Comparative Program on Health and Society Lupina Foundation Working Papers Series 2004–2005

Edited by
Jillian Clare Cohen
and
Jennifer E. Keelan

Comparative Program on
Health and Society
University of Toronto

Comparative Program on Health and Society Lupina Foundation Working Papers Series 2004–2005

Edited by
Jillian Clare Cohen
and
Jennifer E. Keelan

Comparative Program on
Health and Society
University of Toronto

MUNK CENTRE FOR INTERNATIONAL STUDIES
UNIVERSITY OF TORONTO

The Munk Centre for International Studies at the University of Toronto seeks to be an internationally recognised leader in interdisciplinary academic research on global issues and to integrate research with teaching and public education. We place special emphasis on the fostering of innovative interdisciplinary knowledge through the exchange of ideas and research among academics as well as the public, private, and voluntary sectors.

The CPHS Working Papers Series

The Comparative Program on Health and Society maintains a collection of academic papers which we call our Lupina Foundation Working Papers Series. These works can range from research papers to thought pieces; and from statistical analyses to historical case studies. Our series represents a snap-shot of the work being done by our Lupina Fellows, past and present. Taken together, our Working Papers Series encapsulates the wide-ranging approaches to the study of the social determinants of health. We hope that you will find the individual papers in our series thought-provoking and helpful.

Copyright

MCIS Briefings

1. *Bound to Follow? US Foreign Policy, International Reactions, and the New Complexities of Sovereignty.* By Louis W. Pauly.
   September 2005 ISBN 0-7727-0825-5.

2. *The Multilateral Agenda: Moving Trade Negotiations Forward.* By Sylvia Ostry.
   November 2005 ISBN 0-7727-0822-3.

3. *Comparative Program on Health and Society Lupina Foundation Working Papers Series 2004–2005.* Edited by Jillian Clare Cohen and Jennifer E. Keelan.
   December 2005 ISBN 0-7727-0818-5.

TABLE OF CONTENTS

Foreword

Embracing a Comparative Approach towards Understanding the Social Determinants of Health Research

The Comparative Program on Health and Society (CPHS) was founded in 2000 and is a vibrant Program which continues to evolve and thrive to meet the challenges of health research needs. The CPHS is based at the Munk Centre for International Studies at the University of Toronto. Supported by the Lupina Foundation, the CPHS supports innovative, interdisciplinary, comparative research on health, broadly defined through an extensive range of fellowships. The Program builds on the scholarly strengths of the University of Toronto in the social sciences, humanities and public health. Three themes underpin the research which the CPHS supports. These are:

1) **Socioeconomic Status and Health Outcomes**: The role of public policy and civil society groups in mediating the relationship between income inequality, socio-economic status and health effects.

2) **Socioeconomic Status and Access to Health and Health-related Services:** The impact of the relationship between demographic, gender or socio-economic factors, the level and distribution of public and private investment in these areas and the design of programs on access to health and health related-services by various groups in society.

3) **Accountability Mechanisms in the Governance of Health and Health-related Services:** Research will consider lines of accountability linking patients, providers, payers, investors and citizens under different institutional arrangements.

The papers contained in this Working Paper Series offer us innovative ways of thinking about health research issues; in many cases, pushing us to re-examine our current methods of thinking.

McGuire's paper outlines the competing methodologies involved in the assessment of evidence in public health. Beyond critiquing evidence based medicine (EBM) as a foundation for assessing evidence to set public health policy, McGuire offers a new synthesis based on critical realism: a theoretical model for envisioning the multiple layers of social, institutional and cultural factors from which diseases emerge, and are socially reified. Critical realism hinges on making epistemological distinctions between the open systems of the realm of policy making and the closed systems of pure medical research. Medical research is thus liable to profound distortion if the co-production of health is ignored.

Complementing McGuire's analysis, Aguinaldo's paper tackles the debates within social science on the epistemology of qualitative research evidence. Aguinaldo offers a new discursive approach to analysing subject's testimony that both acknowledge the role of interviewer, subject, and the artificial medium of the interview itself. His model offers concrete tools for analysis and a rich interpretive framework for critical

review of qualitative research. The applications of this analytic model may reach beyond qualitative research into the realm of clinical and empirical research by making physicians and medical researchers more aware of their role as participant-observers collecting data, even in routine clinical settings.

Borgerson's critique of EBM takes a different approach by exploring the conceptual boundaries imposed by the dominant system of medicine on alternative medical practitioners. Alternative medical practitioners face an essentialist conundrum when they attempt to achieve legitimacy by adhering to the possibly incommensurable normative values of orthodox medicine, for example by submitting their practices to randomised, controlled experiments. By drawing analogies to feminist writings, Borgerson argues that alternative practitioners are currently confronted with a choice between appeals to sameness or difference in the design of research programs. Borgerson argues that both paths are self-limiting, and that alternative practitioners should actively engage to reshape the normative standards themselves to include their perspective.

Normative category formation is also the subject of Daley's paper. Daley highlights the barriers that gays and lesbians experience when trying to access health care in Canada. She makes a case for using the existing rights-based notions surrounding citizenship and introducing into this framework the concept of sexual citizenship. Daley makes a strong argument that this analytic tool offers the most seamless opportunity to integrate the health issues of gays and lesbians into existing health policy structures while encouraging better representation in medical research and equitable access to heath resources.

In a different approach, Kapiriri also deals with issues of access but her analysis focuses on models of priority setting. Like Borgerson and McGuire, Kapiriri is skeptical that a model focusing on specific technical fixes, such as EBM, is either sufficient or desirable in such a value-driven and culturally contingent process. Madden et al.'s paper explores new territory by investigating, from a participants' perspective, what the current priority-setting goals are in diverse Canadian health care settings, and the means of evaluating, constructing and achieving an ideal priority setting process. Their work identifies the boundaries and contexts of priority-setting in government, public clinics, and hospitals providing a framework for future empirical and qualitative research. They found that participants equated good priority setting with formal systems and processes that would help identify and address gaps between the priorities of different stakeholders, whether they are part of institutions, or a broader community. Participants were also dissatisfied by the eclectic approach to priority setting in Canada suggesting that health policy researchers can play a key role in creating new models and providing a much needed professional discourse to systematise priority setting in Canada.

Powell provides us with a history of priority setting by examining the institutional and political features of public health care in Ontario. Staking out the political and institutional loci of public health expertise and power in Ontario, Powell sets the scene for the Walkerton water contamination scandal and the SARS epidemic, both watersheds in the history of public health in Ontario. Coelho's paper also describes a significant public health issue, the obesity epidemic, and surveys the underlying mechanisms and strategies used by people to lose weight.

Lastly, Keelan and McLeod look at two very different cases of conscientious objection. In Keelan's case study of resistance to compulsory vaccination, individuals are currently permitted to decide, based on personal conscience, to opt out of routine childhood immunizations. Keelan argues that historical research points to an upper boundary for a realistic level of vaccination compliance in a modern democracy. This small but persistent level of resistance will likely prevent a significant improvement in vaccination coverage and will be resistant to educational campaigns that focus on explaining the risk/benefits of vaccination from the expert's perspective. Finally, McLeod's research explores the right of physicians to refuse to refer patients for abortion based on the principles of conscientious objection. McLeod argues that the physicians, under obligation to provide care, should be required to make referrals for abortion even against their conscience. However, the complex tension between conscience and professional obligations as defined by social norms defied unilateral philosophical models.

We hope that you find this series of papers thought-provoking and sources of knowledge that can help us understand better the social determinants of health.

Jillian Clare Cohen and Jennifer E. Keelan.

Acknowledgements

Contributors

Jeffrey P. Aguinaldo is a Lupina/OGS doctoral fellow working towards a Ph.D. in the Department of Public Health Sciences at the University of Toronto, and has a M.A. and B.A. in Psychology from Wilfrid Laurier University and York University respectively. His research focuses on the interpretive practices of gay and bisexual men in talking about their health as well as talk and texts of social scientists influential in the domain of gay men's health.

Kirstin Borgerson is a Lupina/OGS doctoral fellow and is a PhD candidate in the Department of Philosophy at the University of Toronto. She completed her MA at the U of T, and received her BA from the University of Saskatchewan. Kirstin's research interests include philosophy of science, bioethics, feminist philosophy, and epistemology.

Jennifer Coelho is a Lupina/OGS doctoral fellow and is a PhD candidate in the Department of Psychology at the University of Toronto. She is specializing in Social/Personality/Abnormal Psychology, and is completing a clinical extension in Eating & its Disorders. As part of this clinical extension, Jennifer recently completed a clinical research practicum in Australia, working as a psychology intern at an eating disorders clinic affiliated with the University of Sydney. Her doctoral research is currently focusing on food cravings in chronic dieters and non-dieters, and the effects of exposure to food cues on mood and eating behaviour.

Jillian Clare Cohen, PhD is the Director of the Comparative Program on Health and Society and Assistant Professor at the Leslie Dan Faculty of Pharmacy, University of Toronto. Dr. Cohen is involved in research which focuses on access to medicines for the poor and has advised Governments world-wide on drug policy issues. She is the author of numerous articles and opinion pieces on drug policy issues.

Andrea Daley is a Lupina/OGS doctoral fellow and is a PhD candidate in the Faculty of Social Work and is enrolled in the Collaborative Program in Health Services and Policy Research at the University of Toronto. Andrea received her B.S.W. and M.S.W. at York University (Toronto). Her research interests include the invisibility of marginalised groups within health care policy and program delivery, women and serious mental illness, and sexuality and identity.

Sarah Downey is the director of the Planning and Performance Measurement, University Health Network.

Lydia Kapiriri is a CPHS & Joint Centre for Bioethics Post-Doctoral Fellow at the University of Toronto. She is a medical doctor with a Masters degree in Public health from the Royal Tropical Institute, Amsterdam and a PhD from the University of Bergen, Norway. Her research focuses on various aspects of health systems and of late specifically on priority setting. Her PhD thesis entitled "Opportunities for Public Accountability in Priority Setting in Health: the Case of Uganda" addresses various dilemmas faced by decision makers in developing countries and also looks at the values, the actors and processes of priority setting in Uganda. Her postdoctoral project focuses on a comparative study of priority setting processes at three levels in Uganda, Canada and Norway.

Jennifer Keelan is a CPHS and SSHRC Postdoctoral Fellow affiliated with the University of Toronto, The Wellcome Trust Centre for the History of Medicine, UCL, London and The Kennedy School of Government, Harvard University. She completed her PhD in History of Science and Medicine at the University of Toronto's Institute

for the History and Philosophy of Science and Technology, under Professor Pauline Mazumdar. Her research focuses on how diverse social groups experience and understand risk and how this intersects with popular notions of medical authority and expertise in the public sphere.

Shannon Madden is a Lupina/OGS doctoral fellow and a PhD candidate in the Department of Health Policy Management and Evaluation and the Joint Centre for Bioethics (JCB). She completed her MSc at U of T and her BHSc at the University of Western Ontario. Shannon is involved in bioethics as it relates to health policy. More specifically, her interests include priority setting in health care within Canada and around the world.

Douglas K. Martin, PhD is the director of the Collaborative Program in Bioethics, Assistant Professor, Department of Health Policy, Management and Evaluation and the Joint Centre for Bioethics, University of Toronto;
Career Scientist, Ontario Ministry of Health and Long-Term Care, Health Research Personnel Development Program.

Wendy McGuire is a Lupina/OGS doctoral fellow and a PhD candidate in the Department of Public Health Sciences (Social Science and Health Program) at the University of Toronto. She completed her MSW at the University of Toronto and her BA (Sociology) at the University of Guelph. She currently works as the Knowledge Transfer Coordinator in the Department of Health Policy, Management and Evaluation. Her research interests include the social determinants of health and illness, health inequalities, the role of theory in health research, and the relationship between knowledge, policy and practice in the area of public and population health.

Carolyn McLeod is the CPHS New Faculty Fellow for 2004–5 and is an Assistant Professor of Philosophy at the University of Western Ontario. She completed her PhD in Philosophy at Dalhousie University. Carolyn received both her Master's and Bachelor's degrees in philosophy from Queen's University. Throughout her career, she has focused on philosophical issues in health care ethics. Both her PhD and Master's dissertations were in reproductive ethics. With her recently published book, Self-Trust and Reproductive Autonomy (MIT Press 2002), she was able to bring together interests she has in reproductive ethics, philosophical moral psychology, and feminist theory.

Mary Powell PhD is the 2004–5 CPHS Visiting Scholar. She is an associate professor of political science at Laurentian University in Sudbury, Ontario. She teaches in the fields of public administration, gender and politics, and health policy. Mary has also worked in government and non-government organizations as a specialist on social policy and decentralization. With Laurentian colleagues, Mary has conducted research on the experience of women in blue-collar jobs in mining, focusing on Inco, one of the world's largest nickel producers, and a major Sudbury employer. Chief among her research interests is public health in Ontario. In her current project she is analysing recent crises in public health (withdrawal of provincial funding, the Walkerton crisis, and SARS) and their historical antecedents.

Peter A. Singer MD, MPH, FRCPC is the Sun Life Financial Chair in Bioethics and Director, University of Toronto Joint Centre for Bioethics Professor of Medicine, University of Toronto Canadian Institutes of Health Research Investigator Director, PAHO/WHO Collaborating Centre for Bioethics Program Leader, Canadian Program on Genomics and Global Health.

Epistemological Approaches to Qualitative Data Analysis in Gay Men's Health

Jeffrey P. Aguinaldo

This discussion paper compares and contrasts two epistemological approaches to the analysis of self-report data in the health sciences.[1] I consider these approaches within the context of my own research that relies on data derived from in-depth qualitative interviews on gay men's health. While it explores the assumptions upon which health researchers 'do analysis', this paper is broadly linked to concerns with what is taken as evidence of the social world (c.f. McGuire, this volume). In the first section of this paper, I discuss the most pervasive epistemological approach to the analysis of self-report qualitative data in the health sciences. This approach assumes that data gained from qualitative interviews can offer access to the social world. I illustrate the merits of this approach drawing upon my own research in gay men's health — research that was dedicated to privileging the voices of research participants. However, many researchers who adopt this approach must inevitably contend with a host of issues that arise when confronted with data that advance competing or contradictory self-reports. In the second section, then, I briefly discuss these issues. In the final section, I describe an alternative epistemological approach to the status of participants' self-reports. What is emphasised in this approach is the performative qualities of participants' talk. The implications of this approach will be discussed.

In many ways, the concerns I identify in this paper map onto the vast body of feminist and queer theorising (and politicising) of research practice (e.g., Ristock & Taylor 1998, Reinharz 1992, Ristock 1998, Ristock & Pennell 1996, Honeychurch 1996). For example, Kitzinger (2004) states, "listening to women's voices and validating women's experiences remains central to the feminist qualitative research

enterprise" (p. 126). For feminist researchers, such as Kitzinger, descriptive interviewing remains the central method for accomplishing this goal. Likewise, qualitative health researchers and especially those who, like myself, conduct research with marginalised and oppressed social groups have used qualitative interviews 'to gain participants' perspectives', 'to allow participants to speak on their own behalf', and 'to validate their experiences'. For health researchers committed to the principles of community based participatory research (CPBR), privileging the perspectives of the researched and validating their experiences are central to the research agenda. Presenting what research participants say as accurate representations of their social world is one way of privileging participants' interpretations of their experiences above and beyond the interpretations offered by so-called health experts. However, as Kitzinger (2004) argues, understanding what is involved in the process of listening to what participants say is no longer a straightforward task. Feminist social scientists have identified and criticised the competing analytic approaches to qualitative data and the ways in which these approaches guide and direct our research claims (e.g., Wilkinson 2000, Wilkinson 2004). It is from these discussions and the critical insights they provide that I bring to bear on my research on gay men's health.

Self-Reports as Route or Resource

The most pervasive epistemological approach within qualitative health research is to assume that what participants say can be used as a route or resource to a taken-for-granted social world. As Wilkinson (2004) states, it is an approach "in which research participants' talk is taken as providing a 'means of access' to something that lies behind or beyond it" (p. 187). I have taken this epistemological approach in my own research with the goal of providing concrete suggestions for public health interventions or achieving particular political aims.

My most recent example of this comes from a collaborative research project aimed at exploring the role of substance use and HIV infections among gay and bisexual men. The investigators of this project were concerned with substantiating the claims made in the established epidemiological literature. Although it is widely understood that needle sharing is a common route for HIV infection, epidemiological studies continue to implicate alcohol and non-injection drug use as a likely contributor to HIV transmission. These

studies have mainly involved statistical associations between substance use and HIV infections among gay and bisexual men (e.g., Penkower et al. 1991, Burcham et al. 1989, Weber et al. 2001, Page-Shafer et al. 1997, Molitor et al. 1998, Chesney et al. 1998, Ostrow et al. 1995).

Reviews of this literature continue to put forth the notion that substance use inevitably prevents gay and bisexual men from implementing HIV protected sexual behaviours (Stall & Purcell 2000). Some authors have asked us to consider "the value of treatment for alcohol and other substance abuse problems as an HIV-prevention method among MSM [men who have sex with men]" (Shoptaw & Frosch 2000, p. 193). This has led researchers, such as Stall and Purcell (2000) to conclude that, "studies that demonstrate a link between higher levels of non-intravenous substance use and later HIV infection may be most relevant to considering whether the relationship between substance use and high-risk sexual behaviour is an appropriate part of AIDS risk reduction efforts" (p. 187).

The investigators of the study I discuss here chose to conduct in-depth qualitative interviews with gay and bisexual men in the hopes of obtaining self-reports of unsafe sex practices resulting from the men's drug or alcohol use. As the data analyst hired on for this project, my interest in the voices of the men we interviewed were quite different. Based on my commitment to qualitative methodology, I was immediately critical of the types of decontextualised statistics that have been most often used to construct HIV vulnerability among gay and bisexual men who use alcohol and drugs. My scepticism was not so much based on the belief that the existing literature 'got it wrong'. Rather, I was sceptical of the epistemological privilege given to accounts put forth by epidemiologists and social scientists above those from HIV positive gay and bisexual men.

Working from an implicit commitment to harm reduction principles, as well as an acute awareness of the stigmatisation that substance users and gay and bisexual men deal with everyday, my goal was simply to act as a conduit for the voices of participants. By asking and privileging how the men themselves understand the 'cause' of their HIV infection, I wanted to foreground their interpretations rather than presume their experiences based on the epidemiological literature. I felt that validating the accounts the men provided would eventuate in

more effective health promotion campaigns directed towards the gay and bisexual male populations who use recreational substances.

When the majority of men told us that they did not believe that their substance use had any consequences on their risk behaviours, I simply reported this to be the case. Only a small minority of the men interviewed reported affirmative associations between their substance use and their seroconversion. The vast majority reported one of four alternative explanations to account for their seroconversion: (1) they became infected with HIV through sexual behaviours they did not know were unsafe; (2) their negative emotional state lead them to 'not care' and engage in unsafe sex behaviours, which in turn led to their seroconversion; (3) their trust in their partners led them to take risks they otherwise would not have taken; and, (4) they became infected with HIV as a result of a sexual assault.

These findings do not differ substantially from the HIV prevention literature on gay and bisexual men more generally and suggest that gay and bisexual men who use substances for recreational purposes would most likely benefit from HIV prevention campaigns designed for *gay or bisexual men* rather than *substance using* men. Conceivably, then, an explicit focus on the substance use of gay and bisexual men as the sole account of their seroconversion is, at the very least, misleading for effective HIV prevention.

For the purposes of this paper, I bring to light the commitments I adopted for this research and the way in which they resulted in altogether different public health implications than those of the epidemiological literature. Taking participants' talk unproblematically brought to the fore experiential accounts of men whose perspectives have not been typically appreciated in the public health literature on HIV risk and substance use. Whereas the vast bulk of epidemiological HIV research has promoted abstinence, or reduction of substance use, as the most effective strategy for preventing HIV transmission, the men we interviewed told us to refocus prevention efforts to other areas of intervention — unrelated to their substance use — that they believed to be more appropriate.

Competing Versions

Treating participants' talk as a reflection of a taken-for-granted reality more easily translates into concrete suggestions for public health interventions. However, those who commit to this epistemological

approach must necessarily assume that there can be only one account that can 'accurately' reflect the 'true' state of things presumed to lie beyond the talk. Dilemmas arise when the researcher is confronted with two (or more) competing accounts of the same event under investigation. This can occur in two ways. The first arises when participants provide contradictory or inconsistent self-reports of the same event or experience (also known as 'multiple versions', Wilkinson 2000, or 'variability', Potter & Wetherell 1987) during their interviews. The second arises when participants offer accounts of experiences, behaviours, or psychological states (e.g., emotions, attitudes, beliefs) that may not coincide with the accounts of the researcher. These will be discussed in turn.

It is not uncommon for a single participant to offer different and contrasting versions of the same experience, behaviour, or belief that they are asked to report about during health research interviews (Wilkinson 2000). To illustrate this, I draw upon qualitative data from another project I was involved in that explored the relationship between substance use and HIV infections among gay and bisexual men. Similar to the study discussed above, the vast majority of participants vehemently rejected the commonly held view that their seroconversion was the result of risk behaviours caused by drug or alcohol use. However, on occasion, an interviewee would make contradictory reports during the *same* interview. For example, in accounting for his seroconversion, Thomas[2] changes his account two times during his interview (the arrows below indicate relevant sections):

Excerpt 1:

Tina: ... *could you tell me about your alcohol and, or your drug use at the time you become infected with HIV.*

Thomas: → *And so no, ah, no major, ah, ah, role, I think, that*
            → *drinking played in*

Tina: *Right*

Thomas: → *in contributing to my HIV infections*

In Excerpt 1, Thomas reports that his alcohol use played 'no major… role' in contributing to his HIV infection. However, in a subsequent

2. All names used here are pseudonyms.

account (Excerpt 2), Thomas concedes that alcohol 'contributed in some way' to his unsafe sex practice that lead to his HIV infection. In a final account (Excerpt 3) offered near the end of his interview he, again, asserts that substance use did not contribute to his unsafe sex practice at the time of his seroconversion.

Excerpt 2:

Thomas:		*My guard was down at the time [I became infected]. It was unsafe sex*
Tina:		*Hum*
Thomas:		*whichever it was*
Tina:		*Right*
Thomas:	→ →	*And, ahm, it, you know, I'm sure alcohol contributed to some degree in that.*

Excerpt 3:

Thomas:		*Like here I am HIV positive. So obviously*
Tina:		*Sure*
Thomas:		*I wasn't practicing safe sex at the time.*
		[lines deleted]
Thomas:	→ →	*So, ahm, no, it, it didn't. I don't think substance use contributed in that way.*

Self-reports such as these, offered by the same participant, present an analytic problem if one takes the epistemological stance that what people say as unproblematically representing their experiences or behaviours.

Competing version of events can also arise when research participants offer accounts with which the researcher/analyst may be in profound disagreement. Feminist psychologist, Celia Kitzinger (2004) offers a cogent example of this in writing reflexively about her doctoral research in the early eighties on lesbian identities. Wanting to offer representations of lesbians outside of then dominant constructions of homosexuality as pathological and sick, Kitzinger conducted

descriptive interviews as a means to give voice to lesbian experiences. The dilemmas of representing lesbian experience became self evident upon hearing participants' self-reports that simply reinforced representations (of lesbianism) that Kitzinger was attempting to counter. She offers two rather lucid examples of the types of data that she found extremely problematic and that challenged her position as a researcher committed to validating the women's voices.

> I suspect we [lesbians] are in a slightly retarded state. Well 'retarded' is perhaps not quite right. It's a fear, an inability to relate to the opposite sex. There's nothing you can do about it (Jane, quoted in Kitzinger 1987, p. 119).

> Lesbianism is not something you choose: not something anybody in their right mind would choose. But, if you're stuck with it, then you just have to put up with it, and live your life with as much dignity as you can. Certainly things aren't helped by exhibitionists who run around screaming about their lesbianism and somehow link it to politics, as though you could vote Labour Conservative or Lesbian (Lynne, quoted in Kitzinger 1987, p. 141).

These types of data and those reviewed before place the qualitative researcher in the precarious position whereby s/he must dismiss at least some aspect of participants' talk in order to put forth a coherent narrative (see Kitzinger and Wilkinson 1997 for a number of strategies qualitative researchers use to 'invalidate' participants responses). Flicker (2004) articulated this challenge quite succinctly in writing about her process of analysing a set of interviews conducted for a HIV prevention research project focussing on the experiences of positive youths. On the one hand, she wanted to remain faithful to her commitment to CPBR principles, which committed her to honouring what research participants told her. On the other, she was doubtful of the truthfulness of one of her research participants, 'James,' whose self-reports of his experiences with HIV related illness did not coincide with the established HIV/AIDS literature. Flicker writes,

> Given that this interview took place in the context of community-based participatory research project, the question of what to do with James's story was brought back to the larger stakeholder group (of HIV positive youth and community-based organization

representatives) that designed the study. The discussions lead to a lengthy debate about what it means to be the "arbitrator" of truth (Flicker 2004, p. 534).

Ultimately, they decided to 'compromise'. The stakeholder group chose to include James's narrative in the analyses, but was very attentive to which pieces of data would be included in the final report. While this 'resolution' might seem sensible, it raises considerable concerns about the basis upon which data analysts would choose to disregard some aspects of what someone shares with them during the research interviews.

Self-Rreports as Topic: Constructionist Approaches to Talk

In this final section, I present an alternative epistemological approach to self-reports and qualitative data more generally. This approach has developed in light of spirited discussions in the social sciences on constructionist approaches to language (see Edwards et al. 1995, Gergen 1999, Burr 1995). It considers talk as *constituting* the social world (as oppose to *reflecting* it) within the interactional context. By this, it is meant that "participants build the context of their talk *in* and *though* that talk itself, on a moment by moment basis" (Wilkinson 2004, p. 188, italics in original). From this perspective, self-reports are analysed based on their social function within the context from which they are generated rather than their capacity to infer what is presumed to lie beyond the reports. This approach thus side-steps the dilemma of competing accounts because it is premised on the assumption that self-reports are endogenously produced to accomplish particular interactional goals, rather than as a route to what has 'actually' happened.

It has become commonplace to acknowledge that what participants say in research interviews are collaboratively produced. What participants share often depends on the local interactional context (e.g., who they are talking to, what they have been asked, what the research interviews are about). However, while this collaborative process is often acknowledged in qualitative methodology texts and book chapters (e.g., Holstein and Gubrium 2004, Holstein and Gubrium 1995, Gubrium and Holstein 1997), it is less often demonstrated and incorporated empirically into qualitative health studies. In most cases, qualitative researchers present what participants say as though they are produced in a social interactional

vacuum. Yet, as Sue Widdicombe (1995) states, "accounts generated through interviews, whatever else they may be doing, are primarily produced to address the interactional business deemed relevant to the particular circumstances" (p. 110).

In exploring these ideas for my doctoral research on gay men's health, I have adopted a discursive analytic framework that draws from the interdisciplinary fields of discourse (Edwards and Potter 1992, Edwards 1997, Potter 1996) and conversation analysis (Ten Have 1999, Hutchby & Wooffitt 1998). Discursive analysis takes as its empirical focus to assess not how adequate research participants' self-reports map onto what happens 'out there', but how the talk itself is something relevant. In this sense, the research interview context and the self-reports that are generated within it are the objects of observation. To compare and contrast the two epistemological approaches reviewed in this paper, I take as an example self-report data used to illustrate the notion of 'negotiated safety', a health behavioural phenomenon cited commonly in the gay men's health literature.

According to Kippax (2003), the return of unprotected anal intercourse, which re-emerged as a common sexual practice among gay men during the nineties, was thought by some researchers to be an indication of relapse. However, other researchers believed the men were exercising a deliberate form of safer sex practice. In rejection of the 'condom use every time' strategy urged by medical models of prevention, gay men willingly engaged in unprotected sexual intercourse with 'regular' sex partners or in committed relationships where their sero-negative (and concordant) status was known and shared. Thus, 'negotiated safety' represents an agreement to dispense with condoms and to engage in sexual decision-making (and risk taking) based on shared knowledge between committed partners. Upon its conception, negotiated safety became a popular explanation among qualitative health researchers to account for self-reports of unprotected anal intercourse among gay men.

