these reflected or supported patient or community values.

Referring to rules and standards (standardising evidence)

The definition, evaluation and application of evidence were often linked to the use of rules, standards and criteria. The standardisation of evidence was also related to a category labeled *measurement*, where the term evidence was related to the process of seeking and quantifying an external reality.

Contextualising

Standardising evidence against rules and criteria was contrasted to a process of relating the concept of evidence to particular situations, applied contexts or practice decisions. This category highlighted how the term evidence was an applied concept that takes into account individual needs and local contexts. A category called *humanising* focused on the considerations that were given to the personal and human dimensions of evidence.

Socialising

Evidence was also considered at a broader social level, where focus was given to the social constructions of science and evidence. This related to definitions of evidence and considerations about the implications of particular constructions of evidence over others.

Professional Territory

Evidence was mostly considered within the professional role. Professional territory included the skills needed to generate good evidence, the art of

critique required to evaluate evidence. Participants also described the independence and autonomy required to interpret evidence and to apply it to professional decisions. These were related to a process of making *professional judgements*: that is *weighing* options based on different types of information, including evidence, and balancing competing agendas.

Positions on EBM

This category included critiques and appraisal of EBM, the debates and controversies about EBM and the professional positions adopted in relation to EBM.

At this stage of analysis I was preparing to combine the key categories into a more integrated formulation and to develop further theoretical propositions.

Figure 4 indicates how I had organised the key categories in March 1999, when I asked the question “what have I got so far”?

Figure 4

In Figure 4, the active construction of evidential frameworks for different professional contexts is identified as a key process (left side of page). People constructed their ideas about evidence within a social and cultural environment where *accessing truth* was given great value and priority (right side of page). The concept of evidence was primarily discussed in relation to the three main domains: that of *knowledge, action and (professional and social) values* (second from top of page).

Another important category was the notion of *weighing up* and *balancing* competing evidence and other issues within, and between, the three domains of knowledge, action and values. The judgement process was placed at the top of the page in Figure 4 to reflect a potential core category. The remaining categories were grouped below. The next step in the analysis was to tease out the properties and dimensions of the key categories and explore their theoretical links.

4.4 6 Developing category dimensions and exploring generalisability

To explore dimensions of the key categories, questions were used to search for the different antecedents and outcomes of these concepts. The following questions used to explore the category of *misusing evidence* are an example: In what different ways is the *misuse of evidence* conceived? In what contexts

and under what conditions does it vary? What leads to misuse of evidence and what are the implications?

To answer these questions I collated all segments of transcript data coded to a category, examined the memos, and sampled the other data sources. I also collected new data to look for supporting and contradicting instances, particularly if there were gaps in the developing propositions. I used “Inspiration” (a flexible computer software package for on-screen brainstorming and concept mapping) to map the key categories and their properties. [Appendix 4](#) is an example of a concept map for the category *misusing evidence*.

Throughout the analysis I iterated between generating categories from interview data, and conceptually elaborating and explaining these categories by drawing contextual and historical details from the other data sources. These data sources helped to determine the wider relevance and generalisability of the concepts developed from the transcript data. Data from participant observation, the Internet and medical literature also provided a perspective on local and international developments in EBM, which formed a backdrop to the in-depth examination of evidence in my interviews.

In summary, categories were reinforced when they were reflected in the literature, on the Internet or had been discussed at conferences or in local

Faculty meetings. Conversely categories emerging in the other data sources were often developed and explained in interviews. This process of triangulation and theoretical “testing out” of ideas was maintained throughout all stages of data collection and analysis.

4.4 7 Exploring theoretical links and the “story so far”

The NUDIST program was developed for use beyond data management and coding and has been promoted by its developers as a package to facilitate theory development (QSR, 1997, p3). I found it to be a valuable tool for developing conceptual categories but rather cumbersome and limiting when it came to thinking about relations among categories. This was primarily due to its emphasis on hierarchical and tree-like coding structures, which limit how conceptual relations can be mapped.

After I switched to the “Inspiration” software my analysis developed beyond the coding frameworks I had stored in NUDIST, and I did not employ its advanced functions aimed at theory building. Rather than computer-generated searches for word, phrase or concept patterns, I adopted a more traditional analysis. I printed out and reviewed all data related to each category, re-read my interview summaries and participant observation memos, and prepared regular summaries of “the story so far”. I used terms from the key conceptual categories in the writing, and generated new categories to reflect my proposals

about the relations between existing categories. These summaries represented early versions of the story line presented in my results.

Strauss and Corbin (1990) recommend selecting a core concept (or process) to link and facilitate reporting on all other categories. After I had developed each of the key categories more fully, the three central categories of relating evidence to *knowledge*, *action* and *values* (personal or social) were retained as priority concepts. I still needed however, to tease out the links between them and to identify a core conceptual spine for writing up the results.

I returned to the interview transcripts and prepared a new format to summarise these using a cross-tabulation of the key categories. Each category was considered under the three domains of *knowledge*, *action* and *values* and a fourth column for *weighing up / balancing*. My aim was to get a fresh look at relationships between key categories and to compare their properties under each of the three central domains. [Appendix 5](#) contains an interview summarised into this table format as illustration.

At the end of these summaries in table format, I added a brief overview of the evidential framework that had been constructed in each interview and the participant's position on EBM. This was a means of drawing the various concepts together and linking participants' generic approaches to evidence to their particular approaches to evidence-based medicine.

From these tabulated overviews, I identified the process of *judgement* as the core category. Judgement linked the domains of evidence and knowledge, evidence and action and evidence and values. This relationship is outlined in Figure 5 (below) and explained overleaf.

Figure 5

The left and right sides of the diagram distinguish between different aspects of *judgement*; the left side represents how evidence is *applied to make judgements*, while the right side is about stepping back (*refocusing judgement*) to critique the rules and criteria used to define and appraise of evidence.

The top row in Figure 5 represents how evidence related to the generation and substantiation of *knowledge*. The bottom row reflects the application of evidence in *decisions and action*.

The middle row identifies how participants reflected on evidence in relation to professional and social *values*. To the left, values underpinned judgements about knowledge or action. To the right, values were considered in relation to how they determined the construction of rules and criteria for evidence.

After I identified the process of *judgement* as the core category (and teased out its properties and dimensions) I re-examined all the other categories in the light of that analytical decision. This included exploring relations between aspects of judgement and each of the existing categories. I also examined the implications of the core category for new relationships among these key categories. Hence the notion of judgement was used as a central concept for the purpose of structuring my results and writing about all the other key categories. In the findings I have included the *dynamic* construction of participants' concept of evidence, their *definitions* and *evaluations* of

evidence, the reflections on evidence in relation to *knowledge, action and values*, and the influences of EBM. The structure adopted to present my findings is summarised in the introduction of the next chapter.

Chapter 5

RESULTS AND REFLECTIONS

“For both the medical practitioner and the medical researcher, the amorphous concept of “medical judgement” (whether that judgement is executed at the bedside or the laboratory bench) cannot be reduced to a set of explicit rules.”

Matthews JR (Quantification and the Quest for Medical Certainty, 1995)

5.1 INTRODUCTION

This chapter presents the main conceptual categories and their theoretical links in the form of description and interpretation. Each section begins with a brief introduction and a summary of my propositions (grounded theory) about concepts of evidence. These are followed by illustrative and explanatory details of the conceptual categories in that section and reflections specific to the concepts presented. Final reflections and conclusions are contained in chapter 6.

Quotes are used throughout the results to illustrate the concepts presented.

Some sections i.e. 5.2 3 (Shifting sands: developments in the evidence-based

movement) contain quotes from published sources such as medical literature and newspapers, which have been identified and referenced. All anonymous quotes are from the interviews, participant observation (words or phrases) or the Internet email discussion list (where indicated). The quotes selected reflect ideas and concepts that recurred across the different sources of data. As such they are representative of the analytical category to which they have been linked. Individual terms or short phrases have also been highlighted throughout the body of the text (in *italics*), and these signify concepts that were used as categories or codes (sub-categories) during my analysis.

Identifying the quotes by participants' profession or discipline was not appropriate. Many straddled two or even three areas of training and professional activity and reflected on evidence accordingly. Local and international advocates and teachers of EBM had training in both clinical medicine and clinical epidemiology or public health. Several participants combined clinical practice with basic science research. Those trained in public health were often involved in clinical research, as well community-based health services and health policy.

My proposition based on this research is that the concepts presented apply across many medical and public health disciplines, although some concepts are highlighted in relation to certain professional roles. The sampling in this

study however was directed at generating hypotheses, and a different study would be required to verify these propositions.

My report on concepts of evidence and perspectives on evidence-based medicine has been integrated throughout this chapter. In the data, the amount of interplay between the topic of evidence and EBM was influenced by the context in which evidence was discussed. In relation to clinical practice 80-90% of interview data and nearly all data from other sources reflected overlap between discussions of evidence and EBM. Whether in support or challenge, EBM was raised in the definition of evidence, and its potential implications for decision making were considered. For a public health context, overlap between the topic of evidence and EBM was around 50-60%. This centered on comparing clinical and public health research and using evidence for clinical, versus policy, decisions. Evidence was related to EBM in only about 10-20% of the data that addressed medical science research.

5.1 2 Matters of judgement: a structure for chapter 5

Chapter 5 has been divided into 6 sections (including this introduction and overview as section 5.1).

The core conceptual category to emerge from my analysis was the process of *judgement*. Glaser and Strauss referred to this as the “core social process”, which I have adopted to present a proposed grounded theory on evidence.

First, judgement reflects the process in which my interview participants engaged in order to determine their positions on evidence. Their judgements were mirrored in interactions observed from the other data sources, as various health care researchers and professionals negotiated their positions on evidence.

Judgement also reflects the social process or “mode of reasoning” in which concepts of evidence were invoked or used. Thus in medicine and public health concepts of evidence were primarily defined in the context of *forming* and *influencing judgements* in research and practice.

Finally judgement represents the step-back and refocusing in which participants engaged as they reflected on EBM and its “rules of evidence”; or the criteria adopted to guide evidence-based conclusions and decisions. This

reflected judgement on the social construction of evidence and the dominant frameworks that were emerging in the light of EBM.

The results chapter has been structured so that the other sections represent these different elements of judgement as the core category:

Section 5.2: Dynamic and evolving evidence

Concepts of evidence were the products of a judgment process that continued to evolve and change. The theoretical explications of evidence presented here were not established or fixed professional territory. Discerning and judgement were required by those participating in the study to determine the definition, boundaries and scope of the topic of evidence. Section 5.2 portrays the fluid and actively constructed nature of evidence. Also presented are developments in the evidence-based health care phenomenon, which provided an influential and changing context to this research.

Section 5.3: A quest for “always latent” truth

Concepts of evidence were influenced by value judgments about what is important in research and practice. In the academic and scientific environment of this study, interactions about evidence revealed some shared principles and values, which are described in section 5.3. These include significant value given to the notion of seeking and describing “latent truth”, and to principles of substantiation, critical appraisal, pragmatic compromise and managing uncertainty. These values are presented as a common ground

for interactions about the topic of evidence. They also reveal a domain of potential disagreement if these underpinning values and principles are not shared.

Section 5.4: Defining evidence: description and appraisal

As matters of judgement, definitions of evidence varied. Definitions or concepts of evidence formulated by those participating in this research involved two key elements: descriptions of evidence and appraisals of evidence. Section 5.4 presents how the descriptions of evidence were approached from three different “points of entry”. These points of entry were mirrored by appraisals of evidence along three “dimensions”.

Section 5.5: Forming and influencing judgements

The *context* in which evidence was described and appraised was that of using evidence as a basis for one’s own judgment, or invoking evidence to influence the judgement of others. Descriptions and appraisals of evidence varied according to the judgement context. Section 5.5 outlines two different types of judgement: forming epistemic judgements and making decisions about action. Also outlined is the process of using evidence in argument and persuasion to influence the judgements of others. These generic functions of evidence applied across many disciplines, but they were highlighted in particular professional settings and professional roles.

Section 5.6: Refocusing judgement in the light of EBM

In medicine and public health, concepts of evidence have been reconstructed or modified in the light of EBM. Many researchers and practitioners have reflected on the way EBM has influenced professional definitions, appraisals and uses of evidence in health care. Section 5.6 describes how, in the light of EBM, judgements about evidence partly refocused from a methodological and practical perspective to the social construction of evidence and its social implications. This included reflections on what type of evidence is valuable and important, and the potential for misusing the concept of evidence.

5.1 3 A note on the use of terms

These findings present the variety of ways in which evidence was described and appraised, and the different contexts in which the concept of evidence was invoked and examined. Thus to reflect the diversity of concepts of evidence represented in the data, I have used the term evidence in varying ways. For example, I present evidence as a concept that is the product of judgement in section 5.2. In section 5.5 it is discussed as an entity on which judgements are based. Evidence is also discussed as a mode of reasoning and as a social construct. My aim has been to map the various ways in which the concept of evidence was explored and invoked, and to present some of my own theoretical propositions based on the data.

Another feature of the interactions examined for this research was the diverse way in which terms such as *facts*, *data*, *results*, *information* and *knowledge* were used. Some participants used these terms relatively casually and interchangeably with each other. Often these words were also used interchangeably with the term *evidence*. Others focused on making comparisons between evidence and some of these terms to distinguish them and form a position on the concept of evidence.

For the purpose of clarity in presenting the results, I have listed some terms below and given their common meanings in the data. Where possible I have attempted to use the terms in the text in a way that is consistent with the definitions presented here. However, these meanings were certainly not universal and some diversity is reflected in the quotes used.

The terms *truth* and *reality* were often used interchangeably. In relation to evidence, these terms were mainly used to represent the concept of “latent truth” i.e. a reality about the external world that is hidden from direct observation, but which can be researched and estimated. Participants often referred to *facts* as reflecting the closest and most reliable estimates of reality or “latent truth”. *Data* and *results* were mostly used to mean the findings of empirical research. *Information* was used very generically and unlike data or results, information could have numerous sources. *Knowledge* was mostly used to mean a belief that was scientifically or empirically substantiated.

This use of these terms reflects an epistemological hierarchy, with “latent truth” or reality at the base. Facts sit just above this reality because they reflect that truth with the greatest certainty. Above facts are data or results, and the (broad) concept of information sits at the top, and thus the furthest from reality at the base. Hence knowledge supported by facts can be seen as having the greatest degree of certainty in terms of whether it reflects “latent truth”.

Concepts of *evidence* were placed in different positions in relation to this hierarchy of terms, depending on how evidence was defined. Variations in how evidence was described and appraised by participants are described in this chapter. The position given to concepts of evidence in relation to the epistemological hierarchy described above is discussed.

5.2 DYNAMIC AND EVOLVING EVIDENCE

5.2 1 Introduction

Before starting this research I expected that in addition to accepted standards and criteria for research and practice, professional disciplines would have relatively established, although possibly varied, concepts of evidence. I anticipated that I would be able to identify and contrast the existing differences and thus gain insight into the confusion and disagreement surrounding the emergence of evidence-based medicine.

My findings challenged the assumption that the concept of evidence was an established professional norm, with defined boundaries. What I found were dynamic reasoning processes, often evolving and sometimes in stages of significant re-formulation.

5.2 2 Summary of my grounded theory on evidence

Concepts of evidence were the products of individual and group judgements that were negotiated and determined in the various levels of interaction reflected in my data. Evidence based medicine was a highly influential social phenomenon, that had triggered and intensified these interactions about evidence. Many researchers' and practitioners' concepts of evidence were

influenced by or modified as a result of their observation, critique and interaction with EBM. At the time of study and writing, professional concepts of evidence continued to evolve and change.

Details of these concepts are presented in the sections below.

5.2 3 Evidence as uncharted territory

“...I hadn't really thought about what evidence was ‘till you asked me that question, ‘cos I've never looked it up in the dictionary or thought about it”

Many of the senior and distinguished academics and practitioners whose views I tapped for this research reflected on the degree to which the term evidence was undefined. They perceived its range of potential meanings as relatively uncharted territory. That is not to imply that they were short of ideas. On the contrary, questions of evidence stimulated many ideas. But people often required a period of reflection and dialogue to explore the various dimensions of evidence. They needed the process of reflection and judgement to arrive at what they thought, rather than tapping in to a ready-made position.

“Tricky one isn't it. I haven't really thought this through before...when you do think about it, you realise it's not a simple issue at all [laughs].

“...it makes me realise that, sort of, it's a much wider topic. I used to think of evidence as the p value basically, and of course it's in everything we do.”

Discussions about the topic of evidence took place within a broader awareness of evidence-based medicine, which grew throughout the time of the study. Thus interactions about evidence included comment, interpretation, support or critique of EBM. The perspectives developed in my interviews were mirrored in local discussions within the Faculty. They were also represented in broader interactions about evidence on the Internet and in the medical and public health literature.

Developments within EBM were often identified as a primary trigger for reflection about evidence, or for changes in perspective on evidence. Many ideas about what evidence was, or wasn't, were formed as a direct result of, or in reaction to, local or global EBM-related initiatives. These included formation of the Cochrane Collaboration, new journals such as “Evidence-Based Medicine”, and EBM conferences (Cochrane Colloquiums or the Scientific Basis for Health Services).

Thus EBM was a catalyst for generic and philosophical reflection on evidence per se. This process was also triggered, magnified and tapped in my interviews. In clinical medicine particularly, new ideas were contrasted to

those held prior to the advent of EBM, sometimes expressed as a “before and after” or “conversion” experience.

Even those who had many years of association with EBM, and were familiar with debates about its meaning, described their ideas as “evolving”.

Therefore a well-formed personal interpretation of EBM did not preclude fluidity and ongoing development in the approach taken to the broader notion of evidence.

The study was conducted at a time when discussions about evidence in public forums such as the medical literature, Internet and conferences were growing in frequency and intensity. Debates were also taking place between people from various disciplinary and professional perspectives. For example clinical medicine, epidemiology, basic science, social science, public health and health promotion, health service management and medical education. Each of these interactions with alternative points of view, personal reflections, and the interview experience, teased out new dimensions of evidence and raised new challenges.

“Every time I talk about some of these things, you know what it's like, you move forward a little bit and the veil starts to lift and you can see a little bit further, but it's still hazy out there...”

“...in the back of my mind a whole lot of questions have been raised by just this discussion that I will continue to pursue about the nature of evidence...”

5.2 4 Shifting sands: developments in the evidence-based movement

Chapter 2 outlined the antecedents and growth of evidence-based medicine. This social phenomenon continued to develop and change throughout the time of this study. Just prior to and during the time of this research, broad debates about EBM moved back and forth through several stages. I have summarised these developments as the “shifting sands” of the evidence-based movement. They include the proposition and heralding of a new approach to evidence in clinical practice; controversy and critique of this approach; and a re-definition and modification of arguments about EBM. This was followed by a moderation in the reaction to EBM; a greater focus on evidence about the better application of evidence; implementing structural and organisational changes in the health system; and a wider adoption of the principles of EBM across all of health care.

In all the data sources examined for this research, EBM was primarily presented as a more quantified and probabilistic way of thinking about clinical practice. The term “evidence-based” was meant to distinguish the quantified approach from the more “qualitative” and “traditional” one. Traditional practice referred to treatment outcomes that were predicted from knowledge

of basic science such as patho-physiology, combined with the evidence of clinical expertise. This approach to evidence relied on the ability of individual doctors to translate the experience of their previous patients into predictions for new patients. EBM was presented as placing more emphasis on probabilities obtained from systematic clinical research, with clinical expertise used to interpret and apply this evidence to individual patients.

A majority of the leading EBM proponents had a medical background. Yet it seemed that they did not anticipate the strength of controversy and emotional backlash (particularly from clinicians) that followed introduction of this phrase. Critics particularly retaliated from the perceived implication that medicine had only just discovered science, and thus evidence. They asked: what had medicine been before, if not evidence-based?

“The presumption is made that the practice of medicine was previously based on a direct communication with God or by tossing a coin.” (Fowler, 1995)

There was also antagonism to the primary focus on quantification and measurement. This was seen to reduce medical care to simplistic algorithms without taking heed of the varying needs of individuals and their social circumstances.

“...they only accepted as ‘evidence’ that which could be summarised on one graph...complexity was devalued and context denied...” (Carr-Hill, 1995)

The critique of EBM also focussed on the way the message was presented, and the style and personality of its leaders or *“evidence-based medicine’s elite” (Lancet, 1995a).*

“The difficulty with evidence-based medicine lies with its exponents – their arrogance, their jargon, and their penchant for denigrating others” (Morgan, 1995)

“...will the enthusiasts of evidence-based medicine do us a favour and nominate an advocate who is more of a scientist and less of a prophet?” (Shahar, 1995)

In the interviews, local supporters of evidence-based medicine said they found the phrase more a hindrance than help when promoting critical review of established practice and its evidential base.

“...I think the name [EBM] just gets up peoples' noses. It has the connotation that we know what evidence is and you don't...”

Evidence became “a dirty word”, which was being perceived by antagonistic colleagues as having “sullied” the clarity of clinical practice. The issue of

evidence became “toxic” for some in the Faculty as they developed “EBM-antibodies”. Those involved in curriculum development in the new Graduate Medical Program juggled promoting the new dimension of medical education, and fighting perceptions of “EBM-creep” among the critics, who saw it as infiltrating and taking up disproportionate time with students.

A commonly cited explanation of what it means to be evidence-based was published in a BMJ editorial in January 1996. It emerged from a posting on the international evidence-based health Email list at the end of 1995, requesting feedback and assistance in formulating an “updated definition of EBM (for one-on-one clinical practitioners)”. The aim was also to address the “misconceptions” that had emerged in continuing discussions about what it was and its implications.

Replies focused on a need explicitly to highlight the role of the clinician and their expertise. To focus on the clinician’s judgement in the process of applying epidemiological research evidence to individual patients. The resulting definition was regularly cited at local teaching sessions, international conferences and in future publications.

“Evidence-based medicine is the conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients. The practice of evidence-based medicine means integrating

individual clinical expertise with the best available external clinical evidence from systematic research.” (Sackett et al, 1996)

Two notable developments unfolded over the next three years; some lessening of angst and controversy about what is meant by evidence in EBM (at least within a clinical setting), and emergence of the term as a primary currency in health care rhetoric. Calls for explicit and quantifiable evidence on diagnostic tests, aetiology, prognosis and intervention effectiveness shifted from radical fringe to new orthodoxy. It “...acquired the kind of sanctity often accorded to motherhood, home and the flag” (Feinstein, 1997).

Recognition and use of the “evidence-based” prefix became widespread as it was picked up in the health literature, health service and health policy debate and in the media.

“To me evidence-based medicine is one of the most important movements in the practice of medicine in recent times...” (Speech by Michael Wooldridge, Federal Health Minister, 1996)

“...explicit consideration of formal ‘research evidence’ is now expected in the decision making of politicians, bureaucrats, clinicians and industry.”
(Lomas, 1997b)

The move into mainstream thinking resulted in additional significance given to the interpretation and application of the principles of EBM.

“Evidence-based medicine is an elegant idea whose time has come. ...But it may be, to quote that Yeats poem again, that a terrible beauty is being born”.
(Financial Times, UK, July 1998)

Thus reactions to EBM as a central and potentially dominant framework for health care varied from positive to cautionary and sometimes critical. This depended on one’s starting point in its interpretation and the perceived underlying agenda.

“...authoritarian clinicians may see evidence based medicine as a threat. ...At times it will alter the dynamics of the team, removing hierarchical distinctions that are based in seniority.” (Rosenberg and Donald, 1995).

“They see your beloved evidence based medicine as a means to shackle the doctors and bend them to their will. ...Beware, Enthusiasticus, that you are not used as a dupe in a political game of health economics”.
(Grahame-Smith, 1995)

The promises of EBM and concerns about its implications set up a dichotomy in many people’s minds. This dichotomy included the advantages of quantifying the benefits and harms of interventions and explicit evidence-

based decision making. It also included however, the disadvantages from a potential misappropriation of the rhetoric and zeal about evidence and measurement. Thus EBM was seen as either a step towards democracy, liberated junior practitioners and better informed patients, or a step to limit clinical autonomy, increase profits for the pharmaceutical industry and gain commercial control by managed care organisations. The concerns are considered further in section 5.6

5.2 5 Reflections

Dynamic and evolving evidence

In section 3.6, I described my *a priori* perspective on this research. This included the anticipation that my study would identify and describe established positions on evidence among researchers and practitioners in medicine and public health. The dynamic and rapidly evolving nature of participants' concepts of evidence was unexpected and at first surprising. It emerged, however, that the rapid rise of EBM just prior to and during the time of the study provided a previously unprecedented impetus for researchers and practitioners to debate and to negotiate the meaning and nature of evidence.