Adam, Husband, Murray, and Maxwell, (2003) conducted 70 in-depth interviews with men about their sexual behaviours and their HIV risk protective behaviours. Fifteen "followed the general prescriptions of 'negotiating safety'" (p. 17). To illustrate this, Adam et al. wrote, "negotiating safety means permitting unprotected sex within the relationship but maintaining a cordon sanitaire around it by applying

protective measures outside of it" (p. 17). They use the following participant excerpt to support their claim:

> When I said to you that I never have sex without a condom, this is if I do it outside of my relationship with my partner, but with my partner I do not use a condom. (40s, Latin American, HIV-, 4 years[3])

Offering only a simple introductory statement, Adam et al. (2003) treat the excerpt as self-evident of 'negotiated safety.' The differential implementation of condom use reported by the speaker is assumed to represent an agreement between the speaker and his partner that condoms need only be implemented in sexual encounters outside their relationship. Presumably, such an agreement would render unprotected sexual intercourse with the speaker's partner as 'safe'. Moreover, the presentation of this excerpt is treated as a testament to how the participant behaves during his sexual encounters. We are left to assume that 'outside my relationship with my partner', the participant will indeed 'never have sex without a condom'. By contrast, condoms are not used within the relationship. It is in this way that the data are treated as 'transparent' as a medium through which the participants' condom use (i.e., experience) can be inferred. Within the qualitative health literature, these sorts of reports are collected and presented in aggregate form as evidence for what Adam et al. refer to as 'negotiated safety'. Of course, my intention here is not to dismiss or discount the findings of Adam et al. In fact, I believe it very likely that gay men do enter into negotiations about their condom use with their partners. However, to leave the data analysis at that is to commit an interpretative gloss, one that skims over important details of the talk. From a discursive analytic perspective, the goal, instead, is to explore the ways in which people manage their accounts in relation to the interactional demands at hand.

A closer fine-grained (i.e., discursive) analysis indicates that this stretch of talk was produced within an interaction, but also allows for initial observations as to the function it may have served within that interaction. The 'When I said to you' displays the speaker's orientation to his previous utterance as a shared localised interaction with the interviewer ('to *you*'). The previous utterance referred to here by the speaker ('I never have sex without a condom') is an instance of reported speech — its actual occurrence is not provided by the data

3. Indicating the duration of the participant's primary relationship

analysts. Here the speaker constructs this utterance as an extreme case formulation[4] ('*never* have sex without a condom'). The phrase followed is marked by some ambiguity: The 'this' may refer to the previous utterance (i.e., the extreme case formulation) however, the referent 'it' is unclear (though given its context may refer to 'having sex'). The 'if' constructs the 'I do it outside my relationship with my partner' as a conditional ('If I do it outside of my relationship', then 'I never have sex without a condom'). Through the introduction of the contrast ('but with my partner I do not use a condom'), the speaker offers when he would, in fact, forego condom use. From these initial observations, further analytic claims can be made.

Reporting previous utterances projects what the speaker will subsequently modify.[5] In the excerpt, the speaker treats the extreme case formulation of his condom use as one in need of clarification. An extreme case formulation can work up a compelling statement and it is presumably used here to substantiate the speaker's consistent use of condoms during sexual intercourse. However, at the same time extreme case formulations can be open to challenge as implausible (Consider, 'You never had sex without a condom?'). The speaker thus offers a more moderate version of his condom use through the conditional as well as with contrasting circumstances when he would in fact have sex without a condom. In other words, he modifies his account from an extreme to a more hedged version that acknowledges instances of unprotected sex. A tentative analytic claim may be that this stretch of talk (that Adam et al., take as evidence of 'negotiated safety') is designed as a concession, one that can head off a provoked or anticipated challenge to the speaker's previously reported (i.e., challengeable) version of his condom use.

This is not to suggest that the speaker previously falsified his accounts of condom use. To make such a claim is to consider oneself in a position to know the speakers 'actual' condom use behaviours. Moreover, this analytic work-up is tentative. As is usually the case in many presentations of qualitative research findings, we are not given any of the immediate (discursive) context to establish how the utterance was provoked or to what interactional effect. In the Adam et al.'s (2003) study, the interviewer is asking the interviewees to explain their safer

4. Pomerantz (1986) studied the conversational use of words such as "always", "every", and "never", (e.g., "I always do that", "Everybody does it", or "That never happens to me") to which she applied the technical term "extreme case formulations".

5. I would like to thank Linda Wood for this observation.

sex practices. Within such a context, reports of unprotected sex are an accountable matter — such reports are in need of explanation or rationalisation. However, reports of 'always having sex with a condom' (or conversely, 'never having sex without a condom') can be challenged as implausible. The concession above could be invoked to manage the speaker's identity as 'one who practices safe sex' in the face of a potential challenge to his 'never having sex without a condom'. Such a challenge would necessitate a modification in the speaker's account; one that concedes to having sex without a condom, but in such a way that still constructs the speaker's identity as 'safe'.

The analysis so far offers some implications for public health practice. For example, traditional medical models of HIV prevention targeting gay and bisexual men have constructed 'safer sex' as consistent condom use 'always', 'all the time' and 'every time'. This discourse practice creates a central irony: Gay and bisexual men must report unfailing and unwavering condom use during sexual situations that pose some risk for HIV infection (or else be deemed 'irresponsible', 'unsafe', or 'risky'). The irony is that by doing so, gay men's reports of their condom use might be met with scepticism, derision, and disbelief. Although these insights should not be overstated, the discussion here warrants a detailed exploration of the ways in which gay men talk about substance use and HIV risk and how this talk can sustain broader social practices.

In sum, the epistemological approach to qualitative data discussed here and discursive analysis offers an altogether alternative framework to conceptualise and analyse self-reports gained from qualitative research interviews. This approach considers participants' talk as action-oriented within the local context it is generated and thus, sidesteps the dilemma of competing or contradictory self-reports that might arise. Moreover, it focuses our attention to the types of interactional concerns that participants themselves are orienting to and managing in and through their talk — talk that may sustain broader social practices.

Conclusion

In this paper, I have discussed two epistemological approaches to qualitative data in the health sciences. The first takes participants' self-reports as a 'transparent' window to something presumed to exist beyond the talk. I then briefly described some dilemmas that arise when researchers who take this approach are confronted by particular

types of data. In the final section, I offered an alternative epistemological approach, one that considers research participants self-reports as a localised social practice to accomplish particular interaction goals. What does this mean for qualitative health research more generally? I am not advocating one particular methodological perspective above all others. As I argue, there are particular advantages and disadvantages to the approaches I reviewed and qualitative health researchers might well be aware of the assumptions to which they commit and the dilemmas that they may face when adopting a particular epistemological approach to qualitative health research.

References

Adam, B. D., Husband, W., Murray, J. and Maxwell, J. 2003, *Renewing HIV prevention for gay and bisexual men*, AIDS Committee of Toronto and the University of Windsor, Toronto, Canada, pp. 1–43.

Burcham, J. L., Tindall, B., Marmor, M., Cooper, D. A., Berry, G. and Penny, R. 1989, 'Incidence and risk-factors for human immunodeficiency virus seroconversion in a cohort of Sydney homosexual men', *Medical Journal of Australia*, vol. 150, no. 11, pp. 634–639.

Burr, V. 1995, *An introduction to social constructionism,* Routledge, New York.

Chesney, M. A., Barrett, D. C. and Stall, R. 1998, 'Histories of substance use and risk behavior: Precursors to HIV seroconversion in homosexual men', *American Journal of Public Health*, vol. 88, no. 1, pp. 113–116.

Edwards, D. 1997, *Discourse and cognition*, Sage Publications, London, UK.

Edwards, D., Ashmore, M. and Potter, J. 1995, 'Death and furniture: the rhetoric, politics and theology of bottom line arguments against relativism', *History of the Human Sciences*, vol. 8, no. 2, pp. 25–49.

Edwards, D. and Potter, J. 1992, *Discursive psychology*, Sage Publications, London, UK.

Flicker, S. 2004, '"Ask me no secrets, I'll tell you no lies": What happens when a respondent's story makes no sense', *The Qualitative Report*, vol. 9, no. 3, pp. 528–537.

Frith, H. and Kitzinger, C. 1998, '"Emotion work" as a participant resource: A feminist analysis of young women's talk-in-interaction', *Sociology*, vol. 32, no. 2, pp. 299–320.

Gergen, K. J. 1999, *An invitation to social construction*, Sage Publications, London, UK.

Gubrium, J. F. and Holstein, J. A. 1997, *The new language of qualitative method*, Oxford University Press, New York.

Holstein, J. A. and Gubrium, J. F. 1995, *The active interview*, Sage Publications, Thousand Oaks, CA.

Holstein, J. A. and Gubrium, J. F. 2004, 'The active interview', *In Qualitative research: theory, method and practice*, ed. Silverman, D., Sage Publications, London, UK, pp. 140–161.

Honeychurch, K. G. 1996, 'Researching dissident subjectivities: Queering the grounds of theory and practice', *Harvard Educational Review*, vol. 66, no. 2, pp. 339–355.

Hutchby, I. and Wooffitt, R. 1998, *Conversation analysis: Principles, practices and applications*, Polity, Cambridge.

Kippax, S. and Race, K. 2003, 'Sustaining safe practice: twenty years on', *Social Science & Medicine*, vol. 57, no. 1, pp. 1–12.

Kitzinger, C. 2004, 'Feminist approaches', In *Qualitative research practice*, eds. Seal, C., Gobo, G., Gubrium, J. F., and Silverman, D., Sage Publications, London, UK, pp. 125–140.

Kitzinger, C. and Wilkinson, S. 1997, 'Validating women's experience? Dilemmas in feminist research', *Feminism & Psychology*, vol. 7, no. 4, pp. 566–574.

Molitor, F., Truax, S. R., Ruiz, J. D. and Sun, R. K. 1998, 'Association of methamphetamine use during sex with risky sexual behaviors and HIV infection among non-injection drug users', *Western Journal of Medicine*, vol. 168, no. 2, pp. 93–97.

Ostrow, D. G., DiFranceisco, W. J., Chmiel, J. S., Wagstaff, D. A. and Wesch, J. 1995, 'A case-control study of human immunodeficiency virus type-1 seroconversion and risk-related behaviors in the Chicago MACS/CCS cohort, 1984–1992', *American Journal of Epidemiology*, vol. 142, no. 8, pp. 875–883.

Page-Shafer, K., Veugelers, P. J., Moss, A. R., Strathdee, S., Kaldor, J. M. and van Griensven, G. J. P. 1997, 'Sexual risk behavior and risk factors for HIV-1 seroconversion in homosexual men participating in the tricontinental seroconverter study, 1982–1994', *American Journal of Epidemiology*, vol. 146, no. 12, pp. 1076–1076.

Penkower, L., Dew, M. A., Kingsley, L., Becker, J. T., Satz, P., Schaerf, F. W. and Sheridan, K. 1991, 'Behavioral, health and psychosocial factors and risk for HIV infection among sexually active homosexual men: The multicenter AIDS cohort study', *American Journal of Public Health*, vol. 81, no. 2, pp. 194–196.

Pomerantz, A. M. 1986, 'Extreme case formulations: a way of legitimizing claims', *Human studies*, vol. 9, pp. 219–229.

Potter, J. 1996, *Representing reality: Discourse, rhetoric and social construction*, Sage Publications, Thousand Oaks, CA.

Potter, J. and Wetherell, M. 1987, *Discourse and social psychology: Beyond attitudes and behaviour*, Sage Publications, London, UK.

Reinharz, S. 1992, *Feminist methods in social research*, Oxford University Press, New York.

Ristock, J. L. 1998, 'Community-based research: Lesbian abuse and other telling tales', In *Inside the academy and out: Lesbian/gay/queer studies and social action*, eds.

Ristock, J. L. and Taylor, C. G., University of Toronto Press, Toronto, ON, pp. 137–154.

Ristock, J. L. and Pennell, J. 1996, *Research as empowerment: Feminist links, postmodern interruptions*, Oxford University Press, New York.

Ristock, J. L. and Taylor, C. G. (eds.) 1998, *Inside the academy and out: Lesbian/gay/queer studies and social action*, University of Toronto Press, Toronto.

Shoptaw, S. and Frosch, D. 2000, 'Substance abuse treatment as HIV prevention for men who have sex with men', *AIDS & Behavior*, vol. 4, no. 2, pp. 193–203.

Stall, R. and Purcell, D. W. 2000, 'Intertwining epidemics: A review of research on substance use among men who have sex with men and its connection to the AIDS epidemic', *AIDS & Behavior*, vol. 4, no. 2, pp. 181–192.

Ten Have, P. 1999, *Doing conversation analysis*, Sage Publications, London, UK.

Weber, A. E., Chan, K., George, C., Hogg, R. S., Remis, R. S., Martindale, S., Otis, J., Miller, M. L., Vincelette, J., Craib, K. J. P., Masse, B., Schechter, M. T., LeClerc, R., Lavoie, R., Turmel, B., Parent, R. and Alary, M. 2001, 'Risk factors associated with HIV infection among young gay and bisexual men in Canada', *Journal of Acquired Immune Deficiency Syndrome*, vol. 28, no. 1, pp. 81–88.

Widdicombe, S. 1995, 'Identity, politics and talk: A case for the mundane and the everyday', In *Feminism and discourse: Psychological perspectives*, eds. Wilkinson, S. and Kitzinger, C., Sage Publications, London, UK, pp. 106–127.

Wilkinson, S. 2000, 'Women with breast cancer talking causes: Comparing content, biographical and discursive analyses', *Feminism & Psychology*, vol. 10, no. 4, pp. 431–460.

Wilkinson, S. 2004, 'Focus group research', In *Qualitative research: theory, method and practice*, ed. Silverman, D., Sage Publications, London, UK, pp. 177–199.

Wilkinson, S. and Kitzinger, C. (eds.) 1996, *Representing the other: A feminism & psychology reader*, Sage Publications, London, UK.

Playing by the Rules: Feminism, Alternative Medicine and Standards of Evidence

Kirstin Borgerson

In her book, *Toward a Feminist Theory of the State* (1991), Catharine MacKinnon analyses the different paths to sex equality that are available to women under the liberal state. She argues that although the state appears to be objective and value-neutral it actually has a strong male bias and, given that the state is male, women are faced with two options in their attempts to gain equality. Women must either find a way to make themselves the same as men, in order to be considered equal on their terms, or emphasise the significant ways in which they differ from men, and request special treatment on those grounds. In this paper, I borrow the analytic tool of sameness and difference in order to shed light on the situation facing researchers of alternative medicine.[1] I suggest that researchers in alternative medicine are currently confronted with a choice between appeals to sameness or difference in the design of research programs. Drawing on lessons learned from MacKinnon's work on sex equality, I will consider the benefits and drawbacks of these two paths for alternative medical researchers. Following this, I will consider MacKinnon's suggestion that what is necessary is a re-evaluation of the standard which underlies both the sameness and difference approaches. I argue that alternative medical researchers need to critically engage with the standard to which they are being held: the evidence hierarchy of evidence-based medicine.

Understanding Sameness and Difference: Feminist Theory

Liberal feminist theory, from its earliest origins in the writings of Wollstonecraft, has pursued the equality of women through the establishment of legal rights such as the right to vote and the right to

1. The shared history between women and alternative healers makes the use of this analytic tool particularly appropriate.

education and employment. This 'first wave' feminism required that the law recognise that women are 'just like men' in important respects and therefore that there was no reason to deny women the same rights as men. There could and should be a single standard for men and women. Women have the right to vote because they are the same as men in their abilities to act as citizens of a state. Women have the right to education because they appear to be relevantly similar to men in their capacity to learn.[2] As MacKinnon points out, for these early stages in the feminist movement, equality means equivalence: 'equal to' means 'the same as'.

The importance of this move to identify women as the same as men should not be underestimated. It is because of powerful speeches and essays on the need for women to be recognised as competent and intelligent 'just like men', and the emphasis placed on the clear injustice of treating people of similar moral standing in different ways, that women were able to earn the legal standing of citizens. Once these rights were granted, however, the challenging project of living up to the male standard began. Women who chose to pursue non-traditional education and employment felt (and continue to feel) the burden of this new responsibility to prove themselves 'as good as men'. Living according to the male standard of excellence has come with many benefits but, as a strategy, it has also raised its fair share of new problems.

The clear and unambiguous statement that 'we are all the same' seems at first glance like a good foundation for sex equality. After all, it did manage to earn women the foundational rights just outlined, as well as others. Yet, as time passes and more cases come before the courts, women are finding that legislators use the sameness standard to grant men access to the few privileges women historically had. One such example arises in custody battles – which, under new gender-neutral rules shaped by the sameness approach to equality, are now seen as a choice between two equal parents. Yet, as MacKinnon argues,

> Men often look like better parents under gender-neutral rules like level of income and presence of nuclear family, because men make more money and (as it is termed) initiate the building of family units. They also have greater credibility and authority in court (p. 221).

2. Many women in developing countries around the world continue to fight for these 'first wave' rights. References to the second wave of feminism in this paper are meant to reflect the women's movement in Canada (where I am writing) and the United States (where MacKinnon writes).

Social advantages shared by men cannot be taken into account because this would violate the appearance of gender neutrality in the court's decision. Conditions that make women, as a group, more in need of alimony are not relevant under the gender-neutral account. These and other decisions of this sort make the sameness approach less than ideal in the pursuit of full sex equality.

In the second half of the twentieth century, women became aware of the fact that despite having their legal rights enshrined, they were still facing tremendous challenges in the workplace and at home. The responsibilities of childrearing and domestic labour, in particular, were creating the infamous 'double day' for women who also worked outside the home. Discrimination and job expectations that were tailored to men's lives impeded women's efforts to enter the public domain. The sameness approach had been partly successful, but seemed to leave a number of serious problems untouched. Women began to call for proper recognition of the differences between men and women, especially those differences with serious social and economic implications. And so the double standard was re-introduced, despite obvious 'slippery slope' concerns raised by the decision to enshrine different treatment for the sexes within the law. To sum up:

> The moral thrust of the sameness branch of the doctrine conforms normative rules to empirical reality by granting women access to what men have. The differences branch, which is generally regarded as patronizing and unprincipled but necessary to avoid absurdity, exists to value or compensate women for what they are or have become distinctively as women (MacKinnon 1991, p.220).

The situated differences attributed to women under the latter approach are compensated for with legislation on maternity leave, and in affirmative action programs. Social science evidence on the number of women denied employment or further employment because of pregnancy, and on the lack of women entering university programs for non-traditional training, supported these claims of difference (see for example Davies, Avison and Cassidy 2001).

As one might expect, and as alluded to in MacKinnon's description above, the difference approach makes many feminists deeply uneasy. Because the difference approach appears to rely upon many of the same sorts of arguments that were used *against* women for hundreds

of years (women are different therefore deserve 'special' – read oppressive – treatment), feminists are worried about the possible slippery slope as legislators struggle to define the boundaries of what counts as a *legitimate* difference. The possibility that they will choose to define legitimate difference broadly is of considerable concern. The difference approach can just as easily be used to deny certain treatment and protection as to grant it.

MacKinnon offers an example of the failings of the difference approach: in one case, contact jobs in male-only prisons were denied to women in the name of "their very womanhood" (MacKinnon 1991, p. 227). Because of women's 'rapability', which is seen as a stable and legitimate difference, rather than something that is shaped by social conditions, they were denied job opportunities. This clearly makes use of the difference approach to justify limiting women's opportunities, without any recognition of the social forces at work in creating the conditions of women's rapability. If we imagine an extremely violent society, where women's rapability is a real issue for a wide range of jobs where there is contact with men, one can start to imagine the restrictions and limitations that could be 'legitimately' applied, and the justification would probably still be ostensibly based on women's innate biological rapability, rather than on any socialization processes in society which create the conditions under which men are violently attacking women. The difference approach has been somewhat helpful in achieving the goal of sex equality for women, but, as this example demonstrates, it is also dangerous in that appeals to difference can just as easily be used to defend inequality as equality. In addition, when feminists use this approach they open themselves to the charge that the special exceptions to the standard unfairly advantage women. Why should women be allowed to receive special (unequal) treatment? This appears to fit awkwardly, if at all, within the greater project of equality.

The situation looks unpromising for feminists who work for full sex equality through either the sameness or difference approach, or even some combination of the two. This is because, in either approach, the referent is male. Women are either the same *as men*, or different *from men*. The standard against which women's progress is measured is, in both cases, male. Approaching the debate over sex equality from either of these perspectives, "merely provides two ways for the law to hold women to a male standard and to call that sex equality" (MacKinnon 1991, p. 221). For example, a female academic would be expected to

either live up to strict criteria for tenure (originally designed under the assumption that the academic is not the primary caretaker of children), or appeal for a special exception in order to receive permission to take time off to have or raise children. The tenure standards remain unchanged in either of these two scenarios.

MacKinnon argues that superficial approaches to equality only end up reinforcing the standard they are trying to overcome. Under the sameness approach women are granted the same rights and privileges as men. But the fact that what men have is the standard goes unremarked. Under the difference approach, women are seen as exceptions to the standard rule and treated as special cases. Again, this never challenges or changes the standard. The possibility, in either of these cases, that the standard is less than ideal, is not seriously considered or critically evaluated. The persistent problems encountered by attempts to apply the sameness and difference approaches are proof that the foundations of inequality have not yet been addressed. MacKinnon stresses that the problem we face in attempting to apply these tactics in pursuit of sex equality is not merely a transitional one; it will arise as long as our approaches to equality presuppose a male standard.

The solution to these problems is offered primarily by example in MacKinnon's work. She suggests the need for us to be "questioning the principledness of neutral principles" (Mackinnon 1991, p. 232). In other words, we need to be careful in choosing our 'neutral standard' because it might actually be shaped by social values that we are failing to identify. In the case of sex equality, it is clear that MacKinnon believes the discussion is being limited by an appeal to a male standard as neutral. MacKinnon argues that people misidentify the real problem in the quest for sex equality. The relevant underlying issue, according to MacKinnon, is the gender hierarchy: "the social meaning of the sexuality and gender of men and women" (MacKinnon 1991, p. 232–3). It is unhelpful to talk about gender differences without first engaging with the social circumstances in which gender differences are constructed and valued. Failing to critically engage with these underlying assumptions amounts to conceding that the roots of gender inequality are somehow natural and allows that the "central epistemological pillars of gender as a system of power are permitted to remain standing" (MacKinnon 1991, p. 233).

MacKinnon does not go into detail on the next steps of this project, but it is reasonable to assume that once the critical process has

occurred, and we have a better understanding of the biases and value-judgments underlying what were previously presented as neutral standards and goals, we should take this new knowledge and apply it to the development of a new and improved standard of sex equality. Throughout this process (and hopefully continuing beyond it), the emphasis will be on a critical examination and re-evaluation of the current standard.

Alternative Medicine

The National Institutes of Health (NIH) Panel on Definition and Description defines complementary and alternative medicine (CAM) as:

> A broad domain of healing resources that encompasses all health systems, modalities and practices and their accompanying theories and beliefs, other than those intrinsic to the politically dominant health system of a particular society or culture in a given historical period (NIH 1997, p.49).[3]

This broad definition, as applied in the USA and Canada, encompasses: biologically based treatments (herbs, special diets, and vitamins), manipulative and body-based treatments (chiropractic, massage, osteopathy), energy therapies (reiki, magnet therapy, qi gong), mind-body treatments (yoga, spirituality, relaxation/ meditation), and entire alternative medical systems (traditional Chinese medicine, naturopathy ayurveda). Alternative medicine is often contrasted with: 'mainstream', 'conventional', 'allopathic', 'orthodox', or 'Western' medicine.

Alternative medicine has a tremendous amount of public support in North America, and this popularity appears to be growing. Americans spent between $37 and $47 billion on alternative medicine in 1997, and these numbers are increasing. According to the most recent survey released by the National Center for Health Statistics (NCHS) of the United States in May 2004, approximately 36% of American adults are currently using some form of alternative medicine (NIH 2004).

---

3. CAM has also been defined as "a group of diverse medical and health care systems, practices, and products that are not presently considered to be part of conventional medicine—that is, medicine as practiced by holders of M.D. (medical doctor) or D.O. (doctor of osteopathy) degrees and their allied health professionals, such as physical therapists, psychologists, and registered nurses" (NIH, 1997). Given recent attention to CAM in medical schools, and the integration of some CAM practices with mainstream medicine, this definition is becoming less accurate. The designation 'complementary' or 'alternative' simply signifies whether the medicine is being used in conjunction with conventional medicine or in place of it.

Similar polls in Canada indicated that between 42% and 50% of the population have used some form of alternative medicine in the past year. This was a more than 80% increase when compared with a poll conducted five years earlier (Canadian Poll 1997). Despite what critics continue to regard as a serious lack of scientific evidence, alternative medicine appears to be increasingly gaining acceptance within Canada and the United States.

The most common mainstream reaction to alternative medicine can be read in two of the top medical journals:

> There is no alternative medicine. There is only scientifically proven, evidence-based medicine supported by solid data or unproven medicine, for which scientific evidence is lacking. Whether a therapeutic practice is "Eastern" or "Western", is unconventional or mainstream or involves mind-body techniques or molecular genetics is largely irrelevant except for historical purposes and cultural interest. (Fontanarosa and Lundberg 1998, p. 1619).

> It is time for the scientific community to stop giving alternative medicine a free ride. There cannot be two kinds of medicine – conventional and alternative. There is only medicine that has been adequately tested and medicine that has not, medicine that works and medicine that may or may not work. Once a treatment has been tested rigorously, it no longer matters whether it was considered alternative at the outset. If it is found to be reasonably safe and effective, it will be accepted. But assertions, speculation, and testimonials do not substitute for evidence. Alternative treatments should be subjected to scientific testing no less rigorous than that required for conventional treatments. (Angell and Kassirer 1998, p. 839).

According to these demands, alternative medical researchers should strive to conduct the same sorts of meta-analyses of large-scale randomised controlled trials currently regarded as the gold standard of medical research.

The current standards in medicine are set by a 'hierarchy of evidence' developed in the last 15 years by the evidence-based medicine (EBM) movement, some version of which is now adopted in most teaching hospitals and research institutes across North America. Under the EBM approach to medical decision-making, physicians are advised to

critically assess the best available evidence (for a presented illness) with the assistance of an evidence hierarchy, and apply those results judiciously to individual patients. The hierarchy, which provides the 'evidence base' for clinical decisions, ranks research methodologies according to the generalisability of their results, as well as their perceived ability to eliminate bias and establish clear causal connections between treatments and effects. The top of the hierarchy of best evidence is the meta-analysis of randomised controlled trials (RCTs). Qualitative research, outcomes research, case studies, case-series and other small-scale studies are considered to be of lower quality in the evidence hierarchy, and are therefore less likely to earn medical treatments and practices respect within the mainstream medical community.

Some alternative medical researchers have taken the demands for high quality research seriously and, especially in the last decade or so, there have been numerous studies evaluating a variety of alternative medical treatments. Some results have been positive, for example, the use of moxibustion (burning herbs at an acupuncture point) to correct breech presentation in late pregnancy (Cardini and Weixin 1998), traditional Chinese medical herbs for irritable bowel syndrome (Bensoussan et al. 1998), glucosamine for the treatment of osteo-arthritis (McAlindon et al. 2000), and acupuncture for nausea and vomiting (NIH 1997). These are examples of alternative medical research that has 'passed the test' and lived up to the standards outlined by evidence-based medicine. Many of these successful treatments, despite achieving the status of 'best evidence', are not accepted into mainstream medical practice.

Alternative medicine must prove itself equal to conventional medical treatments by, at the very least, meeting current standards; in some cases, alternative medical research is even required to exceed current standards, based on certain Bayesian ideas in medicine (which advise extra scrutiny for those practices and treatments that have low prior probabilities).[4] In such a situation, a few alternative medical

4. Of course, alternative medical research is not alone in failing to live up to current standards of evidence. As Morreim (2003) points out, "Much of actual clinical practice does not and never can measure up to the scientific standards to which critics of CAM would like to hold alternative modalities" (p. 222). A large percentage (roughly estimated to be around 80%) of medical treatments currently offered in hospitals and clinics across North America have never been tested by an RCT (University of Sheffield, 2004). Research into alternative medicine is required to meet the highest standards even though many currently accepted medical practices have not met (and likely never will meet) those same standards. For example, most surgical interventions (which share methodological difficulties with alternative treatments such as acupuncture) are well-respected and continue to be performed on patients despite a lack of RCT evidence. This demonstrates some of the power imbalances at work in the demands for evidence.

researchers will jump through the designated hoops, prove the efficacy of select therapies, and earn legitimacy and respect for those treatments. Treatments such as herbs (which resemble, and are testable like, drugs) are best candidates for selective incorporation into mainstream medical care.

Much research into alternative medicine has failed to meet the methodological requirements of the EBM evidence hierarchy; research designs commonly consist of individual case-studies or other small-scale or qualitative studies. There are several commonly cited explanations for this failure to produce RCT evidence within the alternative medical literature. According to Anthony (1987), alternative medical treatments are often highly individualised (as compared to the generalised treatments offered in an RCT), complex (have a number of therapeutic components), require physical treatment to which it is difficult to 'blind' patients and practitioners (for example, needling in acupuncture), actively involve the therapist as an integral part of treatment (making randomisation difficult), set different (and multiple) end points based on different philosophies of medicine (for example, having endpoints such as high quality of life and balanced 'chi' in addition to alleviation of symptoms), and rely on principles of self-healing and mind-body control (which incorporate, rather than rule out, the placebo effect). These and other problems (including social and economic difficulties in funding and organizing large scale studies on often un-patentable treatments) are often raised as explanations for the lack of gold standard evidence in alternative medicine. These differences are foundational to alternative medicine and make meeting the EBM standard very difficult if not impossible.

In light of these difficulties, alternative medical practitioners and researchers could choose to insist that because they practice in 'a different paradigm', from 'a different world view', and hold 'a different philosophy', there should be special exceptions to the EBM standards made for alternative medical research. The EBM hierarchy would be accepted as legitimate, but alternative medical researchers would ask to be special exceptions to the rules.

The motivation for this option is clear. Even when best evidence exists, successful high quality RCTs of alternative medical treatments are not affecting the treatment decisions of many physicians. And there do appear to be some important differences in metaphysical views on the nature of health and illness between alternative and

mainstream medical researchers that provide some ground for an appeal to difference. In addition, there is much historical, cultural, and sociological evidence demonstrating that what is considered to be 'mainstream' medicine in North America today is in part the result of social, political and economic forces in the past century. By no means are the boundaries of mainstream medicine in perfect correspondence with any formal distinction between science and pseudo-science. Many alternative medical practitioners are aware of historical data about the development and use of standards as a method of keeping outsiders from infiltrating the Western medical profession: sociologists Timmermans and Berg point out that, "Professions have relied on credentialing, registration and licensing mechanisms to safeguard their jurisdiction against competitors" (Timmermans and Berg 2003, p. 85). There is no reason to think that mainstream medical methodology (as, in part, a product of these social forces) has any exclusive grasp of the true nature of health or disease or any special claim to epistemological superiority in the assumptions of the evidence hierarchy.

In response to these sorts of appeals to difference, critics decry the 'quackery' of alternative medicine and caution patients to avoid alternative medical practitioners. Alternative medical researchers have found that there is no guarantee that labeling yourself different will earn you any respect from the mainstream medical community. In fact, it is just as likely, if not more likely, to lead to worse treatment. In the eyes of the mainstream medical community, there is an epistemological standard that alternative medicine simply fails to meet. When it is assumed that everyone should meet the same standard, an appeal to difference comes across as a sign of weakness – an indication of failure and an appeal for special treatment.

Appealing to the differences of alternative medicine may also too heavily shortchange the common ground between conventional and alternative medicine (including agreement on basic scientific principles such as falsifiability), making it difficult for alternative medicine to appeal for funding and other privileges granted to conventional medicine. As well, appeals to difference may, because they presuppose the cohesive group 'alternative medicine', undermine our ability to distinguish amongst those treatments that are effective and those that are not. Presumably, at least some of the interest in conducting research in the first place is to make these sorts of distinctions.

Resolving a False Dilemma: Solutions

We saw earlier how women were measured according to their ability to be the same as men (therefore equal) and different from men (therefore deserving of special treatment). In choosing a strategy for gaining equality, women found that both approaches, while able to provide some gains, had significant limitations. In the end, MacKinnon pointed out that each approach, because it presumes the male standard, fails to address the underlying inequality inherent in the assignment of the standard.

In a parallel case, alternative medicine is currently evaluated as either the same as or different from conventional medicine.[5] Alternative medical researchers are confronted with a choice. On the one hand they could abide by the demands for 'better evidence of efficacy' and perform research according to the evidence hierarchy of mainstream medicine. On the other hand, they could choose to emphasise the ways in which alternative medicine differs from mainstream medicine, and to take pride in these 'special differences' by arguing that because of their special status, exceptions to current standards of evidence should be permitted.

The analogy with feminism suggests that while there are gains to be made by pursuing either of the two approaches, in the end neither will be entirely successful. What is necessary for full equality or legitimacy is serious critical engagement with the standard. In the case of sex equality, we have to be aware of and question the continued assumption of the male standard. In medicine, we have to question the current standard of evidence: the evidence hierarchy designed by the EBM movement. Relevant questions would include, though by no means be limited to: Where did this standard of evidence come from? Is this standard best designed to answer all questions of medical significance? What are the assumptions underlying this approach to medical evidence? Does this standard reflect existing power imbalances?

According to my analysis, alternative medical practitioners should join those in the medical community who criticize and attempt to

5. This designation as the 'other' is probably most strikingly evident in the names assigned to the two types of medicine. Some alternative medical practices have been around longer (some thousands of years) than so-called conventional medicine, yet they retain the name 'alternative'. This reflects the power imbalance in the construction of the term 'alternative' medicine. MacKinnon would argue that gender has been constructed in a similar way.

modify the current standards. Many of the sorts of concerns raised by alternative medical researchers are already fiercely debate within the conventional medical community (Feinstein & Horwitz 1997; Black 1998; Upshur 2000). The extraordinarily influential evidence-based medicine movement of the last decade not only rewrote the standards of evidence required of physicians and researchers, it also galvanised debate within the medical community on the nature of good evidence in medicine. Many of the concerns with the evidence hierarchy raised by alternative medical researchers are similar to those already raised in medical journals. For example, critics argue that evidence produced by meta-analyses of RCTs is not particularly helpful for guiding treatment decisions at the level of individual patients (RCTs were, after all, originally designed for use in agriculture), that evidence supplied by qualitative research or case-studies can be, at times, even more helpful and relevant than the 'gold standard' evidence, that the direction of health care research is biased toward pharmaceuticals because the RCT methodology is most appropriate for evaluating pills, and that it does not make sense to have one evidence hierarchy, or one 'best methodology' in a field as diverse as medicine. These practical and epistemological concerns with the current hierarchy of evidence are gaining support within the mainstream medical community, and provide further support for a critique of current standards.

Case Study

In order to make this argument more concrete, consider a case where feminists have identified the need for critique of the male standard. The typical work week in Canada and the USA, set by a number of factors and (loosely) a result of the industrial revolution, is 40 hours, and is usually met by working 9 am–5 pm, 5 days a week. These hours suited the workers (most, if not all, men with no child-rearing responsibilities) at the time of development of the standard. In the past half century, women have called into question the need for workplaces to be organised strictly according to these common hours. As a result of persistent demands for a more flexible full-time work schedule that would allow women to pick children up from school or care for elderly family members, some companies have now adopted 'flex-time' scheduling. This has turned out to have advantages for the companies as well, as resources can be shared and jobs can be covered during a broader set of hours. In addition, men have benefited from flextime scheduling and have been able to increase time spent with

their families. By challenging the accepted standard, women were able to improve upon it in a way that is responsive to the needs and goals of modern couples and families (Christensen and Staines 1990).

Now let us consider a possible parallel case in alternative medical research. A recent proposal arising out of the alternative medical literature, but also poised to spread into mainstream medical research, is something called whole systems research. "The new discipline of whole systems research (WSR) targets the study of complex CAM therapies as system-level phenomena, as opposed to single-agent or uni-dimensional effects" (Ritenbaugh et al. 2003, p. 32). This new approach to research is described as an innovation of the conventional RCT. The main question asked in this new approach is 'does system x facilitate healing?' replacing the familiar research question, 'does treatment x, on average, affect condition y?' Designers of this new approach aim to maintain the emphasis on rigourous standards in research while modifying the methodology to better fit a range of modalities – in particular those that are complex and otherwise difficult to research. What the researchers proposing this new tool have done is to call into question the RCT standard, and offer an improved research design that might, with some support, even end up significantly improving the sort of conventional research that is commonly done.

Objection and Reply

I must take a moment to respond to a predictable objection to this argument, raised by those who are concerned about an overly-enthusiastic application of the analytic tool of sameness and difference. Whenever there is a standard, it appears as though those who fail to meet the standard have some grounds on the sameness and difference approach for re-evaluating it. And yet, in many, perhaps even most cases, we think that standards are a good thing with plenty of reasonable support; we should not be constantly engaging in re-evaluation of our standards just because there are those who fail to meet them. Consider, for instance, the Olympic standards. In order to make it to the Olympics, you have to meet certain standards of fitness and skill in your sport. Does the fact that an athlete might fail to meet, say, the standards for running the 100 meter dash mean that they have a right to 'critically evaluate' and call into question those standards? Are they, like women and researchers in alternative medicine, in a

position to critique the neutrality of the standards to which they are being held?

I believe that the response to this question is 'yes'. Everyone should be involved in the process of critically discussing and evaluating any social standards (especially ones, such as these, that may be grounded in tradition and convention). Do I think that the Olympic standards will change because one athlete challenges them? Of course not. But that is not because the critique was inappropriate in any way. Rather, it is that, given the fact that we do have good reasons underlying our choice of Olympic standards, we will be able to provide those reasons to the satisfaction of any reasonable critique. If it were to turn out that we do not have good reasons for setting the Olympic standards as we do, then perhaps they should change. One could imagine a case where the whole sport has evolved over time and there is need for changing regulations to accommodate the new form of the sport.

Olympic standards, and so many of the other excellent standards employed in human life, are supported by good reasons. It is this that protects them from crumbling under scrutiny, not immunity to critical discussion. In the case we are discussing, people tend to assume that evidence-based medicine is supported by similar excellence of reasoning. Surely, the thought goes, the evidence-based approach to medicine must be based on sound reasoning and that is why we accept it as our standard and why we should continue to hold all medical research to its demands. A full exploration of this objection would go far beyond the scope of this paper, but suffice it to say that evidence-based medicine is not, in fact, based on solid evidence (it is a theoretical approach that has not yet been empirically tested), nor is it based on appropriate and philosophically sound assumptions about the nature of medical evidence. It is an approach to medicine that, as I argue in detail elsewhere (see my Lupina 2003/4 paper), is based on a number of assumptions about medical care with shaky epistemological founding.

Conclusion

MacKinnon's analysis is meant to provide one plausible explanation of the problem of continued disparity between the sexes. Regardless of whether the analytic tool of sameness and difference was best at explaining this disparity in the case of sex equality, it may in fact be the best analysis of the situation confronting alternative medical

researchers. I suggest that alternative medical researchers face many of the same sorts of hurdles originally identified by MacKinnon.

In the same way that women achieved some elements of equality through appeals to sameness, alternative medical researchers will find some success by performing large scale RCTs. In the same way that women found appeals to difference necessary in order to gain recognition for the unique challenges of pregnancy, alternative medical researchers will find appeals to difference useful in explaining the need for and value of individualized case-studies. At the end of the day, however, alternative medical researchers would do a great disservice to all patients if they were to limit their strategies to these two. Alternative medical researchers will only ever achieve the most significant goals of integration, respect, and legitimacy within the medical community if they involve themselves in critique and re-evaluation of the standards to which they are being held. This is reinforced by models of scientific communities (such as those proposed by social epistemologists) whereby objectivity is established and upheld only by active critical participation of all diverse community members.[6] Recognizing the need for critical discussion and engaging mainstream physicians in a collaborative evaluation of current evidential standards will, I suggest, lead to new and better definitions of good evidence, that are scientifically rigourous and yet responsive to the real needs of practitioners of all types.

References

Angell M., Kassirer J. P. 1998, 'Alternative Medicine - The Risks of Untested and Unregulated Remedies', *N Engl J Med*, vol. 339, pp. 839–841.

Anthony H. M. 1987, 'Some Methodological Problems in the Assessment of Complementary Therapy', *Statistics in Medicine*, vol. 6, no. 7, pp. 761–771.

Bensoussan, A. et al. 1998, 'Treatment of irritable bowel syndrome with Chinese herbal medicine: a randomized controlled trial', *JAMA*, vol. 280, no. 18, pp. 1585–1589.

Berman, B. et al. 2000, 'The public debate over alternative medicine: the importance of finding a middle ground', *Altern Ther Health Med*, vol. 6, no. 1, pp. 98–101.

Black, D. 1998, 'The Limitations of Evidence', *Persp Biol Med*, vol. 42, no. 1, pp.23–6.

Canadian College of Naturopathic Medicine (CCNM). 2004, 'Naturopathic Oath', [Online], Available from: www.ccnm.edu/about.html [Accessed October 26, 2004].

6. See Helen Longino's book *The Fate of Knowledge* (2002) for a recent defense and full explanation of this view.

Canadian Poll Details 2004, '" Alternative Medicine" Use (Angus Reid Poll conducted Aug. 22–24, 1997.)',[online], *Dynamic Chiropractic*, vol. 15, no. 25, Available from: http://www.chiroweb.com/archives/15/25/23.html [Accessed Aug. 26, 2004].

Cardini, F., Weixin, H. 1998, 'Moxibustion for Correction of Breech Presentation: a randomized controlled trial', *JAMA*, vol. 280, no. 18, pp. 1580–1584.

Christensen, K. E. Staines, G. L. 1990, 'A Viable Solution to Work/Family Conflict?' *Journal of Family Issues*, vol. 11, no. 4, pp. 455–467.

Status of Women Canada, [online], Available from: http://www.swc-cfc.gc.ca/pubs/0662653327/ index_e.html [Accessed May 16, 2005].

Feinstein, A., Horwitz, R. 1997, 'Problems in the 'Evidence' of 'Evidence-based Medicine', *Am J Med*, vol. 103, pp. 529–535.

Fontanarosa, P. B., G. D. Lundberg. 1998, 'Alternative medicine meets science', *JAMA*, vol. 280, no. 18, pp. 1618–19.

Heymsfield, S. et al. 'Garcinia cambogia (Hydroxycitric Acid) as a Potential Antiobesity Agent: A Randomized Controlled Trial', *JAMA*, vol., 280, no. 18, pp. 1596–1600.

Lazarou, J., Pomeranz, B., Corey, P. 1998, 'Incidence of Adverse Drug Reactions in Hospitalized Patients: a meta-analysis of prospective studies', *JAMA*, vol. 279, no. 15, pp. 1200–1205.

Longino, H. 2002, *The Fate of Knowledge* Princeton University Press, New Jersey.

MacKinnon, C. 1991, 'Chapter Twelve - Sex Equality: On Difference and Domination,' In *Toward a Feminist Theory of the State*, Harvard University Press, Cambridge.

McAlindon, T.E., LaValley, M.P., Gulin, J.P., Felson, D.T. 2000, 'Glucosamine and chondroitin for treatment of osteo-arthritis: a systematic quality assessment and meta-analysis', *JAMA*, vol. 283, pp. 1469–1475.

Morreim, E. H., 'A dose of our own medicine: alternative medicine, conventional medicine, and the standards of science', *J Law Med Ethics*, vol. 31, no. 2, pp. 222–235.

National Institutes of Health. 1997, 'NIH Consensus Statement', [online], Nov.3–5 vol. 15, no. 5, pp. 1–34. Available from: http://odp.od.nih.gov/consensus/cons/107/107_statement.htm [Accessed Sept.1, 2004].

National Institutes of Health. 2004, 'The Use of Complementary and Alternative Medicine in the United States,', [online], Available from: http://nccam.nih.gov/news/camsurvey_fs1.htm [Accessed Aug. 24, 2004].

National Institutes of Health: Panel on Definition and Description. 1997, 'Defining and Describing Complementary and Alternative Medicine', *Altern Ther Health Med*, vol. 3, no. 2, pp. 49–57.

The Fraser Institute 2004, 'Canadians Spend $3.8 Billion on Alternative Medicine To Cure What Ails Them', [online], Available from:

http://www.fraserinstitute.ca/shared/readmore.asp?sNav=nr&id=392   [Accessed Aug. 26, 2004].

Ritenbaugh, C, et al. 2003, 'Whole Systems Research: a discipline for studying complementary and alternative medicine', *Altern Ther Health Med*, vol. 9, no. 4, pp. 32–36.

Shlay, J. et al. 1998, 'Acupuncture and Amitriptyline for Pain Due to HIV-Related Peripheral Neuropathy: A Randomized Controlled Trial', *JAMA*, vol. 280. no. 18, pp. 1590–1595.

Timmermans, S., Berg, M. 2003, *The Gold Standard: The Challenge of Evidence-based Medicine and Standardization in Health Care*, Temple University Press, Philadelphia.

Tonelli, M. 1998, 'The Philosophical Limits of Evidence-based Medicine', *Acad Med*, vol. 73, no. 12, 1234–1240.

University of Sheffield. 2004, *What Proportion of Healthcare is Evidence-based?* [online], Available from: http://www.shef.ac.uk/scharr/ir/percent.html. [Accessed Sept. 5, 2004].

Upshur, R.E.G. 2000, 'Seven Characteristics of Medical Evidence', *J Eval Clin Pract*, vol. 6, no. 2, pp. 93–97.

Chronic Dieting and Eating Behaviour

Jennifer S. Coelho

Obesity is a growing concern in North America, with recent reports indicating that more than 50% of Canada's population is either overweight or obese. Given that research has determined that being overweight leads to a variety of negative physical and mental health consequences (e.g., Larsson, Karlsson & Sullivan 2002), researchers need to address issues related to obesity. There are numerous societal implications of the obesity epidemic, including the financial burden of obesity-related illnesses. Katzmarzyk and Janssen (2004) recently estimated the financial impact of obesity in Canada. These researchers estimated both the direct expenditures for obesity-related illness, as well as the indirect costs (including lost income due to illness, injury, or premature death). It was estimated that the cost associated with obesity is $4.3 billion, with $1.6 billion in direct costs and $2.7 billion in indirect costs. Public policy is being influenced by this epidemic, as governments in both the United States and Canada are taking measures in order to address obesity. As indicated by Stephen Samis (Director of Health Policy at the Heart and Stroke Foundation of Canada), the obesity epidemic is a complex problem, stemming from both environmental and individual factors (Heart and Stroke Foundation, 2004). Clearly, however, research that investigates factors that may lead to overeating and weight gain within individuals is necessary, in order to gain insight into factors that may underlie the obesity epidemic.

Dietary restraint is a key individual difference variable to consider in research on overeating and weight status. For example, Heatherton, Polivy and Herman (1991) suggested that chronic dieters (restrained eaters) seem to exhibit a cycle of dieting and subsequent overeating, which precludes actual weight loss. Furthermore, in our laboratory we consistently find that the body mass index of restrained eaters is higher than that of non-dieters (unrestrained eaters). Hence, it seems that engaging in chronic dieting may backfire, as restrained eaters

typically fail to lose weight, and may even weigh more than non-dieters. However, despite the fact that the majority of individuals who engage in chronic dieting behaviours fail to achieve significant weight loss (Heatherton et al. 1991), dieting is pervasive throughout Western culture, particularly in women.

In order to gain insight into some of the issues underlying overeating and weight gain, and understand the factors that make it so difficult to achieve weight loss, we must first consider the process of self-change. Attempting self-change, such as trying to lose weight, is a complicated process involving a sequence of steps. I will review this process, and outline why attempts at self-change may fail, particularly attempts at weight loss. Finally, I will outline a program of research focusing on studying the eating behaviour of chronic dieters, as compared to non-dieters, which can provide insight into some of the triggers for episodes of overeating.

The Process of Self-Change

Self-change is a process that is frequently attempted by people in our society; people often engage in this process in order to reduce habits and characteristics that they may consider to be unhealthy or unattractive. Every year millions initiate attempts at self-change, specifically by trying to quit smoking, lose weight, or rid themselves of an addictive behaviour such as alcohol or drug abuse, among other things. Empirical evidence indicates that success at these attempts is relatively hard to attain. Furthermore, self-changers who do not seek therapy or formal treatment in order to achieve this success do not seem to have a differential likelihood of success at an attempt in comparison to those who do receive formal guidance in their attempts (see Cohen et al. 1989).

In contrast to the bleak forecasts for success at self-change based on the empirical literature, however, is evidence presented by Schachter (1982) indicating a high success rate for quitting smoking and weight loss in a non-self-selected population of subjects. Anecdotally, there also is an impression that success at self-change is very possible — many people are able to name a friend or acquaintance who has succeeded at losing a significant amount of weight, quitting smoking, or overcoming alcoholism. This contrast raises some questions as to how much is known about self-change, and if the empirical literature on treatment programs is either accurate or representative of the processes and successes associated with self-change.

It is logical that enduring self-change follows from success at self-regulation, in that people who are able to successfully self-regulate and control their behaviours will consequently be more successful at self-change of these behaviours in comparison to those who fail at self-regulation. As a result, common themes within the literature on self-regulation may provide an indication of processes that may increase successes at stable self-change.

Self-Regulation

Key Concepts

The concept of self-regulation implies a process controlled by the self in order to achieve and maintain certain standards or states. Carver and Scheier (1982) discuss self-regulation as a process similar to a negative feedback loop, in which behaviours are performed in order to reduce deviations of a current state from a comparison value. Higgins (1997) also discusses this concept of discrepancies between current and desired end-states, such that self-regulators either focus on approaching matches with a desired end-state, or avoiding mismatches with this end-state. Hence, a comparison value (Carver & Scheier 1982) or desired end-state (Higgins 1997) is a key component in the process of self-regulation, such that behaviours are performed in order to achieve goals and reduce deviations from these goals.

The choice of goals associated with this process is not a simple task in and of itself, however. A goal may be proximal (in that relatively short-term regulation of behaviour can result in reaching the goal), or distal, in which case behaviour regulation is more long-term. Similarly, a goal can be either concrete or more abstract. The type of goal sought after in regulation can influence the degree of success one will experience. Carver and Scheier (1982) indicate that abstract goals are attained more gradually than concrete goals, yet these abstract, higher-level goals are also associated with more commitment to the goal (Mischel, Cantor & Feldman 1996).

Factors Associated with Failure in Self-Regulation

Self-regulation is not a process that is inherently associated with success. A desired goal is not necessarily achieved either quickly or easily, and hence self-regulation may consequently fail. Performance of behaviours to reach a desired goal may be an on-going process, particularly if the goal is abstract or distal. This factor of the time frame associated with goal achievement is pervasive across the literature on failures in self-regulation.

Whereas goal achievement has thus far been framed mainly in terms of performance of behaviour, achievement of a goal may also be associated with behavioural inhibition. For example, consider dieters' common goal of losing weight. In this example, the inhibition of behaviour (eating/overeating) is required, as is the performance of behaviours in accordance with this goal (e.g., exercising, choosing foods that are considered healthy or low-fat). Polivy (1998) indicates that behavioural inhibitions can have negative consequences, including maladaptive behaviour, behavioural excesses, and cognitive disruptions. Longer-term inhibitions seem to be associated with more negative consequences, indicating that striving for abstract, distal goals may be associated with more negative consequences than striving for concrete, proximal goals given that these distal goals take longer periods of time to achieve.

Maintenance of self-regulation is also problematic, in that self-regulation may be a limited resource that is depleted as efforts at self-regulation endure (Muraven & Baumeister 2000; Baumeister & Heatherton 1996). Hence, persistent efforts at self-regulation will deplete this resource and increase the likelihood of failure at the attempt. Similarly, multiple simultaneous attempts at self-regulation are more likely to lead to failure because there is a depleted resource of self-regulatory power in reserve for any one attempt. As a result, individuals who are trying to lose weight may be less successful at their dieting attempts if, at the same time, they are also attempting to quit smoking.

Self-Regulation – Summary

Successful self-regulation is dependent upon the time frame associated with maintaining behavioural regulation in order to achieve a goal, as well as the specific goals towards which an individual is striving. Long-term self-regulation involving an end goal that differs from the initial state (i.e., regulation towards change as opposed to regulation towards maintenance) can be considered to be a form of self-change, such that an individual has achieved and is maintaining a change in a behaviour as a result of self-regulatory processes.

Extending Self-Regulation to Enduring Self-Change

While self-regulation is clearly associated with self-change, the specific manner in which self-regulation progresses towards enduring self-change is unclear. However, it is clear that self-change is not a

dichotomous entity (i.e., behaviours either changed or not changed) but rather a dynamic process that involves maintenance of the changed behaviours. It is possible for maintenance to fail, resulting in relapse and consequently a failed effort at self-change. In fact, relapse seems to be relatively common, even in individuals who have maintained a degree of behavioural change for a period of a few months (Cohen et al. 1989). Extending the summarised literature on self-regulation, however, may provide indications on how to increase successes at achieving and maintaining self-change, and also decreasing relapses.

Prochaska, DiClemente and Norcross (1992) indicate that a series of steps are involved in self-change: precontemplation, contemplation, preparation, action, and maintenance. Individuals in the 'precontemplative' stage typically are unaware that problems with their behaviour exist, and have limited or no intentions to change their behaviour. Individuals in the 'contemplative' stage are considering taking steps towards changing their behaviour, yet have not committed to the process of making these changes. In contrast, those in the 'preparation' stage are preparing to make behavioural changes within the next month, and have had unsuccessful attempts at changing their behaviour within the last year. Those in the 'action' stage are working towards behavioural change. Finally, individuals reach a maintenance stage, in which they work to maintain the changes they have made and prevent relapse to their earlier behaviours. Prochaska and colleagues suggest that individuals making self-change attempts work through these stages in a spiral pattern, in which relapse often occurs and individuals regress to earlier stages in the model. For example, an individual who has committed to quitting smoking, and has gone several weeks without a cigarette, may have a cigarette and relapse. As a result of this relapse, regression occurs, from an 'action' or 'maintenance' stage, to an earlier stage, such as 'preparation', or even spiraling back to the first stage, 'precontemplation'.

However, the definition of these stages lacks some clarity. For example, Prochaska and colleagues (1992) describe individuals within the preparation stage as having had a recent failure at a self-change attempt. However, this does not follow given that the action stage (i.e., making behavioural changes towards a goal) follows this preparation stage. Hence, it seems somewhat impossible for one to have failed an attempt without having any actions towards that attempt. Similarly, there is inadequate attention paid to the process of goal setting within

these stages, given the importance this appears to have within the self-regulation literature. Due to the deficiencies in the literature on the processes involved in self-change, it is a valuable course of action to extend findings and assertions from the self-regulation literature to these processes. Based on this extension, three clear stages emerge: goal-setting (somewhat equivalent to the contemplation/preparation stage of Prochaska et al.), self-regulatory processes (the 'action' stage in which self-regulation and behavioural changes occur), and maintenance of self-change (a specified goal has been achieved and individuals now maintain their behavioural changes in order to maintain their goal).