As described in chapter 2, traditional epistemological debates in medical or public health literature had often focused on comparison between scientific

and 'alternative' research paradigms, or the relative merits of quantitative versus qualitative research. Such debates often occurred between people trained in different disciplines (i.e. science versus anthropology) or between post-positivist and constructivist partisans within the social sciences (Baum, 1995). The EBM challenge was notable in that it emerged from within the scientific ranks of medicine and public health. It was also perceived to have significant implications for the allocation of research and health service resources.

Prior to evidence-based health care, mainstream approaches to substantiation had not been significantly challenged. Thus researchers and practitioners had not felt a pressing need to defend their individual perspectives on substantiation. It became clear that contemporary debates about evidence, EBM, and evidence-based health care took many people beyond their established research and practice territories. For example, those accustomed to an orientation towards primary research participated in discussions on research findings as evidence for clinical or public health practice.

Concurrently, practitioners reported a new awareness of research design, the rigours of research method, and the notion of bias. Debates about evidence and EBM also led researchers and practitioners into broader discussions about the evidentiary basis of health care management and health policy. Hence EBM, and the resulting unprecedented level of interaction between people

with different perspectives on substantiation, resulted in people thinking and debating beyond their usual domains of evidence.

The open format of the study interviews also contributed to the development of people's ideas. Several study participants commented that only rarely had they had the time for the type of free association and unstructured thinking that was provided by the interviews. Their professional time was instead fully spent on solving technical or practical problems, clinical and / or teaching loads, administrative tasks, publishing demands and deadlines for grant submissions. The interviews appealed especially to those for whom they were a window of opportunity to engage in unpressured creativity, a process of discovery.

Group discussions about evidence were also described as creative and enjoyable when they allowed time for unstructured debate about the meaning of the term evidence, challenged participants' established ideas, and provided opportunity for interaction between people with different perspectives. Examples of such meetings were EBM staff development seminars or evidence-based project meetings, such as the National Public Health Partnership's project on developing criteria for evaluating evidence on public health interventions (Rychetnik and Frommer, 2000). It seems that the intellectual friction created by alternative interpretations of evidence stimulated personal reflection that was perceived as rewarding by many

researchers and practitioners. In addition, it was acknowledged that the potential for evidence to determine resource allocation decisions added significance to reflections on what “counts” as evidence (a topic discussed in section 5.6).

At the end of 1999, the dynamic and evolving nature of the concepts of evidence within local and international forums had not subsided. For example, three interview participants requested another look at their transcripts a year after their second interview “to see how my ideas have changed”. This degree of engagement with the topic, and ongoing reflection, was matched by a continued emphasis on evidence in medical and public health literature and at health related conferences.

The third international conference on the Scientific Basis for Health Services was held in Toronto in November 1999. A central theme in the presentations, and the discussions that followed, was the increasingly eclectic interpretation of the term evidence. In the early to mid 1990s, conferences about evidence-based practice had primarily presented evidence as defined by clinical epidemiologists and statisticians. By 1999 the international meetings included presentations and challenges from allied health and nursing, experts in management and behavioural science, and from politicians and bureaucrats who discussed health policy and the allocation of health service resources.

In summary, the interactions examined in this study reflected an intensive, and for many a relatively new, focus on the epistemology and meaning of evidence. They also reflected an uncommon degree of interaction between diverse disciplines and professions. For example, clinical medicine borrowed from social science to identify effective management and behaviour change strategies (Bero et al, 1998; NHS Center for Reviews and Dissemination, 1999). Allied health, public health and health promotion considered the implications of EBM defining their evidence (Ziglio, 1997). The main concepts of evidence discussed at international forums were also discussed and examined locally with the participants of this study, and are outlined in sections 5.3 –5.6.

Shifting sands in EBM

In 1991, Mechanic described the profession of medicine as moving from a position of dominance to one of “countervailing power”: “a historical dynamic that begins with one party accumulating such power that it prompts other parties to muster their forces and attempt to control the first” (Mechanic, 1991, p 500). Between 1992 and the end of 1999, EBM developed from a probabilistic approach to clinical medicine into a broader and more overarching philosophy about evidence-based health care and evidence-based health policy.

Although evidence-based decision making has been advocated by clinical epidemiologists and some prominent clinicians, it has also been interpreted as a tool of health bureaucracies and managed care interests (Short, 2000; McDonald and Daly, 2000). This is partly due to the social and political context from which EBM emerged and which has reinforced its development. These social determinants included fund holder demands for efficiency and cost containment; concerns over practice variations; and the rise of consumerism with calls for medical and government accountability (Fox, 1990; Melhado, 1998; Saltman and Figueras, 1998). Hence EBM has been described as a social movement (Davis, 1996) and has been considered in the context of social change that has impacted on all aspects of medicine (McDonald, 1998).

Sociological explanations for EBM were also considered by a few of the participants in this study. Such perspectives, however, were balanced with often more immediate considerations of the epistemology of scientific research, and the pragmatics of using research findings and other forms of evidence in medical or public health decisions. Thus evidence was primarily defined in relation to participants' direct spheres of work and influence. Their orientation is reflected in the findings presented in this chapter.

5.3 THE QUEST FOR “ALWAYS LATENT” TRUTH

5.3 1 Introduction

The interactions on evidence examined in this study reflected several commonly shared perspectives on the natural world, and on medical and public health research and practice. The views were relatively constant and can be described as the key values and principles that underpinned judgements on evidence. These shared values and principles are presented here under the umbrella of a *quest* for “*always latent*” truth. This quest was an ideal to which one could aspire and a principle against which the concept of evidence was then described and appraised.

5.3 2 Summary of my grounded theory on evidence

Although people’s concepts of evidence were dynamic and evolving, they were founded on several constants. These constants were the values and principles that underpinned perspectives on evidence, and which were widely shared among those trained and working in medical and public health research and practice. They included a primarily realist view of the natural world. There also existed an ideal of a quest for the “always latent” truth; that is, seeking to describe and explain the elusive reality.

The quest for latent truth had to comply with rigorous, primarily scientific standards. All claims had to be substantiated. Significant value was given to the science and art of critique, and the ability to judge the quality of that substantiation. Where successful, the quest resulted in a reduction of uncertainty about the natural world. Concepts of evidence were underpinned by the value given to the management of uncertainty in medical and public health research and practice.

In an applied context however, the quest for truth and the need for substantiation were balanced against other prevailing agendas. This resulted in the notion of pragmatic compromise in relation to evidence.

Details of these concepts are presented in the sections below.

5.3.3 Realism, substantiation and importance of criteria and standards

“...we are constructing meaning, but we are constructing meaning around something that does actually exist, otherwise we are all wasting our time”

“It’s part of the scientific religion which would make us believe that that’s the way in which we tap into a true underlying reality which we can’t directly see.”

The interactions examined revealed that a widely shared working perspective underpinning ideas about evidence was a realist view of the natural world. This included a view that the world exists independently of its observers and has stable relationships that can be sought, described, explained and predicted.

For these mostly scientifically trained academics, to investigate and portray this reality was highly valued and formalised as research. Research perceived as rigorous, believable and able to answer important questions about the natural world was clearly prized. Hence in this social and cultural context, peer reviewed publications based on empirical research indicated success in the quest for truth and the generation of knowledge. This success was seen to be rewarded with esteem and funding.

All my data sources reflected an almost universal respect for science as a mode of inquiry, particularly for the contribution of scientific methods to generating accurate and reliable estimates of the natural, physical and biological worlds. For those who defined the concept of evidence as any peer-reviewed, scientific research, the fuss over EBM was bemusing. For example, if “evidence-based” meant “substantiated by science” then medicine had been based on science for centuries.

Thus the notion of substantiation was sometimes used synonymously with the notion of claims being “evidence-based”. Some researchers however,

preferred to use terms like “data” or “results” when discussing their findings. These individuals perceived experimental data to be more neutral than their concept of evidence, which they described as more value-laden than scientific research. This is described in section 5.4.

The widely adopted realism, combined with a scientific mode of inquiry directed at physical and concrete phenomena, allowed for the possibility of distinctions to be made between being *closer* or *further* from the latent *truth*. Also the respect for science included a focus on research method as the key criterion for reliable and valid substantiation.

“...what we do when we're doing scientific research is trying to find some way of tapping into or getting as close as possible to what that truth is. And because we will never really know what it is, we have to design a set of guidelines or rules which on the basis of our experience to date makes us believe that that's going to be the best way of getting close...”

Thus there were *right* and *wrong* ways of doing things and people often had definitive views on the way to determine a best estimate of reality. Data (research results) were judged by the method, or process, by which they were generated. They were also judged by their source, where credibility was determined by whether the person or group could be trusted to apply appropriate methods and to implement them correctly and honestly.

The importance given to *standards* and *criteria* revealed that this was not a relative or radically post-modern world where anything goes. The various disciplines had established standards for research quality. Criteria were used to determine the adequacy of research that substantiated professional knowledge or the information that supported decisions about practice. For example, there were standards requiring clearly defined and controlled conditions in a laboratory experiment, and epidemiology required an appropriate study design and validated measurement tools.

The principle of invoking standards was inherent to evidence-based medicine, which identified explicit critical appraisal criteria to determine the reliability and truth-value of clinical research. The EBM “rules of evidence” were derived from scientific criteria, such as minimising error and bias through research design and method execution. An important emphasis was placed on the skills required to apply these criteria in order to scrutinise evidence and to identify the potential for bias.

The adoption of criteria extended beyond the scientific domain to include standards of rigour for qualitative research, such as its amenability to audit and adopting steps to obtain participant feedback and check the researcher’s interpretations. In both clinical and community settings some participants regarded hermeneutic and qualitative forms of inquiry very highly for

providing different, but important forms of knowledge. During the time of this study, critical appraisal guides for evaluating qualitative research were published in the medical literature. A comprehensive guide was also developed locally by the scientific committee at Westmead Hospital (part of the Faculty of Medicine) in order to assist their scientifically trained members to assess (and value) qualitative research proposals.

These criteria for appraising qualitative research were influenced by the critical appraisal approaches adopted in EBM, and mostly focused on assessing the quality and rigour of research methods. Some interpreted the appraisal of qualitative research as determining its truth-value, while others focused on how adequately the process was able to capture and portray the perspective and meaning of the subject. Qualitative research was often discussed in the context of debates about what are appropriate inclusion or exclusion criteria for valid evidence in order to practice evidence-based medicine and evidence-based public health. This topic is addressed in section 5.6.

5.3 4 The science and art of critique

“I think it's the methodology of evaluation ... that is what the role of a University ought to be and I don't think universities these days succeed

particularly well...they are much more highly geared to information transfer, rather than the skills of critique.”

Researchers, and practitioners committed to EBM, often stressed the *science and art of critique* as pivotal to dealing with research data as a means of substantiating claims. The science and art of critique was about applying existing criteria to assess the quality of information and interpreting the findings to formulate sound conclusions. The notion of critique as an essential professional skill was identified across all disciplines within basic science, clinical medicine and public health, although the standards and criteria adopted would vary.

“I think it is very, very dangerous when people leap to conclusions and take far reaching actions on the basis of fairly flimsy evidence. Often what that means is that there is an emotional or conceptual appeal to it...they might have skipped the bit about gathering the facts and evidence properly, before taking a big leap.”

Participants expressed concern if a critique of research findings, or other forms of information, was not undertaken or was done poorly; i.e. if critique was based on the wrong criteria or inappropriately influenced and biased by undeclared agendas. This importance given to the process of critique was matched with a common perception that the skills or their application were

often inadequate, not only in the community but also among students and professionals.

The notion of critique as a science and an art drew attention to the use of professional *standards*, as well as judgement based on personal *skill* and *experience*. Scientific criteria were emphasised in the critical appraisal literature that underpinned EBM, which focused on critique as a means of identifying the potential for bias in clinical research. Thus standards for research methods were invoked to assess the validity and reliability of research as evidence. The art of critique was about being able to apply the appraisal to a particular context. Experience and judgement determined what degree of substantiation was appropriate or adequate for a given decision. Substantiating knowledge and decision making are discussed further in section 5.5.

5.3 5 Uncertainty and pragmatic compromise

“I think that is the whole essence of scientific inquiry, that uncomfortable but committed suspension of belief, the committed refusal to ever regard anything as so self evident, so established, that it might never be challenged...I mean I may have firmly held beliefs that I'm very loath to relinquish and that I'll try very hard to hold onto, but I'd have to say that in my very depths I hold to the

commitment to always suspending judgement and never going that tiny final step”

There was wide agreement on the value of obtaining a “firm” basis for disciplinary knowledge and practice decisions. On a philosophical level, however, there was near universal caution about the notion of *proof* and claims of certainty. Hence, the truth about the world is out there, but it remains “always latent”. Research was described as generating estimates of reality, and what appears true about the world change as methods improve or perspectives alter.

Hence the realist view of seeking to discover truths about the natural and physical world was balanced with the notion of *sceptical inquiry*. This encompassed a critical perspective on all forms of information, including scientific research, and our ability to represent that reality. None of the participants engaged in a fundamentally post-modern critique of scientific knowledge. Yet the research process and specifically generating and presenting data as a form of evidence were often described as socially influenced and inherently limited.

“I’m not completely persuaded by the arguments of the people who talk about science being a social construct, in a way that leads to its denigration. And that any value system is as good as any other...creation science versus the

theory of evolution... I can't accept that these value systems are of equal weight, I'm sorry, it's all to do with your training and everything isn't it [laugh], but I can't. But on the other hand, you'd be an absolute fool if you didn't believe that science is influenced by a whole range of social factors."

This led to a common view that there was always scope for reassessing even the most convincing evidence, and of prevailing fundamental *uncertainty*.

Uncertainty about the world was a condition that was eased through research by obtaining data to substantiate claims and conclusions. However, a sense of uncertainty was also described as a consequence of the inherent failings of all research and thus mankind's capacity to reflect the latent truth. This was not a nihilistic post-modern critique of all knowledge, but a rather more pragmatic scepticism about proof and certainty.

"...they say you're always sitting on the fence, you can never make a decision, you always say 'well, on balance' or 'is consistent with' or these sorts of weasel words, why don't you come out and say what you think."

The fallibility of knowledge and thus fundamental uncertainty about the world appeared to have two components. These were the limitations in methods, technologies and theoretical understanding that were always improving; and the inevitable influences of personal and social biases, which could be minimised but not eliminated.

On a more pragmatic level, these epistemological positions on the nature of inquiry and truth-value of all evidence were weighed up against the need for decisions and action. In practice, people said they worked with the best estimates available and according to the needs of their environment.

“...if you are an honest academic, then your whole life is fraught with ‘I don't know anything’...easy to do that if you are sitting in the University somewhere, but it's harder to do that when...you've got to make a decision here about this patient...”

This resulted in a commonly drawn distinction between academia with the luxury of time and self-doubt, and the demands of the practice setting. Clinicians particularly discussed the need for prompt (sometimes immediate) action. Their views on substantiation were influenced by the requirement for daily practice decisions, despite prevailing situations of uncertainty. These circumstances required some pragmatism and often resulted in what was described as a form of *compromise* in relation to substantiation or evidence. The notion of a pragmatic compromise was about weighing the needs of the context, against the ideal encompassed in the quest for truth.

Hence discussions about evidence were often underpinned by an ongoing interplay between the importance of substantiation and the inevitable presence

of uncertainty. Assessment of what was *adequate substantiation*, and what was an *acceptable compromise* of the importance given to substantiation was highly context dependent. Whether primacy was given to seeking latent truth or “getting things done” depended on the professional context. Many researchers and practitioners straddled several professional roles and modified their approach to, and application of, evidence accordingly. This is discussed in more detail in sections 5.4 and 5.5.

5.3 6 Managing uncertainty

“I seek to control my universe, so in knowing more things I can control my universe better...”

Discussions about the importance of substantiation, critique and pragmatic compromise were underpinned by the value given to *managing uncertainty*. This included recognising uncertainty, assessing and measuring degrees of certainty and if possible, minimising uncertainty.

Several participants distinguished between professional and non-professional dealings with evidence as a means of substantiation. In a situation where one was not required to assume a professional role, it was acceptable not to worry about uncertainty; i.e. to make “leaps of faith” between evidence and conclusions and to make greater compromises. In these situations participants

felt free to adopt any criteria they wished for evaluating evidence or not to think about evidence at all.

“...working in my professional capacity, where lets say the money to pay for that is coming out of a public purse, then the situation is very different.”

The flexibility of situations where one was not required to manage uncertainty was distinguished from the research and practice setting. Here uncertainty had to be managed explicitly and professional standards of evidence were often described as more conservative, rigorous and objective than personal ones. The professional role in relation to evidence is described further in section 5.5

A research or practice setting determined the degree of uncertainty in which one had to form conclusions or make decisions. There were some professional contexts where good evidence was available, the rules and criteria for evidence were established and relationships between hypotheses, evidence and conclusions were clear. Other research or practice settings required management of a significant and ongoing degree of uncertainty.

Areas of research with low uncertainty were described as tightly controlled situations with high levels of confidence in the available data. This was found especially in clearly defined and focused areas of research e.g. laboratory-

based, scientific experiments where it was possible to identify and control all potential variables and categorically make a statement about the nature of the relationships observed.

In clinical medicine and public health research, uncertainty was seen to be more manageable if the questions had a specific or narrow focus, particularly if this was associated with sharply defined variables, quantifiable objectives or outcomes and clearly defined pathways between variables. This type of research question increased the potential to obtain results with explicit (quantifiable) degrees of certainty. Uncertainty increased and was less manageable in areas where the pertinent questions had not yet been explored or defined, where the questions were inherently broad and open, or involved multiple and complex relationships between variables.

Some areas of professional practice required comfort with a significant lack of research as a basis for knowledge, or highly tenuous or speculative relationships between that research and decisions about appropriate action. This applied especially to new clinical specialties, general practice and in relation to many public health or social interventions. In these settings, decisions often involved a consideration of multiple styles of research, as well as forming judgements based on speculation rather than research findings.

Personal comfort with working in situations of high uncertainty was sometimes linked to career choices, such as those between different forms of basic research or specialised and general clinical practice. Hence some people (at least partly) managed uncertainty by self-selecting into areas of research and practice that suited their uncertainty and risk comfort zone.

“If you are not comfortable with uncertainty you shouldn't be in general practice, because general practice is based on uncertainty.”

“I don't like working with low support or systems where results have to have statistics applied to them to make sense. I find it makes me worry when you have to apply probability to something.”

5.3 7 Uncertainty and EBM

“there is no such thing as certainty, there is only $p = 0.05$ ”.

For many of its supporters, EBM was a functional way of managing the inevitable uncertainty of health care. It involved examining and making explicit the relationship between health care decisions and the available evidence. The probabilistic approach and the rules of evidence provided a means of assessing, grading and applying research to applied contexts.

EBM was also perceived to encompass a trend towards an explicit acknowledgment of the unknown. This involved open discussions among colleagues, with students and with patients, about the uncertain basis of many clinical decisions. For some practitioners this relieved the burden of portraying inflated confidence, which had always felt rather tenuous and dishonest.

Conversely many controversies about EBM were attributed to the way it had opened a Pandora's Box of uncertainty. The need to contain this uncertainty was posited as the basis for some of the negative emotional response to EBM.

"...it's hard to maintain a consistently sceptical view of the world...they reflect human anxiety in the sense of being overwhelmed. I think that is the biggest problem that EBM faces."

Thus while for some EBM managed uncertainty, for others EBM increased uncertainty and made it appear less manageable. Critical appraisal was a means of making the unknown explicit, but it was also described as an unmanageable deconstruction of the evidential base of existing practice.

"...dirtying the practice, literally dirtying my clinical practice. I thought I had everything in place or that's how I perceive it...you're now making

everything too difficult...we're telling people to practice based on this evidence and yet then we pull it apart and completely undermine it."

EBM challenged traditional authority and the substantiation of practice based on professional hierarchies that were derived from years of practical experience. The potential threat to previously accepted standards of evidence were seen to undermine the basis of professional practice "*...the water has been muddied...things we were sure of we aren't sure of any more...*".

Senior clinicians, who had been socialised to "*sound like a consultant, not a registrar...to speak with conviction*", felt they had to reverse gear and change their language that had legitimated evidence as a derivative of expertise. Their expressed confidence had not only been for the benefit of their patients and colleagues, but also for themselves. Speaking and acting with conviction was an expected way of functioning and managing the stress of working within the uncertainty of clinical practice. Even practitioners who supported EBM described the process of acknowledging what they don't know as "*disempowering...it's a frightening thing and it's easier to be omnipotent and God-like, because it reassures yourself.*"

Hence the uncertainty described as the result of EBM was due to the deconstruction of previously valued evidence, the requirement for unfamiliar

skills, and the switch of authority from senior colleagues to a new external authority in the form of systematic reviews.

Uncertainty also stemmed from concerns about the potential impact of applying an unfamiliar approach to evidence. EBM represented new rules and criteria for weighing which evidence should influence decisions about patient care. Yet there was no assurance (evidence) that this would lead to an overall improvement in care for patients.² The goal posts of substantiation were seen as having been moved, and thus could be moved again, without assessment of the ultimate impact.

5.3 8 Reflections

Principles and values reflected in the quest for latent truth

The notion of a *quest for latent truth* encompasses both an epistemological position and a statement about professional values. These formed the cultural context in which concepts of evidence were examined and defined (definitions and appraisals of evidence are described in section 5.4).

² The argument that the introduction of EBM was itself not evidence-based was often raised in debates about the uncertainty of EBM i.e. that the practice of EBM was being promoted and implemented before its benefits and harms had been empirically assessed.

The epistemological perspective reflected in the quest for latent truth is one that has been defined by others as “entity” realism. This identifies science as a fallible attempt to describe entities (and their relationships) that have a reality that is independent of their investigation (Harre, 1986; Hacking, 1983). This combination of realism and fallibility of inquiry is also incorporated into the post-positivist world-view; that is, no conclusion is beyond the reach of criticism or modification (Medawar, 1984). Hence, the study participants’ articulated belief that the product of research is a “best-estimate” of reality reflects a post-positivist epistemology.

Epistemological perspectives that address the influence of social factors on the generation, presentation and interpretation of scientific findings have also been named “post-Kuhnian” philosophies of science. From this viewpoint “science is both a useful method of creating knowledge” and “one in which the influence of value and worldview of those within it is continually present. It is objective...only to the extent that it is an open and negotiated process in which bias can be detected...” (Bayley, 1995).

When researchers and practitioners conceptualised evidence, they formulated a stance on its relationship to the external reality in which they conducted research and practice. Their reflections on evidence addressed pertinent questions such as; what is the link between evidence and reality? Is it representative of a “latent truth” or someone’s perspective? Is evidence

objective or does it reflect a particular agenda? Answers to these questions depended on whether their concept of evidence encompassed a *reduction* or an *enhancement* of the impact of bias and social influence. These different approaches are described in section 5.4.

The findings described in section 5.3 also embrace a set of principles or values. These can be described as the cultural or professional values that underpinned participants' concepts of evidence. They included the merit given to describing and understanding reality, minimising bias, and adherence to strict disciplinary standards of inquiry and substantiation. It is in relation to these values that decisions made in situations of uncertainty, although inevitable, were regarded as a form of compromise with regard to evidence.

The degree to which the values described were given top priority, and acted as drivers of decision making, was somewhat variable. The values did however reflect an underlying commonality that encompassed the traditions of academic scientific inquiry. This has been broadly defined elsewhere as respect for "reason, evidence and logic" (Sokal, 1997). It is to these principles that proponents of EBM regularly appealed when they challenged a reliance on expertise at the expense of empirical scientific evidence: "...new ideas and new investigators are thwarted by experts, and progress toward the truth is slowed" (Sackett, 2000).

The values described above reveal areas of potential conflict (about evidence) with people whose ideas do not share the same foundations. I had noted the potential for such a disagreement following a comment by a scholar who was visiting our Faculty and attended a seminar on the topic of evidence. At a meeting he remarked that what he really valued in qualitative research, as a form of evidence, was the degree to which it told an interesting story. He added that he wasn't especially interested in the methods that generated that story. At the time I was struck by how this statement sat in juxtaposition to the values reflected in my data. The comparison brought into focus quite distinct approaches to evaluating research findings as evidence.