Goal Setting

Of utmost importance in the realm of self-change is the facet that an individual is endeavouring to change. The goal towards which a self-changer is striving can predict to some degree the success this individual will experience in relation to the goal. If a goal is very proximal and concrete, it is a faster and perhaps easier goal to attain. However, such a goal may not entail the same degree of commitment to achievement as a more important, abstract goal (Mischel et al. 1996) and ultimately the success at the smaller goal will not necessarily be higher than that for the more abstract goal that involves more commitment.

It can be difficult to set a goal before a task is undertaken, given that the achievability of a task is dependent on both the characteristics of the task and the individual (Polivy & Herman 2002). Hence, while an abstract goal is likely to have a more desirable outcome for an individual than a more proximal goal, the feasibility of achieving the abstract goal may be unclear at the outset. As a result, goal flexibility may increase the likelihood of success (e.g., Mischel et al. 1996). If a goal is set which, after behavioural attempts to achieve this goal are made, turns out to be somewhat unrealistic, goal flexibility will allow for revision of the goal to a more realistic level. While this flexibility may not increase success for the initial goal, subsequent failures are less likely because individuals are more realistic in their capabilities, and hence a cycle of relapse and further behavioural changes to achieve an unrealistic goal will be averted (Polivy & Herman 2002).

It thus follows that successful self-change is more likely with a choice of a goal that is both meaningful to the individual (in order to increase goal commitment) and reasonable to achieve. If the goal turns out to

be somewhat unreasonable after attempts to change behaviour have been initiated, maintaining flexibility with regards to revising the goal will increase the likelihood of a degree of success, rather than having an individual give up on the goal altogether. This implies, then, that the relation between goal setting and action can be represented with a double-side arrow — action does not always follow from goal setting and the person moves on from there, but rather action may provide some indications that goal-revision is necessary. Goal setting and action towards the goal are not fixed events that have only one possible representation in a program of self-change. Rather, success may involve some wavering and trying out different goals, and individuals will be better served to recognise this from the outset as opposed to feeling like a failure for not hitting their goal for self-change on the first attempt.

Self-Regulatory Actions and Efficiency

After settling on a goal, individuals need to focus on maintaining their self-regulatory strength in order to succeed at self-regulation and ultimately self-change by effectively changing their behaviour in order to meet their goal. Individuals must understand that self-regulatory strength is not an unlimited resource (Muraven & Baumeister 2000), so they are more likely to succeed at self-change if they undertake a minimum of tasks at any one time (as opposed to attempting multiple self-change goals at once). Optimising the self-regulatory strength and efficiency will in turn enhance efforts in the self-change process.

Given that the length of time self-regulation endures is a common concept within the literature on self-regulatory failure, individuals attempting a program of self-change should heed this fact and incorporate it into their plans for behavioural change in order to increase their successes. A great deal of self-change involves a relatively lengthy time commitment, however. In order to reduce the likelihood of failure, self-changers should incorporate rewards and reinforcement into their programs in order reduce the likelihood of failure of the self-regulatory processes (e.g., Polivy & Herman 2002). This reinforcement can potentially offset some of the accrued negative consequences of long-term self-regulation.

The methods involved in self-regulation and self-change often involve behavioural inhibition (as previously indicated), which in and of itself could increase failures in attempts at self-change. Polivy (1998) hypothesises that behavioural inhibition can lead to behavioural

excess. This concept is demonstrated particularly well with dieters, who tend to disinhibit their eating as a result of their self-imposed dietary restrictions. It thus seems that dieters would fare better by approaching food in moderation, and avoid classifying certain foods as dietary forbidden. Research has demonstrated that when dieters eat a preload of food that they perceive to be a forbidden food (i.e., a milkshake), they subsequently eat more than after consuming a food they don't consider to be a forbidden food (i.e., cottage cheese), or after no preload. In contrast, the food intake of non-dieters is not influenced by consumption of a milkshake or cottage cheese (Knight & Boland 1989). These results demonstrate the all-or-none phenomenon that occurs when individuals restrain their eating — chronic dieters attempt to avoid consumption of forbidden foods, such as milkshakes, yet when presented with a preload of a forbidden food, they subsequently overeat. Knight and Boland suggest that labeling foods as dietary forbidden can lead to feelings of failure upon consumption of these foods, and can also lead to abandoning the diet. Thus, in order to achieve weight-loss, approaching food in moderation and avoiding labeling foods as dietary forbidden may prevent subsequent cycles of dietary restraint and overeating.

However, consideration of smoking presents a different perspective than eating. Whereas eating is a fundamental aspect of each and every day, and eating in moderation is advisable, there seems to be few benefits of smoking in moderation. Smokers trying to quit will attain more health benefits by cutting out their habit altogether. This, at first glance, may seem difficult to reconcile with Polivy's (1998) article on the negative consequences of behavioural inhibition. However, it is possible that substitution of another activity may decrease these negative consequences and allow the smoker to cease the habit with fewer difficulties.

Maintenance of Self-Change and Prevention of Relapse

Upon reaching the specific goal set out in the beginning of the self-regulatory process, self-change has occurred. However, the key to ensuring this change endures centres around prevention of relapse. Continuance of self-regulatory efforts to maintain behaviour at the particular goal is necessary in this stage, whereas in the previous stage behavioural changes served to reach the goal. Many of the indications for avoiding failure and ensuring self-regulatory efficiency in the previous (action) stage are applicable also to the maintenance stage. It thus becomes important for individuals wishing to undertake another

self-change attempt for a different goal to consider the efforts they must still exert to maintain the current goal, in order to avoid depleting their self-regulatory strength and subsequently failing at both endeavours (i.e. the maintained goal and the newly undertaken goal).

Maintenance, as mentioned, is a stage in which individuals must maintain self-regulation of behaviours in order to avoid relapse. It is possible that a lapse will occur in this stage; however, if an individual considers it just that (a lapse) as opposed to a complete breach of their efforts, then total failure is not inherent. Individuals who have met their goal and undergo a minor setback can maintain their efforts and return to the goal, rather than giving up on their efforts altogether.

Conclusions

An amalgamation of the research on self-regulation and self-change contributes some added insight into the processes involved in self-change and methods to increase success at this endeavour; however, many unanswered questions remain. Some future research could investigate the division between maintaining self-change versus achievement of a change (i.e., maintenance of self-regulatory efforts is no longer required, because the change can now be incorporated as part of the self). Such an investigation could provide more insight into the processes involved in self-change, as the research cited in the current paper is unable to fully account for these processes. Furthermore, increased understanding as to the similarities and differences between self-change and change under formal guidance (e.g., treatment programs) can perhaps also provide researchers with increased insight into the processes and mechanisms involved in regulation and change of behaviours.

Research on Chronic Dieting and Eating Behaviours

As indicated in the above review of self-regulation and self-change, this process often involves behavioural inhibition. As Polivy (1998) has suggested, behavioural inhibition seems to lead to behavioural excess. One area of behavioural inhibition of particular interest is dieting, and the inhibition of food intake. There is an abundance of research indicating that individuals who are chronically dieting actually increase their food intake in a variety of circumstances, thus suggesting that food restriction produces episodes of overeating (see Polivy 1996).

My Master's research, conducted under the supervision of Dr. Janet Polivy, investigated the effects of carbohydrate or protein restriction of food intake and food cravings. I found that carbohydrate restriction was associated with cravings for carbohydrates, and also resulted in increases in subsequent intake of croissants (a food that is high in carbohydrates). This research has implications for the current popularity of fad diets that recommend selectively restricting particular types of food in order to achieve weight loss. This work on food restriction and food cravings led to my doctoral research investigating the role of food cues on inducing food cravings and overeating.

My research is currently focusing on eating behaviours in female dieters and non-dieters. In particular, I am investigating whether exposure to a food cue (e.g., the smell of cookies baking) increases food intake, depending on whether an individual is a chronic dieter or a non-dieter. Previous research conducted in our laboratory has demonstrated that chronic dieters increase their food intake after exposure to food cues (e.g., Fedoroff, Polivy, & Herman 1997). In contrast, it seems that non-dieters are not influenced by these food cues, as their food intake does not change in the presence of a food cue. Researchers have postulated that this pattern of results stems from psychological and behavioural differences between non-dieters and chronic dieters. It seems that non-dieters regulate in accordance with internal cues, following normal energy regulation. In contrast, chronic dieters have to ignore their internal signals in order to eat less and achieve their goals of weight loss. However, this leaves dieters susceptible to other external signals, or cues to eat (see Polivy 1996). One such signal is exposure to food cues in the environment — as Fedoroff and her colleagues demonstrated, chronic dieters, but not non-dieters, are susceptible to these cues, and increase their food intake in the presence of these external cues. I am extending this line of research by investigating the particular circumstances in which chronic dieters may increase their food intake, in order to attempt to further elucidate the underlying mechanisms involved in this response.

More specifically, I am following the methodology employed by Fedoroff and colleagues by randomly assigning participants to either a condition in which a food cue is present (the smell of chocolate chip cookies baking), or in which this food cue is absent. I am then assessing their food intake, in addition to their self-reported mood.

Finally, I am assessing their dietary restraint, in order to classify the participants as either chronic dieters or non-dieters. This line of research provides insights into triggers for episodes of overeating, which ultimately may provide researchers and clinicians with techniques for preventing these episodes.

It is evident that chronic attempts at dieting do not necessarily lead to successful weight loss. In fact, chronic food restriction appears to lead to episodes of overeating. In terms of the implications of this area of research for the obesity epidemic, it seems that severe behavioural inhibition is unlikely to be the most successful approach to losing excess weight. As Polivy (1996) suggests, some restriction of food intake may be necessary for obese individuals who are at risk for weight-related health problems, yet a more long-term approach of moderation rather than serious restriction is likely to be more successful than the cycle of deprivation and overindulgence associated with chronic dieting.

References

Baumeister, R. F. & Heatherton, T. F. 1996, 'Self-regulation failure: An overview', *Psychological Inquiry*, vol. 7, pp. 1–15.

Carver, C. S. & Scheier, M. F. 1982, 'Control theory: A useful conceptual framework for personality-social, clinical, and health psychology', *Psychological Bulletin*, vol. 92, pp. 111–135.

Cohen, S., Lichtenstein, E., Prochaska, J. O., Rossi, J. S., Gritz, E. R., Carr, C. R., Orleans, C. T., Schoenbach, V. J., Biener, L., Abrams, D., DiClemente, C., Curry, S., Marlatt, G. A., Cummings, K. M., Emont, S. L., Giovino, G., & Ossip-Klein, D. 1989, 'Debunking Myths About Self-Quitting: Evidence from 10 prospective studies of persons who attempt to quit smoking by themselves', *American Psychologist*, vol. 44, no. 11, pp. 1355–1365.

Fedoroff, I.C., Polivy, J. & Herman, C.P. 1997, 'The effect of pre-exposure to food cues on the eating behavior of restrained and unrestrained eaters', *Appetite*, vol. 28, no. 1, pp. 33–47.

Heart and Stroke Foundation 2004, 'Press Releases — Obesity — a public or private issue?' [online], Available from: http://ww2.heartandstroke.ca/Page.asp?PageID=33&ArticleID=3531&Src=news&From=SubCategory, [Accessed February 6, 2005].

Heatherton, T. F., Polivy, J. & Herman, C.P. 1991, 'Restraint, weight loss, and variability of body weight', *Journal of Abnormal Psychology*, vol. 100, no. 1, pp. 78-83.

Higgins, E. T. 1997, 'Beyond pleasure and pain', *American Psychologist*, vol. 52, pp. 1280-1300.

Katzmarzyk, P. T. & Janssen, I. 2004, 'The economic costs associated with physical inactivity and obesity in Canada: An update', *Canadian Journal of Applied Physiology*, vol. 29, no. 1, pp. 90–115.

Knight, L. J. & Boland, F. J. 1989, 'Restrained eating: An experimental disentanglement of the disinhibiting variables of perceived calories and food type', *Journal of Abnormal Psychology*, vol. 98, no. 4, pp. 412–420.

Larsson, U., Karlsson, J. & Sullivan, M. 2002, 'Impact of overweight and obesity on health-related quality of life — A Swedish population study', *International Journal of Obesity*, vol. 26, pp. 417–424.

Mischel, W., Cantor, N., & Feldman, S. 1996, 'Principles of self-regulations: The nature of willpower and self-control', In *Social Psychology: Handbook of Basic Principles*, eds. Higgins, E. T. & Kruglanski, A. W., Guilford, New York, 329–360.

Muraven, M., & Baumeister, R. F. 2000, 'Self-regulation and depletion of limited resources: Does self-control resemble a muscle?', *Psychological Bulletin*, vol. 126, pp. 247–259.

Polivy, J. 1996, 'Psychological consequences of food restriction', *Journal of the American Dietetic Association*, vol. 96, pp. 589–592.

Polivy, J. 1998, 'The effects of behavioural inhibition: Integrating internal cues, cognition, behavior, and affect', *Psychological Inquiry*, vol. 9, pp. 181–204.

Polivy, J., & Herman, C. P. 2002, 'If at first you don't succeed: False hopes of self-change', *American Psychologist*, vol. 57, no. 9, pp. 677–689.

Prochaska, J. O., DiClemente, C. C., & Norcross, J. C. 1992, 'In search of how people change: Applications to addictive behaviors', *American Psychologist*, vol. 47, pp. 1102–1114.

Schachter, S. 1982, 'Recidivism and Self-Cure of Smoking and Obesity', *America Psychologist*, vol. 37, no. 4, pp. 436–444.

Lesbian and Gay Health Issues: OUTside of the Health Policy Arena

Andrea Daley

Introduction

Recent policy and legislative gains have improved the lives of lesbian women and gay men in Canada. However, despite gains in civil rights, for example, same-sex workplace benefits, survivor benefits, and a change in the definition of spouse, there continues to be a failure by health policy-makers to recognise sexuality as a relevant issue within the health policy arena. In this way, lesbian women and gay men, as distinct and vulnerable health populations, are invisible to health policy makers and, specific health and wellness needs related to lesbian and gay sexualities are not addressed through health service delivery. In the light of their invisibility, lesbian and gay activists are increasingly demanding forms of inclusion within the institutions of health. For example, reports produced by lesbian, gay, bisexual, and transgender (LGBT) health coalitions recommend the inclusion of lesbian and gay populations, and lesbian and gay health issues in health research policy and health care delivery policy (Ontario Association of Public Health 2000; Gay and Lesbian Health Services Saskatoon 2003 a, b; Provincial Health Council: Gender Committee, Nova Scotia 2003). In this way, lesbian and gay activists are calling attention to their exclusion from the social right to the provision of health services in accordance with the standards of the Canada Health Act.

This discussion paper will explore the invisibility of lesbian and gay sexualities in the health care policy arena by situating the struggle for the social right to health care provisions within the Canada Health Act. In doing so, a basic premise to this discussion is being identified — that the Canada Health Act represents the space where the struggle for civil rights by lesbian and gay activists meets health care. I will argue that inherent in the Canada Health Act is an ideology of

heterosexuality that is reflective of the assumptions underlying current ideas of citizenship, and that shapes a specific model of health care delivery that promotes heterosexuality while ignoring or disregarding lesbian and gay sexualities. Consequently, I will suggest that lesbian women and gay men are unable to exercise their full right to health care provisions as guaranteed by the Canada Health Act.

I will advance this argument in a series of steps. First, I will provide a brief overview of the ideology of heterosexuality in the discipline and practice of social policy. Second, I will explore how the ideology of heterosexuality is reflected in the Canadian health care delivery model using the notion of 'sexual citizenship' as an analytical lens. Finally, I will consider the implications for 'how' lesbian and gay health issues might get 'taken up' by health policy-makers when the concept of 'sexual citizenship' is privileged within the context of health policy debates. In doing so, I will introduce the notion of 'state-centred' health activism as a strategy towards the visibility and inclusion of lesbian women and gay men in the health research policy and health care delivery arena.

Social Policy and the Ideology of Heterosexuality

Carabine (1996) defines the ideology of heterosexuality within the discipline and practice of social policy as "an institutionalized system that openly promotes the belief that only heterosexuality and the heterosexual family is 'normal' and 'natural'" (p.119). The privileging of heterosexuality as the 'normal' and 'natural' sexuality consequently leads to the invisibility of or disregard for lesbian and gay sexualities. Moreover, the invisibility of lesbian and gay sexualities is promoted by a non-specificity or rather a 'not naming' within social policy that *appears* to include all but that, in fact, excludes the lives of lesbian and gay people, among others. Conceivably, the ideology of heterosexuality in social policy, and its non-specificity are largely reflective of an underlying notion of citizenship that assumes a heterosexual citizen.

In Canada, the phenomenon of non-specificity, in relation to an assumed heterosexual citizen within the practice of social policy, is most observable by underscoring the need for legislative reform towards the 'writing in' of lesbian and gay lives. To this end, LGBT activists in the province of Ontario, for example, successfully challenged the notion of an assumed heterosexual citizen underlying provincial social policy. In 1999, Bill 5 – 'An Act to Amend Certain

Statutes Because of the Supreme Court of Canada Decision in M v. H.' was passed in response to pressure from LGBT activists, and allied community groups and Members of Provincial Parliament (Legislative Assembly of the Province of Ontario 2005). Bill 5 introduced the term 'same-sex partner' into 67 provincial acts and established for same-sex relationships almost all the rights and responsibilities of opposite-sex common-law relationships (Coalition for Lesbian and Gay Rights in Ontario). Some of the acts impacted by Bill 5 include the Child and Family Service Act, the Education Act, and the Family Law Act, among others.

The following discussion offers to challenge the ideology of heterosexuality within the health policy arena, in a similar manner, by taking up a 'rights-based' approach that incorporates the notion of sexual citizenship. The 'rights-based' approach is also a response to the partial success/partial failure of current strategies to challenges heterosexism within health agencies. That is, notwithstanding local and organisational educational strategies for reducing heterosexism and homophobia, and therapeutic strategies that focus on the needs of LGBT people both heterosexism and homophobia persist within the area of health care (Bella & Yetman 2000). It is my desire that a policy oriented challenge to heterosexism within the arena of health policy will offer a possibility for systematic change.

The Sexual Citizen and Social Rights

As a more recent contribution to discourses on citizenship the notion of 'sexual citizenship' is currently being taken up in different ways by different authors. Richardson's (1998) critique of the Marshallian notion of citizenship as defined by a set of civil, political and social rights within a nation-state, for example, is useful for locating the notion of the sexual citizen within contemporary discourses on citizenship. Richardson's (1998) critique identifies new understandings of the basis for belonging that go beyond the boundaries of a nation-state largely as a function of globalization and new modes of electronic communication. Integral to conceptualising new ways of belonging, or rather, conceptualising citizenship beyond the boundaries of nation-states, requires an understanding of 'nations' as "systems of cultural representation whereby we come to imagine a shared experience of belonging to a particular community" (Richardson 1998, p.85). In this way, alternative notions of citizenship are constructed based on culture, consumerism, and sexuality, among other representations.

Following this, Richardson (1998) applies a sexuality lens to critique the traditional notion of citizenship in a manner similar to other theorists who have critiqued citizenship as both gendered and racialized. Ruth Lister (1996), for example, states that the traditional notion of citizenship was established on women's exclusion and similarly, Williams (1996) emphasis the social exclusion of Blacks through limited notions of citizenship. Richardson's critique (1998) uncovers an assumed heterosexual citizen in a way that Lister and Williams uncover an assumed male and white citizen, respectively. The revelation of an assumed heterosexual citizen creates a space whereby we can imagine and include representations of the citizen as gendered, racialized and with a particular sexuality (Taylor 1996).

In this way, Weeks (1998) suggests that the notion of sexual citizenship as a new understanding of the basis of belonging is "about enfranchisement, about inclusion, about belonging, about equity and justice, about rights balanced by responsibilities" (Weeks 1998, p. 39). To some extent, Plummer (1995) operationalises this statement by stating that sexual (intimate) citizenship is about "the *control (or not) over* one's body, feelings, relationships: *access (or not)* to representations, relationships, public spaces, etc; and s*ocial grounded choices (or not)* about identities, gender experiences" (Plummer 1995, p.151, emphasis in original).

And finally, although not exhaustively, Phelan (1995) posits that sexual citizenship is about the access to rights more generally. That is, the notion of sexual citizenship is about the extent to which a person's sexual status restricts access to citizenship in terms of social, civic, and political rights. It is this latter concept of sexual citizenship that will inform the following analysis of lesbian and gay invisibility within the Canada Health Act. In this way, the Canada Health Act represents a 'public space' or a 'social institution' where access to citizenship for lesbian women and gay men can be explored in relation to the ideology of heterosexuality and in relation to the social right to health provisions, as reflected in the Canadian health care delivery model.

The Rights Struggle: Outside of the Canada Health Act

In Canada, there is a well-developed history of rights struggle by lesbian and gay activists. During the past two decades lesbian and gay activists have successfully fought for the advancement of civil rights through legislative change. Successful struggles have included the protection against discrimination through amendments to the Canadian Charter of

Rights and Freedoms and the Canadian Human Rights Act, the right to adopt a partner's biological children through changes to family law legislation, and the right to survivor benefits through amendments to the Income Tax Act, among others (Petersen 1996). The advancement of civil rights such as those identified above is important, and has served to improve the lives of many lesbian women and gay men. However, with the exception of demands by lesbian and gay activists for the recognition of gay men and their health needs related to HIV/AIDS the struggle for rights by lesbian and gay activists has not included the full right to health provisions.

To some degree, the invisibility of the issue may be a function of the Canadian health care delivery model that positions health services as a 'right' or entitlement for all citizens. That is, in Canada, health care delivery occurs within a universal health care system that guarantees health care for all citizens defined by national identity. Health care delivery is guided by five criteria including public administration, comprehensiveness, universality, portability, and accessibility. It is the criterion of universality and accessibility to which this discussion is concerned. In terms of the issue of universality the Act states that "the health care insurance plan of a province must entitle one hundred percent of the insured persons of the province to the insured health services provided for by the plan on *uniform* terms and conditions" (Canada Health Act 1984, c.6, s.10, my emphasis).

And, In terms of the issue of accessibility the Act states that:

... the health care insurance plan of a province must provide for insured health services on uniform terms and conditions and on a basis that does not *impede or preclude*, either directly or indirectly whether by charges made to insured persons or otherwise, *reasonable access* to those services by insured persons (Canada Health Act 1984, c.6, s.12, my emphasis).

Conceivably then, lesbian women and gay men, as members of the nation-state *have already* the right and entitlement to health provisions through the Canada Health Act. However, Canadian research on the health care experiences and patterns of service use of lesbian women and gay men suggest that, despite the criterion of universality and accessibility, lesbian women and gay men do not receive health services on uniform terms and conditions. Nor are health services provided on a basis that does not impede or preclude

reasonable access. Mathieson, Bailey & Gurevich (2000), for example, state that:

> Lesbian and bisexual women continue to report barriers to care: assumptions of heterosexuality in the medical interaction; health care environments that are not reflective of alternative sexualities, such as the exclusive use of brochures/posters that do not depict lesbians or bisexuals; and outright hostility and refusal of care as expressions of homophobia in the health care system (p.186).

Similarly, the Ontario Public Health Association (2000) reports that "when lesbian and gay men engage the health care system or social services, they face biased, insensitive and inadequate practices with respect to their treatment or assessment" (p.8). And finally, Systems Failure: A Report on the Experiences of Sexual Minorities in Ontario's Health-Care and Social Service Systems (1997) states that:

> Lesbian, gay, bisexual, transgendered, and transsexual people have always been aware of the feelings of fear, anxiety, and anger that arise within the mainstream health-care and social-service organizations they contact for service. Organizations and institutions whose mandate is to help them become unsafe places where they are misunderstood, and definitely not helped. They go away fearful, angry, and mistrustful of a system that has interpreted them as non-compliant, unreachable, and resistant (Coalition for Lesbian and Gay Rights in Ontario, Executive Summary 1997, p.4).

These findings are echoed by several U.S. studies that have explored the health care experiences of lesbian women. Several authors have described lesbian women as often being ignored, dismissed, subordinated, silenced, shamed, and denigrated during their interactions with health care providers (Stevens 1994; Robertson 1992; Denenberg 1992). Additionally, research literature on the health care experiences of lesbian women cite participants as being handled in a 'rough' manner during physical examinations or having the physician overtly refuse to complete the examination by leaving the room (Simkin , 1992; Denenberg 1992).

These experiences often translate into the delay of preventative care, the failure to return for follow-up appointments, and a general reluctance to report health issues for lesbian and gay people. As importantly, lesbians and gay men are at greater risk than the general population for receiving missed diagnosis and for potentially poorer

treatment outcomes (Provincial Health Council, Nova Scotia 2003; Ontario Public Health Association 2000). Consequently, lesbian women and gay men are more likely to orient their health care around a specific health crisis.

In these ways, lesbian women and gay men do not receive *uniform* care as determined by the criteria of universality. Moreover, the assumption of heterosexuality and practices of homophobia within the realm of health care provision can be conceptualised as impeding and precluding reasonable access to health services as determined by the criteria of accessibility. This constitutes the inequitable treatment of lesbian women and gay men within the health care arena. However, this is not an inequity that can be understood within the context of contemporary health care issues, for example, waiting lists, urban vs. rural services, and the privatisation of services. Instead, this is an inequity that is produced, promoted and maintained by a health care delivery model that assumes heterosexuality as 'natural' and 'normal', and consequently, that constructs lesbian and gay sexualities as either invisible (assumption of heterosexuality) or problematic (practices of homophobia). In this way, health care delivery in Canada reflects the traditional notion of the citizen as 'naturally' heterosexual.

This analysis used the notion of sexual citizenship as an analytical tool to uncover the ideology of heterosexuality within the traditional notion of citizenship, and to highlight how this ideology as reflected in the Canadian health care delivery model is embedded in the Canada Health Act as a health policy. In this way, I have indicated that sexualizing citizenship focuses attention on the monolithic construct of the heterosexual citizen while, paradoxically creating a space for lesbian and gay visibility within the health care arena. Their visibility, however, is based on a partial citizenship, in that the ideology of heterosexuality (the assumption of heterosexuality and practices of homophobia) often prevents lesbian women and gay men from exercising their *full* right as citizens to health provisions as guaranteed by the Canada Health Act. In this way, lesbian and gay visibility is characterized by their exclusion.

The following discussion will explore how the notion of sexual citizenship can function as a remedy against the exclusion of lesbian and gay lives in health care policy and delivery. To this end, I will argue that the notion of sexual citizenship facilitates the inclusion of lesbian women and gay men by shifting their visibility from the private sphere into the public sphere.

Sexual Citizenship as a Transformative Strategy

Carabine (1996) suggests that in the absence of sexuality as an analytical lens, sexuality and social policy are positioned in relation to private relations and public policy, respectively. Consequently, sexuality in general, and lesbian and gay sexualities specifically, are relegated to the private sphere and disregarded as irrelevant for social and health policy.

The notion of sexuality as privatized or as existing within the private sphere has been conceptualised by Kitzinger (1987) as a process of individualisation. Kitzinger (1987) identifies individualisation as the "personalisation of the political through an insistent focus on the individual and internal as opposed to the institutional and sociopolitical" (p.34). Individualisation is an oppressive process that relies on the liberal tradition of 'person-blaming' and, in doing so, silences individuals and ignores the responsibilities of social institutions for social problems (Kitzinger 1987). The focus on the individual and internal depoliticizes lesbian and gay sexuality and constructs lesbian and gay health issues as independent and discrete from the social institutions that promote the ideology of heterosexuality. In this way, the health of lesbian women and gay men is often viewed in direct relation to their sexuality rather than in relation to the impact of assumed heterosexuality and practices of homophobia (the ideology of heterosexuality) within health policy and the health service delivery system.

Alternatively, using the notion of sexual citizenship to uncover of the ideology of heterosexuality within health policy (i.e. the Canada Health Act) shifts sexuality into the public sphere. This is transformative, in that, locating lesbian and gay sexualities within the public sphere facilitates an understanding of lesbian and gay health issues in relation to social institutions rather than in relation to characteristics that are intrinsic to 'homosexual' life.

Reports on the economic and human costs of homophobia by Gay and Lesbian Health Services of Saskatoon (2003a, b) for example, identify the principle determinant of increased rates in suicide, smoking, alcohol abuse, illicit drug use and depression in lesbian and gay populations as homophobia (social) rather than sexual orientation, per se (personal). Conceivably, uncovering the ideology of heterosexuality not only shifts lesbian and gay health issues into the public sphere but also suggests a social rather than a personal remedy

for its effects. To this end, health activists are demanding the inclusion of lesbian and gay lives in health research policy and health care delivery policy by advocating for the development of a national lesbian and gay research strategy, the funding of Provincial Advisory Panels on lesbian and gay health issues, and the inclusion of lesbian and gay health issues within the program standards of Provincial Mandatory Health Programs and Services Guidelines (Ontario Public Health Association, 2000). In this way, locating the source (the ideology of heterosexism) of lesbian and gay health issues outside of the person has a de-essentializing effect, and hence, the potential to reduce the likelihood of (re)pathologizing lesbian and gay sexualities through the health policy process.

Notwithstanding the transformative power of the notion of sexual citizenship in the advancement of lesbian and gay visibility within the health care policy and delivery arena there are some implications that pose potential threats to the health and well-being of lesbian and gay communities. The following discussion will outline two implications of adopting the position of sexual citizenship within health policy debates as a cautionary note to lesbian and gay health activists.

The Implications of Sexual Citizenship

Epstein (2003) defines the strategies for the inclusion of lesbian women and gay men in health research policy and delivery previously identified as a 'state-centred' type of health activism. More specifically, he defines state-centred lesbian and gay health politics as involving the:

> ... concerted efforts by advocates and researchers to make demands on the state for inclusion and incorporation — demands to institutionalize LGBT (or, often, just lesbian and gay) health as a formal concern of public health and health research bureaucracies (Epstein 2003, p.132).

Epstein (2003) acknowledges the usefulness of adopting this strategy for inclusion but cautions that an explicit integration of lesbian and gay lives into health research policy and health care delivery policy may present important consequences for lesbian and gay communities and their health issues. For the purpose of this discussion two implications will be explored. These include the potential normalizing effect of a state-centred health activism and the potential threat of identity politics.

The Normalising Effect

The potential normalising effect of a state-centred approach to the inclusion of lesbian women and gay men in health service provisions is most explicit when juxtaposed to the grass-roots feminist health movement of the 1970s and gay health activism related to the AIDS epidemic during the 1980s. Most notably, in the feminist health movement, lesbians actively critiqued the medicalisation of women's bodies as a means of social control, and advocated for the reclaiming of one's body through a local, group consciousness-raising process (Terry 1999). In 1987 the AIDS Coalition to Unleash Power (ACT UP) was formed within the Lesbian and Gay Community Services Center in Manhattan in response to the U.S. government's mismanagement of the AIDS crisis. The *modus operatus* of ACT UP was to challenge governmental resistance to funding AIDS-related research and to the approval of experimental therapies by engaging in non-violent direct action — "noisy, disruptive, and 'media-genic' street activism" (Epstein 2003, p.137). In these ways, both the feminist health movement and gay health activism employed strategies outside of the social institutions of health to focus attention on the health needs of their communities. Inherent to these strategies was the freedom to self-define, as lesbian women and gay men, what constitutes health and wellness — based on notions of identity and behaviour, and a commitment to broad-based participation.

Conversely, a state-centred health activism that seeks the inclusion of lesbian and gay sexualities in health policy engages with and relies on the principles and funding of health institutions. This poses a potential risk to the self-definition of health and wellness. In the context of state-centred health activism, groups with a critical approach to the institutions of health run the risk of being marginalised whereas groups falling in-line with these institutions are invited 'to the table'. Conceivably, this means that advocacy groups that consist of professional membership are more likely to represent the best interests of lesbian and gay communities within health policy debates (Mayer 2000). From this perspective, community-based knowledge about health and sexuality is divorced from the health policy process. In this case, accessing community-based knowledge is particularly important, and often difficult given the history of the pathologisation of LGBT people by researcher and medical communities, and the resultant mistrust of these communities by lesbian and gay communities (Bauer & Wayne 2005).

In the absence of local knowledge and representation there is a risk that lesbian and gay health will be narrowly defined within the constraints of "normal cultural behaviour" (Epstein 2003, p.156). In this way, explicit discussions of sexuality and the 'sex-positive' ideology of lesbian and gay grassroots movements may take a back seat to other health issues as defined in relation to sexual identity (Epstein 2003). For example, when lesbian women are visible within health care they are often portrayed as engaging in low-risk sex that is "chaste, dry, and monogamous (Dennenberg 1996, p. 16; Gentry 1992). This leads to the idea that lesbian women are not at risk for HIV or sexually transmitted infection, while ignoring their sexual experiences that exhibit a wide range of attitudes and behaviours (Dennenberg 1992). Carabine (in Richardson 1996) refers to this phenomenon as the normalising effect of social policy, in that, it enforces and regulates "appropriate and acceptable sexuality – hetero - and homo – sexuality" (p.61). Similarly, Meyer (2001) states:

> ... we may see that for every sensitive effort to include the target population in decision making, there may be another program that seeks to restore health by eliminating practices essential to self-expression and identity (p.858).

Conceivably then, health research policy and funding may be available to the extent that research "conforms to government prescriptions about health, sexuality, and identity" (Epstein 2003, p. 163). This represents the social reproduction of sexual minorities, in that, some members of lesbian and gay communities will be (re)marginalised or erased by institutional aspects of the health policy process (Morgan & Maskovsky 2003).

The Threat of Identity Politics

Epstein (2003) asserts that state-centred activism towards the inclusion of lesbian and gay sexualities in the health policy and delivery arena requires the transformation of "social identities into categorical identities that can be 'operationalized' and measured" (p.158). This is problematic in several ways. First, it links specific health risks to the specific practices of specific categorical identities, and therefore, assumes that all members of the group are equally at risk. Second, it assumes that all individuals within a given categorical identity accept the set of attributes that have been determined to constitute its identity as a group, and consequently ignores differences within the group. Third, it suggests an essentialist and fixed nature to

lesbian and gay identities. In these ways, state-centred health activism incorporates an identity politics strategy that conceptualises groups according to a substanstialist logic (Young 2000).

There is some risk to adopting an identity politics strategy towards the recognition of lesbian women and gay men within the context of the health policy and delivery arena. This risk is embedded in the history of the medicalisation and pathologising of lesbian and gay sexualities. That is, the need for categorical identities and its association with health risks supports the idea that certain groups are susceptible to illness as a result of their biological differences (Epstein 2003). In this way, state-centred activism that depends upon categorical identities has the potential to (re)pathologize lesbian and gay lives. Indeed, identity politics as a strategy towards recognition has been, and continues to be, critiqued by queer theorists and activists (Butler 1990).

However, the cautious use of categorical identities may prove necessary within a health policy context that bases action on "the provisional stability of categories of identity" (Phelan 1995, p. 194). Indeed, Young (2002) asserts that identity politics is both useful and effective for groups making "claims for political equality, inclusion, and appeals to justice directed at a wider public which they claim that public ought to accept" (p.86). Similarly, Phelan (1995) states that "if lesbians are to claim any interest at all, the first interest must be recognition as members of the political community" (p.203).

In consideration of the potential to (re)pathologise lesbian and gay lives, Phelan (1995) argues for an identity politics strategy that is tempered by a recognition of the provisional nature of categories of identity. To this end, the categories of 'lesbian' and 'gay' are not conceptualised as an essence but rather reflect an awareness that the lives of women and men are structured in important ways by their relationship to these categories (Phelan 1993). The use of categorical identities in this way is implicit in the previous discussion that highlighted the transformative power of the notion of sexual citizenship. That is, by focusing on the effects of assumed heterosexuality and practices of homophobia women and men are located socially in relation to the categories of 'lesbian' and 'gay'. In this way, the (re)pathologising of lesbian and gay sexualities is avoided through a focus on lesbian and gay health issues in relation to the social force of homophobia rather than in relation to characteristics that are intrinsic to 'homosexual' life.

This is beginning to happen in small ways through the insistent and tireless work of lesbian and gay health activists. For example, in Canada current health care services and practices that link the impact of heterosexism and homophobia to lesbian and gay health issues include lesbian-and gay-specific addiction programs, and hospital training programs in domestic violence that acknowledge and address violence between same-sex partners.

Conclusion

This discussion paper has explored sexuality as a relevant issue in the health policy arena. More specifically, the invisibility of lesbian and gay sexualities was examined in relation to the traditional notion of a nation-state citizenship as reflected in the Canada's health service delivery model and as embedded in the Canada Health Act, as a health policy. I argued that, as an analytical lens, the notion of sexual citizenship functions to uncover the assumption of heterosexuality underlying the traditional notion of citizenship. And, that the uncovering of this assumption brings into focus the privatisation of lesbian and gay sexualities, and hence, the exclusion of lesbian women and gay men from exercising their full social right to health provisions. I introduced the notion of 'state-centred' health activism as a concept associated with the strategy of sexual citizenship while outlining some of the potential implications for lesbian and gay communities of adopting this strategy.

Clearly, the risks to the health and well-being of lesbian women and gay men as a result of assumed heterosexuality and practices of homophobia highlight the need to address the issue of invisibility in the health policy arena. However, in consideration of the identified implications, activism that relies on the values and funding of health institutions must proceed with caution. Specifically, health activists must strongly recommend community-based representations that reflect the diversity of lesbian and gay communities based on class, race, ethnicity, age, ability, religion, and professions as well as on expressions of sexuality. Additionally, activists must demand that research strategies include and support existing lesbian and gay health researchers and networks, and promote the development of community-based knowledge. In this way, the visibility and inclusion of lesbian and gay sexualities in health research policy and health care delivery policy will function to support the full rights of lesbian women and gay men as sexual citizens.

References

Bauer, G. R., & Wayne, L.D. 2005, 'Cultural sensitivity and research involving sexual minorities', *Perspectives on Sexual and Reproductive Health*, vol. 37, no. 1, pp. 45–47.

Bella, L., & Yetman, L. 2000, *Challenging heterosexism: towards non-heterosexist policy and social security agencies for women's health*, Maritime Centre of Excellence for Women's Health, Halifax, NS.

Butler, J. 1990, *Gender Trouble*, Routledge, New York.

Carabine, J. 1996, 'Heterosexuality and social policy', In *Theorizing Heterosexuality*, ed. Richardson, D., Open University Press, Buckingham, UK, pp. 55–74.

Carabine, J. 1996, "Constructing women': women's sexuality and social policy', *Critical Social Policy: A Reader*, ed. Taylor, D., SAGE Publications Ltd., London, pp. 113 – 126.

Coalition for Lesbian & Gay Rights in Ontario 1997, *Systems Failure: A Report on the Experiences of Sexual Minorities in Ontario's Health-care and Social-services Systems.* CLGRO.

Coalition for Lesbian & Gay Rights in Ontario [online]. Available from: http://www.web.net/~clgro/home_cat.htm.

Denenberg, R. 1992, 'Invisible women: lesbians and health care', *Health/PAC Bulletin*, Spring, pp. 14 – 21.

Epstein, S. 2003, 'Sexualizing governance and medicalizing identities: the emergence of 'state-centered' LGBT health politics in the United States', *Sexualities*, vol. 6, no. 2, pp. 131 – 171.

Gay and Lesbian Health Services 2001a, *The Cost of Homophobia: Literature Review of the Economic Impact of Homophobia on Canada*, Christopher Banks, Rochon Associated Human Resource Management Consulting, Inc.

Gay and Lesbian Health Services 2001b, *The Cost of Homophobia: Literature Review of the Human Impact of Homophobia on Canada*, Christopher Banks, Rochon Associated Human Resource Management Consulting, Inc.

Gentry, S. 1992, 'Caring for lesbians in a homophobic society', *Health Care for Women International*, vol. 13, pp. 173–180.

Kitzinger, C. 1987, *The Social Construction of Lesbianism*, SAGE Publications Inc., London.

Lister, R. 1996, 'Citizenship Engendered', In *Critical Social Policy: A Reader*, ed. Taylor, D., SAGE Publications Ltd., London, pp. 168 – 174.

Mathieson, C., Bailey, N., & Gurevich, M. 2002, 'Health care services for lesbian and bisexual women: some Canadian data', *Health Care for Women International*, vol. 23, pp. 185 – 196.

Mayer, K. 2000, 'Beyond Boulder: a glance back and the road ahead', *Journal of the Gay and Lesbian Medical Association*, vol. 4, no. 1, pp. 1 – 2.

Meyer, I. 2001, 'Why lesbian, gay, bisexual, and transgender public health?', *American Journal of Public Health*, vol. 91, no. 6, pp. 865 – 859.

Morgan, S. & Maskovsky, J. 2003, 'The anthropology of welfare "reform": new perspectives on U.S. urban poverty in the post-welfare era', *Annual Review of Anthropology*, vol. 32, pp. 315 – 38.

Ontario Public Health Association 2000, *Improving the Access To and Quality of Public Health Services for Lesbians and Gay men: A Position Paper for the Ontario Public Health Association*.

Petersen, C. 1996, 'Canada', In *Unspoken Rules: Sexual Orientation and Women's Human Rights*, ed. Rosenbloom, R., Cassell, London, pp. 35 – 48.

Phelan, S. 1995, 'The space of justice: lesbians and democratic politics', In *A Simple Matter of Justice?*, ed. Wilson A. R., Cassell, New York, pp. 193 – 220.

Phelan, S. 1993, '(Be)Coming out: lesbian identity and politics', S*igns: Journal of Women in Culture and Society*, vol. 18, no. 4, pp. 765 – 790.

Plummer, K. 1995, *Telling Sexual Stories: Power, Change, and Social Worlds*, Routledge, New York.

Provincial Health Council 2003, *Provincial Health Council: Gender Committee*, Nova Scotia.

Richardson, D. 1998, 'Sexuality and citizenship', *Sociology*, vol. 32, no. 1, pp. 83 – 101.

Robertson, M. 1992, 'Lesbians as an invisible minority in the health service arena', *Health Care for Women International*, vol. 13, pp. 155 – 163.

Simkin, R. 1992, 'Lesbians face unique health care problems', *Health Sharing*, Spring/Summer, pp. 40 – 42.

Stevens, P. 1994, 'Protective strategies of lesbian clients in health care environments', *Research in Nursing and Health*, vol. 17, pp. 217 – 229.

Taylor, D. 1996, 'Social policy and social relations: introduction', In *Critical Social Policy: A Reader*, ed. Taylor, D., SAGE Publications, London, pp. 1 – 12.

Terry, J. 1999, 'Agendas for lesbian health: countering the ills of homophobia', *Revisioning Women, Health, and Healing: Feminist Cultural, and Technoscience Perspectives*, eds. Clarke, A. E. and Olesen, V. L., Routledge, New York, pp. 324 – 42.

Weeks, J. 1998, 'The sexual citizen', *Theory, Culture & Society*, vol. 15, nos. 3 – 4, pp. 35 – 52.

Williams, F. 1996, 'Racism and the discipline of social policy: a critique of welfare theory', In *Critical Social Policy: A Reader*, ed. Taylor, D., SAGE Publications, London, pp. 48 – 76.

Young, I. M. 2000, *Inclusion and Democracy*. Oxford University Press, Oxford.

Legislation

Bill 5. Legislative Assembly of the Province of Ontario [online], Available from: http://www.ontla.on.ca/Documents/StatusofLegOUT/b005_e.htm [Accessed April 2005].

Canada Health Act. [online]: http://www.hc-sc.gc.ca/medicare/chaover.htm. [Accessed June 2004].

Priority Setting in Health Care: Current Approaches and a Way Forward

Lydia Kapiriri

Introduction

In the world today, the demand for health care usually outstrips the supply of the resources to finance it. This necessitates making painful decisions about who gets what and at whose expense, with consequences that are bound to be unfavourable to one person or another. At its best, Klein described the process as "muddling through the mud" because there is lack of agreement about how best it can be done (Klein 1998). Hence, priority setting is one of the greatest challenges faced by health planners in today's health care systems (Van der Grinten 2000, Klein 1998, Williams 1988). In this paper, I present some of the arguments for and against priority setting in health, and some approaches used by industrialized and non- industrialized countries in priority setting. From these experiences I point out that focusing on technical fixes, such as strict criteria, cannot work in a value laden process such as priority setting and argue that more focus should be on the process. At the end, I present a current way of thinking about priority setting where I present the theoretical model of 'Accountability for Reasonableness'.

Background

Priority setting refers to the allocation of resources in health care, both in terms of the relative priority attached to different demands and needs and to decisions that are made not to fund treatment for individuals, groups or specific disease conditions. Ham and Coulter (2000) use priority setting interchangeably with rationing (which definition I adopt in this paper), while others differentiate the two and use priority setting to refer to limit setting decision making at the macro- level and rationing to refer to decisions at the micro or bedside

level. Priority setting occurs at national (macro- level) or regional or institution (meso-level) or the bedside (micro-level) (Ham &Coulter 2000). Priority setting is a continuous, messy, and conflict ridden process involving value judgements which vary according to the individuals and groups involved. As such, there are different perspectives about how priorities should be set and these emerge from five major disciplines: philosophy, law, political science, medicine and economics.

1) *Philosophy*: In priority setting, the discipline of philosophy contributes theories of distributive justice such as the utilitarian, egalitarian and libertarian theories. The utilitarian theory recommends doing the greatest good for the greatest number; the egalitarian theory emphasises need and equality of opportunity and the libertarian emphasises the process by which resource allocation decisions are made. The different theories lead to different conclusions.

2) *Law:* The discipline of law contributes the legal aspects of priority setting such as the legal right to health care; prohibitions against discriminations as per international conventions; and the obligations of the physician to prioritise the patient relative to treatment costs.

3) *Political Science:* The key contribution from political science is fairness of the priority setting process and participation of stakeholders, especially the public.

4) *Medicine:* Health providers and scientists provide evidence-based medicine (EBM) to contribute to the understanding of the benefits, harm and effectiveness of an intervention, thus contributing to rational decision-making.

5) *Economics*: The discipline of economics brings in the concern for efficiency, cost-effectiveness and sometimes equity in resource distribution. They tend to rely heavily on the utilitarian reasoning.

To date, the discipline of medicine providing EBM and economics, efficiency and cost-effective analysis, dominate priority setting (Martin & Singer 2000). If priority setting is based on one discipline, it may not necessarily agree with the other disciplines, or may not be all together be acceptable; countries and different institutions have

invariably used a combination of these disciplines to develop a way of thinking about priority setting.

Is It Necessary to Set Priorities?

Some economists have argued that if all markets, including health care, were perfect, priority setting would not only be unnecessary, but would also result in distorted solutions and reduced welfare. Others have argued that improvements in efficiency would avail more resources, which would bridge the gap between demand and need and hence render priority setting useless. Yet others have argued, that there's no need to set priorities in contexts of extreme resource constraints. However, in the real world, there are no perfect markets, especially within health care, and resources are never adequate, even with improved efficiency. These and several other practical reasons make priority setting inevitable for industrialised and non-industrialised countries (Ubel 2000).

Industrialised countries have a growing number of elderly people who have seemingly limitless needs and ceaseless demands, which strain the available resources. Moreover the advancement in medical technology and access to information through the internet and other media, escalates the public's expectations and demands from the health system. This leads to increase in labour intensity and the costs of health care provision. Since not all new technologies are necessary or beneficial to the patients, decisions have to be made about whether or not these should be publicly financed (Søren 2000).

Non-industrialised countries are characterised by a relatively younger population, high infant, child and adult mortality rates. Common causes of morbidity and mortality are infectious diseases. These countries are now experiencing the epidemiological transition whereby non-communicable diseases are also beginning to compete for the meagre resources available for health care. Many of these countries are poor with low GDP and a low per capita expenditure on health. For example, Canada has a GDP of 29, 235 international dollars and a per capita health expenditure of 2,792, whereas Uganda has a GPD of only 964 international dollars and a dismal per capita health expenditure of 57 dollars (WHO 2002). Additional contextual factors complicate the implementation of systematic priority setting processes. Political instability, corruption, poorly developed social

sectors, marked inefficient and inequitable resource allocation processes, and weak judicial systems make it impossible to apply the methods required for confident decision making. Moreover, these countries also lack credible information and institutions that are necessary for priority setting (Bryant 2000).

Priority Setting Experiences

Experiences from Industrialised Countries

Experiences from eight countries which pioneered systematising their priority setting processes, are presented.

The Health Commission from the State of Oregon: The American style of priority setting is described as dynamic, diverse and rarely drab *"… just like Americans…"* (Clancy & Danis, 2000). In the state of Oregon (1984), a health services commission made the first attempt to systematise rationing and to base resource allocation on a set of explicit criteria. A special appointed commission was given the duty of proposing a package of health services that should be covered by Medicaid. The commission elicited public views on conditions and values, used the cost-benefit ratio, reasonableness, influence on public, number of cases, social costs of treating or not treating and incorporated this information in deciding the priority rank order. They placed diagnoses and treatment combinations in categories according to their value to the community, to the patient and if needed for the package. In each category, the treatment and diagnoses were placed in order of their added net value and the results weighed using a number of reasonableness criteria. The output, however, gave intuitively unreasonable ranks e.g. trivial conditions like tooth capping ranked higher than appendicitis. It was also thought not to reflect the values of the public, especially their concern for the severely ill. The report was not implemented and the commission had to re-define the package (Eddy 1991 & Tengs 1996).

Holland also elected a national priority setting committee in 1991, the "Dunning Committee", which was charged with a similar responsibility. The committee developed a four tiered filter system to define a basic package for universal access:

1st tier: *Necessary care-* Defined as services that are useful to all members of the community, services aimed at restoring normal

functioning in society, services necessary because of the severity and prevalence of the disease

2ⁿᵈ tier: *Effectiveness-* Services with confirmed and documented effectiveness

3ʳᵈ tier: *Efficiency-* Services with high effectiveness and low costs

4ᵗʰ tier: *Individual responsibility-* Services that cannot be left to individual's responsibility.

These were considered under a commitment to broad solidarity and limits to rights. Services that went through the funnel were included in the basic health care package (Rijswijk 1992).

In Sweden, the parliamentary priorities commission, in 1992, defined principles that should guide priority setting. The principles included; *human dignity* where all people should be treated as equal in spite of their personal characteristics, need where resources should be devoted to those in greatest need, and *social solidarity* where the vulnerable are given more priority. They rejected cost-effectiveness or efficiency and recommended that it should only be considered at the level of the patient when deciding between treatment choices for the same disease (and severity) but not for selecting between broad categories of services. Some additional considerations included: health gain, usefulness, medical result, risk, cost/resources, quality of life and available evidence. The recommendations introduced a way of thinking about priority setting that assists the decision-maker (Einhorn 1995).

In New Zealand, the Core Services Committee, in 1992, proposed approaches to thinking about priority setting these have evolved over time. They also proposed that a list of health services should be defined. They recommended the definition of a philosophical framework to inform priority setting decisions. They advised that increased emphasis should be put on health technology assessments, cost-benefit and marginal analyses. Such information would inform health investments at macro, meso, and micro levels. On the side of the patients, the committee recommended analyses of patient's benefit from an intervention, to assess improvements and limits to survival and functioning. They also proposed that evidence-based best practice

guidelines should be developed to inform clinical decisions. Explicit tools to rank people's clinical priority and their urgency for elective surgical and other procedures should be developed. They also recommended the use of explicit public information and debate on priority questions (STAKES 1995).

Finland appointed a multidisciplinary working group on prioritisation in health in 1993. The group developed a report which was not implemented, since they used a simplistic view of the purpose of the health care system and lacked experience in priority setting. In 1996, the medical association re-activated the process and formed a committee with wide representation to form a board. They produced guidelines, following the principles of *dignity, autonomy, equality and equity*; which were publicly discussed and implemented. The priority for services provided was founded on a universal respect for human dignity, which means that each person has a right to live, to live without pain or suffering, to receive assistance in the face of a threat to life or health and to self-determination and autonomy. They made recommendations regarding ethical principles of priority setting, principles of examination and patient treatment, EBM, cost containment and public participation (Ihre 2002).

Denmark appointed a council of ethics in 1996. It quickly realised that the health care system had complex goals of disease treatment, to meet health care needs and ensure equality in health status but also to maximise population health. Hence they proposed that the focus should be on the process which should be democratic and should involve *transparency* and *accountability*. A major input to the process was to understand societal values. They developed guidelines and criteria for priority setting with the guiding principles of freedom and self-expression, safety and security, equal worth and solidarity.

Norway also elected a national priority setting committee, the Lonning's Committee, in 1987. They described five levels of priorities arranged according to severity of the condition and consequences of not treating it. The main guiding principles were *severity of disease* and *benefit of treatment*. The report was never implemented and was revised in 1996. The new report focused more on the priority setting process. They recommended that the goals of the health care system should be clearly defined and that the priority setting process should

be fair, legitimate, and transparent. They also proposed that the reasons for the choices made should be explicit, written and presented to the public, hence ensuring accountability. To facilitate this, bottom up approaches were proposed whereby specialists from the various disciplines proposed priority interventions in their speciality using the following groupings: core services, supplementary services, low priority services and services that should not be financed or reimbursed by the public health care system. In defining the core services the health state and expected benefit were used. The list developed should then be presented to the politicians who should make the final priority decisions. Decision-makers are expected to give reasons for their choices. The Norwegian priority policy gives highest priority to basic health services for patients with the most *urgent needs* and to *treatment with documented effect* and with an acceptable cost-benefit ratio (Ihre 2002).

Unlike the proceeding descriptions which involved committees at the national level, Canada has never set up a national committee for priority setting. Canada's size and level of decentralization, has resulted in priority setting occurring in different institutions and levels and hence qualifies to be called a laboratory for priority learning. Several institutions use different approaches to allocate resources. For example, the Canadian Council on Health Technology Assessment, which is responsible for identifying new technologies that should be prioritised, and the drug benefit management schemes which depend heavily on EBM and cost-effectiveness analysis. There are also diagnosis-specific groups such as Cancer Care Ontario and the Cardiac Care Network of Ontario and the Western Canada Waiting List Project (involving seven regional health authorities) (Martin & Singer 2000). The Western Canada Waiting List Project (WCWL) is founded on the assumption that *fairness in access* to health care would be enhanced if there were a *transparent* and standardised way to assess the urgency and priority of patients on waiting lists. There ought to be standard maximum waiting times which correspond to *patient urgency*. The focus of the WCWL project has been to develop and refine practical tools for prioritising patients on scheduled waiting lists. Scoring tools for priority setting were developed through extensive clinical input and highly iterative exchange by clinical panels constituted in five clinical areas: cataract surgery; general

surgery procedures; hip and knee replacement; magnetic resonance imaging (MRI) scanning, and children's mental health. Although the WCWL project may not solve the problem of waiting lists and times, they believe that having a standardized, reliable means of assigning priority for services is an important step towards improved management in Canada and elsewhere (Hadorn 2003).

Priority Setting Experiences from Non-industrialised Countries

While there is a wealth of literature about priority setting in industrialised countries, few studies have described priority setting in non-industrialised countries. Available literature indicates general lack of explicit and systematic ways for setting priorities in these contexts. As a result, drawing on international experiences with priority setting, the World Bank and the World Health Organisation recommend a way of thinking about priority setting in these contexts: the Burden of Disease (BOD) and Cost-Effectiveness Analysis (CEA). It is important to note that these are recommendations from institutions of the industrialised world, and not practices that have evolved from the non-industrialised countries themselves.