Evaluations of research have been classified according to the theoretical orientation or "paradigm" that is adopted by the assessor (Denzin and Lincoln, 1994). For example, a post-Positivist paradigm has criteria of internal and external validity, while a Constructivist paradigm gives value to trustworthiness, credibility, transferability and confirmability of the findings. A Feminist paradigm gives priority to lived experience, dialogue, caring, reflexivity, emotion, praxis and concrete grounding, and Cultural Studies focus on cultural practices, praxis, social texts and multiple subjectivities (p109). Of the four paradigms described the post-Positivist gives most value to limiting bias and uncertainty in evidence, an emphasis derived from the relative degree to which it also incorporates a notion of "external truth" about which one could be "more" or "less" certain. The "quest for latent truth"

identified in this study was significant in that it was discussed and valued (although participants' gave it different degrees of emphasis in their concepts of evidence). It also revealed a point of likely dispute about evidence with those who challenged the encompassing principles described in section 5.3, or did not give them value.

Uncertainty

The predicament of uncertainty has been linked to the concept of evidence in public health and medicine by other authors. They have, however, approached the topic from varied worldviews and have thus presented a diverse analysis. Lupton (1998) described uncertainty as an inherent condition of the post-modern era. This has led to distrust of traditional authorities and a questioning of “the medical, scientific, epidemiological and social scientific knowledges (that) are routinely employed as unchallengeable truths”. Alternatively the editors of a leading medical journal proposed that the inevitable uncertainty of clinical practice was a source of doctors' ongoing power because the community will always rely on them to sort, understand and lead through the maze of evidence (Lancet, 1995b).

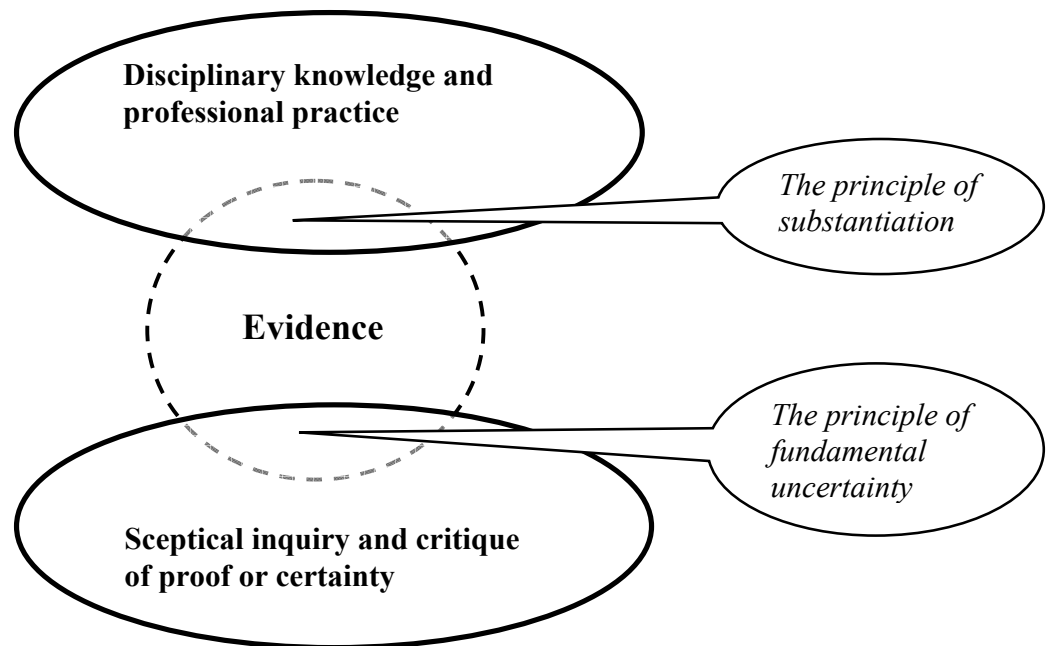
Bauman et al (1991) identified a tendency for practitioners to have certainty about individual decisions (micro-certainty), but contrasted this to the existing wide variations in practice (macro-uncertainty). They also found that

expertise was associated with declarations of confidence about one's decisions. My findings indicate that in the context of clinical practice, expertise as the basis for individual certainty was sometimes undermined by the advent of EBM. The data also revealed a perception that EBM could potentially be used to limit practice variations. For example, if evidence-based guidelines replaced individual doctors' interpretation of the "maze of evidence", and were used to restrict doctors' freedom to determine their practice, this would impact on a source of doctors' authority and power. The controversy surrounding the emergence of EBM must be interpreted in the light of such concerns.

The role of judgement

The principles and values described in this section (such as that of managing uncertainty and pragmatic compromise) point towards a concept of evidence that required a balance between potentially divergent inclinations. For example, Figure 6 (below) illustrates how discussions about evidence encompassed the principle of substantiating professional knowledge and practice, as well as the principle of fundamental uncertainty (based on scepticism about attaining proof or certainty).

Figure 6: The concept of evidence as a bridge between the principles of substantiation and fundamental uncertainty



The other, potentially competing, components of the concept of evidence (such as substantiation versus taking action or disciplinary standards of evidence versus contextual flexibility) also encouraged me to identify *judgement* as the core process in this study. Participants described judgements that encompassed criterion-based notions such as “valid” and “reliable” evidence. They also described explicitly subjective judgements, such as “appropriate” substantiation and an “acceptable” degree of compromise.

The judgements made were influenced by the principles and values described in this section (5.3) of the results. Different forms of the core process of judgement are described and explored in the remaining sections (5.4 - 5.6).

5.4 DEFINING EVIDENCE: DESCRIPTION AND APPRAISAL

5.4 1 Introduction

Section 5.2 outlined how approaches to evidence were evolving formulations rather than established professional norms. Interview participants often required a period of reflection and dialogue to tease out the scope and key domains of the topic. The process of exploring and defining the term evidence usually involved an element of description and an element of critique or appraisal. For example, evidence was described as scientifically derived results that met particular criteria for reliability or truthfulness.

The shared principles outlined in section 5.3 as the “quest for latent truth” underlay diversity in how evidence was described and appraised. This diversity resulted in different perspectives on the relationship between the concept of evidence and the underlying reality or truth about the world.

5.4 2 Summary of my grounded theory on evidence

The topic of evidence was explored and defined through a combined process of describing and appraising (making judgements about) its qualities and properties.

Evidence was described from three key “points of entry”. The first was to describe evidence as an *observation* or *measure* of reality. These descriptions focused on the entity of evidence, such as data or research findings. The second point of entry was to define evidence by describing its *functional* status and the role that it played. These descriptions focused on the context and the reasoning process in which evidence was used, such as scientific hypothesis testing. The third point of entry was to describe evidence as a crafted and *constructed* product: either an entity or a concept. Descriptions of evidence as a constructed entity focused on the interpretation that was inherent in its generation, and the potential for bias. Descriptions of evidence as a mental construct focused on the values and agendas underpinning its construction.

The three points of entry to describing evidence were accompanied by three modes of *appraisal*. I have presented these appraisals as dimensions along which the topic of evidence was evaluated and critiqued. In the *benchmarked* dimension, evidence was judged against criteria and standards. These appraisals primarily focused on the inherent quality of evidence as an observation or measure, where judgement determined whether it was trustworthy, credible or convincing. In the *applied* dimension, evidence was appraised in the context of its application to a particular purpose or function. The appraisals focused on meeting objectives and the relevance and utility of the evidence. These judgements were thus highly context dependent.

Evidence was also appraised along a *social* dimension. Here the social impact of a piece (or body) of evidence was examined. Alternatively, judgements were made about the prevalent or dominant social constructions of evidence and their implications.

Hence people's concepts of evidence (as reflected in the descriptions and appraisals) varied according to whether they directed attention to the latent truth, to process and action, or the "human factor". During the interviews and over the study period, people's concepts of evidence became increasingly multi-faceted and multi-dimensional.

Details of these concepts are presented in the sections below.

5.4 3 Three points of entry to describing evidence

Prima facie approaches to exploring and defining the term evidence appeared to be diverse. After distinguishing the elements of description and appraisal, I was able to identify in the different perspectives some common elements. I combined the different approaches to describing evidence into three key groups. The groups can be seen as different *points of entry* that were adopted to enter a broad and complex topic. A metaphor could be of three alternative gates by which to enter and explore a large walled garden. Details of each point of entry are given below.

Importantly, these descriptions of evidence were not always mutually exclusive. The metaphor of three gates through a garden wall meant that although a person entered via one, it occurred that in their subsequent exploration of the *garden of evidence* they described similar properties of evidence to those who entered via another gate. Hence, where there was time for extended reflection, such as in interviews or at conference debates or workshops, it was quite common to start by giving emphasis to one description of evidence, but then to allude to, or even address in detail, one or both of the others.

5.4 4 Describing evidence as an observation or measure of reality

“concrete information...something that can be measured, or weighed...I see evidence as something that you can either see or hear or feel or read, that describes what happened...”

A common way to approach the topic of evidence was to describe it as an observation or a measure of reality. This focus on observations or measures often led to descriptions of evidence as an entity, such as data or research findings. They also tended to be about data on tangible phenomena in the natural or physical world, such as the cause of ill health or the effects of an intervention.

Several interviewees highlighted their enjoyment of working with phenomena on which they could collect observable or “concrete” evidence i.e. “I’ve always enjoyed the visual science that I can actually see....” In these descriptions of evidence as an observation or measure, evidence was also often described as an estimate of the “latent truth” about the natural world. Specifically, some people described the concept of evidence as closer to the underlying reality than other types of information. This was about conceiving of evidence as “hard data” or “the facts”.

Describing evidence as an observation or measure was not solely related to tangible phenomena however. Sometimes evidence was also defined in terms of measures that could be made on intangible, psychological or social phenomena.

“...I think most things are measurable. I can't think of an example of something that matters that isn't measurable. Some things are difficult to measure, but they can be measured...anxieties, fear, they can be measured, happiness, satisfaction...”

Thus evidence was also described as the reproducible data that may be collected about intangible, but never-the-less real, or at least perceived, phenomena. Some respondents discussed the limitations of conceiving

evidence only in terms of reproducible measures. This is presented in section 5.4 10, as the social dimension for appraising evidence.

5.4 5 Describing evidence by its role and function

“...evidence is gathered or examined in the light of a hypothesis in order to confirm or disconfirm”

The second way of defining evidence was to describe its role or function. In this group of descriptions, which I have labeled the *functional* point of entry, the term evidence was an operational concept, rather than a passive entity. The notion of evidence required an essential infrastructure of an agenda or an intention. Facts became evidence when they were invoked to achieve something, such as to answer a question or to weigh options for a decision. Hence the identifying properties of evidence were that it was “activated” information that had a “purpose”.

When evidence was described in terms of its function, the descriptions focused on the context in which that evidence was “grounds for action”. There was also a description of the reasoning process in which the evidence had been invoked. This outweighed descriptions of the physical attributes or inherent qualities of the evidence. At the generic level, evidence defined by its functional role could comprise of information that came from any source

and could be applied to any form of inquiry. This led to broad inclusive criteria for what *counted* as evidence.

Other definitions that described the function of evidence were more specific. Here the *type of process* determined the label of evidence. For example, information was described as evidence if it had a function in a “formal” or “structured” process, where it was used to make “important” or “significant” judgements. This could be buying a house, hypothesis testing or diagnostic problem solving. Thus “real evidence” was distinguished from the casual type of information that would be used for a minor, domestic decision.

Some participants carefully identified differences between scientific and legal processes. They described the evidence invoked in a legal process as artfully selected and constructed. The aim was to win the case. The legal process was distinguished from scientific experimental design, where the aim was to eliminate personal bias. One person thought evidence belonged only to the legal process, and said he wouldn’t use the term evidence within scientific discourse.

“...the evidence I view in this legalistic [way], whereas data to me is something which if you like is independent, you've got to make the observation, but it is an observation structured in such a way that it's almost

independent of the observer. I mean anybody could make it. Whereas most [legal] evidence has to be constructed.”

These definitions of evidence as information invoked in a legal process represented an overlap between functional descriptions of evidence, and descriptions of evidence as a constructed product, which was the third point of entry.

5.4 6 Describing evidence as crafted and constructed

“It's value laden in terms of what evidence is collected (and) in the sense of what evidence is not. ...there has to be some sort of selection process and that selection process is value laden.”

The third way of defining evidence was to describe it as crafted and constructed. These descriptions focused on the actual process of construction and the influences of that process on the product. Evidence was described as a product of the perspectives, bias and agendas of the people who created the evidence. Hence the distinguishing feature of describing evidence as constructed was to highlight its “selected” and “value-laden” nature.

Some participants focused on evidence as a constructed entity to distinguish the term evidence from “facts” or “data”. As a constructed product, evidence was a composition of “aggregated” and artfully “selected” facts.

“...the same facts could be collected as evidence in different ways by different people and the same evidence of course could be interpreted in different ways by different people.”

Many intended the description of evidence as constructed to connote “fallible” and “unreliable” qualities. Because evidence was constructed, it was less likely to be “neutral”, “objective” or a “reliable” representation of truth than the product of an experiment.

“... data to me is something solid, evidence is an interpretation. ...it's almost a sort of 'when did you stop beating your wife' situation. You have almost pre-judged the situation. The thing is there is no other side of the coin.”

Thus the notion of evidence was contrasted to scientific observations. This was a different perspective to that reflected in the first point of entry, where evidence was described as a measure of reality. From the constructed point of entry, the concept of evidence entailed a higher level of interpretation than observations or measures. Thus evidence was data that had been interpreted and presented in a crafted form.

“...you can argue that the papers we write are no more than stories that might have been talked about by Aborigines around the campfire, thousands of years

ago...that these things that we write... isn't it the same as me going out and finding a place where there is honey, a beehive, and saying that there is a lot of honey there and coming back and singing a song..."

When evidence was described as a constructed product, its properties were attributed to the "human factor". This factor ranged from human error to social influences or direct agendas.

"...you could certainly contest evidence, because it's usually got that human factor in it, which is slightly sullied, the purity of the fact. ...there's a lot of flexibility and rubberiness in evidence"

By discussing evidence in terms of this "human factor", some participants identified it as a source of "bias" that detracted from the potential truth-value of the evidence. Others described the "human factor" as the ideas, motivations and interpretations that were inherent to the production of evidence.

"...it lets you get into a set of values that belong to somebody"

For a few, the definition of evidence as a constructed product extended to describing a social constructivist epistemology. Here evidence was defined as a socially determined mental construct that changed over time and was

determined by context, culture and history. These descriptions highlighted the attributes of the social or professional groups who defined evidence (e.g. their values, priorities and agendas), rather than the attributes of evidence as an entity (e.g. its validity or reliability).

5.4 7 Appraising evidence along three dimensions

Definitions of evidence also included an element of *appraisal*, where a description of evidence was accompanied by an evaluation of that evidence. For example, evidence described as a measure of reality was accompanied by an appraisal of its validity, according to the methods used. Thus the process of appraising the quality of a measure was inherent in its classification as evidence.

I grouped the ways in which evidence was appraised into three categories, and details of each type of appraisal are given below. As with the descriptions of evidence, these appraisals were not necessarily mutually exclusive. During the time of an interview and over the time of the study, there was a tendency towards describing and appraising evidence from multiple perspectives. This led me to reflect that the topic of evidence could be conceived as a multi-dimensional construct. Thus the three ways of appraising evidence can also be presented as three *dimensions* of evidence.

A metaphor for these dimensions of evidence could be the features of the metaphorical “garden of evidence” that may be examined (e.g. types of plant, colours and scents). When time was available for reflection and dialogue, participants often appraised two, and sometimes all three, of the dimensions of evidence.

The relations between the three descriptive points of entry and the three dimensions of appraisal encompassed in the construct of evidence are illustrated in Figure 7 below.

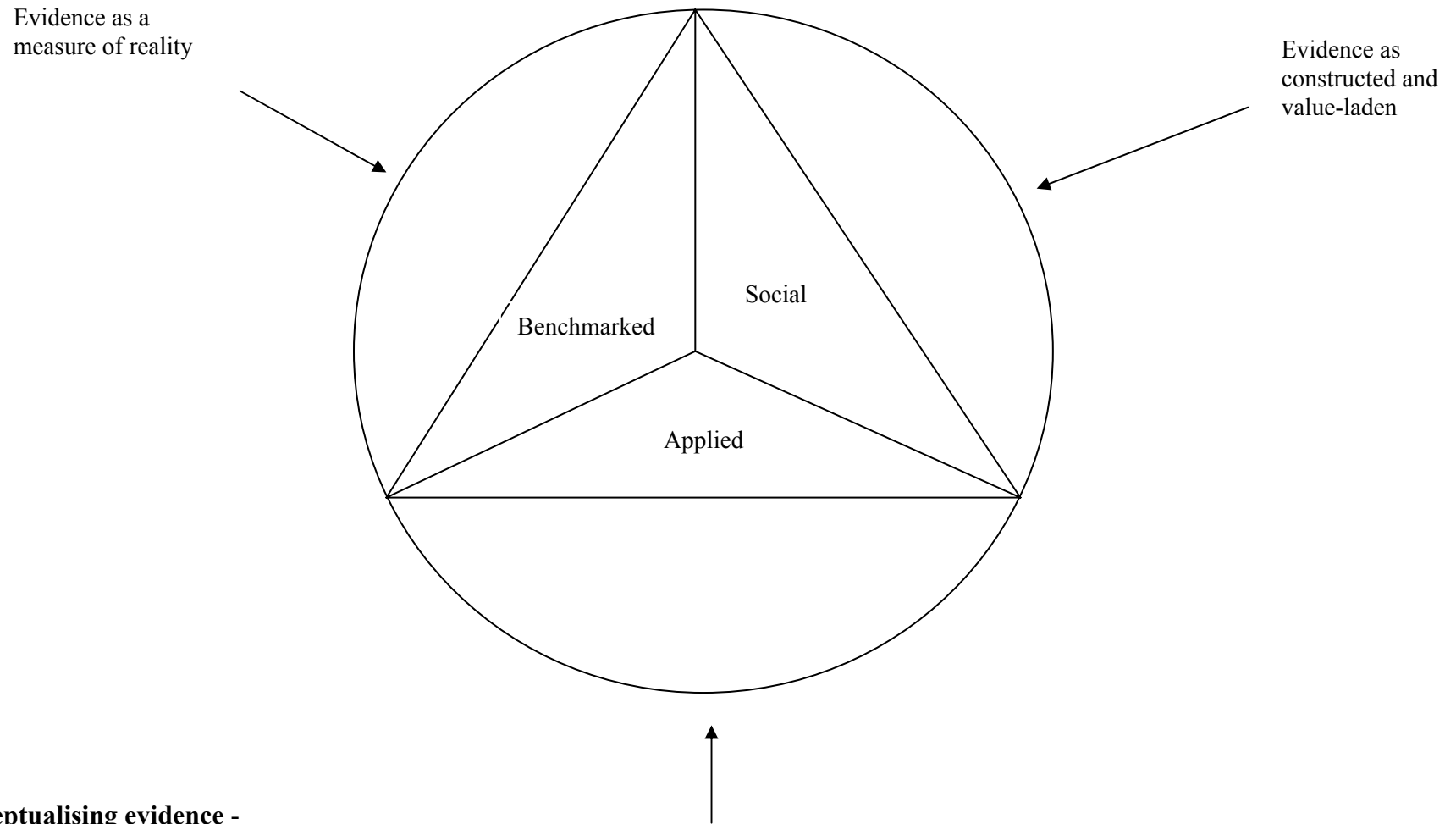


Figure 7: Conceptualising evidence -

Three points of entry to describing evidence and three dimensions of appraisal

Evidence defined by its functional role

5.4 8 The benchmarked dimension of evidence

“...reproducible, scientifically-based, rationally supportable...”

The first way in which evidence was appraised was to judge it against standards or criteria. These were appraisals of evidence in the *benchmarked* dimension. Most of the benchmarks related to the notion of evidence as a measure or estimate of reality. Thus in the benchmarked dimension, participants were appraising evidence as an entity, rather than a social construct.

In some people’s definition of evidence, data were classed as evidence only after they had been appraised and found to meet minimum benchmarks for quality and credibility. Hence appraising an observation or measure against benchmarks was inherent to determining if it “counted” as evidence.

Evidence needed to be convincing or even conclusive to count; it was “...what I believe in, what I believe to be true”. This link between defining and appraising evidence did not imply that evidence was always a reflection of the “whole truth”, but if it was not credible or believable, it was not classified as “real” evidence.

“...my concept of evidence is hard evidence, if we’re talking about hard evidence, my concept is something that is conclusive - the conclusion you can come to is the only one.”

There were many benchmarks associated with the notion of evidence. For example, evidence was “valid”, “objective”, “accurate”, “precise”, “reproducible” and “demonstrably replicated”, “convincing”, “most likely to represent truth”, “concrete”, “measurable”, “quantified”, “scientifically-based”, “rational”, “logical”, “generated by explicit methods”, from a “trustworthy source”, “transferable” (between contexts), “relevant”, “appropriate”, “insightful” or “meaningful”.

The criteria often applied to the method or process that generated that evidence. There were established standards for each particular type of observation or measure, e.g. controlling extraneous variables was a benchmark for laboratory measures, or study design was a benchmark for epidemiological measures. Some participants reported that they read only the methods and results sections of journal publications. These were primarily interested in the “accuracy”, “validity”, “significance” and “magnitude” of the findings, which were assessed against criteria for scientific methods. They preferred to make their own judgements about the value and implications of those findings, and were less interested in an author’s opinions. This led to a dislike of journal formats that relegated study methods to a small font and

inadequate space, and a preference for keeping discussion sections to a minimum.

People also adopted criteria for evidence that related to the professional source of the observations (e.g. whether they were “trust-worthy”). Others included benchmarks related to the potential application of the evidence (e.g. was it “transferable”) which represented some overlap with the applied dimension of appraisal, which is described below.

Occasionally information was appraised against other benchmarks in order to be classified as evidence. These criteria were not related to the scientific credibility of evidence as a measure of reality or its truth-value. Rather the benchmarks were related to a broader social credibility. Evidence had to meet criteria of “formality” or “authority”: qualities that were derived from an official or certified body. Hence evidence was defined by an appraisal of its social “respectability” and “weightiness”. Appraisal against these benchmarks represented some overlap with the social dimension of appraisal described in section 5.4 10.

5.4.9 The applied dimension of evidence

“Context is the most important way to interpret data...”

The second way that evidence was appraised was to consider the context in which it was used and the purpose for which it was used. These were appraisals of evidence in the *applied* dimension. Evidence was appraised in *relation* to its application and how well it met particular objectives.

The applied dimension of appraisal accompanied definitions of evidence as an operational concept, where the concept of evidence was described in terms of its function. Evidence was defined as data that are *invoked* to make a decision about clinical practice. This concept of evidence was appraised according to how relevant and useful it was to that clinical decision. Judgements about evidence in the applied dimension were context dependent, rather than assessments against external criteria or standardised benchmarks. For example, a clinician would consider a systematic review as a piece of evidence if, and when, it was relevant to her patient’s problem and assisted her consideration of treatment options. If the review was not relevant, it was unlikely to be described as evidence when appraised in the applied dimension.

Many participants interviewed in this study reflected on and discussed the way they used data or information as evidence in their professional roles as

researchers and practitioners. Hence their discourse about evidence nearly always included an element of appraisal in the applied dimension. The application of data (or other information) as evidence in different types of judgement is discussed in section 5.5 of this chapter.

5.4 10 The social dimension of evidence

The third way that evidence was appraised was first to consider the social determinants of its construction or interpretation, and then to assess the social implications. This entailed assessing the social impact of a piece or body of evidence when it was interpreted and applied. Other participants appraised the social construction of the term evidence, and the social impact of prevalent or dominant definitions. I judged these appraisals as belonging to the *social* dimension of evidence. Appraisal in this dimension often followed those descriptions of evidence as a crafted or constructed product (as outlined in section 5.4 6).

Appraisals of the impact of a piece of evidence, or a body of evidence, included considering how it was accepted and adopted in professional practice or policy. Participants reflected on whether the evidence was acceptable to health professionals and the wider community. They judged the interpretation of the evidence, and the conclusions that were reached, from a social rather

than a methodological or technical perspective. They also examined the resulting action and the social implications of those actions.