Recommended Approach to Priority Setting in Non-industrialised Countries
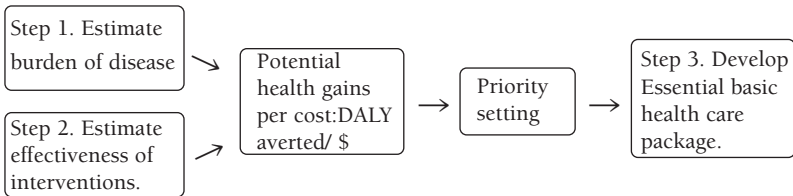
The Global Burden of Disease (GBD) study and cost-effectiveness analysis, whose results were first published in the World Bank's World Development Report of 1993, is one of the approaches proposed for use in health planning and priority setting world wide but has been mostly recommended for developing countries (Jamison, Saxenian, and Bergevin 1995). The approach is said to facilitate the inclusion of non-fatal health outcomes in debates of international health policy, and the separation of epidemiological assessments from advocacy by developing objective estimates of mortality and disability and quantification of the burden of disease in a single measure- " *Disability Adjusted Life Years* (DALYs)"; that can also be used for cost-effective analysis. DALYs lost, a measure of the burden of disease (BOD), is based on estimates of morbidity by cause, incidence, average age of onset, duration and degree of disability and time lost due to premature mortality. DALYs also incorporate social values, namely the number of years of life lost due to death at different ages, comparison of time

lived with a non-fatal health outcome and time lost due to premature death, the value of life at the different ages and discounting of future health (Murray 1996).

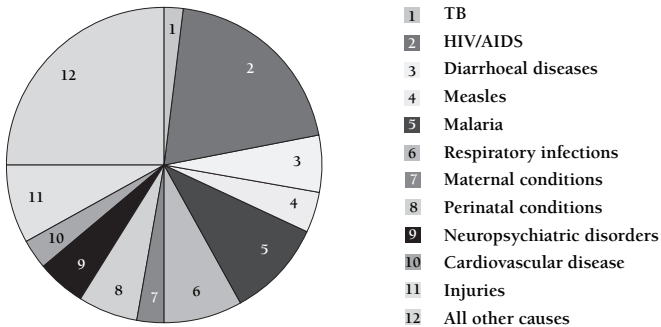## Burden of Disease (DALYs) and Priority Setting

In priority setting, leading causes of DALYs lost are identified, and cost-effectiveness of possible interventions against the leading causes are calculated. When the BOD is large, and both the cost-effectiveness of the intervention against it and the potential gains in population health are high, the intervention is considered a priority. Such interventions should comprise the essential package (Bobadilla 1992). These steps are summarised below.

Figure 1 Summary of the priority setting methodology

The leading causes of DALYs lost for developing countries are given in Figure 2.

Figure 2: Leading causes of DALYs lost for developing countries

1	TB
2	HIV/AIDS
3	Diarrhoeal diseases
4	Measles
5	Malaria
6	Respiratory infections
7	Maternal conditions
8	Perinatal conditions
9	Neuropsychiatric disorders
10	Cardiovascular disease
11	Injuries
12	All other causes

Source: WHO, 2002 graphed by author.

The cost-effectiveness ratios of possible interventions against the leading causes of DALYs lost, are then calculated. However, this is applied to individual interventions, not broadly against disease or causes which necessitates the evaluation of a wide variety of possible

interventions. The costs of an intervention may vary from country to country. The most cost-effective interventions, those that maximise the benefit (in terms of DALYs saved per unit cost) are then selected to comprise the essential basic health care package. The package is perceived as a vehicle for managing demand and improving referral. It is also thought to simplify planning and provide basis for priority setting, helping governments to focus on doing what they are capable of doing and are responsible for. At low local health expenditures, one intervention may be the best option. However, as more resources are available, an intervention that prevents more cases but costs more per unit of health benefit gained may be added. This approach is said to encourage maximum use of available information, favour systemisation of information and to make resource allocation fair (Lozano 1997).

Limitations of the BOD/ CEA Approach

The limitations of the BOD/ CEA approach come from two sources: those identified in the literature and those actually experienced by the users, some health planners in Uganda.

In the literature, we found that although the robustness and novelty of the DALY approach is hailed, concerns have been expressed with regards to the approach's conceptual, and technical basis, the limited public involvement, and the value choices used (Paalman and Bekedam 1998). These concerns were also raised by the Ugandan planners. In Uganda, the health planners appreciated the fact that it helped them systematise priority setting and its re-activating interest in evidence-based planning. However they raised three main concerns as outlined below (Kapiriri 2003).

1) The method lacks transparency

There was concern that the method lacked transparency and the components of the formulae used were not clear and too complicated for planners to understand. As such, it is poorly understood by many of the non-technical people involved in planning.

2) The method is complicated

There was a feeling that the approach was too complicated to be understood by all stakeholders who should participate in priority setting and as such, it is non-participatory. A recent study carried out in Uganda asking about the current leading actor in priority setting, and the ideal leading actors, found that health professionals,

politicians and donors (who may have the skills to understand the approach) play a leading role in priority setting. Ideally, the respondents wanted to the public and patients to play a more leading role compared to the politicians and donors.

3) It does not directly address local values and contexts of priority setting

There was concern that the values used in the burden of disease approach does not reflect values held by local populations and that it does not capture local values. The values used include disability weights, age weights and discounting which we examine in detail.

Disability weighting involves assigning of weights to years spent in different health states and allows for the incorporation of time lost due to morbidity. Disability weights are elicited using various methods. A set of 22 indicator conditions are given weights ranging from 0 (perfect health) and 1 (death) and consequently the other health conditions are fitted into this framework. Table 1 shows the weights used for some of the indicator conditions.
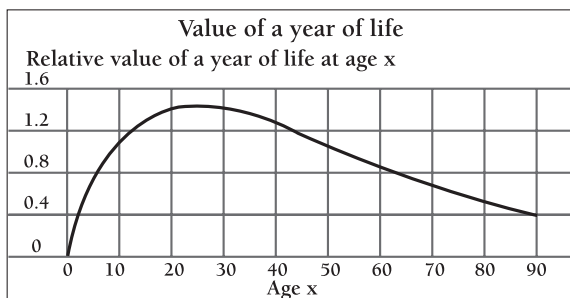
Table 1: Severity Weights of Seven Indicator Conditions

Disability class	Severity weights	Some Indicator conditions
1	0.00–0.02	Vitilago on face
2	0.02–0.12	Watery diarrhoea
3	0.12–0.24	Infertility, Erectile dysfunction
4	0.24–0.36	Below knee amputation
5	0.36–0.50	Down's syndrome, RVF
6	0.50–0.70	Unipolar major depression
7	0.70–1.00	Active psychosis, quadriplegia

(*Source: Murray & Lopez, 1997*). *Table configured by author.*

At the level of an individual; if someone lives to be 70 years, but developed schizophrenia at the age of twenty five (disability weight of 0.627), and lived with it for the rest of his life, the 55 years he lives with schizophrenia are adjusted, using the disability weight for schizophrenia, to reflect the severity of the condition using its disability weight. (Murray 1997).

Age weighting measures healthy life years lived at different ages. The value attached to years of life lived at extremes of the age curve is less than that for the mid-life years. When applied, a year of life lived at 2 years counts for only 22% of a year lived at 25 years and at 70 years, a year counts for 46% of a year lived at 25 years.

Value of a year of life
Relative value of a year of life at age x

Source: World Bank data. Table compiled by author.

Discounting of future benefit: People were asked if they would prefer a year of healthy life now, compared to a healthy year of life 10 years later. The fact that many people would prefer a healthy life year now justifies discounting of future health. When applied, one life saved today is worth more than 5 saved in 55 years, and a healthy life year, bought for 10 years will be 24% less worth that a healthy life year bought for now (Murray 1997).

Controversies with these values include their validity, ethical implications, and their generalisability. It is thought that these values may vary with the cultural, economic and social context and hence cannot be generalised. Moreover, some people have intimated that these values may not be acceptable in some local contexts and that the approach leaves out some other locally relevant value (Anand 1995 & Sayers 1997). This was also found in a study carried out in Uganda where social stigma, concern for the vulnerable populations (women and children), equity and community views were important values held by the study population yet these are not well captured in the values incorporated in the BOD approach (Kapiriri, and Arnesen 2004 & Kapiriri and Norheim 2003).

Priority Setting Lessons from the Industrialized and Non-industrialized Countries

The experiences from the industrialized countries show that there are no simple or straight forward solutions to the problem of priority setting. It is governed by many disciplines, is value laden and countries that have tried to develop explicit criteria soon find that the criteria cannot be generalised. For example, cost-effectiveness analysis is just one of many criteria and approaches, and was rejected as the

main criteria in some contexts. Commenting on this, Goold said, "*…CEA can be a valuable source of information, but is a poor "technologic fix" for a thorny problem of allocating limited health care resources…*"(Goold et al. 1998).

Since it is a political process, the context of priority setting is important, hence the limitations observed in the use of a general approach such as the BOD/CEA. While the evidence is vital, there are, as we have seen, other factors such as the political environment, power, influence, culture, etc., that may drive priority setting and these may prove stronger than the evidence; hence, the need to consider the context where priority setting occurs:"*…there is no alternative to the 'messy' process of working out solutions in the ordinary political and policy framework of a given country…*"(Jayasinghe 1998).

Currently, there is a deliberate move from the simple and technical solutions to priority setting to focus more on:

1) *The institutions which should set priorities:* These should be strengthened and have clear mechanisms for priority setting. While some countries have set up specific standing committees on priority setting (e.g. New Zealand) others work within existing structures. The focus is on wide representation from the committees and emphasis on public consultations. There is need for a conducive political and social environment.

2) *The evidence necessary for confident decision making:* There is an undeniable need for evidence based decision making. While epidemiological and scientific evidence and evidence about costs and effectiveness of interventions is very vital in priority setting, there is increasing realisation for the need to incorporate other kinds of information in priority setting. Information on the public's views on values in priority setting and also their response to decisions made should be elicited and also considered when setting priorities.

3) *The process of priority setting.* Recognising that priority setting is value laden and may have serious consequences for one person or another. There is consensus that the priority setting should be fair. It is assumed that fair processes produce fair results. "Accountability for Reasonableness" is a conceptual framework that has been used to evaluate fair processes in priority setting.

Accountability for Reasonableness

Accountability for reasonableness is theoretically grounded in justice theories emphasising democratic deliberation. It was developed in the context of real-world priority setting processes, and is therefore able to give practical guidance to decision makers. It has emerged over the past five years as a leading framework for priority setting research. According to the model, a fair priority setting process meets four conditions: relevance, publicity, appeals, and enforcement.

1) **The Publicity Condition:** Decisions regarding both direct and indirect limits to care and their rationales must be publicly accessible.

2) **The Relevance Condition:** The rationales for limit-setting decisions should aim to provide reasonable explanation of how the organisation seeks to provide "value for money" in meeting the varied health needs of a defined population under reasonable resource constraints. Specifically, a rationale will be reasonable if it appeals to evidence, reason, and principles that are accepted as relevant by fair-minded people who are disposed to finding mutually justifiable terms of co-operation. 'Fair-minded' people seek to cooperate according to terms they can justify to each other — this narrows, though does not eliminate, the scope of controversy, which is further narrowed by specifying that reasons must be relevant to the specific priority setting context.

3) **The Revision and Appeals Condition:** There must be mechanisms for challenging and for resolving disputes regarding limit-setting decisions, and more broadly, opportunities for revision and improvement of policies in the light of new evidence or arguments.

4) **The Regulative Condition:** There is either voluntary or public regulation of the process to ensure that conditions 1–3 are met (Daniels 2002).

According to Daniels (2002), 'accountability for reasonableness' provides a common language for discussing priority setting and in doing so facilitates the public's understanding for the need for limits and conditions for making decisions about them. *Relevance* and *Publicity* are conditions that serve as 'case law' in establishing the means to set limits over time in a transparent and responsive process.

It is thought that if these four conditions were met the choices thus made would be fair and legitimate. The procedure also avails opportunities, for public participation at the different stages of the decision making process. Since many countries subscribe to the concepts highlighted in this approach, Accountability for Reasonableness is increasingly being accepted as a leading framework for assessing fairness in priority setting and has been used in several contexts in both industrialized and non-industrialized countries (Singer et al., 2000 & Martin et al. 2003 & Reeleder et al. 2005 & Mielke and Kalangu 2003 & Daniels et al. 2000).

Conclusions

Priority setting remains a challenge for both industrialized and non-industrialized countries. Different disciplines govern priority setting and these have been invariably used to develop approaches. Based on these, countries and/or organisations have developed explicit criteria, others have general guidelines and principles that should govern priority setting. Yet others have used some criteria to develop a basic care package which excludes all treatment against certain diagnoses. Since several of these approaches have been criticized, and there is lack of agreement of which criteria or how much weight to give to which criteria, this leaves decision makers with a challenge. There is an urgent need for devising strategies for improving priority setting processes in the different contexts.

Normative approaches (i.e. philosophical theories of distributive justice) are necessary because they help identify key values that clarify policy choices, but they are insufficient because different approaches lead to different conclusions and there is no consensus about which ones are correct. Empirical approaches are necessary because they help to identify what is being done and what can be done, but they too are insufficient because they cannot identify what should be done. Moreover, to be really helpful an improvement strategy must utilise rigorous research methods that are able to analyze and capture experience so that past problems are corrected and lessons can be shared with others.

Therefore, a constructive, practical, and accessible improvement strategy must be research-based and combine both normative and empirical methods. This involves combining two linked methods: case study research to describe priority setting and interdisciplinary research to evaluate the description using a conceptual framework for

fair priority setting processes, leading to evidence-based and context sensitive and context specific recommendations for improvements.

Comparative studies would facilitate the sharing of experiences across organisations, regions and nations, and enrich the current knowledge and stimulate debate on priority setting This, with strong institutions, availability of relevant evidence and a conducive political, cultural and social environment , would contribute to the improvement of priority setting and strengthening of health care systems.

References

Anand, S., Hanson K. 1995, *Disability-Adjusted life Years: a critical review*, Harvard School of Public Health, Cambridge.

Bobadilla, J. L., Cowley, P., Musgrove, P., Saxenian, H. 1992, 'Design, content and financing of an essential national package of health services', *Bulletin of the World Health Organisation*, vol. 72, no. 4, pp. 653–662.

Bryant, J. H. 2000, 'Health priority dilemmas in developing countries', In *The Global Challenge of Health Care rationing*, eds. Coulter, A., and Ham, C., Open University Press, Philadelphia, pp. 63–73.

Daniels, N., Bryant, J. 2000, Castano, R. A., Dantes, O. G., Khan, K. S., Pannarunothai, S.. 'Benchmarks of fairness for health care reform: a policy tool for developing countries', *Bull World Health Organ*, vol. 78, no. 6, pp. 740–50.

Daniels, N., Sabin, J. E., 2002, *Setting limits fairly. Can we learn to share medical resources?*, Oxford University Press, New York.

Eddy, D. M. 1991, 'Oregon's Methods. Did cost- effectiveness analysis fail?', *Journal of American Medical Association*, vol. 266, no. 15, pp. 2135–2141.

Edgar, W. 2002, *Successes and failures in priority setting. 4th International conference on priority setting in health*. Oslo.

Einhorn, J. 1995, Vårdens Svårå val. Stockholm: Statens offentligar; social departementet, Stockholm 123 pgs.

Hadorn D. 2003, 'Steering Committee of the Western Canada Waiting List Project. Setting priorities on waiting lists: point-count systems as linear models', *Journal of Health Services Research Policy*, vol. 8, no. 1, pp. 48–54.

Ham, C., Coulter, A. 2000, 'International experience of rationing (or priority setting)', In: *The Global Challenge of Health Care Rationing*, eds. Coulter, A. and Ham, C., Open University Press, Philadelphia, pp. 1–12.

Ihre, T. 2002, *Report from the committee for prioritisation. 4th International conference on priority setting in health*. Oslo.

Jamison, D. T., Saxenian, H., Bergevin, Y. 1995, 'Investing in Health wisely: The role of needs-based technology assessment', *International Journal of technology assessment in health care*, vol. 11, no. 4, pp. 673– 684.

Kapiriri, L., Norheim, O. F., Heggenhouugen, K. 2003, 'Using burden of disease information for health planning in developing countries: the experiences from Uganda', *Social Science and Medicine* 2003, vol. , 56 (12): 2433–2441

Kapiriri, L., Arnesen, T., Norheim, O. F. 2004, 'Is cost-effectiveness analysis preferred to severity of disease as the main guiding principle in priority setting in resource poor settings? The case of Uganda', *Cost Effectiveness and Resource Allocation*, vol. 2, no. 1, p. 1.

Kapiriri, L. Norheim, O. F. 2004, 'Criteria for priority setting in health care in Uganda: exploration of stakeholders' values', *Bulletin of the World Health Organization*, vol. 82, pp. 172–179.

Klein, R. 1998, 'Puzzling out priorities. Why we must acknowledge that rationing is a political process', *British Medical Journal*, vol. 317, pp. 959–960.

Lozano, R. 1997, 'The burden of disease in Mexico in 1994. Advances and challenges. Funsalud, Mexico: *El Observatorio de la Salud, Necesidades, Servicios, Politicas.*

Martin, D. K., Hollenberg, D., MacRae, S., Madden, S., Singer, P. 2003 'Priority setting in a hospital drug formulary: a qualitative case study and evaluation', *Health Policy*, vol. 66, no. 3, pp. 295–303.

Martin D. K., Singer P. A. 2000, 'Priority setting and health technology assessment: beyond evidence-based medicine and cost-effectiveness analysis', In: *The Global Challenge of Health Care rationing*, eds. Coulter, A. and Ham, C., Open University Press, Philadelphia, pp. 135–145.

Mielke, J., Kalangu, K. K. 2003, 'The surgeon and human immunodeficiency virus,' *World Journal of Surgery*, vol. 27, no. 8, pp. 967–71.

Murray, C. J. L., Lopez, A. D. 1997, 'Point of view. The Utility of DALYs for public health policy and research: a reply', *Bulletin of the World Health Organisation* vol. 75, no 4, pp. 377–381.

Murray, C. J. L. 1996, 'Rethinking DALYs,' In *The Global burden of disease*, eds. Murray, C. J. L., and Lopez, A. D., Harvard University Press, Cambridge, pp. 1– 98.

Paalman, M., Bekedam, H., Hawken, L., Nyheim, D. 1998, 'A critical review of priority setting in the health sector: the methodology of the 1993 World Development Report', *Health Policy and Planning*, vol. 13, no. 1, pp. 13– 31.

Reeleder, D., Martin, D. K., Keresztes, C., Singer, P. A. 2005, 'what do hospital decision-makers in Ontario, Canada, have to say about the fairness of priority setting in their institutions?', *Health Services Research*, vol. 5, no. 1, p. 8.

Sayers, B., Fliedner, T. M. 1997, 'The critique of DALYs: a counter- reply', *Bulletin of the World Health Organisation*, vol. 75, no. 4, pp. 383–384.

Singer, P. A., Martin, D. K., Giacomini, M., Purdy, L. 2000, 'Priority setting for new technologies in medicine: qualitative case study', *British Medical Journal*, vol. 321, pp. 1316–1318.

Søren, H. 2000, 'Developments in the Nordic countries-goodbye to the simple solutions', In *The Global Challenge of Health Care rationing*, eds. Coulter, A., and Ham, C., Open University Press, Philadelphia, pp. 29–37.

STAKES 1995, *From values to choices*, Helsinki: National Research and Development Centre for Welfare and Health.

Tengs, T. O., Meyer, G., Siegel, J. E., Pliskin, J. S., Graham, J. D., Weinstein, M .C. 1996, 'Oregon's Medicaid Ranking and cost-effectiveness: Is there any relationship?', *Medical Decision Making*, vol. 16, no. 2, pp. 99–107.

The Government Committee on choices in health care (The Netherlands) 1992, *Choices in Health Care,* Rijswijk: Ministry of Welfare, Health and cultural affairs.

Ubel, P. A. 2000, *Pricing Life. Why It's Time for Health Care Rationing*, The MIT Press, (McGee, Caplan A, eds).

Van der Grinten, T. E. D. 2000, *Actors in priority setting: Intended roles and actual behaviour. 3rd International conference on priorities in health care*, Amsterdam.

Williams, A., 1988, 'Priority setting in public and private health care systems: a guide through the ideological jungle', *Journal of Health Economics*, vol. 7, no. 2, pp. 173–183.

World Health Organisation 2002, *The World Health Report, 2002. Reducing risks, promoting healthy life*, Geneva: World Health Organisation.

Biopolitics and the Body Politic: Anti-vaccinationism in Canada from a Historical Perspective

Jennifer E. Keelan

Introduction: From Access and Equity to Compliance

In the early 1980s a series of new vaccines became available for a variety of infectious diseases and issues of equity and access to these new vaccines dominated government and professional public health discourse on immunization. Some provinces adopted new vaccines quickly whilst others lagged behind, leaving a patchwork of vaccine coverage against various vaccine-preventable diseases: "Despite an expert committee's recommendation that there be routine immunization for 13 infectious diseases, coverage remains uneven" (Sibbald 2003, p. 598). In regions where governments did not cover the costs for new vaccines, parents were able to purchase them for their children, but the costs, often up to $800 per child, made new vaccinations inaccessible for working-class families and made it awkward for family practitioners to advocate for a complete immunization schedule (Paterson et al. 2004). In 2003, the federal government became directly involved in immunization delivery by promising provinces and territories access to $300 million dollars to bring their programs up to the standards recommended by the National Advisory Committee on Immunization (NACI) and the newly formulated National Immunization Strategy (NIS). After only a few years of federal funding (2003–5), the disparity in access to new vaccines dramatically decreased, though differences in the provincial and territorial schedules still remain (Government of Canada Department of Finance 2004, PHAC 2005).

While the NIS has, to some degree, addressed the ongoing issues of access to new vaccines, and made vaccination delivery more equitable, childhood vaccination rates continue to be checked by parents refusing routine immunization for their children. The vigorous resurgence of resistance to vaccination in the late 1980s and the

revival of active anti-vaccination societies in Canada, the US and Britain, took many vaccination advocates by surprise. While there have been few studies on the prevalence of resistance, the trend is troubling and it remains a significant barrier to the goal to achieve the level of vaccination required for herd immunity. For example, in a 1987 survey of the uptake of Haemophilus influenza type b vaccine, one third of respondents chose not to vaccinate their children (Shawn & Gold 1987). In a more recent survey in Montreal, one third of two year olds in the survey had an incomplete series of childhood immunizations (Pinker 1999).

Parents explained their reluctance to vaccinate in a variety of ways. They reported that they were overwhelmed by the sheer number of routine vaccinations and were concerned about the short and long-term safety. The study's authors also suspected that there were other socio-economic factors that contributed to under-immunization, such as surcharges for delivery of vaccination, and the inconvenience and financial burden for parents who lost hours of work by complying with routine immunization (25% of parents missed work because their children were sick after vaccination) (Pinker 1999).

Parents choosing not to vaccinate their children, or delaying vaccination, are often seen as instances where the public profoundly misunderstands risk and the technical decision making of medical scientists. Unfortunately, this model of the public's understanding of science, described by Brian Wynne (1995) as the knowledge 'deficit' model, limits the realm of possible explanations for resistance to a lack of scientific fluency. Critiques of resistance also portray the public as prey to unscientific alternative medical movements that have historically been strongly associated with anti-vaccinationism (Gross 1994, Kaufman 1967, Leask 2002). In other words, people resist vaccination because they cannot distinguish good science from bad science. Framing the problem in this way also directs researchers to undertake intensive educational campaigns which make strenuous attempts to *translate* expert medical knowledge and risk assessment for a lay audience. Researchers have noted, however, that information campaigns targeting resisters often have little impact on their target audience (for example see Shawn & Gold 1987).

This article will attempt to illustrate why information campaigns designed to *translate* medical expertise and authority largely failed and still fail to counteract anti-vaccination sentiments. A historical examination of resistance to vaccination in Montreal, during the 1885

smallpox epidemic, yields interesting parallels to the modern anti-vaccination movement and provides a helpful vantage point to assess nascent anti-vaccinationism. This case study shows that the attempts to compel citizens to be vaccinated invoked a complex web of social, political and cultural commitments that went far beyond any simple calculus of risk (risk of dying from vaccination versus the risk of dying from smallpox). These commitments cannot be easily disentangled from the technological or scientific arguments on *either* side of the vaccination debate and hinge on differing perceptions of political enfranchisement, notions of citizenship, and trust in government authorities and medical expertise.

Resistance to Vaccination in Montreal during the 1885 Smallpox Epidemic

Late nineteenth century vaccine technology, following its predecessor inoculation, was a holistic practice that involved an increasingly complex system of interpretation and monitoring to predict whether the individual's vaccination was effective. At every stage of the procedure, faith in the physician's judgment was paramount. Lymph was taken from the pustules of cows suffering from a 'good case of the cowpox' and used as stock vaccine while, in general, the vaccine received by the public was selected from recently vaccinated 'healthy' children whose bodies served as vaccine factories. Good lymph gave good protection but bad lymph could cause painful side-effects, spread disease and even rarely caused the death of the patient. Official records of deaths from smallpox *vaccination* itself were in the hundreds in the 1890s in Britain at the same time the incidence of smallpox sharply declined (Final Report RCV 1898, p. 122).

Once vaccinated, the physician then had to verify that the vaccine actually 'took'. This was the point where the person normally received a certificate of vaccination and in some countries, their vaccination was recorded and the procedure was officially deemed a success. However, during smallpox epidemics in Canada, many of these traditions broke down under the strain of mass vaccination. For example a public vaccinator might vaccinate over 200 people in one day. Poorer parents being vaccinated in the public stations had almost no control over the practice and the results were often left unconfirmed and their vaccination status ambiguous. For example they could not select a 'healthy' or baby from a family they knew to provide their child's vaccine lymph, as the upper classes generally did. Significantly, citizens had to judge whether or not good lymph would

be used based on their faith in the public health officers who were hired by the municipal government. Certificates of vaccination were issued, but the real evidence that a vaccine took was recorded literally on the person's body in the form of one to four scars.
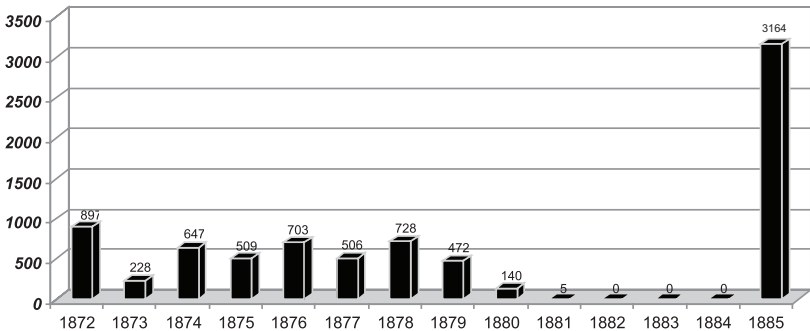
Rejection of smallpox vaccination as a prophylactic measure, and formal resistance to compulsory vaccination laws preceded the great epidemic of 1885. The first Canadian Anti-vaccination League was formed in Montreal in the early 1870s to counter municipal efforts to make infant vaccination compulsory. Anti-vaccinationists argued that vaccine-immunity had not been *demonstrated* and this had serious consequences for how they read and understood empirical data supporting the practice. In turn, they argued that compulsory vaccination was a class-based legislation whose function was to exert control over the working classes. These arguments were successfully taken up by French nationalists in Montreal who used vaccination to promote discord between the French and English. By disseminating reports of serious side-effects from public vaccinators' vaccine, questioning the theoretical and empirical basis for vaccine programs, and by adding the cultural argument that compulsory vaccination was a part of paternalistic and monopolistic medical profession, anti-vaccinationists both stimulated and reflected the scope of popular resistance to compulsory vaccination among the working class.