Appraisal of evidence along the social dimension also scrutinised the values inherent within particular types of evidence. This included reflections on the influence of EBM, and how EBM changed professional definitions of evidence (and its appraisal), and the implications of those changes. For example, accurate quantification of the effects of an intervention was identified as a priority in the evidence for EBM. This was contrasted to more traditional approaches to substantiation in clinical practice. The value given to quantification and eliminating bias was also compared to interpretive or qualitative approaches to gathering data and generating evidence.

Hence, appraisal in the social dimension of evidence considered the strengths and weaknesses of various constructions of the notion of evidence. Several participants commented on the limitations of conceiving evidence only in terms of reproducible measures. This was linked to a perceived risk of undervaluing observations that were difficult to quantify or replicate, and thus disregarding potentially important information, such as patients' emotional experiences of therapy. For example, a weakness of EBM was presented as the "...incommensurability between the linguistic value laden appreciation of what happens during illness and the desire to compute..." which may result in patients' feelings being understudied and thus overlooked. The notion that

ideas or phenomena could be undervalued because they were not measurable, and thus would not count as “real” evidence, is discussed in more detail in section 5.6.

Appraisals in the social dimension also identified the explicit or perceived agendas underlying professional and political discourse about evidence. This related especially to EBM, which was sometimes described as representing the social phenomena of managerialism and economic rationalism.

5.4 11 Descriptions and appraisals of evidence-based medicine

“...if the parameters are unmeasurable then I don't know how you can have an evidence-based approach that is going to work...”

The three *descriptions* of evidence (a measure of reality, by its functional role, or as a constructed product) and the three related approaches to *appraisal* (along benchmarked, applied and social dimensions) were reflected in participants’ interpretations and critiques of evidence-based medicine.

Some people concentrated on the benchmarked dimension of EBM. They distinguished the evidence-based approach to clinical practice by the use of explicit criteria and rules of evidence. Standards were used to detect bias in the measures of effect obtained from clinical research. Systematic reviews of

clinical research were distinguished as evidence from the estimates made by individual clinicians (about the effectiveness of interventions, the accuracy of diagnostic tests or the probability of prognosis). The benchmarked dimension of EBM gave weight to the skills of *critical appraisal* against *standards* and *criteria* of study design and research method.

A focus on the benchmarked dimension of EBM was not always compatible with the orientation of those who viewed evidence more from a functional, applied perspective. For some practitioners, the rigid benchmarks for appraising evidence appeared limiting and impractical. That interpretation of EBM placed too much emphasis on “gold standard” evidence (in the form of clinical trials and systematic reviews) and not enough on the clinical decision process in which it was applied.

To overcome the perception of EBM as impractical, its supporters went on to identify it as a process of critical thinking and a new way of clinical reasoning. This correlated with the descriptions of evidence by its functional role and appraisal along the applied dimension. They described how they used research as evidence to assess and weigh diagnostic or treatment options in defined settings. In their discussions the contextual factors that determined a clinical judgement (such as a patient’s underlying risk of disease or preferences) were given greater emphasis than the qualities of the information they used as evidence.

In their description of evidence-based decision making they talked about using a wide range of information, that was classed and appraised as evidence. It included patient history, the diagnosis, their assessment of a patient's risk factors and prognosis, published research, their clinical experience and colleague's opinions about what "works", the patient's expectations and preferences, and even "gut instinct" and "intuition". The various types of evidence were sometimes distinguished into "hard" evidence (quantifiable and biomedical measures) and "soft" evidence (qualitative clinical observations). Focus on the applied setting however, was a more palatable discourse than the comparative bias of opinion against standards of experimental research. Thus the applied dimension of EBM gave most emphasis to the skill of *weighing* and *combining* different forms of evidence, and the ability to make a correct decision for that particular *context*.

"I'd hate to work in a medical setting, where we did have very rigidly imposed guidelines based on the evidence for what we can, how we can treat things, how we manage things and how we prescribe..."

Retaining autonomy of judgement was important to clinicians. Conversely influencing, and possibly mandating, the direction of clinical judgement was considered an option by those who took on a health service manager's

perspective. Evidence-based decision making is discussed further in section 5.5.

Those who described evidence as constructed and value-laden sometimes highlighted these characteristics as “flaws” and “biases”. In their view, the concept of evidence did not comply with scientific requirements of objectivity, and they extended this critique to the notion of EBM. Thus evidence-based medicine was sometimes interpreted as a potential move away from scientific thinking and sound judgement. This was more likely if the construction of evidence was perceived to increase the potential for errors i.e. evidence was defined in a legalistic sense or because data for the evidence were inappropriately selected and combined.

The constructed description of evidence was also used by those who believed EBM represented only a narrow, and thus limited, component of what was important in health care decisions. This led to appraisals of EBM along the social dimension of evidence that considered the social origins of EBM and assessed the social impact of this movement.

The social critique often came from those interested in areas of health care not well served by prevailing EBM models. The dominant EBM rules and criteria for evidence were seen as influential determinants of resource allocation decisions. There were concerns that the benchmarking standards, and

constructions of evidence from the biomedical, clinical setting would be inappropriately imposed on public health, health promotion and social medicine. Hence appraisal of EBM on the social dimension of evidence gave emphasis to the origins of EBM, and the social impact of its influence on the generation, appraisal and application of evidence in policy and practice. This appraisal of EBM is discussed further in section 5.6.

5.4 12 Reflections

The descriptions and appraisals outlined here each draw attention to different priorities when conceiving the topic of evidence: i.e. determining “latent truth”, process and action, or the “human factor”. These differences can be interpreted on two levels.

First, they can reflect alternative and distinct ways that evidence is conceived and defined among researchers and practitioners. The various descriptions and dimensions of appraisal can thus be treated as separate and potentially conflicting perspectives on evidence. These different perspectives may explain the confusion that occurred when some researchers and practitioners first began to debate evidence and found they were at cross-purposes.

A second and possibly more sophisticated interpretation is that the notion of evidence is actually a multi-dimensional construct (a composite), with at least

three points of entry and three dimensions of appraisal. This conclusion was supported by the fact that when interview participants engaged in a comprehensive examination of the topic of evidence, their reflections were often multi-faceted. Hence in the time of an interview people often described evidence from several perspectives and addressed multiple dimensions of evidence.

These two interpretations of the three points of entry and three dimensions of evidence are discussed in more detail below.

Three distinct concepts of evidence

Prima facie the three points of entry to describing evidence can reflect different worldviews. That is, evidence as a measure of reality and the benchmarked dimension of evidence share many similarities with the post-positivist/scientific research paradigm described by Denzin and Lincoln (1994). The post-positivist worldview also requires standards, rules and criteria as a means of appraising critically and judging a piece of evidence. The criteria highlight the qualities imbued by the method by which the evidence was generated, such as objectivity, reliability and validity.

The benchmarked dimension for appraising evidence was often discussed by proponents of EBM. It was also invoked by those who sought to promote the

value of alternative (qualitative) research to other researchers and practitioners who attributed credibility to explicit benchmarking criteria. This dimension of evidence inspired the publication of standards for the implementation of qualitative research (Mays and Pope, 1995; Greenhalgh, 1997; Popay and Williams, 1998), and guides to its critical appraisal (Popay et al, 1998; Mays and Pope, 2000, Giacomini et al, 2000).

Aspects of the functional description of evidence and the applied dimension of appraisal can be described as reflecting the worldview of practitioners or policy makers. This worldview has been described by Orosz (1994) and Lomas (1997a) as more pragmatic than that of researchers, and they have attributed the “gap” between research and practice to their different orientations to evidence.

Finally, descriptions of evidence as a constructed product and its appraisal on a social dimension share attributes with the social/constructivist research paradigm. Here each “story” is told from a particular point of view rather than reflecting objective “truth” (Denzin and Lincoln, 1994; Rosenau, 1996). Constructivism has not only been applied to qualitative research. Veatch and Stempsey (1995) have written about the incommensurabilities of different constructions of scientific research findings. Hence constructivist philosophies of science consider the influences of “...human beings who act, think, and speak in terms made available to them by the culture, and

who...have not only created these terms, but also act out the meanings they inscribe” (Fox Keller, 1994, p328).

Especially in public health, leading writers have examined the notion of evidence along a social dimension and explicitly argued that research data are political products (Kreigar, 1992). Sociologists have described how frames of reference determine the construction and interpretation of evidence. They influence the questions formulated, the evidence collected and thus the “answers” found (Peterson and Lupton, 1996). Medicine and public health have also been critiqued, however, for failing to adapt to, and reflect, the post-modern world and its notions of constructed realities (Lupton, 1998; Chan and Chan, 2000).

Evidence as a multi-dimensional composite

Consideration of the descriptions and appraisals of evidence as distinct worldviews, however, does not do full justice to the scope and complexity of my data. Interactions on evidence were rarely one-dimensional, and the second interpretation of the model presented in [Figure 7](#) is that the three points of entry and three dimensions of evidence are components of a composite construct of evidence. This interpretation reflects more closely the way debates about evidence and evidence-based health care evolved and became increasingly multi-dimensional over time. This development

occurred during individual interviews as participants reflected on philosophical, methodological, applied and social aspects of the concept of evidence. The development also emerged over the study period, as interactions about evidence on the Internet, at conferences and in the health care literature increasingly addressed multiple, rather than single, dimensions of evidence. (The progression towards three-dimensional evidence is discussed in section 5.6.)

Evidence as a multi-dimensional construct presents a model for collaboration and (potentially) better communication between science and social science researchers, and those who apply research to practice. Hence, rather than highlighting the differences between traditionally conflicting epistemologies, I am more inclined to adopt the notion of evidence as a multi-dimensional construct, and to formulate the following hypothesis. It seems that explicit discussion of each point of entry and each dimension of appraisal is essential for an inclusive and comprehensive debate about evidence in medical and public health research and practice.

A comparable argument, which also promotes multiple dimensions of evidence, has been introduced in social, education and health promotion literature on program evaluation (Rutman, 1977; Hawe, 1990). It is encompassed in phrase “evaluability assessment” where appraisal criteria, the

application of evaluative research, and prevailing values are each made explicit, before a program evaluation is conducted.

My own perception, as a participant observer in discussions about EBM, has been that communication improved when the multiple components of evidence were addressed and clarified. This occurred when the measures required and the evaluation criteria were made explicit; how the evidence was to be used and the context was clear; and when the constructed limits of the evidence were openly acknowledged and the implications were considered. This did not necessarily result in a convergence in thinking, but it did appear to lead to more informed disagreement.

Thus incorporating a multi-dimensional approach to evidence into mainstream medicine, public health and health services research may promote a closer relation between traditionally distinct disciplinary or professional perspectives. However, the different dimensions of evidence will undoubtedly have different significance and value, and be given different priority, when it is invoked in different roles and settings. The application of evidence in formulating and influencing different types of judgement is discussed in section 5.5.

5.5 FORMING AND INFLUENCING JUDGEMENTS

5.5 1 Introduction

“I don't think I can evaluate evidence without knowing what the purpose of the evidence is...”

Descriptions and appraisals of evidence were rarely maintained at the level of theory or philosophical abstraction. Most often they were grounded in discussions about research or clinical and public health practice. The notion of evidence was invoked and critiqued in the context of *forming* and *influencing judgements*. Thus participants discussed evidence as a basis for their own judgement in research or practice, or how they (or colleagues) used evidence to influence the judgement of others.

5.5 2 Summary of my grounded theory on evidence

Descriptions and appraisals of evidence are determined by the type of judgement to be made. They are also influenced by the judgement context: i.e. whether a person is making their own judgement or if evidence is used to influence the judgement of others.

In research and practice settings, evidence was invoked to form two categories of professional judgement. First there were *epistemic* judgements, where evidence was used to *substantiate knowledge*. In this context, evidence was mostly defined as an observation or measure of reality, and these measures were assessed in the benchmarked dimension. Evidence was appraised to determine how much is known and to critique knowledge claims.

The second types of professional judgement were *decisions* about *action*, e.g. decisions about clinical practice. In this context, evidence was often described by its functional role and appraised in the applied dimension. The influence of EBM was also to highlight the benchmarked dimension of evidence. This led to appraisals of evidence along two dimensions (applied and benchmarked) when forming decisions about health care.

Evidence was also invoked to *influence* the judgement of others, a process I labeled *argument and persuasion*. In argument and persuasion evidence was used to influence another person's knowledge and to influence their decisions about action. For example, proponents of EBM invoked the concept of evidence to form an argument and persuade others about "best practice" and to influence their clinical decision-making. Argument and persuasion were also apparent in discussions about public health practice and health policy. Here researchers and academic practitioners invoked evidence to influence the

judgement of decision-makers who were often from another discipline or profession e.g. health service managers or politicians.

When discussing the process of argument and persuasion, evidence was often described as a constructed product and appraised along the social dimension. Attention to the social dimension, however, supplemented rather than excluded attention to the benchmarked and applied dimensions of evidence.

Details of these concepts are presented in the sections below.

5.5 3 The judgement context and different types of judgement

I developed the category of *forming and influencing judgement* to represent the core social process that was described by participants as they defined the term evidence. Hence forming one's own judgement or influencing others are presented here as different aspects of the *judgement context* in which evidence was described and appraised.

Where participants discussed evidence in relation to professional judgements made in order to generate, establish or challenge knowledge, I labeled these *epistemic* judgements. The second types of professional judgement were *decisions about action*, e.g. clinical practice.

The judgement context and type of judgement being made were distinguishable by their primary aim or “mode of reasoning”. In forming epistemic judgements, evidence was invoked to determine a *proximity to latent truth*. In forming decisions about action, evidence was invoked to determine a *sound choice*. In the social process of argument and persuasion, where the primary aim or mode of reasoning was to impact on the judgement of others, evidence was invoked to determine *influence and effect*.

Forming epistemic judgements, making decisions about action, and influencing others were not mutually exclusive. For example, research involved iterative cycles of establishing knowledge (literature review, assessing your own research), making practical decisions (planning and implementing studies) and persuading others to one’s point of view (grant submissions and publications). Participants also described progressive relationships between these concepts i.e. evidence was assessed to determine what was known about the effects of an intervention, decisions were made about “best practice”, and other practitioners were persuaded to follow the recommendations.

Formulating and influencing judgements also encompassed a dimension of *time*. Evidence could be invoked forward in time to guide judgement, and also retrospectively to examine previous judgements. For example evidence was used to generate hypotheses or test hypotheses; to make a decision about

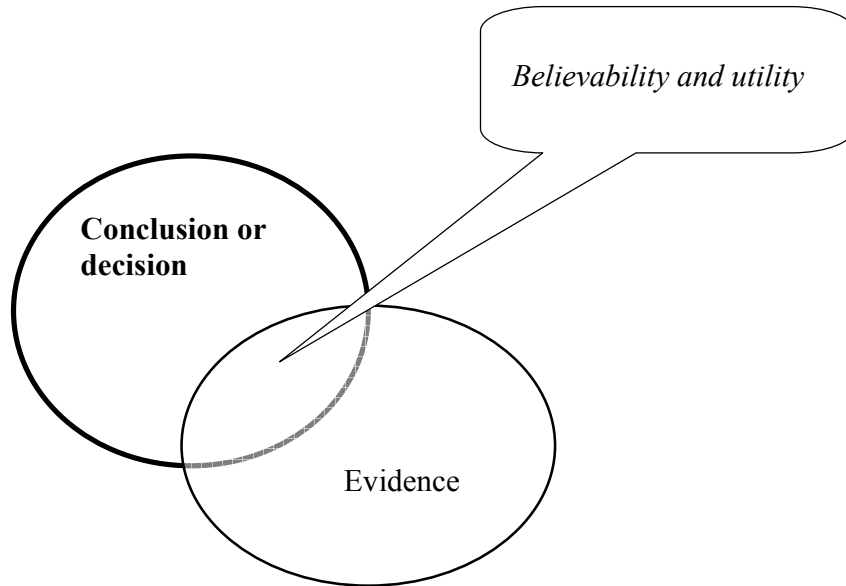
practice or assess past practice; to persuade others to take action or justify a decision already made.

In the process of forming judgments, the relationship between evidence and the final conclusions was seen as critical in both research and professional practice. In epistemic judgements, evidence was primarily appraised on its *believability*. This was conceived as its likely approximation to latent truth.

To make decisions about what to do in practice, evidence was also appraised on its believability, and as well as its *utility*: that is how relevant and useful it was in guiding that decision. Assessment of believability and utility correlated with appraisals on the benchmarked and applied dimensions of evidence. In the context of argument and persuasion, where people invoked the concept of evidence to influence the judgement of others, evidence was also considered on the social dimension of evidence.

The relationship between evidence and a conclusion (or decision) in the process of forming judgments is illustrated in [Figure 8](#). The characteristics of evidence believability and utility are presented as the shared territory between the evidence and the conclusions.

Figure 8: The relationship between evidence and conclusions



In epistemic judgements, because truth was described as latent it could only be approximated using evidence. The believability of the evidence primarily determined the conclusions, and the relationship between the evidence and the conclusion was perceived to be quite direct. Highly believable evidence led to greater confidence in what is known about the latent truth. Believable evidence can be represented in Figure 8 as a high degree of overlap between evidence and conclusions.

The believability and utility of evidence also determined how much weight it was given in making a decision about practice. When believable and useful evidence was available, this led to greater confidence in the decisions made in an applied setting. The relationship between evidence and a clinical decision, however, was often described as indirect. Evidence mostly guided the decision rather than serving as a primary determinant of the decision. This was partly because evidence was often incomplete, yet a decision had to be made. This can be illustrated in Figure 8 as a smaller area of believability and utility and thus less of overlap between evidence and decisions.

In argument and persuasion, the social dimension of evidence further blurred the relationship between evidence and conclusions (or decisions). The processes of forming epistemic judgements, making decisions about action, and argument and persuasion are discussed in more detail in sections 5.5 5 – 5.5 7 below.

5.5 4 Professional judgements

The concept of evidence was discussed in a context of *professional judgement*, in contrast to lay judgements. Professional dealings with evidence were related to research, teaching, practice or policy, where judgements were undertaken in the capacity of an academic or health care provider. Hence, professional activities were either those funded from public sources, or where

judgements were undertaken on behalf of the public or as a representative of the University or a professional body.

In the professional context, objectivity and detachment from personal interests were ideals to which one should at least aspire, even if they were not totally achievable.

“...once you tie something to your ego, you are stuffed, you really are, because you have to lose face to change, and I think that is a cardinal error, and a lot of people make it”

Participants also described a requirement to set limits on unacceptable external influences and encroachments on professional judgements, and to retain appropriate degrees of autonomy and independence. Professional judgements on evidence were described as based on “rationality”, “logic”, “scientific principles”, “critical appraisal”, and a “balanced” consideration and “weighing” of all available evidence.

They were also described as “explicit”, “open”, “systematic” and based on *standards* and *criteria* and in line with professional training. Thus a medical practitioner, who has public credibility on the basis of their medical qualifications, was described as duty bound to examine evidence within the scientific paradigm of their medical degree. However, professional

judgements were also described as based on “practical experience”, “tacit knowledge”, “intuition”, and “gut instinct”, which all relied on finely tuned personal skills and “acumen”.

Professional assessments of evidence were distinguished from those where one had no training or experience. Several people contrasted their professional role to one in which they felt like a novice, like buying a house or a car. In the latter, they lacked the skills to evaluate, weigh, and appropriately judge the evidence. Thus professional judgments on evidence were more sound than those merely based on common sense and emotional appeal. Professional judgements retained a certain order and structure so as to identify and address all the important issues.

Retaining *independence* to formulate one’s own judgement about evidence was highly valued, both in the role of an academic or researcher and as a health professional. In the research context, this referred to the freedom to interpret results as evidence and publish findings. Restrictions on this independence were often linked to the source of research funding e.g. commercial sponsors of research that may influence how the findings are reported.

In the clinical setting there was an ideal of *professional autonomy* to accept or reject evidence when making decisions with (or on behalf of) patients. From

the independent practitioner perspective, the evidence of EBM was accepted as *informing* practice, and rejected as *prescribing* practice. Concerns about EBM were that it is “cookbook medicine”, where judgements about the implications of evidence are removed from the individual practitioner. This was deeply worrying if decisions were perceived as heavily influenced by management or government, rather than the clinical profession.

Evidence was also discussed in relation to professionals (in research or practice) having an influence on other or external decision-makers, i.e. those who manage health services or determine policy and practice more broadly. Academic or health professional judgements were seen as playing only a minor role in determining social and political activities, including formulating health policy.

Successful argument and persuasion required consideration of the social, structural, political and commercial influences on stakeholders’ judgements. Professional lobbying required participants to determine the acceptable lengths to which they would go to determine the desired effect. Here researchers and practitioners considered the implications of the construction of evidence and its appraisal on the social dimension.

5.5 5 Epistemic judgements: establishing proximity to latent truth

“...having your hypothesis confirmed is actually pretty boring. It's having it denied...that's where you see something new...I always tell students it's like seeing into the eyes of God really, despite the fact that I'm an atheist.... You see how something is done before anyone else has seen it... that is real explorer stuff...that's David Livingstone...”

Epistemic judgements were often made in the context of formal research, which was a key component of many peoples' work. Discussions about evidence in relation to this topic often inspired personal enthusiasm and passion.

Epistemic judgements were also made in relation to abstract discussions on what was known about clinical medicine or public health (as distinct from making actual decisions about what to do in a particular context). This included judgements that evaluated propositions or claims about seemingly established or “common” knowledge about effective practice.

In forming epistemic judgements, evidence was mostly described in terms of an observation or measure of reality i.e. data collected to describe, understand, quantify or explain phenomena. Appraisals on the benchmarked dimension of evidence applied criteria and standards to assess the quality of observations or

measures, and their likely proximity to latent truth. Hence evidence was appraised to formulate conclusions about what was known, and as a substrate in the process of determining if existing conclusions were convincing and if so, how convincing.

Epistemic judgements in the context of research were often described as a form of hypothesis testing (both as falsification or confirmation). Thus evidence was invoked to substantiate or challenge propositions. Observations and measures were also used to develop or examine theories and exclude alternatives. Evidence was also described as the observations that triggered new insights, stimulated intuition and creativity, and led to new ideas for directions in research.

In epistemic judgements, the durability of professional independence was emphasised. Participants valued the freedom to collect data, and analyse and report that data as an independent or “objective” form of evidence. Some academics described increasing pressure to obtain private or commercial funding for research. There was concern that this potentially coloured how research findings were reported and discussed. These were often subtle influences: for example, the funding source for a project may impact on one’s sense of freedom to express scepticism or caution in the discussion section of a scientific paper.

When professional independence was stipulated within funding contracts, some remained concerned about perceptions of bias from peers or the community, and the loss of prestige associated with one's role as an independent arbitrator. Government or commercially independent research funding bodies tended to be associated with perceptions of greater professional independence. Grants from these independent bodies carried high levels of prestige and peer recognition.

The ideal professional role was that of a skilled, independent arbiter of the data collected and presented as evidence. The skills identified as essential included knowing which questions to ask and thus the type of evidence to be examined, as well as collecting and appraising this evidence. As described above, appraising the believability of the evidence was a key requirement. An appraisal of believability determined whether propositions remained as hypotheses, if they were completely rejected, or if they were classified as having been substantiated.

"...if the evidence was surrounded by a dirty great fog of uncertainty then I would tend to discount it."

Evidence was conceived along a range of believability. It could include data that suggested a hypothesis or data that led to its verification; preliminary results or conclusive results; evidence in the form of a seminal paper or a body

of widely published and established research. Provisional or “soft” evidence was associated with single studies, first-time experiments and small sample sizes. Soft evidence was also associated with less formal processes, such as “quick and dirty” experiments, early explorations, personal experiences, uncontrolled or informal observations, consultations and anecdotes, and observations based on intuition.

Appraisals of evidence believability were mostly based on established standards, rules and criteria for conducting scientific research.

“...the idea of evidence is absolutely immutably mixed with the concept of controls, experimental evidence at any rate, which is really what I'm talking about. If you do not have the requisite controls then you don't have evidence basically, the controls are critical to the hardness of the evidence...”

These evaluation criteria were based on theoretical concepts (such as the notion of controlling external variables, control groups or randomisation). They were also related to standards for the implementation of data collection methods and procedures, like analysis and reporting. These appraisals were conducted in formal steps such as the peer review of journal articles. Also students in basic science, public health and clinical medicine were taught explicit critical appraisal of research against standards of method.

Evidence invoked to establish knowledge was less believable if it conflicted with *a priori* expectations, theory or logic, and if the conclusions had low “self-evidence”. In such situations evidence was not discarded, but conclusions were more likely to remain provisional until repeated or corroborated from other sources. Perceived vested interests (commercial or political) that went in the same direction as the evidence reduced its believability (e.g. tobacco companies producing evidence that smoking was not harmful, political parties producing evidence to support their policies).

Conversely, the believability of evidence was reinforced by accumulated data from repeated studies, which point in the same direction. Therefore commensurable and reinforcing evidence of different types and from different sources, and evidence that conformed to current models or paradigms were more believable in establishing knowledge. The credibility of the source of the evidence was also important. This was often judged on the perceived skills and experience of an individual or group, and the reputation and independence of their institution.

Overall there was emphasis on making a formal and direct link between evidence in the form of research, and conclusions about medical and public health knowledge. Participants described a close and direct relationship between the amount and quality of the evidence and their conclusions. A judgement about the scope and certainty of current knowledge was based

primarily on an assessment of the supporting evidence. As described in section 5.5 4, I have distinguished the perceived *direct* relationship between epistemic judgements and available evidence, from the *indirect* relationship that was described between evidence and decisions about action.

In forming epistemic judgements, the *utility* of evidence also received some attention. The utility of evidence ranged from that which solved pressing problems, provided invaluable insights or assisted in confirming or discounting key hypotheses, through to “the bottom drawer” evidence which had no immediate or obvious relationship to existing ideas. An ongoing area of debate in relation to medical and public health knowledge was the appropriate balance between investing in research that produced evidence with explicit, direct or immediate utility for practice or commercial functions, and basic or exploratory research with unpredictable utility or technical or commercial importance.

5.5 6 Making decisions about action: a sound choice

“...I would be evaluating the evidence in the context of the decision, right; what impact does the evidence have, if any, on my ability to make a decision...”

In decisions about action the primary aim was to make a sound choice.

Evidence was invoked to consider and weigh decision options so as to determine the best course of action at a particular time and within a defined context. These professional judgements balanced evidence with multiple other variables and the competing demands of an applied context.

The process of forming decisions about action was mostly discussed in relation to clinical practice. In this setting, conversations about evidence were strongly influenced by local and international debates about evidence-based medicine. Over time the influence of EBM moved beyond clinical medicine and the terminology was adopted in discussions of evidence as a contributor to decisions about health care in general, public health, health promotion, and health policy.

Making decisions about public health practice and health policy was often described as a group process with competing stakeholders. The study participants used evidence to formulate sound and convincing arguments in such a context. The use of evidence in argument and persuasion to influence the decisions of others is addressed in the next section (5.5.6). This section is about weighing evidence to make independent decisions, especially in clinical practice by autonomous clinicians.

Evidence was also discussed in relation to decisions about the allocation of resources for research. It was anticipated that research, which generated evidence that was useful for practice (and policy) decisions, would be more likely to receive future government funding than basic or exploratory research. This raised the stakes for researchers to ensure that their findings and publications were identified and valued as “useful evidence”. The notions of valued and excluded evidence are discussed in section 5.6

As in epistemic judgements, the art of critiquing evidence was given a high priority in decisions about action. Critique was about determining the credibility of the evidence, and its relevance and applicability to the particular decision at hand. This reflected an appraisal of evidence along the applied dimension, where equal weight was given to the believability and utility of that evidence.

Determining believability and utility often required distinct types of skill and used different criteria, particularly in clinical practice settings where there was a new focus on using probabilistic research as evidence. Many practitioners had only recently acquired the skills required for critically appraising clinical research. This was conceived as a different art to that of assessing the relevance and applicability of the evidence to a particular patient. Appraisals of evidence believability required knowledge of scientific benchmarks for

research methods and data analysis, while appraisals of evidence utility and applicability relied on the art of clinical diagnosis.

Evidence believability and utility each fell along a spectrum. Situations described ranged from one of full confidence in evidence that was highly pertinent and useful, to situations of maximum uncertainty, where very little appropriate evidence was available. Most decisions had some empirical basis, such as believable evidence of partial utility (i.e. physiological evidence about the effects of an intervention, but no research on patient outcomes) or relevant evidence of moderate believability (i.e. practitioner past experience about the effects of an intervention).

Epistemic judgements employed standardised procedures and criteria for appraising evidence (benchmarks), particularly for establishing knowledge in formal research. In contrast, the traditional relationship between evidence and decisions about health care was described as implicit and idiosyncratic.

Physicians "*would tinker with the way the management was going and change this medication for that, I thought often on the basis of no evidence or very scanty evidence, just 'I'm more accustomed to using drug X than drug Y'.*"

Such ambiguity about the relationship between evidence and decisions was often described as commonly accepted prior to systematic critical appraisal as introduced by EBM. The appraisal, interpretation and application of evidence

had used implicit reasoning processes, bound up with the notion of clinical acumen. The components of clinical acumen were described as exclusive to the decision-maker and thus implicit and difficult to challenge. The professional role of senior clinicians had been to rely on their experience as evidence. This clinical experience had traditionally been perceived as the most believable and useful evidence when making clinical decisions.

EBM was perceived as changing the traditionally implicit and unstructured approaches to assessing evidence believability and utility and its influence on decisions. Clinical epidemiology (the founding discipline of EBM) was presented as the “the basic science of clinical medicine” (Sackett et al, 1991) because it incorporated scientific critical appraisal criteria into the relationship between evidence and decisions. I have presented this as introducing appraisals along the benchmarked dimension of evidence into the clinical setting.

Explicit criteria were intended to reduce the impact of supposition and human error in determining what was believable evidence, what type of evidence was given most weight, and how it was used in clinical decisions. The contemporary professional role in relation to evidence (in a clinical setting) has been conceived as that of a skilled negotiator of many forms of evidence, who can interpret its relevance for each individual context and use it to assist sound choice.

Decisions about practice that were perceived to be in conflict with research evidence were occasionally put down to an unwillingness to admit to being wrong, conflicting structural or remuneration incentives to act otherwise, or an unwillingness to accept evidence that contradicted experience. These were discussed for both practice settings (e.g. using treatments despite good evidence that they are ineffective) and in relation to decisions on directions in research (e.g. refusing to let go of a favourite, but widely disqualified, hypothesis).

However, even good research evidence was rarely conceived as making obvious a final decision. Rather different forms of evidence were combined to inform a more complex judgement process in which research evidence was explicitly only a part.

“...you still have to rely a lot on the art of medicine and you have to rely on your clinical acumen and the history taking and the fact that for some reason sometimes the alarm bells ring and you don't really know why”

Decisions about action in a clinical setting were described as including rational assessments of scientific evidence, but also “fuzzy logic” with its intangible components, professional and personal values, patient values and

preferences and multiple competing external agendas such as costs and accountability, legal implications and remuneration.

Hence the nature of a relationship between evidence and a decision about action was conceived as less direct and linear than that between evidence and knowledge. Decisions about action were mostly described as judgements based on multiple variables, of which many forms of evidence were invoked and each played only a partial role in predicting the decision made. The importance given to the believability and utility of research evidence was also dependent on the significance of the decision; the potential benefits, harms, risks and costs and whether there was conflicting evidence. Costly action that involved high potential risk required more believable and directly relevant scientific research and greater confidence in the overlap between the evidence and the decision to be made.

The value placed on professional judgement, particularly in determining the utility of evidence, was recognised and harnessed by many clinicians promoting an evidence-based approach to colleagues. Initially EBM offended those who interpreted it as a challenge to their clinical judgement and in the mid-1990s EBM was explicitly reframed as enhancing, rather than replacing that judgement. In various settings, such as conferences, seminars, training sessions and in published articles, EBM proponents clarified that the evidence (and its critical appraisal) was used to “assist in making decisions, it was not

the decision itself'. The EBM message was acceptable to the degree that it did not imply the relationships between research evidence and clinical decisions were direct and linear, and the complexity of medical judgment was emphasised and valued.

Although supporters of EBM saw it as a justification of skilled critique, there was also some concern that glib interpretations of EBM could hinder and limit critique. They could give a false sense of security about decisions that were presented as evidence-based. The constructed nature of EBM reviews and summaries caused concern about a “blind faith” in clinical experts being replaced with blind faith in statistical or EBM experts.

“Someone has to look at the evidence-based data to make sure that the conclusions that people come to are actually correct...and none of us check it, or I don't, I don't know anyone who actually does the sums...it puts an enormous responsibility on the people who are producing the data...”

Some thought the new critical appraisal criteria for research evidence helped to address questions of believability, but that there was some way to go in determining standards for questions about the relevance of evidence to the decision context. Hence disagreements about evidence partly stemmed from people applying different benchmarks for what was relevant or useful to the decisions at hand.

There were different opinions on the relative utility of evidence that provided an explanation of the patho-physiological mechanisms of an intervention, compared to evidence that provided average, probabilistic information about its outcomes (based on evaluation of prior interventions). Supporters of EBM gave more weight to evidence from systematic clinical research (with its own internal hierarchy of quality) over reasoning from basic science, uncontrolled experience and theoretical, qualitative approaches. This was often based on the principle that “experts are wrong more often than big randomised trials...”.

Thus for those who wished to base decisions on probabilistic evidence, accurate quantification was an important component of evidence believability and of its utility.

“...knowing that some women have this anxiety and some women have that concern and some women found it dreadful and some didn't find it dreadful, without actually knowing something more quantifiable about it isn't helping the decision. So (I'm) thinking of evidence as something that's going to be useful for decision making in the future for someone else...”

This was challenged by those for whom evidence that was believable “on average” was not perceived as helpful in making decisions for individual

patients. Those promoting greater use of probabilistic evidence had to overcome the hurdle that many practitioners were unfamiliar or uncomfortable with thinking probabilistically or converting such evidence into predictions that were meaningful for their patients. Even those experienced with dealing with probabilities described difficulty in applying them to make decisions about individual patients.

“...its like walking into a football field and saying well how many people do you think there are here, I mean I've never got a clue. I think that we're not sort of built to be very numerate... we often have trouble at a deeper level...”

The most challenging situations were those where different pieces of evidence conflicted, and difficult judgements had to be made about where to place most weight. Making the shift from relying on practical experience and “received wisdom”, to giving more weight to the external body of research, or “big E evidence” sometimes required significant modifications to personal decision-making criteria.

“...evidence from the literature... is one which will get closer to the truth and that is a learning phase, it's a big jump. ...individuals have actually got to realise for themselves why other approaches have got potentially much greater bias and so you are much less likely to end up with the truth, or near the truth.”

Even those who were highly involved and supportive of the emerging EBM movement were not always ready to accept new research evidence that conflicted with personal experience. The EBM principle was that “the “default position” should be to let formal research evidence prevail over the anecdotal information, unless the research is shown to be irrelevant, inconclusive, biased or imprecise”, (Enkin and Jadad, 1998). Yet because most evidence came in degrees of believability and utility, when and how to put the EBM principles into practice remained a matter of individual professional judgement.

“...you sort of say, yeah, well that’s the evidence based on that group of little trials, but here’s what I’ve found actually works in practice and until you can actually give me more, maybe I’ll continue doing what I’m doing and what seems to work for me and...for people who I’m being invited to provide medical care for...”

Therefore although EBM introduced new explicit and standardised approaches to the concept of evidence, the complexity of the relationship between evidence and the final decisions continued to be emphasised. The art of critique focused on interpretation of evidence for each context. The professional role was about balancing and weighing different and often conflicting evidence to make sound decisions.

5.5 7 Argument and persuasion: having an influence and effect

“...if you don't address the politics then you are not doing public health. I'm quite amazed at some of our colleagues who seem to think they are purely in search of the truth”

Evidence also figured in the formulation of persuasive arguments. Here it was used to influence the conclusions or decisions of others, to sway them towards one's own conclusion. Argument and persuasion required the selection of appropriate evidence, skilled verbal engineering and presenting an interpretation of the implications of that evidence.

Evidence that was presented in argument and persuasion was sometimes described as a constructed product. Discussions about influencing others, particularly where this was related to an allocation or redistribution of resources, often included appraisals of evidence on the social dimension (this is discussed in section 5.6). This was in addition to appraisals of evidence on the benchmarked and applied dimensions of evidence.

Using evidence to persuade others applied across disciplines. Those who were to be convinced included peer researchers or practitioners, research fund granting bodies, bureaucrats, policy makers, the media or the public. In the

social process of argument and persuasion, evidence was used to achieve a predetermined, desired effect. This included getting published, winning research funds, changing medical practice or having an influence on a health policy.

Evidence was also known to be used retrospectively, particularly in policy settings, where the desired effect was to justify decisions to others.

Participants considered this to be acceptable as long as it was done openly and the evidence truly supported the decision, rather than being purposefully selected or skewed.

“...sometimes evidence will feed into a decision but I suspect it's more likely the other way around, that evidence is used post-hoc to justify decisions. And that provides the credibility and that provides part of the power...but I'm not too precious about it because if the decision is in line with expectations that you think the evidence would promote, then which way it happens I'm not too fussed...”

Making a persuasive argument to peers and colleagues, one-to-one, at seminars or in publication was a familiar process. In these forms of argument, the rules and criteria for evidence were often firmly established. There were clear benchmarks for presenting data as evidence in scientific papers or in conference presentations. In other areas of argument and persuasion, the rules

of evidence were not well or explicitly defined and were often implicit e.g. traditional clinical practice or lobbying for health policies.

Challenges to clinical practice had traditionally been exchanges about experience and individual skill. EBM introduced a new set of rules and criteria for arguments about professional practice in the clinical setting. In promoting EBM, participants argued for changes to evidential frameworks, as well influencing specific practice. Modifying the evidential framework included changes in the way evidence was defined, the rules and criteria for its critique, and changes to the way outcomes-focused research was used in health care decisions.

Supporters of EBM took a leading role by making changes to the medical curriculum and establishing new courses for professional development and training, particularly in clinical epidemiology. They were also involved in a form of argument and persuasion by example, acting as leaders in the way they critiqued and applied evidence in their own practice and discussed evidence with colleagues. The effect of EBM was that formal research was given greater weight as evidence. As a result clinicians reported being more likely to refer to systematic reviews as a basis for a sound argument among colleagues or to students.

The use of evidence to make a successful bid for funding or to influence decisions about health service resource allocation was also familiar to researchers and practitioners, although sometimes it was uncomfortable. Some interviewees saw this use as potentially compromising their notion of professional independence in relation to evidence. It required them to make interpretations, which, if accepted, would lead to gain for themselves or their team.

Researchers emphasised the professional expectations to cite and address conflicting evidence, although winning grants inevitably involved emphasising the supportive evidence. Some felt a tension between remaining “objective”, “conservative” and “neutral” in a submission, and the pressure to claim great benefits and potential profits from an investment in their research program.

Hence, although mounting arguments and forms of persuasion were seen as inevitable and important in most research and practice settings, the degree to which individuals felt comfortable about personally advocating particular interpretations of evidence varied. Approaches ranged from detached summarising and reporting of data (where one’s own view remained in the background) to adopting the role of an active lobbyist and advocate for a particular view (to which one had explicit personal commitment).

The skills of argument and persuasion were particularly highlighted in discussion about public health advocacy and health policy. Here the rules for using evidence were very fluid and the outcomes were unpredictable, irrespective of the benchmarked qualities of the evidence. Some interview participants had extensive experience and knowledge of this role, while others preferred to focus on generating knowledge or personal decision making as a practitioner.

Participants felt that many researchers and practitioners were inadequately equipped to present evidence to bureaucrats and politicians, the media and the general public. These professional skills included preparing media stories, networking with decision-makers, and using evidence strategically in (often unpredictable) windows of opportunity. There was significant variation in how comfortable people felt about this. It was seen by some to take one beyond the role of an independent arbiter of knowledge, into affiliations with explicit social or political agendas.

Hence there were quite diverse perspectives about making explicit recommendations or lobbying for the application of research evidence. Sometimes taking a public stance in a particular direction was seen as potentially hindering future professional or academic independence. This resulted in self-awareness about presenting and interpreting evidence appropriately. Some people preferred to present their research as evidence in

the name of establishing new knowledge, but not to engage in persuading others about its interpretation or implications for practice or policy. Here they maintained the stance of an “independent” and refrained from adopting strong advocacy positions beyond the veracity of their findings.

If one presented a particular interpretation of evidence, it was important to have been seen to declare one’s interests. This was related to the value given to personal and professional integrity, and institutional credibility, when mounting a persuasive argument among academic and health professional peers. Conflicting interests were explicitly taken into account when evidence was presented to support an argument. The notion of maintaining independence was also related to the acceptance of funding, and the potential strings attached.

Thus researchers needed to avoid appearances of working to someone else’s agenda. This was seen as having the potential to damage one’s future credibility and persuasiveness as an independent arbiter of evidence.

Researchers and health professionals developed reputations on having a particular stance or *position* on evidence: they became known for the criteria they used, the forms of evidence to which they gave most weight, for the direction of their interpretations and thus their anticipated conclusions or decisions.

In summary, from the position of examining and interpreting an argument, the art of critique focussed on the person or group pursuing the agenda, and the agenda itself, as much as on the evidence. There was assessment of their credibility and honesty, the validity of their role in relation to their arguments, and of the degree to which their position and intentions had been made explicit. It included examination of the declared and unspoken vested interests and the perceived openness about underlying agendas.

Therefore evidence presented in argument and persuasion was treated as a constructed product, and was often appraised along the social dimension. For example, arguments to maintain autonomous modes of practice were assessed in terms of the financial benefits for the medical profession. Pharmaceutical companies, lobbying to have their products accepted into a pharmaceutical benefit scheme or to win endorsement by health professional organisations, were judged on their explicit commercial interest in the decisions.

Personal and social factors coloured the perceived relationship between evidence and conclusions. They influenced perceptions of the degree to which the evidence was admissible. Professional colleagues were far more credible to clinicians than managers or politicians in promoting practice guidelines, irrespective of the quality of the supporting evidence. Indeed most arguments presented by politicians rated very low, primarily because of the

perceived lack of priority given to seeking truth, or balanced argument, compared to pursuing a political agenda.

“I'd say you talk about evidence to a politician, you are really talking in the legalistic sense, you aren't talking in the scientific sense at all. That worries me”

The pragmatic compromise of argument and persuasion was seen as slipping into a *misuse* of evidence when evidence was selectively interpreted and the real agenda was hidden or deliberately concealed. This broke the fundamental code of staying within the professional bounds of truth. Issues related to misusing evidence are discussed further in section 5.6.

5.5 8 Two exceptions to the proposed model on evidence

In the sections above I have proposed a theoretical model on the relations between the process of forming and influencing judgements in medical and public health research and practice, and various concepts of evidence. This model is also illustrated in Table 1 below.

Table 1: Proposed relations between forming and influencing judgements and concepts of evidence

| Judgement context | Type of judgement & mode of reasoning | Common descriptions of evidence | Dimensions of evidence given attention in appraisal and debate |
|--------------------------|---|---|---|
| Forming judgements | Epistemic judgements, evidence invoked to determine proximity to latent truth | As an observation or measure of reality | Benchmarked |
| | Decisions about action, evidence invoked to make a sound choice | By its functional role | Applied Benchmarked |
| Influencing judgements | Argument and persuasion, evidence invoked to have influence and effect | As a constructed product | Social Applied Benchmarked |

In Table 1 the proposed relations between forming and influencing judgements and concepts of evidence are as follows. In the context of forming epistemic judgements, evidence was invoked to determine proximity to latent truth. It was described as an observation or measure of reality and benchmarked against established standards or criteria. In the context of forming one's own decisions about action, evidence was invoked to make a sound choice. It was described by its functional role and appraised on applied and benchmarked dimensions.

In the context of influencing the judgements of others, evidence was invoked in argument and persuasion to have the desired effect of convincing others to one's own point of view. In argument and persuasion, evidence was often described as a constructed product and appraised along social, applied and benchmarked dimensions.

There were two exceptions to this theoretical proposition on evidence. Both exceptions emerged from those who described evidence as a constructed product.

First (as described in section 5.4 6), some researchers defined evidence as a constructed product in the legalistic sense. They saw the term evidence as inappropriate for describing the more objective data they collected in scientific research. In the context of forming epistemic judgements they

critiqued this constructed view of evidence along the benchmarked dimension. They concluded that because the term evidence reflected a constructed and thus value-laden product, it reflected a “biased” or poor measure of reality.

Second, evidence was also described from a social constructivist epistemological position. Here evidence was not only a constructed product (an entity), but also a socially constructed concept. These participants addressed the social dimension of evidence not only in the context of the social processes of argument and persuasion, but also in relation to forming their own judgements or decisions.

5.5 9 Reflections and discussion

Implications for debates about evidence-based practice

On reviewing this section I was struck by parallels between the judgements and judgement contexts described in this study, and the key components of Precautionary Principle (a structured approach to analysing risk). The Precautionary Principle comprises of risk *assessment* (evidence is used to assess the degree of scientific uncertainty about an issue), risk *management* (decision-makers use evidence to weigh options and decide on action), and risk *communication* (presenting the evidence and conclusions to others) (FAO/WHO, 2000).

Although each component of the Precautionary Principle invokes the concept of evidence, their objectives are different and they involve different types of reasoning and judgement. For example, risk assessment is mostly technical, based on scientific standards. Risk management uses this information in negotiations that are determined by local priorities. Risk communication often involves presenting arguments about the conclusions reached, and is dependent on the evidence that is acceptable to the community.

The generic use of the “evidence-based” phrase can conflate and confuse, rather than clarify, the differences between these various forms of judgement. It is pertinent, in the light of the widespread evidence-based rhetoric, to make explicit the alternative modes of reasoning in which researchers and practitioners invoke the concept of evidence, and to identify how that can determine their definition and appraisal of evidence. The theory proposed in section 5.5 provides a framework for diagnosing confusions and controversies that often emerge around the topic of evidence, and for exploring potential solutions.

Disputes about evidence-based health care can also be classified according to the type of judgements that are at issue and the judgement context. For example, debates occur about epistemic judgements, and whether various modalities of research are the best means of corroborating or challenging

knowledge (e.g. Herman, 1995). The model in section 5.5 (summarised in Table 1) clarifies the distinctions between epistemic judgements and practice decisions, and leads one to acknowledge up-front the potential for different descriptions and appraisals of evidence.

My findings add to the work of others who have discussed the differences between “data” people and “action” people (Sauerborn et al, 1999), and those who have examined the “gaps” between research and practice. For example, the distinction between using standardised evaluation benchmarks versus appraising evidence on an applied dimension supports Savitz (1994). He compared the evaluation of individual pieces of evidence and the “comprehensive evaluation of all relevant evidence” pertaining to a question. He also distinguished between critical appraisal and decision-making where the “goal of that process is not to judge the validity of the epidemiological study but rather to make a wise decision”.

Lomas (1997a) distinguished between rational decisions about health care (which may seem obvious when considered solely on the available evidence) and sensible decisions (which take into account ideologies, beliefs, interests and other political and institutional factors). This distinction is reinforced by my findings of the direct relationship between research evidence and conclusions within epistemic judgements, but only an indirect relationship between research evidence and practice decisions.

EBM has been a conscious integration between the domains of research and practice and between the benchmarked and applied dimensions of evidence appraisal. EBM demanded greater emphasis (than traditional practice) on judging the methods-based, believability of clinical evidence. As a result, explicit benchmarks for evidence from the research setting have been translated into criterion-based appraisals of evidence in the context of health care decision making. Over time, explicit, external benchmarks for evidence have been accepted and adopted by clinicians as a scientific and rational approach to health care decision making, and “adequate skilling in this area is now an essential part of modern physician training” (Byrne, 2000).

Hence it seems that much of the debate about EBM was about working out and negotiating the merger between the different modes of reasoning about evidence. An ongoing and important sticking point is the potential for EBM critical appraisal benchmarks or similar “rules of evidence” to determine final decisions about clinical practice. There remain varying perspectives on the degree to which the final decisions of clinical practice can be standardised and guided, let alone prescribed. It seems, however, that this is an important issue to be debated further and resolved from a health service management, as well as a clinical, perspective.