In 1885, the debate over vaccination become even more material when a smallpox epidemic struck Montreal killing over 3000 people, mostly children under the age of fifteen. Montreal (a city of approximately 167, 000 people) was a city divided religiously and linguistically: the Catholic French majority formed most of the city's industrial workers and the Protestant English minority was over represented in the merchant and banking industries (LaBerge 1885, p. 71). The class construction of late nineteenth century Montreal is richly described by Herbert Brown Ames in his sociological study of Montreal's classes,

'The city above the hill' is the home of the classes. Within its well built residences will be found the captains of industry, the owners of real estate, and those who labor with brain rather than hand…It is the exclusive habitat of the rich and of the well-to-do … The 'city below the hill,' on the other hand, is the dwelling place of the masses …The 'city below the hill' is the home of the craftsman, of the manual wage-earner, of the mechanic and the clerk, and three-quarters of its population belong to this, the real industrial class (Brown Ames 1972 [1897], p. 6).

Approximately two-thirds of Montreal workers lived among the various factories. Tanneries, mills, private slaughter houses, and boot and cigar factories dominated the French part of the city. There was an antiquated system of wooden drains that led to an outdated and unreliable sewer system. It was thought that the summer's heat, in combination with unsanitary state of the city streets, lack of good drainage and murky water provided ample breeding grounds for germs — this was the common explanation for the root cause of smallpox. Smallpox was a fairly familiar disease; it was endemic in Montreal between 1872 and 1880. Until 1880, it claimed between 228 and 897 lives every year. Most of the deaths occurred in very young children, a pattern that mirrored the mortality of other diseases such as diphtheria (LaBerge 1885).

Table 1: Deaths from Smallpox in Montreal 1870–1885 (Montreal Board of Health Annual Reports 1882–1885). Table compiled by author.

The Montreal Civic Health Committee's mandate was to improve the sanitary state of the city to ward off a variety of infectious diseases that were thought to be spread by filth and dirty water. From 1850 on it sat sporadically as a sub-committee of the city's council. After 1875 the Committee began publishing regular annual reports detailing the unsanitary state of Montreal. Over the next five years, while smallpox remained a regular feature of city life, repeated attempts failed to make vaccination compulsory and it was under-utilised, especially in French children under the age of five, the age group hit hardest by smallpox.

By early September of 1885, reports in the international press claimed that smallpox was raging unchecked through the city because of French obstinacy and poor governance by the City Council (*Montreal Gazette* 19 August 1885). An infamous article appearing in the radical Protestant newspaper, the *Montreal Herald*, argued that the smallpox

epidemic was the fault of the French Canadians who would not be vaccinated, would *not* obey the sanitary laws, and were generally unclean.

> It is the French part of the community who are responsible for the present condition of things; call a spade a spade and put the blame properly where it belongs. It is everywhere the cry, Your French operatives, they are dirty, they do not vaccinate, and you have the pestilence disease always with you, and always will so long as your Council and English-speaking people act as they do ... Let English capitalists, manufacturers and employers of labor, drop off all the French help, have only English speaking people who are vaccinated, and who are not afraid to use soap and water, and it will soon be seen how it will stir action (*Montreal Herald* 2 September 1885).

Attempts to impose compulsory vaccination were extremely unpopular among French working class. A new municipal by-law stipulated that all infants over the age of four months and all un-vaccinated citizens must undergo immediate vaccination. Citizens who would not cooperate and be vaccinated or have their children vaccinated faced dismissal from their work in the factories, up to a twenty dollar fine (a month's wages for a worker)(Bradbury 1993, p. 234), and jail time (Keelan 2004, p. 239). Court issued vaccinations were performed in front of the Sanitary Court by an approved vaccinator to avoid any possibility of mischief or deceit (Keelan 2004, p. 241). Convictions under the Health Laws peaked in January of 1886. However, Montreal's experiment with rigorously enforced compulsion was short lived. Only a few months after the dreaded epidemic, there was a municipal election and the newly constituted Montreal Board of Health single-handedly dismantled the compulsory vaccination program.

Invoking compulsory vaccination put the evidence surrounding vaccination under enormous scrutiny since it was the first compulsory medical procedure, billed as a demonstrated scientific truth, and described as illustrating a *natural law* of immunity. Belief in this natural law led to a very particular (and tautological) understanding of the clinical epidemiology. A 'perfect' vaccine *prevented* smallpox, or severely *mitigated* the clinical expression of the disease. Those with smallpox who showed signs of good vaccination would have had a more severe case had they not been vaccinated. Physician Charles Killack Millard wrote in the late 1880s that any patient suffering a

mild case of smallpox, who claimed to be unvaccinated, must have simply forgotten that they had been vaccinated (Dixon 1962, p. 2). Under this theoretical structure, vaccination that failed to protect was a failure of the technology or the practice not the principle. For those not committed to vaccination, the epidemiology and the disease were experienced and understood differently and the interpretive framework needed to support the practice became increasingly incommensurable with the anti-vaccinationists' viewpoint.

The classic *demonstration* of vaccine's protective power was a challenge with inoculation of live smallpox. Since this was illegal in the late nineteenth century, a straight forward *demonstration* of immunity was impossible and the analysis had to shift to the very imperfect population data and the new proto-statistical sciences. Research projects in the UK and elsewhere suggested that much more sophisticated data collection and analysis were required to make the kinds of causal claims made for vaccine-induced immunity. As the century progressed, so too did the anomalous data. Some of the best vaccinators reported cases where patients had a confirmed cowpox infection, then caught a severe form of smallpox and the reverse scenario was also reported. More troubling were cases where individuals whose vaccination immunity was confirmed with a failed inoculation of live smallpox, but caught a serious form of smallpox later in life. There were reports of people naturally immune to both smallpox and cowpox. For example, one Canadian physician working at L'Hôtel Dieu attended hundreds of cases of smallpox without ever having been vaccinated himself. He reported this in 1871 to the Medico-Chirurgical Society of Montreal (Plante 1871).

Vaccination seemed to *sometimes* provide *specific* protection against smallpox. Wild smallpox itself was mutable and different forms of the disease could be transmitted from one unvaccinated person to another. For example, an unvaccinated person infected with one form of smallpox, such as a mild but distinct smallpox case, could be shown to infect another unvaccinated person with a more serious form of smallpox. Thus, how could the clinical data, largely smallpox hospital data, be used to judge vaccination's effects?

An example of how theoretical understanding of vaccination shaped the reading of epidemiological data is demonstrated in this excerpt from the *Canadian Journal of Medical Sciences* in 1876:

> We have recently passed through a pretty severe epidemic, in which a large number have been attacked; and we think that two

things have been amply demonstrated. First, the great majority of those who have passed through critical attacks have been *unvaccinated, indifferently vaccinated or not successfully vaccinated, for many years previously*. Secondly, it has been clearly shown that where persons recently vaccinated successfully have been attacked, they have passed through a *modified form of the disease*. It has been further shown pretty conclusively that most persons exposed, but recently protected, have escaped altogether (emphasis added, Wright 1876, p. 59).

At first glance this stance seems perfectly supported and reasonable. The critical elements in the above quote however relate to the tricky concepts of what defined a successful and unsuccessful vaccination. Physicians had to develop a system for predicting whether or not a vaccination was real or spurious.

Though it was often argued that a good vaccination provided perfect protection against smallpox, or at least protected as much as a primary infection, smallpox hospital data did not support this. Merely having been vaccinated neither prevented a person from catching smallpox nor dying from the disease, though it was argued that the relative risk of death was much lower among the vaccinated. However, data from Montreal showed that over fifty per cent of patients admitted to smallpox hospitals were vaccinated at a time when the overall vaccination coverage in the population was unknown but probably not much higher. Even data of esteemed pro-vaccinationists did not always support vaccine's efficacy without a number of imposed qualifiers.

By the 1870s, the number of vaccine scars was seen as a critical marker for efficacy. The number of scars correlated with the number of discrete colonies raised during a primary vaccination. It did not usually indicate serial vaccinations as serial vaccinations rarely raised good vaccine scars. Thus someone who had four pronounced smallpox scars likely had four distinct vaccinia colonies raised during their primary vaccination rather than being vaccinated four times. This is important information when trying to decipher nineteenth-century smallpox epidemiology. Hospital data showed that there was a marked difference in the sign of protection against smallpox given respectively by four, three, two and one cicatrices. In one data set, the incidence of patients admitted with four scars differed from those with only one scar by a factor of fourteen, and the incidence of those admitted who had been vaccinated but did not have a clear scar was

forty-two times larger than the population admitted with four clear vaccine scars. Thus, the categorisation of vaccine status or mark as 'true' and 'false' helped stabilize the definition of clinical syndromes of smallpox which in turn became an index of vaccine's efficacy. However, anti-vaccinationists argued that the four scar technique was not widespread and the low rate of admission to the smallpox hospital of those with four scars did not represent any more protection against smallpox, rather it represented the low frequency of the four scar method.

Vaccine scars had to be interpreted as true or false based on the patient's history and the physician's judgment.

> It is often very hard to determine whether the patient has been vaccinated or not. In many cases there is no mark or hardly any mark. The people, when they come in, are generally delirious, and unable to give any information. Even when they are able, the information is not always reliable. We may think that there should be a good typical mark as the evidence of vaccination. Sometimes the vaccine does not take. Sometimes there is nothing to show but the scratch of the lancet. But in a good case there should be a well defined mark. The preparation of the statistics referred to is, therefore, rendered a matter of some difficulty (*Montreal Gazette* 15 August 1885).

As the Montreal epidemic of 1885 progressed it was argued repeatedly that a single vaccine scar afforded scarcely any real protection, and hence was not a real sign of immunity. Certainly, many physicians felt that during an epidemic, when in doubt, re-vaccinate, and if it did not take, then the original vaccine was sound (*Montreal Gazette* 18 August 1885).

Arguments for compulsory vaccination had always hinged on the proof that it prevented smallpox and nearly always prevented the person from suffering from the most severe forms of smallpox. This was not clearly the case in the Montreal smallpox epidemic, as forty per cent of the admissions to the smallpox hospital were vaccinated, and nearly half of the cases of hemorrhagic smallpox (a highly lethal form) had at least one vaccine mark. From the data reported from all those officially diagnosed with smallpox (including both hospital and non-hospital cases), and reported to the Public Health offices, 2,471 were not vaccinated, 1,113 had *doubtful* vaccinations, and 1,187 were vaccinated (LaBerge 1885, pp. 55–62). While eleven per cent of

English patients admitted to hospital during the epidemic had three vaccine scars, increasingly the 'real' sign of protection, in the same patient grouping, less than two per cent of French Canadians had three marks (LaBerge 1885, pp. 55–62). In addition, in the records of all smallpox cases reported to the city, of which over 85 per cent were identified as French Canadian, nearly 50 per cent of the vaccinations were judged 'doubtful' (LaBerge 1885, pp. 46–47). The difference between the *mortality* rate of the vaccinated and unvaccinated was approximately eighteen per cent—but the comparison relied on the physician's ability to accurately distinguish between a spurious vaccination and a real one. The subjective judgment of clinicians in the field could not be easily purged from the data and the corresponding statistical analyses. Public health officials stressed that the data was tainted by a large percentage of bad vaccines; a failure of the practice, not the principle. At the same time they urged the French working class to submit to vaccination in the public stations—in other words to defer to the collective expertise and judgment of the medical profession and to trust municipal vaccinators.

Beyond the possibility that the French actually received poorer quality vaccinations (they were more likely to have been vaccinated by the public vaccinators in an assembly-line process) there were also a number of examples where the vaccine scar was considered suspect just because the person was French Canadian; the person in some cases was forcibly re-vaccinated, " ... unless there was a fresh scar, the passenger is obliged to again undergo the operation or leave the train" (Ross 1883–1890, volume 2). In the following interview, the examining physician for the United States government described his perception that the French were generally resistant to vaccination and hence, their vaccination certificates required careful perusal.

'From what class of people do you experience the most trouble?'

'From the French-Canadians from the country towns. They would as soon have the smallpox as to be vaccinated, and I believe most of them who object would rather.'

'How about the certificates, are they all right?'

'Well, we can't tell. I presume some of them, from what I have heard, are forged. It is hard to tell, they are of so many kinds. Not half of them are on the printed blanks furnished by the Civic Board of your city. Some are written in French and some are in English, and they contain everything from just what is wanted to a certificate that no disease is, ever was, or ever will be in the

family of the bearer. Others don't certify to anything'.(Ross 1883–1890, volume 2).

Moreover, one journalist argued that most people with 'false' vaccinations were from the French-Canadian community. Since inspection of the vaccine scar in these cases normally only occurred if the vaccine certificate could not be produced or itself was suspect, it is possible that vaccinated French Canadians were targeted for immediate, public, forced re-vaccination.

The French working class was also targeted directly by both English and French anti-vaccinationists' campaigns who provided alternative explanations for the epidemic and recommended their own strategies for protection against the disease; organisers distributed circulars and leaflets that were reportedly influential. Several volumes *L'anti-vaccinateur canadien-français* were published during the epidemic. The journal analysed the statistics from the epidemic and collected reports of complications from vaccine as well as numerous detailed testimonials enumerating smallpox among the vaccinated (to dispute its efficacy). Broadsheets in circulation depicted a mother fleeing with her child from the vaccinator who is closely followed by death. Thousands of these circulars were distributed (Keelan 2004, p. 195).

Without a doubt, the construction of risks from smallpox and the benefits of vaccine were deeply politicized entities. The public debate over compulsory vaccination was driven by political and economic interests of the major media stakeholders. For example the *Montreal Herald's* inflammatory editorials, which suggested purging the factories of French labour, enraged the French conservatives giving the French nationalist movement a powerful propaganda weapon. French nationalist newspapers had an equally vested interest in using compulsory vaccination as a wedge issue. *La Patrie* suggested that public smallpox vaccine was poisoned and was part of an English plot to weaken or exterminate the French race in Canada; the theory is less outrageous if one considers the less than ideal conditions of the public vaccination service (*La Patrie* 9 September 1885).

Distinguishing between inadequate and poorly applied vaccine and deliberate misapplication of the technology, drawing a line between negligence and malice, was only a matter of perspective. Even for those who complied and were vaccinated, their social category coloured the reading of their state of protection. During the epidemic, the vaccine scar became a symbol of a person's political, class, and even religious affiliations, but the vaccine scar itself (as a clear clinical sign of

protection from smallpox) was also trumped by these categories. While anti-vaccinationism was not confined to the French industrial workers, political and cultural factors sharpened the resistance to compulsory vaccination. Deeply rooted suspicion against the local English authorities who ran the Public Health Board, scepticism over the cause and extent of the epidemic, the conflicting evidence of vaccine's efficacy and finally the real logistical problems associated with caring for vaccine wound (often without access to clean linens or water) were all components in a rational assessment of risk and resistance.

Anti-vaccination in Canada post 1980 — Are Comparisons Possible?

Nineteenth century smallpox vaccination, relative to today's technology, was an extremely invasive and unpredictable procedure. It was not standardised for virulence, controlled for bacterial contamination and the actual protection given varied exceedingly. There was no easy laboratory test to demonstrate vaccine-induced immunity and the Canadian data collected to prove its effectiveness were clearly not unassailable. It is still difficult for this historian to gauge the ultimate impact of vaccination on the disease in late nineteenth century Canada as smallpox was particularly susceptible to quarantine and both technologies were adopted simultaneously to check its spread. The dangers that brought together poverty and increased risks of severe and fatal infections after vaccination have also been mitigated by a higher standard of living, clean water and access to better nutrition. While the dangers from vaccination have certainly been reduced, the risks from the diseases themselves have also waned creating an ever-receding comparative risk perception in the public sphere that is associated with less and less tolerance for error.

While many of the features of nineteenth century resistance do not easy inform modern debates, the history of resistance to smallpox in late nineteenth century Canada should serve as a cautionary tale for policy makers. For better or for worse, notions of civic responsibility, shared risk, and citizenship will shape vaccination policy, science and its reception in diverse ways. Evidence of vaccination's efficacy and safety is itself socially and culturally bound to specific formal institutions, such as government, and professional groups including public health officers and physicians. Evidence, produced by these bodies does not translate fluidly across diverse social groups nor does it necessarily represent their interests. Different groups may disagree

fundamentally with the premises upon which the evidence is constructed or are sceptical and suspicious of the data. Even when the diverse publics are convinced that a particular vaccine is effective, they may not be convinced to submit to compulsory vaccination, as in the recent case of Ontario paramedics who in 2002 protested against mandatory flu vaccine. At a rally to support a paramedic who was suspended for refusing flu vaccine, one protester carried a sign reading, "MOHs are just bullies!"(CBC News 4 January 2002). The reasons for refusing vaccination are diverse and include differing opinions on health care priorities, fear of unknown rare and serious complications, a lack of concern about the disease itself (be it reasonable or unreasonable), suspicion of the veracity and robustness of professional medical advice, or a fundamental belief in freedom of choice when it pertains to a medical procedure.

Current vaccination programs hinge precariously on the general population's voluntary participation in the childhood immunization program, which requires a very high, 95%, participation rate to secure herd immunity for the most infectious diseases. In 1984, the Canadian Vaccination Risk Awareness Network (VRAN) successfully lobbied the Ontario government to amend the Ontario Immunization of School Pupils Act (Lazenby-Craig 1983). The amendment allowed for conscientious and religious objections to all children's vaccinations. Currently VRAN hosts a website that provides links to anti-vaccination sources and serves as a clearing house for information on how to exercise the right to refuse vaccination. To do so, VRAN provides online legal advice to conscientious objectors, reminding the public that all vaccinations recommended by the Canadian National Advisory Committee on Immunization (NACI) are voluntary (PHAC 1997, Section 1 and VRAN 2005a). Under current immunization legislation, only three provinces require proof of vaccination for school attendance and in some cases for participation in licensed infant and toddler day-care. While it is often awkward or difficult to legally opt out, no person can be compelled to be routinely vaccinated under Canadian law.

Still, Canada currently enjoys relatively high immunization rates with vaccination coverage of major childhood diseases ranging from 85% to 90% (PAHC 1997, Section 8). However throughout the nineteenth and twentieth centuries, vaccination coverage has oscillated greatly in Western countries. For example, the percentage of children vaccinated against measles, mumps and rubella has declined in the UK in the last

five years, a phenomenon often attributed to the controversy over Andrew Wakefield's claim that there might be a connection between autism and the MMR vaccine (Measles, Mumps and Rubella) (Wakefield 1998). In Canada and the United States, the publication of Harris Coulter and Barbara Loe Fisher's popular *A Shot In The Dark: Why the P in the Dpt Vaccination May Be Hazardous to Your Child's Health*, a critique of the whole-cell pertussis vaccine, was associated with a troubling decline or stagnation of some childhood vaccination rates below the level required for herd immunity. While the contentious whole-cell pertussis vaccine has largely been replaced by an a-cellular form, Canadian uptake of Dpt vaccination is still suffering from the controversy generated by media reports of encephalitis and post-vaccine syndromes linked to the Dpt vaccine.

> Pertussis has the lowest coverage of all the vaccine-preventable diseases. This is mainly because of parental fears of serious adverse reactions to the whole-cell vaccine in addition to the practice of health-care providers who omit pertussis vaccination because of perceived "contraindications" (PAHC 1997, Section 6.6).

More recently, several investigative reports created concerns that routine childhood vaccinations caused neurological damage by exposing children to mercury levels that exceeded some US federal safety guidelines (ISR 2001, 1, Kirby 2005).

Several research articles have linked the recent rise in anti-vaccinationism to a parallel increase in, access to, and public acceptance of, alternative medical practitioners such as chiropractors, naturopaths and homeopaths (groups thought to be generally opposed to routine immunization). Wilson et al.'s survey of Canadian naturopathic students found a correlation with alternative medical training and lack of faith in vaccination (Wilson 2004). This correlation re-enforces the belief that alternative medicine itself is driving resistance to immunization putting the population at risk for outbreaks. One author argued, "The negative attitude of some providers of CAM [complementary and alternative medicine] towards immunization constitutes an important example of indirect risks associated with this form of healthcare" (Ernst 2001, p. S90 and see also Ernst 1997 & Andre 2003). Poland and Jacobson further argued in, "Understanding those who do not understand: a review of the anti-vaccination movement" that the anti-vaccination movement was the most significant threat to the eradication of vaccine preventable

diseases (Poland & Jacobson 2001). Ernst argued for a public information campaign to *clarify* the risk-benefit profile and to counter-educate CAM practitioners to encourage vaccination.

As mentioned in the introduction to this article, there is an implicit argument that anti-vaccination sentiment results from a distorted understanding of risk in public discourse. However, recent research by risk expert Sheila Jasanoff suggests that researchers should be cautious when categorising resistance to a particular technology as ignorance of the benefits or risks of a particular treatment or disease. Diverse groups in society may understand risks differently and this is not necessarily because the experts or authorities have failed to explain the risks clearly.

> Proponents of better risk communication generally assume that conflicts can be lessened or made more manageable if only the experts who understand the nature of technological risk are prepared to lay all of their cards on the table … but to conceptualize risk communication as a one-way street extending from experts to the public is to underestimate the extent to which perceptions about risk are socially constructed (Jasanoff 1987, p. 116).

Anti-vaccination groups in Canada and the United States have mounted serious critiques of the limits of epidemiological studies to accurately measure a phenomenon as complex as immunity. They have also raised legitimate questions of whether vaccine research programs are statistically sensitive enough to detect and monitor rare vaccine-induced reactions (NVIC 2005). In turn, the education and attitudes of health care professionals, their understanding of the nature and risk of vaccine associated adverse events (VAAEs), and their presumed contact with children after vaccination, will also impact the surveillance of adverse effects.

> Although vaccine manufacturers are required by law to submit reports on VAAEs [Vaccine Associated Adverse Events] received, the cornerstone of vaccine surveillance activities is a voluntary system in which health care providers (mainly public health nurses and physicians) report to local, provincial/territorial public health authorities events *they feel* are temporally associated with an immunization (emphasis added, PHAC 2000, p. C).

The evidence of risk for VAAEs is thus co-produced by the theories of experts, institutions and cultural practices involved in the delivery of

immunization. This could conceivably lead to either under-reporting or over-reporting of adverse events, depending on the perceived risks of immunization by those health care professionals doing the monitoring and reporting.

Diane Dutton further problematises any approach to resistance that is unpinned by an expert's belief that they are trying to 'understand *those who do not understand*'. Dutton's cultural study of risk aversion in America demonstrates that modern risk perception has been shaped by complex generational, cultural and social reactions to optimistic claims made for the elimination of disease, such as Nixon's 'War on Cancer', that followed the era of magic bullet therapies in the post world war two period (Dutton 1992). In her case study of DES, a synthetic estrogen prescribed to millions of women to prevent miscarriages, and which produced devastating side effects, she argued that government agencies, physicians, and pharmaceutical companies were often ill prepared to, or negligent in warning patients about risks. Dutton further maintained that these high profile events shook the public's faith in medical expertise and government regulation and revealed a deep gulf between the priorities of medical innovation and the concerns of the general public (Dutton 1992).

Anti-vaccination activists have also successfully driven a wedge between the goals of public health officers and the interests of individual parents. This wedge cannot necessarily be bridged by a better articulation of the current data available to public health officials and vaccinologists. While some claims made by anti-vaccination groups can be deconstructed by available evidence and rebuffed as not credible, the question raised by these groups concerning the vision behind, and the timing of, mass childhood immunization programs are not so easily dismissed. The public health officer's mandate for *eradication* of vaccine preventable infectious diseases and 'population thinking' can, in certain circumstances, be difficult to reconcile with a parent's responsibility to protect and promote the health of their own children. Should children bear the risks of vaccination against rubella to protect pregnant women from contracting the disease? Should we be implementing mass influenza campaigns among school children to protect their elderly grandparents?

The risks from vaccine-preventable diseases are themselves fluid, both socially and culturally contingent, and difficult to accurately quantify.

For example, what is the evidence of the risk of an un-vaccinated child catching measles in rural Ontario? Is there an increased risk in downtown Toronto? What are the risks of serious injury from the disease if the child does catch measles? Do socio-economic variables and access to health care play a mitigating role? What do physicians do when the data to make informed decisions is simply unavailable or impractical? Is it desirable that risk be individualised to reflect the needs of diverse patient populations? If not, how can immunization be promoted for individual children without explicitly linking it to the public good and notions of good citizenry.

Vaccination is delivered in very different settings across Canada, with some provinces providing routine childhood vaccinations exclusively in public health clinics and others in private practice settings. Research is needed to assess how, in these very different contexts, decisions about immunization are made and risks from routine immunization are identified, treated and reported.

Literature also suggests that anecdotal experience, rather than evidence, profoundly shapes the way people perceive risk. However, experts are not immune to this phenomenon. Physicians themselves and other health care workers present the public with extremely diverse opinions on vaccination safety and efficacy: "...a surprising number of health care workers at all levels, who themselves do not understand vaccine safety and efficacy, and are not champions of vaccines" (Poland and Jacobson 2001, p. 2441). While there have been few studies of this phenomena, it appears that physicians themselves have an eclectic approach to vaccination. Physicians who have witnessed severe reactions to vaccination, under their care, may take a risk adverse approach to recommending routine vaccinations especially if they perceive that the risk from a particular disease is negligible. Injury from vaccination does occur and it has a profound and disproportionate impact on those who witness it directly or when the media profiles these cases. It is important to note that many North American anti-vaccinationists became activists because a family member, often a child, had a debilitating reaction to a vaccination (Diodati 1999, p. 1). Those interviewed expressed sincere concern about vaccination safety, and felt that their children's experiences were neither validated by the medical profession nor adequately reported on in the medical literature.

Conclusions

Poland and Jacobson (2001) argue that there is a gold standard of vaccination expertise, and it is the accurate translation and dissemination of this understanding of vaccination to the masses that will resolve resistance to vaccination. However compliance to childhood vaccination schedules takes for granted the expert's own sense of community-shared risk (as in the case of Rubella where an individual child takes the risk from the immunization to protect society, especially pregnant women, from the disease). As the historical case study demonstrated, the major determinants for compliance to vaccination will continue to be a sense of political enfranchisement, willingness to participate in shared-risk, and reasonable trust in the objectivity of health professionals, and the watch-dog government and licensing agencies. The corollary also holds true and can help researchers anticipate resistance to vaccination in politically, culturally or even medically margainalised populations. In the end, any attempt to increase compliance by a reiteration of the expert's interpretation of risk ignores the social, institutional and cultural context in which the data itself is co-produced and this lies at the core of resistance. Resistance lies at the surface of a profound cultural skepticism in the autonomy, accuracy and vested interests of those producing data about vaccination. It is also often the result of a rational decision making process wherein the vaccine poses a series of issues both medical and non medical that are perceived to be equivocal to the risks of the disease itself. This risk perception cannot be easily reduced to a simple calculus of risk from the disease (as calculated by epidemiology) versus risk from vaccination (as calculated by reported and documented cases of injury). Even more intractable, resistance can point to fundamental differences in opinion about priority setting in health care expenditure. History suggests that maintaining compliance to vaccination at the level required for herd immunity and eradication will be very challenging in a modern, pluralistic democracy.

References

Andre, F. 2003, 'Vaccinology: past achievements, present roadblocks and future promises', *Vaccine*, vol. 21, pp 593–95.

Bradbury, Bettina. *Working Families: Age, Gender, and Daily Survival in Industrializing Montreal.* Toronto: McClelland & Stewart Inc, 1993.

Brosseau, A.T. 1871. 'Société médicale de Montréal: Séance du 20 décembre, 1871,' *L'Union Médicale*, 134.

Brown Ames, H. 1972 (1897), *The City Below the Hill*, ed. P.F.W. Rutherford, University of Toronto Press, Toronto and Buffalo.

Coderre, J. E. 1872, *Vaccination étude lue à la Société médicale de Montréal, les 31 janvier, 14 & 28 février 1872*, s.n., Montréal.

CBER (Center for Biologics Evaluation and Research). 2005, *Thimerosal in Vaccines FAQ*, [online], Available from: http://www.fda.gov/cber/vaccine/thimfaq.htm] [Accessed 1 June 2005].