The probabilistic nature of the evidence promoted by EBM has invariably led to considerations about the balance between an individual patient or a population perspective, and thus the relative merit of different types of evidence in decision making. This was succinctly put on the evidence-based health email list in the latter part of 2000.

“...policy makers take a societal perspective and such decisions must maximise cost-effectiveness. However, individuals in most health-care systems are not interested in cost-effectiveness, but in clinical effectiveness and using a different outcome measure may provide a different answer to the same question. This is a difficult dilemma for clinicians who may be torn between their responsibilities to the individual and to society...”

This balance determines the degree that evidence-based health care should be standardised, and even mandated, from a societal perspective and to which it must remain open to interpretation by individual professionals. There has been increasing interest in modifying health delivery structures and information systems as a way to favour evidence-based practice, rather than targeting the decisions of individual practitioners. This is the same philosophy as that adopted in contemporary occupational health and safety measures i.e. to engineer the environment, not the worker. It is unclear however to what degree the organisational changes should impact at the process level only (i.e. influence knowledge by providing evidence

summaries) or whether they should restrict clinical freedom by mandating interventions for specific conditions.

These questions also relate to the best way of promoting evidence-based practice. As described, many of the early disputes about EBM focused on the way the case for change was argued, whose agenda EBM appeared to promote (managers or patients), and whether EBM was being used for political or clinical motives. The perceived threat to clinical autonomy in contexts where judgement has traditionally been defended as sacrosanct professional territory (individual-patient doctor relationships) has been and will remain a key issue.

The level of individual responsibility will determine the point within the health system where a researcher or practitioner will need to switch from using evidence to make one's own decisions to making a persuasive argument to influence the judgement of others relates. Ham (1996, p28) distinguished five levels of priority setting. These were 1) micro decisions about individual patients; 2) choices about which patients should receive access to treatment; 3) the allocation of resources to particular forms of treatment; 4) distributions of the budget between areas and services; 5) the macro allocation of funding to health services. Clinical practitioners have a primary responsibility for decisions at the micro level, although they may need to make persuasive arguments to politicians and bureaucrats about the distribution of budgets. In public health or health policy, decisions are explicitly social and political

processes, irrespective of the value given to evidence. Thus for researchers and clinical or public health practitioners who wish to have an impact on health or public policy, or public health interventions, the function for evidence is primarily about argument and persuasion.

It is significant that among those involved in this study there was a wide range of perspectives on the degree to which such a role is expected and acceptable. It also seems that this function for evidence is the least explored and supported by empirical research. My findings have highlighted concerns about the importance given to retaining independence and thus professional credibility. Credible statements about the status of current knowledge are expected to come from individuals, or organisations, that are seen to have political and financial independence. The credibility of an individual or an organisation is important not only among research and practice peers, but also with policy makers. Trostle et al (1999) found policy makers to be more likely to rely on the reputation of a researcher than on other criteria when judging evidence.

It is likely that misunderstanding, or even conflict, about evidence-based judgements will occur unless expectations about professional roles are made explicit and clarified. For example, is it expected that an individual or organisation will use evidence to determine the state of current knowledge only, to make decisions or recommendations for action, or to actively promote

the implementation of recommendations to policy makers and the wider community? This is important because expectations will determine how well that individual, or group, will be perceived to have fulfilled their professional role in relation to evidence-based practice.

In June 2000 the National Institute for Clinical Effectiveness (NICE) in the UK was at the center of controversy about their judgement against beta-interferon getting funded by the National Health Service for Multiple Sclerosis. The NICE committee (researchers and practitioners who are independent of government) weigh evidence submitted by a range of stakeholders to formulate recommendations for action on various health technologies. The committee chair visited our Medical Faculty mid-2000 and said that they placed emphasis on making “balanced” and “realistic” judgements that took into account the source of evidence, the competing demands of current practice, the potential for change, and the feasibility for their recommendations.

Their recommendation on beta-interferon was one of their first formal recommendations and caused significant controversy in the UK media. It was also the subject of dozens of postings on the evidence-based health Email list. Many of these postings were critical, although the reasons and perspectives varied.

“...instead of telling us whether beta-interferon works or not, NICE is in fact making a judgement whether the health service can afford a treatment which does appear to work to some degree. NICE has therefore taken on the role of rationer and in my view risks losing its credibility as impartial assessors of clinical excellence.

NICE has become a National Institute of Clinical Affordability rather than Excellence”

NICE were criticised both for going beyond technical epistemic judgments and making “value judgements” by recommending a particular mode of action. However, they were also criticised for not addressing the “human factor” and for failing to adequately “craft and position” their message.

“Human beings need messages that work for them and it is for the provider (in this case NICE) to treat the recipients of its conclusions like customers and pay them the respect of couching their (objective) messages in human terms. I don't think it is good enough for them to behave merely like a bunch of scientists. Their conclusions hurt people as well as help them and NICE must be encouraged to address the hurt in some of its bleak messages.”

It seems that there is no current consensus among researchers and practitioners on the scope of their professional roles and the appropriate

balance between the different types of judgement on evidence. This includes diversity of opinion about what is expected from organisations like NICE and the acceptable balance between “scientific” and “value” judgements on evidence. Some believe NICE should focus on evidence to establish current knowledge, while others want explicit judgments about appropriate policy and practice, and even a role in communicating and explaining those judgements to the media and the wider community.

Sauerborn et al (1999) analysed the relations between research and policy in terms of degrees of influence by competing stakeholders. They too distinguished between activities that are influenced by strict rules i.e. research study design, data collection and analysis (establishing knowledge), and those that require interpreting research findings and advocating for change (argument and persuasion). Importantly, they asked the question: to what degree “are researchers willing to assume a stakeholder role” in promoting particular health policies?

It is most likely that the role researchers and practitioners expect to take in argument and persuasion at a social or political levels is influenced by our training and professional socialisation. Thus the degree to which active involvement in lobbying at the policy level is acceptable and credible among peers. It will also be influenced by whether a reputation as a public advocate among peers, ultimately has an influence on the credibility they are able to

retain among policy makers. At present issues related to this dimension of evidence is not subject to mainstream medical or public health research or professional development, and recommendations are discussed in chapter 6.

5.6 REFOCUSING JUDGEMENT

5.6 1 Introduction

Over the time of this study the language of EBM permeated the spectrum of health care. As a result, what counted as believable and useful evidence, and conversely the information that was rejected as inadequate evidence, were perceived to be increasingly significant judgements in terms of their impact on practice and the allocation of resources. Some participants worried that the new dominant benchmarks for evidence derived from EBM were too confining. For example the application of these benchmarks might undervalue other important knowledge or exclude a holistic approach to practice decisions.

5.6 2 Summary of my grounded theory on evidence

Observations that the language and principles of EBM were being widely promoted in health care systems around the world led some researchers and practitioners to refocus their descriptions and appraisals of evidence. In addition to the methodological and functional orientation reflected in the benchmarked and applied dimensions, their appraisals of evidence included an increasingly social perspective. They considered how concepts of evidence in

medicine and public health were being reconstructed as a result of EBM, and reflected on the implications of that reconstruction.

Hence EBM encouraged the examination of evidence as a social construct and appraisal along the social dimension. When the term evidence was attributed only to certain types of data, it was treated as a statement about the “value” of that data relative to other types of information. Also perceptions that the rhetoric of EBM was becoming dominant gave rise to concerns about the “appropriate use” and potential “misuse” of the concept of evidence.

I have proposed that a “post-EBM phenomenon” is a greater integration of the social dimension of evidence into mainstream medical discourse and health care literature.

Details of these concepts are presented in the sections below.

5.6 3 Evidence as a judgement on what is valuable and important

“...on a scale of evidence from meta-analyses of several large randomised trials, down through individual trials, through observational studies, through case experience, through experiences, intuition is at the bottom I think.”

Evidence-based health care introduced an ongoing debate about “what counts”, and what is excluded, as “real evidence”. When participants addressed this question on the benchmarked dimension, it led to discussions about research methods and the reliability and validity of the findings. Further appraisal of what counted as evidence on the applied dimension addressed the relevance and utility of that information in a particular decision making context.

The perception that EBM principles were increasingly dominant in the health care setting led participants also to examine what counts as evidence along the social dimension. Here questions were raised about the relative social “value” or “worth” of various modes of inquiry and different types of information. These appraisals of evidence also considered the professional and social groups who produced and used the different forms of evidence, and their relative spheres of influence.

Hence questions about evidence on the social dimension were to do with the “authority”, “credibility”, “weightiness” and “power” that was attributed to its different forms, and the potential impact of these judgements on policy or practice. Thus the technical and pragmatic judgements about which types of information were believable and useful as evidence were attributed with an additional significance (particularly in the light of EBM) in terms of their social impact.

I developed three conceptual categories to group participants' reflections on the social dimension of evidence in the light of EBM. The first two categories represented classifications of evidence:

- *Undervalued evidence*: evidence which was considered important but undervalued by the prevailing EBM criteria or hierarchy (e.g. clinical expertise)
- *Excluded evidence*: evidence which was considered important but was not addressed by the existing rules of evidence, or for which the dominant evaluation frameworks were not appropriate (e.g. qualitative research)

The third category related to the first two, but represented an overarching social process.

- *Misusing evidence*: where evidence is used inappropriately or abused. A dominant "evidence-based" rhetoric increased the potential for the misuse of evidence.

5.6 4 Undervalued and excluded evidence

Discussion about *undervalued* or *excluded* forms of evidence was primarily precipitated by the EBM “levels” and “rules” of evidence. The levels of evidence based on study design placed intuition and clinical expertise at the bottom. The negative reactions to EBM were attributed to this hierarchy representing a symbolic statement about the “value” of research versus clinician’s skill and acumen.

The position of clinical expertise or expert opinion at the bottom of the hierarchy was described by one senior clinician as “one of the greatest stumbling blocks” for advocating EBM. Although the hierarchy was based on benchmarks to indicate that individual clinicians were not as accurate at predicting the probabilities of health outcomes as randomised studies, it was often interpreted to imply that clinical judgement was to be given the least weight in evidence-based decision-making. The influence of such symbolism was recognised in the mid-1990s and clinical impressions were often removed from this hierarchy of evidence.

The other data that were perceived to be important, but excluded by EBM rules of evidence, were qualitative research and patients’ experiences. These other forms of evidence were discussed in terms of their “visibility” and “credibility”. It was felt to be significant that these data were incorporated into formal definitions of evidence in the context of epistemic judgements and

sound clinical decisions. This was to promote allocation of clinical research funds to qualitative forms of inquiry and to enhance appropriate patient-centred modes of practice.

An image that proved to be less controversial in EBM seminars than the hierarchies of evidence was a decision triangle. This included research evidence (with levels of study design) in one corner, and clinical judgement and patient information positioned at the other points of the triangle. All three sources of information explicitly fed into the final decision about practice. Much of the controversy about EBM (among clinicians) calmed when the value of clinical acumen and patient individuality were reinforced in the messages presented by its advocates.

“...you guys finally got it right! Now you are hitting the source of my pride (my clinical expertise) as well as the source of my guilt (that I’m so far behind in my reading). Before all you made me feel was out-of-date and guilty, and my only defence was to reject the whole idea”

(EBH mail list, a “typical response” to the 1996 Sackett et al editorial in the BMJ)

In discussion about “what counts” as evidence, some individuals identified personal conflict between the rational or professional acceptance of a set of standards, and an emotional desire to value information that seemed inferior according to those standards.

“...I've sort of learnt the hard way, that it is the figures that count. ... (it's) through I suppose these hard lessons of the apparent dichotomy sometimes between the figures and your own experience that you realise the figures are right and you can't trust your experience, unfortunately.”

Some interviewees went further to reflect on broader sociological questions about whose discourse dominates medical and public health evidence. They asked “whose perspectives and values do the EBM rules of evidence reflect and whose perspectives and values do they exclude?” For example, EBM criteria reflected value given to clinical outcomes, quantification and eliminating bias, but did not give value to narratives about disease experience. It was also felt that the EBM discourse was epistemologically one-sided in that it highlighted an epidemiological perspective on evidence and neglected other disciplines, like basic science or sociology.

Caution about EBM benchmarks for evidence was raised in relation to the transfer of criteria developed for appraising evidence in a clinical context to public health, health promotion and health policy settings. This included concern that the prevailing appraisal criteria would inadvertently limit the health issues (and solutions) considered under an evidence-based health agenda. The levels of evidence were perceived to encourage decision-makers to narrow the scope of their thinking. Thus to “undervalue” or “discount”

topics that were considered otherwise (and by others) to be important and deserving more attention or weight.

“...(I’m) not arguing against decision making generally being evidence-based, but worrying that evidence that is readily quantifiable will dominate and will then determine objectives ...there is just that slight worry about what is leading here.”

Thus critique of the EBM rules of evidence centered round concerns that they would be employed as barometers of generic value. Health outcomes that were difficult to measure or quantify, or methods of inquiry that did not produce precise measures, might be dropped from health service agendas. At the core of these debates was a concern that the new rules of evidence would inappropriately determine decisions about health care or health policy priorities.

“...like economic rationalism, the unthinking application of EBM may produce unintended and sometimes unwelcome results.”

(Letter to editor, Sydney Morning Herald, May 1997)

5.6 5 Misusing evidence

The category of *misusing* the concept of evidence encompassed four main sub-categories. I labeled these: the *evidence juggernaut*, *uncritical thinking*,

biomedical dominance, and *hidden agendas*. These sub-categories overlapped, and also related to the classifications of *undervalued* and *excluded* evidence. The four sub-categories of misusing evidence are described below.

- The “*evidence juggernaut*”

The evidence juggernaut was used to describe a scenario where research, health care objectives and policy agendas are hijacked and swept along by EBM rhetoric. This included concerns that judgements on the availability and quality of evidence, rather than patient or community values about what forms of health care were important, may end up driving decision making.

“...practice risks becoming evidence-tyrannised and experimental evidence alone cannot solve many of the problems we confront.”

The lobbying power of groups who have unlimited resources to obtain supporting evidence (commercial industries), compared to the lobbying power of those who do not have resources to generate evidence (community groups), was a deep concern. Thus caution was needed so as not to discount perspectives and ideas that were not supported by the “best” EBM evidence. It highlighted the need to include a consideration of the relative capacities to produce evidence when comparing proposals and competing arguments. This included recognising the complexity and situated nature of evidence about

public health interventions and ensuring that public health advocates have credibility when lobbying with evidence for public health resources.

- *Uncritical thinking*

This described scenarios where essential or appropriate skills of critique are not employed to evaluate evidence and its relationship to conclusions or decisions.

“We really have to think about it very carefully because the same observations could be part of a wider story that actually might turn out to be completely different, depending on which bits of evidence you could see, or which bits of evidence you chose to pursue.”

It included caution about an uncritical acceptance of evidence summaries or claims that something is evidence-based. This related particularly to evidence presented in persuasion and justification, where seemingly convincing arguments can be mounted on incomplete or flawed evidence. It also referred to the unthinking application of rules of evidence or practice guidelines, without careful scrutiny and judgement about their relevance to the context in which they are used.

Finally, uncritical thinking included mistaken judgements, such as interpreting “no evidence” as “evidence of no effect”, or interpreting “poor evidence” from inadequate evaluation as a “evidence of ineffectiveness” for discounting or removing interventions for important health issues from decision agendas.

- *Biomedical dominance*

This was a concern that resource allocation or practice decisions may be skewed by constructions of evidence that reflect only a biomedical world-view. Biomedical dominance was primarily expressed through the application of rules of evidence developed in a scientific or clinical setting, and that focus on experimentation and removal of interactions with the context. It was felt that an inappropriate transfer of EBM rules of evidence could limit discourse about evidence in a social or public health context i.e. that questions about the social and political nature of public health may not be asked. Hence this category reflected an expression of the social impact of the classifications of “undervalued” and “excluded” evidence.

Concerns about biomedical dominance also related to the inferior position of observational studies (quantitative and qualitative and hence the least biomedical) as “lesser forms of evidence” compared to experiments in evidence hierarchies. Participants agreed with the principle of making explicit the potential bias in study design where possible. Their worry about

biomedical dominance in the construction of evidence was in terms of impact on research funding allocations, and comprehensiveness of the medical and public health research portfolios that influence policy and practice decisions.

- *Hidden agendas*

This theme was about the use of evidence-based rhetoric to push hidden agendas.

“...the political process is notorious, isn't it, for making decisions that have got little to do with fact or probably evidence... and everything to do with whether or not this is going to be on balance a good or a bad thing in political terms.”

This included concerns about health service managers, bureaucrats or politicians using the language of evidence to impose agendas perceived to conflict with the “values” of clinical practice or public health, such as meeting patient needs or social equity. The agendas perceived to be shrouded in the language of evidence on effectiveness included cuts in health care funding to certain inefficient services, restrictions on clinical autonomy, and winning favour with voters.

Another area of concern was the pharmaceutical industry and their widespread funding of clinical trials. There were perceptions that a commercial agenda encouraged researchers to look more assiduously for intervention benefits to maximise profits. Rather less attention may be given to detecting side effects and harms. As one clinician remarked “use the drugs while they still work”. This was based on his description of a recurring pattern of new drugs being released with a fanfare of evidence about the benefits. A few years later the companies release further, more complete, evidence about modest long-term benefits and significant side effects.

5.6 6 Reflections

A developing social dimension of evidence

Much of the discourse about evidence-based practice in the health literature, and at conferences or teaching seminars, has been about the science of generating evidence, appraising evidence against benchmarks or criteria, and applying evidence to practice decisions. The perspectives described in the last section of my findings came from those who felt it was also important to view evidence from a social perspective, to examine the construction of the concept of evidence, and to consider the implications.

This was a response to the perception that EBM was changing the way evidence was defined, assessed and used. Some participants felt that the implications of these social changes had not been thoroughly assessed. EBM was seen to have implications for the type of research to be funded, the way decisions were made and for presenting persuasive arguments. For example, Burris (1997) has described how the rhetoric and politics of individualism (as a social value) influence the type of research that is given merit in the US at the expense of public health. There was a sense that the notion of evidence-based health care was (sometimes) interpreted too simplistically, and that the constructed nature of evidence, and the social nature of policy-making, were at risk of being underestimated or negated.

The notion that some forms of evidence may be *undervalued* or *excluded* in an evidence-based approach has become an important topic in mainstream health care literature. In his commentary on the US Guide for Community Preventive Services, Green predicted that “it is around the issue of what constitutes *relevant* and *adequate* (my italics) evidence that the Guide will face its most severe criticisms from academic defenders of emerging and alternative methodologies” (Green, 2000, p8).

A similar perspective was adopted by Ziglio (1997) when he proposed that health promotion must maintain “control over how evidence is defined”. At issue in his argument about evidence was the lobbying power of those who

promoted investment in health promotion research or practice. Health promotion advocates were at risk of losing their influence because the “medical hierarchy of evidence will state that health promotion evidence is weak, its effect equivocal, its rationale for resources unjustified.”

The category of *misusing* concepts of evidence developed in my study was underpinned by a conflict of values. The term “misuse” reflects concerns that the concept of evidence would be applied in a way that did not reflect the values held by the participants in my study. Concerns about the misuse of evidence in this study particularly revealed potential conflict between a scientific value-system that gives priority to the “quest for latent truth”, and “other” perspectives on evidence (e.g. commercial or political). In political or policy models of evidence, scientific research is collected and used as “ammunition” to argue agendas and justify ideological positions (Weiss, 1979; Tesh, 1988; Florio and DeMartini, 1993). Although participants in this study anticipated this use of evidence in the political arena, from their researcher and practitioner perspectives it was viewed with caution and suspicion. This study focused on concepts of evidence as conceived by senior academics and practitioners in medicine and public health. It is likely that for other professional or community groups, the notion of “misusing” evidence would either not arise or would encompass alternative concepts. The generalisability of these findings to other groups is discussed in chapter 6.

I proposed in section 5.4 that in medicine and public health we have been taking a progressively multi-dimensional approach to evidence. The proposition that the label of evidence reflects a judgement on what is valuable and important, and concerns that the concept of evidence could be misused, are distinctly “post-EBM” phenomena. They developed as mainstream issues as a result of growing awareness (among scientifically oriented researchers and practitioners) of the social determinants and social impact of EBM. They are being addressed in addition to the epistemic and pragmatic concerns of the benchmarked and applied dimensions of evidence, and this is a reflection of an increasingly mature and sophisticated debate about evidence.

The grounded theory on relations between forming and influencing judgements and concepts of evidence (summarised in Table 1, section 5.5) can be used to describe the changes that have taken place in professional debates about evidence in health care. I have proposed that we have moved through three distinct *eras of evidence* in health care, and that the issues outlined in section 5.6 fall under the current three-dimensional era.

Changing eras of evidence in health care

The first, one-dimensional era of evidence in health care was a period when concepts of evidence in relation to judgements about the certainty of knowledge (epistemic judgements) and in relation to practice decisions were

quite distinct, particularly in the mainstream health literature. This era encompassed the, often unsuccessful, efforts to close the research-practice gap in the 1960s and early 1970s. Although benchmarks existed for basic science, much of the evaluation research conducted on health technologies was of poor quality or irrelevant to clinical decisions. Practitioners also had no systematic means of assessing research or using the findings. The one-dimensional era was prior to EBM, and prior to the formulation at McMaster University of critical appraisal criteria for clinical research. Thus it was the period before practitioners were equipped with skills to appraise research as evidence or challenged to think probabilistically about evidence in the applied context.

The advent of explicit critical appraisal guides can be interpreted as the beginning of the second, two-dimensional era of evidence in health care. This led to the formulation of “levels of evidence” and “graded” practice guidelines, which represented an increasing overlap between the benchmarked and applied dimensions for evidence. Thus the second era slowly emerged over two decades with the influence of Archie Cochrane’s “Effectiveness and Efficiency” and the Canadian and US Taskforces on preventive clinical care (Cochrane, 1972; CTFPHE, 1979; USPSTF, 1989, 1996) and culminated in the introduction of the phrase evidence-based medicine.

During the second era much was written about what constitutes quality and rigour in both qualitative and quantitative research. These attributes were

related to the credibility of evidence. The early EBM literature promoted explicit benchmarks for reviewing evidence in clinical practice: criteria and rules of evidence to identify bias in clinical, evaluative research (Oxman and Guyatt, 1993; Guyatt et al, 1993 and 1994). But EBM also drew attention to importance of using research in an applied setting, and defined evidence as information that had direct relevance to decisions about practice, with special interest in clinical outcomes. Evidence as a basis for establishing knowledge, and evidence as a foundation for decisions about action drew closer together. The acceptance of evidence-based thinking into the mainstream has been founded on a harmony between “sagacious and intuitive” clinical experience and “logical and rational” EBM (Shaughnessy et al, 2000).

Reflections on the role of evidence in argument and persuasion (section 5.5) and refocusing judgement (section 5.6) assert strongly the constructed definition and social dimension of evidence. In debating these matters, scientifically trained researchers and practitioners began to discuss the values inherent in different notions of evidence, and the agendas that drive them. Thus current debates, and those examined in this study, incorporated the social dimension, while retaining the benchmarked and applied dimensions of evidence. As a result, I have labeled the late 1990s and early 2000s as the three-dimensional (3D) era of evidence.

A 3D view of evidence retains a belief in substantiating knowledge and evidence-based decisions about action. But the 3D era also highlights the “human factor” and the value-laden nature of any benchmarks that may be imposed and the values inherent in application decisions. As a result, judgements focus not only on the believability and utility of evidence and its link to conclusions, but also on its social construction. “The question is therefore, not what is ‘best evidence’, but how is ‘best evidence’ determined and is it really the ‘best’ evidence?” (Chan and Chan, 2000).

These are considerations that have been discussed in the light of, or in response to, EBM. In the 3D era, evidence-based health care had become the new orthodoxy: “A new form of dogmatic authoritarianism may (then) be revived in modern medicine, but the pronouncements will come from Cochranian Oxford rather than Galenic Rome” (Feinstein and Horwitz, 1997). Now it was challenged not as a radical new movement, but as a result of its rapid rise into the mainstream and the perceived dominance of its explicit rules and criteria.

In 1998, EO Wilson posited that “...there have always been two kinds of original thinkers, those who upon viewing disorder try to create order, and those who upon encountering order try to protest it by creating disorder” (Wilson, 1998, p43). The first era of evidence was marked by disorder: overwhelming volumes of research, poor quality research and inconsistent and

idiosyncratic application to practice. The second, two-dimensional era of evidence instilled some order: explicit rules and criteria, critical appraisal guidelines, systematic reviews, organised and accessible databases for those making decisions about action.

The current 3D era of evidence is distinguished by the way many have called into to question and challenged the new order established in the mainstream. It includes critique of the dominant standards emerging from EBM and a call for local contextual and relative judgments. Those debating the construction of evidence and its social dimension particularly do so as a result of their perceptions that EBM may limit diversity, flexibility and further marginalise those without resources to generate best evidence. They stress the impossibility of universal definitions and standardised benchmarks for what counts. They also lift the dialogue from epistemological and pragmatic concerns into social and political arenas, and call for a “broader, socialised definition of science” (Hart, 1997).

Hence a three-dimensional era of evidence reflects increasing integration of sociology and post-modern thought into the science of substantiation in medicine and public health. This may indicate a readiness for historical and sociocultural critical appraisals of evidence becoming mainstream (Berkwit, 1998), and for the anticipated “birth of post-modern medicine” (McDonald, 1998).

Chapter 6

FINAL REFLECTIONS AND CONCLUSIONS

6.1 POSITIONING STUDY FINDINGS

6.1.1 Defining and reconstructing evidence

This study has both examined and been a part of a social evolution in medicine and public health in the 1990s, as researchers and practitioners defined and sometimes reconstructed their concepts of evidence. When the study began, the intense focus on evidence in the context of evidence-based health care was new, and the scope of evidence and its conceptual components emerged as relatively uncharted territory.

As the notion of evidence-based practice began to dominate the health care literature and professional discourse, evidence was increasingly debated and contested, both internationally and within our Medical Faculty. It seems that such discussions were a natural outcome of EBM. Evidence-based medicine marked out new and explicit benchmarks for evidence in health care research and integrated these benchmarks into the applied context of clinical decision making. EBM has also had some influence on the type and quality of information that is considered as evidence for a persuasive argument about

health care practice. It was essential that researchers and practitioners reflected on how EBM impacted on the concept of evidence and how evidence was invoked. It was also important that they identified the limitations of emerging frameworks for evidence and the potential for their misuse.

The results presented in this study reflect and contribute to the evolving concepts of evidence. The purpose was that the findings could inform and facilitate future communication about evidence among research and practice colleagues, particularly those who teach and influence others. It is not yet clear if a commonly shared understanding of evidence can emerge, but based on these findings, I propose that recognition of the varied points of entry to the topic of evidence, and the dimensions of its appraisal, may assist meaningful debate about evidence in health care. We may be further assisted by explicit consideration of how these concepts are likely to be invoked in different types of judgement and judgement contexts.

6.1 2 Scope, generalisability and limitations of findings

The aim of this study was to develop theoretical propositions on the ways that teachers of medicine and public health conceptualised the notion of evidence. The grounded theory method allowed me to develop new concepts, to explore permutations within and among those concepts, and to formulate empirically derived propositions and hypotheses. It is important, however, to

acknowledge the limits of the scope of this research and the findings. For example, I did not seek to describe the prevalence of specific ideas about evidence or EBM, or to assess the veracity of my theories. Rather, the findings in chapter 5 are presented as a conceptual map that can be used in the following ways: to interpret or guide debates about evidence and evidence-based practice; to orientate students or interested seekers to potentially different concepts of evidence; or to plan further research.

The focus of the data collection was on interactions and people's ideas, rather than social structures and systems. As a result the findings do not address the sociological, structural and political factors that were not identified (or possibly even recognised) by those whose ideas were studied, but which may have impact on concepts of evidence. Another limitation to the scope of these findings is that because the concepts presented are based on reflections and interactions, they may not represent the "doing" of evidence in research and practice. Propositions identified in this study can be used in further observational research to examine behaviour. It is possible however that a study focused on behaviour, rather than ideas, would produce a different perspective on evidence.

My findings present concepts of evidence that were important to opinion leaders, senior academics and practitioners in medicine and public health. The core data came from interviews conducted at the Faculty of Medicine,

University of Sydney, at a time of curriculum change, and developments around evidence in the wider national and international research and practice context. Many interview participants reported that their ideas had already been influenced by the surrounding debates on EBM, and that these ideas were often consolidated within the interview process. They also reported further development in the 6-12 months between interviews, and even a year or so later. Many of the developments in concepts of evidence related to a broadening of perspective, which encompassed multiple dimensions of evidence.

The dynamic nature of reflections on evidence and the historical context in which the study was conducted have several implications for the generalisability of the findings. The concepts presented are likely to reflect the broader 1990s debates on the notion of evidence from a perspective of mainstream scientifically-trained research and practice opinion leaders and teachers of medicine and public health. The propositions developed from local interviews were mirrored and well supported by data collected from the other data sources (which reflected the views of people with similar roles and backgrounds in local, national and international settings). However, the timing of the study is significant and it is unlikely that the findings would have been the same if the study had been conducted a decade earlier.

It is not clear how these concepts will transfer into the next decade or more, although I have hypothesised that there is a trend towards a multi-dimensional view and examination of evidence. This is occurring in mainstream medicine and public health research and practice, where debates about the “normal science” of substantiation have begun to encompass the benchmarked, applied and social dimensions of evidence. The academic positions of the study participants and the research orientation of other data sources (published literature, evidence related conferences and Email discussion list) also mean that the generalisability of these findings to non-academic medical or public health practitioners is not warranted without further research in that context. It is feasible that there may be a similarity or overlap of concepts, but this would need to be tested.

Finally, the study has revealed areas of potential disagreement or controversy about evidence with people from a non-scientific or non-academic background, such as infringement of the core principles reflected in the “quest for latent truth” or obstacles to autonomous professional judgement. It is unlikely that the concepts of evidence presented here would be relevant to those who do not value scientific principles or methods of inquiry, or to people who give less priority to their independence or autonomy of judgement.

Also, the two exceptions to the primary theoretical model (section 5.5 8) on the relations between forming and influencing judgements and concepts of evidence, may turn out to be more common than was proposed in this research. They may prove to be primary approaches to the concept of evidence among other academic disciplines (some social scientists or those in Faculties of Arts and Humanities) or other professional groups.

6.1 3 My reorientation on evidence

My involvement in curriculum development, teaching, conferences, debates and evidence-based projects has been an integral part of conducting this research. This experience and that of constructing this thesis have resulted in my professional development and given rise to a reorientation of my views on evidence since 1995 when the project was conceived.

I have moved from a strongly methodological and epistemological perspective towards increasing awareness of a social perspective on the way concepts of evidence are formed and challenged. I articulated this shift in perspective at the end of the study, when I returned to the dictionary definition of evidence that was included in the introduction chapter.

Evidence: *“that which serves to make clear, obvious, certain and conclusive an understanding or judgement”* or *“proves or disproves conclusions”*.

In 1996, my first impression had been to give emphasis to the philosophical dimensions of this definition, such as certainty, conclusiveness or proof. I had reflected on the epistemology of science and the methodological requirements for generating conclusive results.

In the light of this research, I am more inclined to focus on the cognitive and social processes of judgement, and to reflect on who is making those judgements, their priorities, agendas and values, as well as external influences. This orientation is reflected in the conclusions I have reached and my recommendations for further research, which are presented in section 6.2 below.

6.2 CONCLUSION AND RECOMMENDATIONS

6.2.1 Evidence as contested ground

Professional perspectives on what constitutes admissible evidence have social implications. They have implications for the ability to get published, the relative degrees of influence with peers, colleagues and other decision-makers, and can impact on the credibility or demise of a discipline or professional group. Professional differences over evidence provoke misunderstanding and inhibit cooperation between disciplines.

Evidence that is regarded as believable and useful can influence the distribution of resources within medical schools, the allocation of research funds among competing research centres and institutes, and the funding of health technologies. In relation to medicine, Foucault developed and inspired theories about the relationship between power and knowledge (Bunton and Peterson, 1997). From his perspective all negotiations about the construction, peer recognition, and evaluation of various forms of evidence can be examined more broadly as social and political acts.

As the influence of clinical epidemiology and EBM increased, the constructions of evidence they promoted were critiqued from many directions, and the concept of evidence emerged as a new site for older battles. These battles included the struggle for prestige and resources between basic science and clinical medicine, and between medicine and public health. The rise of EBM led to probabilistic, outcome-focused research gaining ascendancy as evidence over basic medical science and qualitative experience in determining clinical practice (Marshall, 1997). These changes were reflected in suggestions that clinical professors become expert synthesisers and disseminators of EBM evidence, rather than experts in a field of basic science (Miettinen, 1998). In 1993, Eddy predicted that questions about “sufficient evidence” and the burden of proof, cost containment, and challenges to physician autonomy would become the key battlegrounds in medicine of the

1990s. His predictions came true in many debates about EBM that were examined in this research.

For the University of Sydney, EBM and clinical epidemiology are taught primarily by the Department of Public Health and Community Medicine and at the clinical schools. Also a greater proportion of the new medical program (University of Sydney Medical Program, or USMP, renamed in 2000 from the GMP) is taught by clinicians, compared to the previous undergraduate medical degree that was primarily taught by medical scientists. The jostling for curriculum time within our Medical Faculty (section 2.9) has been political not only because the outcome has educational implications, but also it impacts on the way resources are distributed within the Medical Faculty.

From the population perspective, clinical epidemiology has been described as an inadequate method to measure of the impact of programs that target the social determinants of health or which are aimed at achieving population (rather than individual) effects (Mackenbach, 1995; Sorensen 1998). And although EBM has brought a new credibility to departments of public health, some public health researchers have thought that the energy of many of the best international epidemiologists has followed the allocation of resources, and thus been siphoned away from social or population health into clinical medicine. This may have a significant impact on the future professional standing of public health research and practice.

Debates about evidence have also reflected the long-standing tug of war between quantitative and qualitative paradigms, as those in the qualitative camp have called for more recognition and credibility for interpretive and narrative forms of evidence. For example, Little (1998) has argued that by privileging epidemiology as evidence we often undervalue narrative forms of knowledge. Similar unresolved debate in public health, and questions about which methods create the “real” evidence-base (rather than the rhetoric), have led to a three-year research project based at La Trobe University (Daly, 1999, personal communication). The aim is to determine patterns of use of a variety of research methods and to assess their perceived contribution (by public health decision-makers) to knowledge and practice.

The common theme to all of these examples is the contested nature of evidence. These debates have included appraisals of evidence along all three dimensions by addressing the topic of measurement, the utility of certain types of evidence over others, and the social impact of different forms of evidence. The concept of evidence has been contested because (in the light of evidence-based health care) its construction was perceived to cause shifts in legitimacy and influence by stakeholders of medicine and public health, and thus to channel the flow of resources.

6.2 2 Recommendations

The recommendations from this research address aspects of the social dimension of evidence. The first recommendation encourages an empirical, evidence-based progression towards the integration of the social dimension of evidence with the more traditional benchmarked and applied concerns. This would be facilitated by good quality social science research into evidence-based judgement. The second recommendation relates to how researchers and practitioners use evidence in argument and persuasion.

The social science of evidence-based judgement

The bulk of the research and literature published about evidence-based health care has been on the science of evidence generation and appraisal. Less, but substantial and increasing, attention has also been given to how to promote and implement sound evidence-based findings in health services. A residual gap in the literature is the social science of evidence-based decision-making. We have begun to debate the social dimension of evidence, but this has primarily been an opinion based, rather than an empirically-based, debate.

If evidence is explicitly regarded as a matter of judgement, great scope is opened up for examining who makes judgements on evidence, what they value, and how, why, when and where their judgements are made. Thus in

addition to generating empirical research to inform judgement, we should also conduct empirical research on how judgements on evidence are influenced by competing priorities, agendas, social structures and social systems. The social science of evidence and decision-making should also examine how competing forms of evidence are weighed by different groups and in various contexts.

A few empirical studies have been conducted on the relationship between evidence and judgement. For example Dixon et al (1997) examined priority setting for a district health authority in the UK, and found that evidence had no correlation with final priority ranking decisions. Elliot and Popay (2000) have examined how policy makers use evidence, and found that evidence is more likely to have an indirect impact on policy debate than directly influence resource allocation decisions. We need more research, however, that will explore the reasons behind the judgements made, and explain existing relations (or lack of them) between evidence and decisions.

Potential for new research is implicit in organisations such as the UK National Institute on Clinical Effectiveness (NICE), where there is an explicit focus on evidence-based judgement. The interpretation of evidence is of primary concern, and evidence from different sources and of different quality is weighed and debated to formulate policy and practice recommendations. Ethnographic research in such settings could examine the social dimension of technical and applied judgments, as well as follow how the group's

conclusions are communicated to others. It should also explore the uptake of the recommendations and explain the reactions by various professional, community and political groups. Such studies would tease out not only the social elements of evidence-based decision-making, but also the determinants of successful (or not) argument and persuasion in the wider social and political arena.

Evidence in argument and persuasion

Lomas (1997a) proposed that the divide between research evidence and practice in public health and health policy is the result of different priorities and modes of working by those generating and promoting evidence and those who make program or policy decisions. Researchers presenting findings which challenge current policy can be regarded as hostile by policy makers, while research showing government success is regarded with suspicion by the research community. Researchers also often expect scientific and rational arguments (which mostly address the benchmarked and applied dimensions of evidence) to be enough to make their proposals acceptable (Orosz, 1994). Conversely, Labonte (1999) proposed that success in influencing policies will not come from arguing facts alone, but that it may come from “alliances” with groups and organisations that share the same “ethical or value base.”

The findings in this study indicate that researchers and practitioners often distinguish between concepts of evidence for different types of judgement (epistemic judgments or practice decisions) and different judgement contexts (independent judgements or argument and persuasion). Some may also feel a tension between their roles as independent researchers or practitioners and acting as an instrument of social change. Yet others have argued, particularly in public health, that decisive evidence-based social and political action has been an essential component of a researcher / practitioners' role since John Snow removed the pump handle from the contaminated water supply (Pearce and McKinlay, 1998; Labonte, 1999; McKinlay and Marceau, 2000).

The professional voices of researchers and practitioners should be heard at policy levels, not least because they produce scientific evidence and understand its strengths and limitations. We must find ways for researchers and practitioners to better engage in the wider social arena, while retaining their valued independence and autonomy (Moss, 2000; Collins and Coates, 2000). An important start would be to redress the imbalance in our training. In public health and medical training, including continuing education, the primary focus of discussions about evidence is on establishing knowledge and making decisions about program implementation. The art of being influential and having an effect on others, particularly beyond our own disciplines or professional groups, is less well understood and rarely addressed in coursework, research projects, or practice placements.

A solution would be closer relations (such as through an exchange of course units) between Masters of Public Health programs and Masters of Social Policy or Politics. Alternatively, sociology, social policy, media and cultural studies, and even tactical training in skills such as “formulating a policy proposal or media release”, should be identified as “core” subjects for both public health and medical education. A longer term objective would be that explicit reinforcement of a three dimensional view of evidence could lead to more effective ways to overcome the barriers between health research, practice and policy.

REFERENCES

- Altheide DL, JM Johnson (1994) "Criteria for assessing interpretive validity in qualitative research", in The Handbook of Qualitative Research, eds NK Denzin & YS Lincoln, Sage Publications, California, p. 485-499.
- Altman DG (1994) "The scandal of poor medical research (editorial)", BMJ **308**: 283-284.
- Anells M. (1996) "Grounded theory method: philosophical perspectives, paradigm of inquiry, and postmodernism", Qualitative Health Research **6**(3): 379-393.
- Anells M (1997) "Grounded Theory method, part I: within the five moments of qualitative research", Nursing Inquiry **4**: 120-129.
- Audi R (ed) (1995) The Cambridge Dictionary of Philosophy, Cambridge University Press, Cambridge.
- Barnes B, Bloor D (1982) "Relativism, rationalism and the sociology of knowledge", in Rationality and Relativism, eds M Hollis & S Lukes, Oxford University Press, Oxford, p. 21-44.
- Baum F (1995) "Researching public health; behind the qualitative-quantitative methodological debate", Social Science and Medicine **40**(4): 459-468.
- Bauman AO, Deber RB, Thompson GG (1991) "Overconfidence among physicians and nurses: the micro-certainty, macro-uncertainty phenomenon", Social Science and Medicine **1991**(32): 167-174.
- Bayley C (1995) "Our world views (may be) incommensurable: now what?", The Journal of Medicine and Philosophy **20**: 271-284.
- Becker HS (1995) "How I learnt what a crock was", in Studying Elites Using Qualitative Methods, eds R Hertz & JB Imber, Sage Publications, California, p. 124-130.
- Begley CM (1996) "Using triangulation in nursing research", Journal of Advanced Nursing **24**(1): 122-128.
- Berkwits M (1998) "From practice to research: the case for criticism in an age of evidence", Social Science and Medicine **47**(10): 1539-45.
- Bernard HR (1988) "Unstructured and semistructured interviewing", in Research Methods in Cultural Anthropology, Sage Publications, Newbury Park, p. 203-224.

Bero LA, Grilli R, Grimshaw JM, Harvey E, Oxman AD, Thomson MA (1998) "Closing the gap between research and practice: an overview of systematic reviews of interventions to promote the implementation of research findings", BMJ **317**: 465-468.

Bero LA, Rennie D (1995) "The Cochrane Collaboration: preparing, maintaining and disseminating systematic reviews of the effects of health care", JAMA **274**: 1935-1938.

Bogdewic SP (1992) "Participant observation", in Doing Qualitative Research, eds BF Crabtree & WL Miller WL, Sage Publications, Newbury Park, p. 45-69.

Buchanan DR (1992) "An uneasy alliance: combining qualitative and quantitative research methods", Health Education Quarterly **19**(1): 117-135.

Buchanan DR (1994) "Reflections on the relationship between theory and practice", Health Education Research **9**: 273-283.

Bunton R, Peterson, A (1997) "Foucault's medicine", in Foucault; Health and Medicine, eds Bunton & Peterson, Routledge, London.

Burris S (1997) "The invisibility of public health: population-level measures in a politics of market individualism", American Journal of Public Health **87**(10): 1607-1610.

Byrne E (2000) "The current state of evidence-based medicine (editorial)," Australian & New Zealand Journal of Medicine **30**: 309-310.

Campbell D, Stanley J (1963) Experimental and Quasi-Experimental Designs for Research, Rand McNally, Chicago.

Canadian Task Force on the Periodic Health Examination, CTFPHE (1979) "The period health examination: Canadian task force on the periodic health examination", Canadian Medical Association Journal **121**: 1193-254.

Carr-Hill R (1995) "Welcome? To the brave new world of evidence-based medicine (editorial)", Social Science and Medicine **41**(11): 1467-1468.

Chalmers A (1999) 1976 What Is This Thing Called Science, University of Queensland Press, Queensland.

Chalmers IM, Enkin M, Keirse MJ (eds) (1989) Effective Care in Pregnancy and Childbirth, Oxford University Press, Oxford.

- Chalmers IM, Haynes B (1994) "Reporting, updating and correcting systematic reviews of the effects of health care", BMJ **309**: 862-865.
- Chan J, Chan J (2000) "Medicine for the millennium: the challenge of postmodernism", Medical Journal of Australia **172**: 332-334.
- Charlton B (1997) "Restoring the balance: evidence-based medicine put in its place", Journal of Evaluation in Clinical Practice **3**(2): 87-98.
- Cochrane A (1972) Effectiveness and Efficiency: Random Reflections on Health Services Nuffield Provincial Hospitals Trust, London.
- Collins C, Coates TJ (2000) "Science and health policy: can they cohabit or should they divorce?", American Journal of Public Health **90**(9):1389-1390.
- Cook DC, Campbell DT (1979) Quasi-Experimentation: Design and Analysis Issues for Field Settings, Rand McNally College Publishing Co., Chicago.
- Corbin J (1991) "Anselm Strauss: An intellectual biography", in Social Organisation and Social Process: Essays in Honour of Anselm Strauss, ed D Maines, Aldine De Gruyter, New York, p. 17-42.
- Couvalis G (1997) The Philosophy of Science: Science and Objectivity, Sage Publications, London.
- Daly J, McDonald IG, Willis E (eds) (1992) Researching Health Care: Designs, Dilemmas, Disciplines, Tavistock/Routledge, London.
- Davis P, Howden-Chapman P (1996) "Translating research findings into health policy", Social Science and Medicine **43**(5): 865-872.
- de Bono E (1967) The Use of Lateral Thinking, Penguin, London.
- de Bono E (1972) Po: Beyond Yes and No, Penguin, London.
- Denzin N (1989) The Research Act: A Theoretical Introduction to Sociological Methods, Prentice Hall, New Jersey.
- Denzin N (1994) "The art and politics of interpretation", in Handbook of Qualitative Research, eds N Denzin N & YS Lincoln, Sage Publications, California, p. 500-515.
- Denzin N, Lincoln Y (1994) "Introduction: Entering the field of qualitative research", in Handbook of Qualitative Research, eds N Denzin & YS Lincoln, Sage Publications, California, p. 1-17.

- Dixon S, Booth A, Perrett K (1997) "The application of evidence-based priority setting in a District Health Authority", Journal of Public Health Medicine **19**(3):307-12
- Doll R (1998) "Controlled trials: the 1948 watershed", BMJ **317**: 1217-1220.
- Douglas B, Trent F, MacKinnon V, St Leger L, Mead C, Alexander K (1995) "Report to the Accreditation Working Party of the Pilot Accreditation Review Panel on its visit to Sydney Consortium for Public Health", unpublished report to the University of Sydney and University of NSW consortium, Sydney.
- Douketis JD, Feightner JW, Attia J, Feldman WF (1999) "Periodic health examination, 1999 update: 1. Detection, prevention and treatment of obesity", Canadian Medical Association Journal **160**(4): 513-525.
- Ebrahim S, Smith GD (1998) "Effects of the Heartbeat Wales programme. Effects of government policies on health behaviour must be studied (letter)", BMJ **317**(7162): 886
- Eddy DM (1993) "Clinical decision making: from theory to practice. Three battles to watch in the 1990s", JAMA **270**(4): 520-526.
- Elliott H, Popay J (2000) "How are policy makers using evidence? Models of research utilisation and local policy making", Journal of Epidemiology and Community Health **54**: 461-468.
- Evidence Base Medicine Working Group (1992). "Evidence-based medicine; a new approach to teaching the practice of medicine." Journal of the American Medical Association **268**(17): 2420-2425.
- FAO/WHO Codex Committee on General Principles (2000) Communication of the European Commission on the Precautionary Principle, Food and Agriculture Organisation of the UN & World Health Organisation, Joint FAO/WHO Food Standards Program, Paris.
- Feinstein A (1968) "Clinical epidemiology. I. The populational experiments of nature and of man in human illness," Annals of Internal Medicine **69**(4): 807-20.
- Feinstein A, Horwitz R (1997) "Problems in the 'evidence' of 'evidence-based' medicine", American Journal of Medicine **103**(6): 529-535.
- Fisher R (1926) "The arrangement of field experiments", Journal of the Ministry of Agriculture **33**: 503-513.

- Fisher R (1966) 1935 The Design of Experiments, Oliver and Boyd, Edinburgh.
- Fleck L (1979) 1935 Genesis and Development of a Scientific Fact, University of Chicago Press, Chicago.
- Fletcher R, Fletcher S, Wagner E (1996) Clinical Epidemiology: The Essentials, Williams & Wilkins, Baltimore.
- Fletcher S, Spitzer W (1980) “Approach of the Canadian Task Force to the Periodic Health Examination (editorial notes)”, Annals of Internal Medicine **92**(2), part 1: 253-254.
- Florio E, DeMartini J (1993) “The use of information by policy makers at the local community level”, Knowledge: Creation, Diffusion, Utilization **15**(1): 106-23.
- Foucault M (1980) Power/Knowledge: Selected Interviews and Other Writings 1972-1977, Harvester, London.
- Fowler PBS (1995) “Evidence-based medicine (letter)”, The Lancet **346**(8978): 838.
- Fox D (1990) “Health policy and the politics of research in the United States”, Journal of Health Politics, Policy and Law **15**(3): 481-499.
- Fox Keller E (1994) “The paradox of scientific subjectivity” in Rethinking Objectivity, ed A Megill, Duke University Press, Durham, p. 313-331.
- Freidson E (1988) Profession of Medicine: A Study of the Sociology of Applied Knowledge, University of Chicago Press, Chicago.
- Fujimura J (1991) “On methods, ontologies, and representation in sociology of science: where do we stand?”, in Social Organisation and Social Process: Essays in Honour of Anselm Strauss, ed DR Maines, Aldine De Gruyter, New York, p. 207-248.
- Giacomini MK, Cook D (2000) “Users' guides to the medical literature: XXII. Qualitative research in health care; A Are the results valid?”, JAMA **284**(3): 357-62.
- Glaser BG (1991) “In honor of Anselm Strauss: collaboration” in Social Organisation and Social Process: Essays in Honour of Anselm Strauss, ed DR Maines, Aldine De Gruyter, New York, p. 11-16.

Glaser BG (1992) Emergence vs Forcing: Basics of Grounded Theory Analysis Sociology Press, California.

Glaser BG (1978) Theoretical Sensitivity: Advances in the Methodology of Grounded Theory Sociology Press, California.

Glaser BG (1994) "The constant comparative method of qualitative analysis", in More Grounded Theory Methodology, Sociology Press, California, p. 182-196.

Glaser BG, Strauss AL (1965) Awareness of dying, Aldine de Gruyter, New York

Glaser BG, Strauss AL (1967) The Discovery Of Grounded Theory: Strategies For Qualitative Research, Aldine de Gruyter, New York.

Grahame-Smith G (1995) "Evidence-based medicine: Socratic dissent", BMJ **310**: 1126-1127.

Green L (2000) "Commentary on the Emerging Guide to Community Preventive Services from a health promotion perspective", Americal Journal of Preventive Medicine **18**(1) Supplement 1: 7-9.

Greenhalgh T, Taylor R (1997) "How to read a paper: Papers that go beyond numbers (qualitative research)", BMJ **315**(7110): 740-743.

Guba EG (1990) The Paradigm Dialog, Sage Publications, California.

Guba EG, Lincoln YS (1994) "Competing paradigms in qualitative research" in Handbook of Qualitative Research, eds NK Denzin & YS Lincoln, Sage Publications, California, p. 105-117.

Gyatt GD, Sackett DL, Cook DJ (1993) "Users' guides to the medical literature. II. How to use an article about therapy or prevention. A. Are the results of the study valid?" JAMA **270**: 2598-2601.

Gyatt GD, Sackett DL, Cook DJ (1994) "Users' guides to the medical literature II. How to use an article about therapy or prevention B. What were the results and will they help me in caring for my patients?" JAMA **271**: 59-63.

Haack S (1996) "Pragmatism" in The Blackwell Companion to Philosophy, eds N Bunnin & E Tsui-James, Blackwell Publishers, Oxford, p. 643-661.

Hacking I (1975) The Emergence of Probability, Cambridge University Press, Cambridge.

Hacking I (1983) Representing and Intervening, Cambridge University Press, Cambridge.

Ham C (1996) "Priority setting in health." in Health Policy and Systems Development: An Agenda for Research, ed K Janovsky, World Health Organisation, Geneva.

Hammersley M, Atkinson P (1995) Ethnography: Principles in Practice, Tavistock Publications, London.

Harre R (1986) Varieties of Realism, Blackwell Publishers, Oxford.

Hart JT (1997) "Cochrane Lecture 1997. What evidence do we need for evidence based medicine", Journal of Epidemiology and Community Health **51**(6):623-9.

Harvey K (1998) "Effects of the Heartbeat Wales programme. Health promotion is a waste of time and money (letter)", BMJ. 317(7162):887

Harvey Brown R (1995) "Postmodern Representation, Postmodern Affirmation", in Postmodern Representations; Truth, Power and Mimesis in the Human Sciences and Public Culture, University of Illinois Press, Chicago, p. 1-19.

Hawe P, Degeling D, Hall J (1990) Evaluating Health Promotion; A Health Workers Guide, MacLennan and Petty, Sydney.

Hawe P (1998) "Making sense of context-level influences on health (editorial)", Health Education Research **13**(4): p. i-iv.

Haynes RB, Hayward RS, Lomas J (1995) "Bridges between health care research evidence and clinical practice", Journal of the American Medical Informatics Association **2**(6): 342-350.

Herman J (1995) "The demise of the randomised controlled trial", Journal of Clinical Epidemiology **48**(7): 985-988.

Hertz R, Imber J (1995) Studying Elites Using Qualitative Methods, Sage Publications, California.

Hill AB (1937) Principles of Medical Statistics, Lancet, London.

Hirschon Weiss C (1991) "Policy research: data, ideas or arguments?" in Social Sciences and Modern States: National Experiences and Theoretical

Crossroads, eds P Wagner P & C Hirschon Weiss, et al. Cambridge University Press, Cambridge.

Hollis M (1996) "The philosophy of social science" in The Blackwell Companion to Philosophy, eds N Bunnin & E Tsui-James, Blackwell Publishers, Oxford, p. 358-387.

Holstein JA, Gubrium JF (1995) The Active Interview, Qualitative Research Methods Volume 37, Sage Publications, California.

Jamerson PA, Scheibmeir M, Bott MJ, Crighton F, Hinton RH, Cobb AK (1996) "The experiences of families with a relative in the intensive care unit." Heart & Lung 25(6): 467-74.

Johnson K, Scott D (1997) "Confessional tales: an exploration of the self and other in two ethnographies" The Australian Journal of Social Research 4: 27-48.

Jordanova L (1995) "The social construction of medical knowledge", Social History of Medicine 8(3): 361-381.

Kincheloe J, McLaren P (1994) "Rethinking critical theory and qualitative research", in Handbook of Qualitative Research, NK Denzin & YS Lincoln, Sage Publications, California, p. 138-157.

Kleijnen J, Gotzsche P, Kunz RA, Oxman AD, Chalmers I (1997) "So what's so special about randomisation?" in Non-random Reflections on Health Services Research: On the 25th anniversary of Archie Cochrane's Effectiveness and Efficiency, A Maynard & I Chalmers, BMJ Publishing Group, London.

Knorr-Cetina K (1981) The Manufacture of Knowledge: An Essay on the Constructivist Nature of Science, Pergamon Press, Oxford.

Kreigar N (1992) "The making of public health data: paradigms, politics and policy", Journal of Public Health Policy 13(4): 412-427.

Kuhn T S (1996) 1962 The Structure of Scientific Revolutions, University of Chicago Press, Chicago.

L'Abbe KA, Detsky AS, O'Rourke K (1987) "Meta-analysis in clinical research", Annals of Internal Medicine 107: 224-233.

Labonte R (1999) "Health promotion in the near future: remembrances of activism past", Health Education Journal 58: 365-377.

- Lancet editors (1995a) "Evidence-based medicine, in its place", Lancet **346**(8978): 785.
- Lancet editors (1995b) "Leap of faith over the data tap", Lancet **345**(8963): 1449-1451.
- Little M (1998) "Assignments of meaning in epidemiology", Social Science & Medicine **47**(9):1135-45
- Lindblom CE (1959) "The science of 'muddling through'", Public Administration Review **19**: 79-88.
- Llewellyn G (1997) "Grounded Theory" in Qualitative Research: Discourse on Methodologies, ed J Higgs, Hampden Press, Sydney, p. 27-37.
- Locke K (1996) "Rewriting the Discovery of Grounded Theory after 25 years", Journal of Management Inquiry **5**(3): 239-245.
- Lomas J (1997a) Beyond the Sound of One Hand Clapping: a Discussion Document on Improving Health Research Dissemination and Uptake, report commissioned by Department of Public Health and Community Medicine, University of Sydney and the Research and Development Centre of the NSW Health Department.
- Lomas J (1997b) "Research and evidence-based decision making (editorial)", Australian and New Zealand Journal of Public Health **21**(5): 439-440.
- Lupton D (1998) "A postmodern public health", Australian and New Zealand Journal of Public Health **22**(1): 3-5.
- Lyotard J (1984) The Post-Modern Condition: a Report on Knowledge, Manchester University Press, Manchester.
- MacCormack C (1992) "Planning and evaluating women's participation in primary health care", Social Science & Medicine **35** (6): 831-837
- MacCormack C (1994) "Ethnological studies of medical sciences", Social Science & Medicine **39**(9):1229-35.
- Macey D (1993) The Lives of Michel Foucault, Random House, London.
- MacKay (1991) A Dictionary of Scientific Quotations, Institute of Physics Publishing, London.
- Mackenbach JP (1995) "Public health epidemiology (editorial)", Journal of Epidemiology and Community Health **49**: 333-334.

Marjone G (1989) Evidence, Argument and Persuasion in the Policy Process, Yale University Press, New Haven.

Marshall T (1997) "Scientific knowledge in medicine: a new clinical epistemology?" Journal of Evaluation in Clinical Practice **3**(2): 133-138.

Matthews JR (1995) Quantification and the Quest for Medical Certainty, Princeton University Press, Princeton

May K (1994) "Abstract knowing: The case for magic in method" in Critical Issues in Qualitative Research Methods ed J Morse, Sage Publications, California, p. 10-21.

Maynard A (1997) "Evidence-based medicine: an incomplete method for informing treatment choices", Lancet **349**: 126-128.

Mays N, Pope C (1995) "Rigour and qualitative research", BMJ **311**: 109-112.

Mays N, Pope C (2000) "Assessing quality in qualitative research", BMJ **320**: 50-52.

McDonald IG (1998) "The last wave; medical practice must respond to cultural change", Arena Magazine **37**: 45-48.

McDonald IG, Daly JM (2000) "The anatomy and relations of evidence-based medicine", Australian and New Zealand Journal of Medicine **30**: 385-391.

McInlay, J. B. (1993) "The promotion of health through planned sociopolitical change: challenges for research and policy", Social Science and Medicine **36**(2): 109-117.

McKinlay J, Marceau L (2000) "Upstream health public policy: lessons from the battle of tobacco", International Journal of Health Services **30**(1): 49-69.

McLachlan G (1997) "Archie and the Nuffield Provincial Hospitals Trust." in Non-random Reflections on Health Services Research: On the 25th anniversary of Archie Cochrane's Effectiveness and Efficiency, A Maynard & I Chalmers, BMJ Publishing Group, London.

McMullin E (1995) "Undetermination", The Journal of Medicine and Philosophy **20**: 233-252.

McPherson K (1990) "International differences in medical care practices", in Health care systems in transition, OECD Social Policy Studies No. 7. Paris, p 17-28.

- Mechanic D (1991) "Sources of countervailing power in medicine", Journal of Health Politics, Policy and Law **16**(3): 485-506.
- Medical Research Council (1948) "Streptomycin treatment of pulmonary tuberculosis", BMJ; **2**: 769-782.
- Medawar P (1979) Advice To A Young Scientist, Pan Books Ltd., London
- Medawar P (1984) The Limits Of Science, Oxford University Press, Oxford.
- Melia K (1996) "Rediscovering Glaser", Qualitative Health Research **6**(3): 368-378.
- Melhado EM (1998) "Economists, public provision, and the market: changing values in policy debate", Journal of Health Politics, Policy and Law **23**(2): 251-263.
- Miettinen OS (1998) "Evidence in medicine: invited commentary", Canadian Medical Association Journal **158**(2): 215-221.
- Miles MB, Huberman AM (1994) Qualitative Data Analysis: an Expanded Sourcebook, Sage Publications, California.
- Minichiello V, Aroni R, Timewell E, Alexander L (1990) In-Depth Interviewing: Researching People, Longman Cheshire Pty Ltd., Melbourne.
- Mooney G (1994) "Medical practice variations", in Key Issues in Health Economics, Harvester Wheatsheaf, London, p. 113-124.
- Morabia A (1996) "PCA Louis and the birth of Clinical Epidemiology", Journal of Clinical Epidemiology **49**(2): 1327-1333.
- Morgan WKC (1997) "On evidence, embellishment and efficacy", Journal of Evaluation in Clinical Practice **3**(2): 117-122.
- Morse J (1994a) "'Emerging from the data': the cognitive processes of analysis in qualitative inquiry", in Critical Issues in Qualitative Research Methods, Sage Publications, California, p. 23-43.
- Morse J (1994b) "Designing funded qualitative research", in Handbook of Qualitative Health Research, eds NK Denzin & YS Lincoln, Sage Publications, California, p. 220-235.
- Moss AR (2000) "Epidemiology and the politics of needle exchange", American Journal of Public Health **90**: 1385-1387.

- Munson R (1981) "Why medicine cannot be a science", The Journal of Medicine and Philosophy **6**:183-208
- National Health and Medical Research Council, NHMRC (1995) Ethical Aspects of Qualitative Methods in Health Research, information paper for institutional ethics committees, Commonwealth Department of Health and Family Services, Canberra
- National Health and Medical Research Council, NHMRC (1999) A Guide to the Development, Implementation and Evaluation of Clinical Practice Guidelines, Commonwealth of Australia, AusInfo, Canberra.
- National Health and Medical Research Council, NHMRC (2000) How To Use The Evidence: Assessment And Application Of Scientific Evidence Commonwealth of Australia, AusInfo, Canberra.
- Newell D (1992) "Randomised controlled trials in health care research", in Researching Health Care: Designs, Dilemmas, Disciplines, eds J Daly, I McDonald & E Willis, Tavistock/Routledge, London, p. 47-61.
- NHS Centre for Reviews and Dissemination (1999) Getting Evidence into Practice: Bulletin on the Effectiveness of Health Service Interventions for Decision Makers, Effective Healthcare Bulletin **5**(1), University of York, York.
- Nutbeam D (1996) "Matching research to practice: fostering evidence-based approaches to health promotion", Health Education Research **11**: 317-326 (?).
- Orosz E (1994) "The impact of social science research on health policy", Social Science and Medicine **39**(9): 1287-1293.
- Osborn J (1996) "Beyond Constructivism", Science Education **80**(1): 53-82.
- Outhwaite W (1996) "The Philosophy of Social Science" in The Blackwell Companion to Social Theory, ed B Turner, Blackwell Publishers, Oxford, p. 83-110.
- Oxman A, Guyatt G (1993) "The science of reviewing research", Annals of New York Academy of Science **703**: 125-134.
- Oxman AD, Sackett DL, Guyatt GH (1993) "Users guides to the medical literature I. How to get started?" JAMA **270**(17): 2093-5.

Papineau D (1996) "Philosophy of science", in The Blackwell Companion to Philosophy, eds N Bunnin & E Tsui-James, Blackwell Publishers, Oxford, p. 290-324.

Patton MQ (1990) Qualitative Evaluation And Research Methods, Sage Publications, California.

Patton MQ (1997) Utilization-Focussed Evaluation; The New Century Text, Sage Publications, California.

Pearce N, McKinlay JB (1998) "Dissent: back to the future in epidemiology and public health: response to Dr. Gori", Journal of Clinical Epidemiology 51(8): 643-6 & 647-9.

Peterson A, Lupton D (1996) The New Public Health; Health and Self in the Age of Risk, Sage, NSW.

Plummer K (1996) "Symbolic Interactionism in the Twentieth Century: The Rise of Empirical Social Theory", in The Blackwell Companion to Social Theory, ed BS Turner, Blackwell Publishers, Oxford, p. 223-251.

Popay J, Rogers A, Williams G (1998) "Rational and standards for the systematic review of qualitative literature in health services research", Qualitative Health Research 8(3): 341-51.

Popay J, Williams G (1998) "Qualitative research and evidence-based healthcare", Journal of the Royal Society of Medicine, Supplement 91(Supplement 35): 32-37.

Popper K (1963) Conjectures and Refutations, Routledge and Kegan Paul, London.

QHR (1996) "Advances in grounded theory", Qualitative Health Research (Special Issue) 6(3)

QSR (1997) QSR NUDIST 4 User Guide: Software for Qualitative Data Analysis, Qualitative Solutions and Research Pty Ltd., Victoria.

Renton A (1994) "Epidemiology and causation; a realist view", Journal of Epidemiology and Community Health 48;79-84

Richardson L (1991) "Postmodern social theory: representational practices", Sociological Theory 9: 173-179.

Richardson L (1994) "Writing; a method of inquiry", in Handbook of Qualitative Research, eds NK Denzin NK & YS Lincoln, Sage Publications, California, p. 516-529.

Rorty R (1991) Objectivity, Relativism and Truth, Cambridge University Press, Cambridge.

Rosenau PV (1996) "Health politics meets post-modernism: its meaning and implications from community health organising", Journal of Health Politics, Policy and Law **19**(2): 303-333.

Rosenberg W, Donald A (1995) "Evidence based medicine: an approach to clinical problem-solving", BMJ **310**: 1122-1126.

Rutman L (1977) Evaluation Research Methods: a Basic Guide, Sage Publications, California.

Rychetnik L, Frommer, M (2000) A Proposed Schema for Evaluating Evidence on Public Health Interventions, discussion paper prepared for the National Public Health Partnership, Melbourne.

Sackett DL, Haynes RB, Taylor DW, Gibson ES, Roberts RS, Johnson AL (1977) "Clinical determinants of the decision to treat primary hypertension", Clinical Research **24**: 648.

Sackett DD, Haynes RB, Guyatt GH, Tugwell P (1991) Clinical Epidemiology: A Basic Science For Clinical Medicine, 2nd ed. Boston, Little Brown & Co.

Sackett DL (1995) "Applying overviews and meta-analyses at the bedside", Journal of Clinical Epidemiology **48**(1): 61-66.

Sackett DL, Rosenberg WMC (1995) "On the need for evidence-based medicine", Journal of Public Health Medicine **17**(3): 330-334.

Sackett DL, Rosenberg WMC, Gray JA, Haynes RB, Richardson WS (1996) "Evidence-based medicine: what it is and what it isn't (editorial)", BMJ **312**: 71-72.

Sackett DL (2000) "The sins of expertness and a proposal for redemption", BMJ **320**: 1283.

Saltman RB, Figueras J (1998) "Analyzing the evidence on European health care reforms", Health Affairs **17**(2): 85-108.

- Sandelowski M (1986) "The problem of rigour in qualitative research", Advances in Nursing Science **8**(3): 27-37.
- Sauerborn R, Nitayarumphong S, Gerhardus A (1999) "Strategies to enhance the use of health systems research for health sector reform", Tropical Medicine and International Health **4**(12): 827-835.
- Savitz D (1994) "In defense of Black Box epidemiology" Epidemiology **5**(5): 550-552.
- Schwandt T (1990) "Methodology; paths to inquiry in the social disciplines" in The Paradigm Dialog, ed E Guba, Sage Publications, California, p. 258-276.
- Schwandt T (1997) Qualitative Inquiry; A Dictionary of Terms. Sage Publications, California.
- Seale C, Silverman D (1997) "Ensuring rigour in qualitative research", European Journal of Public Health **7**(4): 379-384.
- Shaughnessy AF, Slawson DC, Becker L (2000) "Clinical jazz: harmonising clinical experience and evidence-based medicine", The Journal of Family Practice **47**(6): 425-428.
- Smart B (1996) "Postmodern Social Theory", in The Blackwell Companion to Social Theory, ed B Turner, Blackwell Publishers, Oxford, p. 396-428.
- Smith K, Biley F (1997) "Understanding grounded theory; principles and evaluation", Nurse Researcher **4**(3):17-30
- Smith R (1991) "Where is the wisdom? The poverty of medical evidence", BMJ **303**: 798-799.
- Sokal A (1997) "A plea for reason, evidence and logic", New Politics **VI**(2): 126-129.
- Sorensen G, Emmons K, Hunt MK, Johnston D (1998) "Implications of the results of community intervention trials", Annual Review of Public Health **19**: 379-416.
- Sox HC, Blatt MA, Higgins MC, Marton KI (1988) Medical Decision Making, Butterworths, Boston.
- Stern P (1994) "Eroding Grounded Theory" in Critical Issues in Qualitative Research Methods, ed J Morse, Sage Publications, California, p. 212-223.

Strauss A, Corbin J (1990) Basics of Qualitative Research: Grounded Theory Procedures and Techniques, Sage Publications, California.

Strauss A, Corbin J (1994) "Grounded theory methodology; an overview", in Handbook of Qualitative Research, eds NK Denzin & YS Lincoln, Sage Publications, California, p. 273-285.

Susser M (1995) "The tribulations of trials - intervention in communities (editorial)", American Journal of Health Promotion **85**(2): 156-160.

Tarnas R (1991) The Passion of the Western Mind, Pimlico, London.

Tesh S Nobel (1988) Hidden Arguments: Political Ideology and Disease Prevention Policy, Rutgers University Press, New Jersey.

Thomas H (1997) "Some reactions to Effectiveness and Efficiency" in Non-random Reflections on Health Services Research: On the 25th anniversary of Archie Cochrane's Effectiveness and Efficiency, eds A Maynard & I Chalmers, BMJ Publishing Group, London, p. 21-27.

Thomas L (1984) "1933 Medicine" in The Youngest Science: Notes Of A Medicine-Watcher, Oxford University Press, Oxford, p. 26-35.

Trostle J, Bronfman M, Langer A (1999). "How do researchers influence decision-makers", Health Policy and Planning **14**(2): 103-114.

Tudor Hart J (1997) "What evidence do we need for evidence based medicine?" Journal of Epidemiology and Community Health **51**: 623-629.

Turner B (1992) Regulating Bodies: Essays in Medical Sociology, Routledge, London.

USPSTF (1996) 1989 Guide to Clinical Preventive Services, Williams & Wilkins, Baltimore.

Van Maanen J (1988) Tales of the Field: On Writing Ethnography, University of Chicago Press, Chicago.

Van Maanen J (1995) "Trade Secrets: On Writing Ethnography", Postmodern Representations: Truth, Power, and Mimesis in the Human Sciences and Public Culture, ed R. Harvey-Brown, University of Illinois Press, Urbana, p. 60-79.

Veatch R, Stempsey W (1995) "Incommensurability: its implications for the patient/physician relation", The Journal of Medicine and Philosophy **20**: 253-269.

Wartofsky M (1978) "editorial" The Journal of Medicine and Philosophy 3(4): 265-272.

Weiss C (1979) "The many meanings of research utilization", Administration Review 39: 426-431.

Wilson EO (1998) Consilience : The Unity of Knowledge, Knopf, New York.

Winkleby M (1994) "The future of community-based cardiovascular disease intervention", American Journal of Public Health 84(9): 1369-1372.

Wittich D (1981) "New insights on the origins and social role of Kuhnian concepts: Ludwik Fleck, Genesis and development of a scientific fact", Science and Nature Number 4.

Wolcott H (1990) Writing Up Qualitative Research, Sage Publications, California.

Wolcott H (1992) "Posturing in Qualitative Research", in The Handbook of Qualitative Research in Education, eds MD LeCompte MD, WL Millroy, J Preissle J, Academic Press Inc, San Diego, p. 3-52.

Wolcott H (1994) Transforming Qualitative Data: Description, Analysis, Interpretation, Sage Publications, California.

Wooldridge M (1996) "The importance of evidence-based medicine", address by the Federal Minister for Health and Family Services, delivered at the opening of The Royal Australasian College of Physicians Annual Scientific Meeting in Canberra 1996. Published in Fellowship Affairs, Sept. 1996.

Woolf SH, DiGuseppi CG, Atkins D, Kamerow DB (1996) "Developing evidence-based clinical practice guidelines: lessons learned by the US Preventive Services Task Force", Annual Review of Public Health 17:511-38

Woolf SH, Battista RN, Anderson GM, Logan AG, Wang E (1990) "Assessing the clinical effectiveness of preventive maneuvers: analytic principles and systematic methods in reviewing evidence and developing clinical practice recommendations. A report by the Canadian Task Force on the Periodic Health Examination", Journal of Clinical Epidemiology 43(9):891-905

Wulff H (1976) Rational Diagnosis and Treatment, Blackwell Scientific Publications, Oxford.

Wulff H (1986) "Rational diagnosis and treatment", The Journal of Medicine and Philosophy **11**: 123-134.

Yoshioka A (1998) "Use of randomisation in the Medical Research Council's clinical trial of streptomycin in pulmonary tuberculosis in the 1940s", BMJ **317**: 1220-1223.

Young A (1980) "An anthropological perspective on medical knowledge", The Journal of Medicine and Philosophy **5**(2): 102-116.

Ziglio E (1997) "How to move towards evidence-based health promotion interventions", Promotion and Education **2** (IV): 29-33.

APPENDIX 1

APPENDIX 2

APPENDIX 3

APPENDIX 4

APPENDIX 5