Colley, F. and Haas, M. 1994, 'Attitudes on immunisation. A survey of American Chiropractors,' *J. Manip. Physiol. Ther.* vol. 17, pp. 584–590.

Coulter, H. and Loe Fisher, B. 1991, *A Shot In The Dark: Why the P in the Dpt Vaccination May Be Hazardous to Your Child's Health*, Avery, New York.

Diodati, C. 1999. *Immunization : history, ethics, law and health*, Integral Aspects, Windsor.

Dutton, D. 1992 *Worse than the Disease. Pitfalls of Medical Progress*, Cambridge University Press, Cambridge, UK.

Ernst, E. 2001, 'Rise in popularity of complementary and alternative medicine: reasons and consequences for vaccination', *Vaccine*, vol. 20, supplement 1, pp. S90–93.

Ernst, E. 1997, 'The attitude against immunisation within some branches of complementary medicine', *Eur. J. Paediatr*, vol. 156, pp. 513–15.

Government of Canada, Department of Finance 2004, 'Budget 2004-Budget Plan-Chapter 4-Moving Forward on the Priorities of Canadians-The Importance of Health-Highlights', 14–15.

Gross, P. 1994, *Higher Superstitions: The Academic Left and Its Quarrels with Science*, Johns Hopkins University Press, Maryland.

Immunization Safety Review (ISR). 2001, *Thimerosal - Containing Vaccines and Neurodevelopmental Disorders*, Institute of Medicine (IOM).

Jasanoff, S. 1987, 'EPA's regulation of Daminozide: Unscrambling the messages of risk', *Science, Technology and Human Values: Special Issue on the Technical and Ethical Aspects of Risk Communication*, vol. 12, no. 3–4, pp. 116–24.

Jenner, E. 1798, *An Inquiry into the Causes and Effects of the Variolae Vaccinae, a Disease discovered in some of the Western Counties of England, Particularly Gloucestershire, and known by the Name of the Cow pox*. first ed. Sampson Low, London, UK.

Kaufman, M. 1967, 'The American anti-vaccinationists and their arguments,' *Bulletin for the History of Medicine*, vol. 41, no. 5, pp.463–78.

Keelan, J. E. 2004, *The Canadian Anti-vaccination Leagues*, 1872–1892, PhD Dissertation University of Toronto, Toronto.

Kirby, D. 2005, *Evidence of Harm. Mercury in Vaccines and the Autism Epidemic: A Medical Controversy*, St. Martin's Press, New York.

Leask, J. 2002, 'Vaccination and Risk Communication: summary of a workshop, Arlington Virginia, USA, 5–6 October 2000', *Journal of Paediatrics & Child Health*, vol. 38, no. 2, pp. 124–8.

LaBerge, L. 1885, *Report of the Sanitary State of the City of Montreal for the Year 1885*, Montreal Board of Health, Montreal.

Lazenby-Craig, B. 1982, "Smallpox in Ontario: Public and professional perceptions of the disease 1884–1885," *Health, Disease and Medicine: Essays in Canadian History*, ed. Charles Roland, Hannah Institute for History of Medicine, Toronto, pp. 215–249.

Millard, C Killack "Clinical notes on Smallpox" (M.D. thesis, University of Edinburgh, 1896). cited in Cyril William Dixon, *Smallpox* (London: J. & A. Churchill, 1962).

Paterson, J. M., Neimanis, I. M., Goebel, C. R., Kraftcheck, D. J. 2004, 'Informed consent for uninsured services: a primary care perspective on the new childhood vaccines', *CMAJ*, vol. 171, no. 8, 877–79.

PHAC (Public Health Agency of Canada). 1997, *Canadian National Report on Immunization, 1996, Government of Canada Communicable Disease Report*, Vol. 23S4–May1997, Section 1.

PHAC (Public Health Agency of Canada). 2000, *Guidelines for Reporting Adverse Events Associated with Vaccine Products. Supplementary guidelines for the Canadian Pharmaceutical Industry*, Government of Canada Communicable Disease Report, Vol. 26S1–Feb 2000.

PHAC (Public Health Agency of Canada). 2005, 'Publicly funded immunization Programs in Canada-Routine Schedule for Infants and Children', [online], Available from: http://www.phac-aspc.gc.ca/im/ptimprog-progimpt/table-1_e.html, [Accessed 1 June 2005].

Plante, P. E. 1871. 'Société médicale de Montréal: Séance du 20 décembre, 1871,' *L'Union Médicale*, pp. 132–3.

Pinker, S. 1999, 'Physicians may have to 'sell' benefits of immunization to sceptical parents', CMAJ, vol. 161, no. 6, 737–8.

Poland GA, Jacobson RM. 2001, 'Understanding those who do not understand: A brief review of the antivaccine movement', *Vaccine*, vol. 19, pp. 2440–45.

Ross, A. M., Anti-vaccination League Scapbooks, "Vaccination on the Trains" s.n., 19 October –newspaper clipping, Volume 2.

Rutherford, P. 1982, *A Victorian Authority: the Daily Press in late Nineteenth Century Canada*, University of Toronto Press, Toronto.

Sibbald, B. 2003, 'One country, 13 immunization programs', *Canadian Medical Association Journal*, vol. 168, no. 5, 598.

Tunis, B. 1982, 'Inoculation for smallpox in the province of Quebec,' *Health, Disease and Medicine : Essays in Canadian history : Proceedings of the first Hannah conference on the history of medicine, McMaster University, June 3–5, 1982*, s.n., Toronto, pp.171–182.

Thomson NP, Montgomery SM, Pounder RE, Wakefield AJ. 1995, 'Is measles vaccination a risk factor for inflammatory bowel disease?', *Lancet*, vol. 345, pp. 1071–74.

VRANa (Vaccine-risk awareness network) 2005, *Vaccination Not Mandatory in Canada* [online], Available from: http://www.vran.org/legal/leg_ont_ltr.htm], [Accessed 1 July 2005].

VRANb (Vaccine-risk awareness network) 2005, Mission Statement, [online], Available from: http://www.vran.org/philosophy/purpose.htm], [Accessed 1 July 2005].

Wakefield AJ, Murch SH, Linnell J et al. 1998, 'Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children', *Lancet*, vol. 351, pp. 637–41.

Wilson K, Mills E, Boon H, Tomlinson G, Ritvo P. 2004, 'A survey of attitudes towards paediatric vaccinations amongst Canadian naturopathic students', *Vaccine*, vol. 22 no. 3–4), pp.329–34.

Wright, G. 1876, 'Vaccination -- Its Efficacy,' *Canadian Journal of Medical Sciences*, vol.1, p. 59.

Wynne, B. & Irwin A. 1995, *Misunderstanding Science? The Public Reconstruction of Science and Technology* (Cambridge: CUP, 1996).

Priority Setting of Hospital Clinical Activities: A Qualitative Case Study and Evaluation

Shannon Madden

Douglas K. Martin

Sarah Downey

Peter A. Singer

Introduction

Priority setting (also known as rationing or resource allocation) can be defined as the distribution of resources among competing programs or people and it occurs at all levels of the health system (McKneally, et al., 1997, p. 157). Most priority setting research though has focused on the macro (health system) or micro (bedside) policy making levels. However, much of the priority setting in a health system occurs at the so-called 'meso' level of policy making, which includes Regional Health Authorities (RHAs) and hospitals. Priority setting is one of the most thorny problems facing any hospitals, because there is no consensus about the 'correct' criteria for selecting priorities. Therefore, hospitals must rely on a fair priority setting process.

To evaluate the fairness of priority setting in hospitals, an explicit ethical framework is required. Daniels and Sabin have developed a framework for fair priority setting that can be used to identify good practices and opportunities for improvement that they call 'accountability for reasonableness' (Daniels and Sabin 1997). 'Accountability for reasonableness', a conceptual framework for fair priority setting, also provides a common framework, or language, that we can use to compare the experiences of different hospitals. Comparing lessons between hospitals can help us to understand the problems faced in different hospital contexts. It is helpful to know how institutions are doing it now, and who are doing it well. 'Accountability for reasonableness' has been recognized internationally as an ethical framework for priority setting in health care institutions.

There exists literature focused on priority setting at the 'meso' level of policy making including hospitals (Singer et al. 2000, Martin, Pater & Singer 2001, Hope, Hicks & Reynolds 1998, Foy et al. 1999, Deber et al. 1994, Deber et al. 1995) but none of these studies have focused priority setting decisions in the form of strategic plan. To date, only one other study has used 'accountability for reasonableness' to evaluate priority setting in the context of hospital priority setting (Singer et al. 2000). Sunnybrook and Women's Health Science Centre (S&W) completed a study in the spring of 2002 describing and evaluating the operational planning priority setting process. The purpose of this study was to describe priority setting in the context of a hospital planning initiative and evaluate it using 'accountability for reasonableness' (see Text Box 1). We also compare the lessons learned from both hospitals.

Methods

To describe priority setting we used qualitative case study methods. A case study is "an empirical inquiry that investigates a contemporary phenomenon within its real-life context" (Yin 1994). This is an appropriate method because priority setting in hospitals is complex, context-dependent, and involves social processes. To evaluate the description resulting from the case study, we used the four conditions of 'accountability for reasonableness' (described below).

The University Health Network (UHN) is a network of three large urban university affiliated teaching hospitals in Toronto, Canada. The focus of our study was the Clinical Activity Target Setting (CATS) process. CATS was the final portion of the strategic planning exercise at the UHN in 2001 whereby all services offered at the hospital were evaluated and their activities for the next five years were determined. The impetus for CATS stemmed from three stresses: insufficient funding, a shortage of staff, and most prominently for UHN, a 'huge capital deficiency' (i.e. space and equipment).

We sampled key documents and people, using a combination of convenient sampling (documents that were available) and theoretical sampling (people who were involved in a significant aspect of the priority setting initiative). There were three primary sources of data for this case study: (1) key documents (e.g. strategic planning documents), (2) interviews with key informants (e.g. administrators, physicians, and nurses), and (3) observations of group deliberations (e.g. planning retreats and meetings). Key documents were obtained in electronic form wherever possible. Key informant interviews were audiotaped and transcribed. An initial interview guide was developed based on relevant literature and previous research (available on request). The interview guide was revised during data collection and analyses to explore

emerging findings (Struass and Corbin 1998). The interview guide contained five questions exploring the CATS process, the fairness of it, and ways to improve it. We conducted interviews with a total of 66 people — 26 individual and 6 focus group interviews — that included members of the Board and Senior Management, clinical leaders, program element leaders, allied health professionals, middle managers and members of the Community Advisory Council. Researchers (Douglas K. Martin and Peter A. Singer) and a key informant (Sarah Downey) observed and/or took part in several meetings throughout the strategic planning process.

Data analysis involved a modified thematic analysis that proceded in two steps: open and axial coding (Strauss and Corbin 1998). In open coding, the data were read and then fractured by identifying sets? of data that relate to a concept or idea. In axial coding, similar ideas were organised into overarching themes. The themes were organised according to the four conditions of the conceptual framework 'accountability for reasonableness'. The 'input' to the evaluation phase of the analysis was the description of priority setting developed in the case study. We compared the descriptions (i.e. what they did) with the conditions of 'accountability for reasonableness' (i.e. what they should do) to identify 'good' practices and "opportunities for improvement."
We addressed the 'validity' of our findings in five ways (Altheide and Johnson 1994). First, the data was triangulated from three different sources (documents, interviews, and observations) to maximize comprehensiveness and diversity (Mays 2000). Second, two researchers (DKM and SM) coded the raw data to ensure accuracy. Third, although the primary researchers collected the data, members of an interdisciplinary research team (also including PAS and others) enhanced the 'reflexivity' in the analysis by becoming familiar with the data and participating in the data analysis. This helped to identify and address prior assumptions. Fourth, 15 participants of the case study in a 'member check' verified the descriptive results of the study. Finally, all research activities were rigorously documented to permit a critical appraisal of the methods (Mays and Pope 1995).

Conceptual Framework

'Accountability for reasonableness', developed by Daniels and Sabin, is a conceptual framework for fair priority setting. It is theoretically grounded in justice theories emphasizing democratic deliberation (Cohen 1994, Rawls 1997). It has become a preferred conceptual framework of priority setting researchers and decision makers internationally (Ham 1999, Ham 2000, Norheim 2000, Martin and Singer 2000, Ham 2001). According to 'accountability for

reasonableness', an institution's priority setting decisions may be considered fair if they satisfy four conditions: *relevance, publicity, revisions*, and *enforcement*, which are described in Table 1.

Table 1. The four conditions of 'accountability for reasonableness'

Relevance	Rationales for limit-setting decisions must rest on reasons (information and values) that fair-minded parties (managers, clinicians, patients, and affected others) can agree are relevant to meeting health care needs under resource constraints in the context.
Publicity	Limit-setting decisions and their rationales must be publicly accessible.
Revision/ Appeals	There is a mechanism for challenge and dispute resolution regarding limit-setting decisions, including the opportunity for revising decisions in light of further evidence or arguments.
Enforcement	There is either voluntary or public regulation of the process to ensure that the first three conditions are met.

UHN agreed to participate in this project and approval for this project was obtained from both the Committee on the use of Human Subjects of the University of Toronto and the UHN Research Ethics Committee. Written informed consent was obtained from each individual before being interviewed. Interviews, transcripts and observations were protected as confidential and available only to the research team. No individuals have been identified in reports without their explicit agreement.

Results

1) Description of CATS process:

In this section the Clinical Activity Target Setting (CATS) process that occurred at University Health Network (UHN) will be described. Since CATS was just one part of the larger strategic planning, it is important to understand the context in which CATS occurred. In November 2000, UHN began a strategic planning process intended to guide its activities for the next 5 to 10 years. The process consisted of four main parts and was completed in January 2002. First, six task forces analysed the major global and environmental factors. Second, seven program groupings were formed. A Program Grouping *"is an organised system of services or interrelated activities designed to address the health needs of a target population, in an efficient, effective and quality driven manner"* (Downey

2001). Third, a reassessment of UHN's mission, vision, and values was completed resulting in a new organizational direction; *Achieving Global Impact* became the new vision. Fourth, the hospital engaged in a Clinical Activity Targets Setting (CATS) process to set 5 year activity targets for each of the 53 elements within the new program groupings. This was the focus of our study.

The goal of the CATS process was "to ensure that there is a balance between excellent leading edge clinical activity and the resources available to carry out this activity" (Downey 2001). It was designed to operationalise the strategic decisions made in the previous planning process. The decision-makers in this process were the Planning and Priorities Council (PPC). The PPC was a group of senior level managers, including both clinician and non-clinician staff. The PPC was to advise the CEO in planning decisions, oversee the development and implementation of facilities, capital budget and information systems plans for the hospital, and recommend the creation, expansion, or downsizing of clinical and academic programs and services consistent with the strategic directions.

To begin the CATS process, workbooks were completed by each of the clinical leadership teams for each of the elements within each program grouping — these were submitted to the PPC on September 7, 2001. Workbooks consisted of a decision tree that focused on six major components: funding, research, education, need, purpose, and relationship to services in other area hospitals. The PPC then participated in two retreat days to develop recommendations on the growth, maintenance or reduction of clinical volumes for each program element by the year 2006 under one of the following graded categories:

A. Significant growth in volumes for the program element >15%

B. Small growth in volumes for the program element 0 to 15%

C. Hold volumes (0 growth)

D. Small decrease in volumes for program element 0 to-15%

E. Significant decrease in volumes for the program element > -15%

NB: 15% represents the anticipated growth in population and impact of aging of the hospitals catchment population over the five years.

At the first Retreat Day (October 4, 2001) the PPC reviewed the workbooks and made preliminary recommendations regarding the clinical activity target volumes for the program elements. They

discovered that every workbook indicated a desire to increase volumes by a total of 25% over the next five years. Since this outcome was impossible to realise for every clinical element of the hospital given fiscal and spatial restraints, it became necessary to develop criteria with which to rank each program element. These criteria needed to encompass the mission and values of UHN and also allow appropriate consideration for the overarching vision of "global impact". In between the retreat days, the Director of Planning and Performance Measurement, the CEO and the CFO created six criteria with which to rank the program elements — Ministry of Health funding, uniqueness of element, interdependence to other programs/elements, use of priority technologies, research, and education. They also created a ten-point scale for each criterion to give guidance and add objectivity to the ranking. Program elements would receive a score of one (lowest) to ten (highest) on each of the six criteria.

The PPC created subgroups to operationalise the criteria and assign scores to each element. The subgroups were created based on expertise and familiarity with the specific criterion and consisted of both PPC and non-PPC members (all were hospital staff). Results were presented to PPC on October 15, 2001 for review and confirmation. Summaries of the results were made available for the program grouping leadership teams. At the second Retreat Day (October 18, 2001) the program grouping leadership made face-to-face presentations to the PPC. The PPC scored the workbooks and presentations against the six criteria and the scores were modified according to the hospitals vision and global trends — for example, globalization and ageing population. PPC then made draft rankings for each program element from highest to lowest. The results were released on October 29, 2001 by e-mail to the Program Grouping Leaders and PPC, and then a couple of days later to the rest of the organization via email and a posting on the hospital's intranet.

An appeals process was launched on October 29, 2001. The appeals process was created to: 1) obtain feedback on the draft five-year clinical activity recommendations, and 2) provide UHN stakeholders with an opportunity to appeal the draft recommendations. Appeals were to be based on: 1) new information or new arguments, or 2) lack of due process. Letters of intent to appeal were submitted and an appeals advisor helped ensure that all necessary data was collected and presented accurately. There was a total of 15 appeals from the program elements and the Community Advisory Committee. All appeals asked for increased scores for one or more of six criteria. No one appealed the

process. On December 13, 2001, 15 appeals involving written submissions and oral presentations were considered by the PPC. As a result, upward changes in the original rankings were made in nine elements, with a significant change to five of them (i.e. the original growth/reduction rank changed). Final results were released the next day on the hospital's intranet, along with a commentary from the CEO. At the January 16, 2002 UHN Board of Trustees meeting, the final recommendations from the PPC were presented and approved.

2) Evaluation of CATS using 'Accountability for Reasonableness'

In this section we will evaluate the UHN CATS process according to the four conditions of 'accountability for reasonableness': relevance, publicity, revisions, and enforcement (described in Table 1). We have included verbatim quotes to illustrate key findings.

Relevance

During the CATS process a large volume of data was collected to support the decision making. Some decision makers found it difficult to become familiar with all of the data,

> I don't think that anybody has actually had the ability to go through the data. I felt a bit guilty initially. And then I realized that I was probably just like everybody else, that I couldn't go through the whole data in a way that would have allowed me to make an informed decision.

and some of the participants felt rushed,

> It was at the last second that I actually saw stuff. I had to respond by noon so I couldn't even circulate it.

The use of subgroups to implement each criterion, and a 10-point scale as a guide, permitted more thorough discussion of each criterion, but also introduced an element of subjectivity between subgroups. Some participants felt there was an over-emphasis on teaching and research at the expense of other facets of patient care, such as patient satisfaction. In commenting on this imbalance one interviewee said:

> That program, albeit is small and is not using hospital resources, carries huge weight in profile within the community. … If you could put it on a balance sheet, if the rate of return on this is huge for the little bit that it costs, that wasn't considered in the equation.

Attempts were made to allow all hospital staff input to the decision-

making. However, there remained a concern about a potential conflict of interest for some decision makers who wore 'two hats' (e.g. being a program leader and a member of the PPC). Some participants suggested all program grouping leaders be included on PPC to diffuse potential advantages.

> I think if you are going to put decision-makers that have conflicts of interest, the least you can do is have it balanced.

> After our first presentation, there were no questions raised. Yet, it was presumably at PPC level, behind closed doors, significant concerns were obviously raised because of the recommendations that came across. Yet, if we were sitting at PPC table and those things had been raised, I am sure that discussion would have been had, and we would have been able to respond to it.

Some participants recommended involvement from an external facilitator.

> If we had used an external facilitator for the clinical activity prioritization process, they would have ensured a more rigourous step by step [process including:] what are you basing that on; tell us a bit more.

Some participants commented that the implications of the decisions were not thoroughly examined. Some suggested the need for "impact analysis" both internally,

> We noticed a significant impact on what the resource requirements will be of our areas, in order to support the service, but no clarity as to how is that resource impact going to be determined.

and externally,

> Does it have an impact on the national scene, then I think you need to have people from the national scene to give you a perspective. I think if you want to know if it has an impact on the local scene, you should look to the surrounding groups and see what they think of what you're doing, both within this region - the other health care institutions as well as the non-academic institutions.

Publicity

CATS decisions and the reasons behind them were readily accessible to the members of the PPC. The decisions and scores were communicated to program grouping leaders and other hospital staff through the UHN

intranet. Some of the communication may have caused fear or misunderstanding.

> But to say, about a program that people have invested time and energy and a patient population that they care about, that there's no strategic advantage to that program, was very, very problematic.

There was little communication of the CATS process outside the hospital. The CEO did present the CATS draft results to the hospital's Community Advisory Committees and they were allowed an opportunity to appeal.

Revision/Appeals

The inclusion of an appeals process was felt by many to be a positive part of the overall CATS process.

> What we had was a resubmission of data, if there was an impression of incorrect or incomplete data. That was very appropriate and given the attempt to quantify, it was inevitable…

Some participants had concerns because the name *'appeals process'* led participants to assume a quasi-judicial process where a different appeals body, other than the original PPC, would rule on the appeals.

> The appeals process was difficult to feel comfortable in because it was the same body that we were going back to make the appeal to. … You are going back to the same people who essentially saw the first sets of arguments. In most appeal mechanisms, you go to a separate body. If that was their opinion from the get-go, it was unlikely you were going to make them change 180 degrees.

The documentation of the appeals process and reasoning was not as strong as it had been throughout the CATS process. Although a formal document outlining the final clinical activity targets was prepared, there remained uncertainty about the reasons why appeals were successful or unsuccessful.

Enforcement

The Senior Management, particularly the CEO, was committed to ensuring the conditions of 'accountability for reasonableness' were met. The planners of the CATS process met with scholars in priority-setting to discuss the components of a fair process. Elements such as the appeals process were added to the original process to enhance fairness. Many participants felt that the process contained many necessary elements that made it fair.

At the end of the day, it was collaborative. It gave people a chance to make their case. It gave people a chance to show what their value has been for this organization over the course of the previous decade. It gave people a chance to think in a visionary way to the future. It gave a lot of people the opportunity to try to be creative and to join with others, join forces with others to form joint programs and make a greater whole through synergies. And it gave people a chance for appeal. So that's why I think it was fair.

Discussion

We have described and evaluated the CATS process at UHN. Martin et al. described priority setting at Sunnybrook and Women's Health Science Centre (S&W). These are the only two existing studies of priority setting at the hospital level. We will now take our analysis one step further by comparing the lessons learned in the study with the lessons from a previous study of a different hospital.

A Comparison of UHN and S&W

Both UHN and S&W underwent a similar priority setting exercises and both processes were studied using 'accountability for reasonableness'. While the UHN and S&W process we named differently, the former being called strategic planning and the later process operational planning, the process were very similar. Seven similarities and differences can be identified. First, both processes at S&W and UHN involved extensive data collection mechanisms, used pre-determined workbooks and a decision tree to help make priority setting decisions. In fact, UHN used the S&W workbook and decision tree as a template for designing their own, illustrating how sharing experience may be beneficial to hospitals.

Second, UHN was more criteria driven than S&W. UHN used a set of six criteria to score and rank the programs elements. S&W did not pre-specify criteria. Third, both the S&W and UHN processes were inclusive whereby many of the stakeholders contributed to final priority setting decisions. The decision making process at S&W involved 70 decision makers, whereas the process at UHN involved 18. Many at S&W felt that, while it was important for the process to be inclusive of a broad range of stakeholders, it was also important to have the 'right' people making the decisions. Involvement should be determined by competence, expertise, access to adequate information, and institutional decision-making authority (Gibson 2002). Fourth,

the attention and the degree to which S&W and UHN paid to context differed. The decision makers at UHN focused greater on the context in which the decisions are being made (e.g. the teaching and research context). Some participants of the UHN process felt that this caused an over-emphasis on certain contexts, while leaving other important areas out (such as the patient and community context). Decision makers at S&W focused on both internal and external contexts in the priority setting process. However, some participants of the S&W process felt that not enough attention was paid to the institutional context in which the decisions were being made. Fifth, the mechanism for reaching agreement of S&W and UHN was different. The voting process used at S&W — open voting and abstentions — was seen by many participants as a major flaw of the process. At UHN, consensus was more actively sought out through the use of pre-determined criteria, a 10-point ranking scale and discussions within subgroups. Sixth, in terms of publicity of the process, both S&W and UHN received the same comments and feedback from participants — communication was felt by participants to be sufficient, however, it could be improved by being more formally organised. Also, neither organization made their decisions making easily assessable by patients of the public. Seventh, perhaps the largest difference between the two processes was the presence of an appeals process to the UHN CATS process. S&W did not have any mechanism for revising their decisions. UHN did include such a mechanism, which was well received by participants. The 'appeals' process used by UHN allowed for a second look and possible revision of original scores. Participants and decision makers felt that this appeals process increased the fairness of the overall process.

Implications for Practice

The framework 'accountability for reasonableness' can be used to help to improve priority setting in hospital strategic planning as is evident by looking at both the S&W and the UHN processes. For example, researchers at S&W have taken steps toward this in the creation of a checklist for decision-making based on the conditions of 'accountability for reasonableness'. They have created a two-page document that outlines 'accountability for reasonableness' in a practical way. By addressing accountability at the institution, challenges of decision-making and the justification for using an ethical framework for decision-making, this document effectively makes fair decision-making practical and comprehensive.

As 'accountability for reasonableness' becomes implemented more frequently, its principles and conditions can become a part of institutional culture, such that decisions made at all levels of care could meet the four conditions of the framework. This was the first time that an appeals process has been described and evaluated using accountability for reasonableness. 'Accountability for reasonableness' is not just a framework used in major decision-making processes, it also fosters a learning organisation for all staff, meaning that the organization is always improving and adapting to change — learning good practices and opportunities for improvement, strategies for good decision-making and organizational involvement throughout the process.

Implications for Research

By comparing the UHN study with the S&W study, we have started to establish a database of cross-institutional learning. Such a database can be useful in all health systems. The process of describing and evaluating priority setting using case-study methodology can help to improve fairness in priority setting at all hospitals. Further, the conceptual framework of 'accountability for reasonableness' provides a means to achieving fair priority setting process. More research similar to this needs to be done to continue to capture and share lessons from hospital priority setting, and to ultimately, improve the fairness in priority setting at all hospitals and health systems nation-wide.

Implications for Theory

While 'accountability for reasonableness' can provide a foundation and a platform on which to base priority setting initiatives, the process itself is shaped and guided by institutional culture and the context in which the process is found. For this very reason, there may be instances where the process will need to be altered to meet demands of the data, of those involved or of external forces. These nuances can help refine the four conditions of the theoretical framework. For example the model should take into account varied agreement and appeals mechanisms at different organizations. In the S&W 'majority rules' voting was the mechanism of agreement, whereas UHN used a consensus and ranking method for reaching agreement. The strategic vision and direction of the organization will also have an effect on the way that priority setting is implemented.

Text Box 1: Sunnybrook and Women's Health Science Centre Case Study (Santiago-Sorrell et al. 2001).

Sunnybrook and Women's Health Science Centre (S&W) is a large, urban, tertiary care teaching hospital in Ontario, Canada. In 2001, the senior management of Sunnybrook and Women's College Health Sciences Centre (Sunnybrook & Women's) launched an innovative priority setting exercise that would guide future decision-making in the organization.

The priority setting exercise took place earlier in the year 2001, in the wake of the amalgamation of three organizations into Sunnybrook & Women's. The amalgamation took place in June 1998, but the organization and management structure was still perceived as fragmented, making it difficult to make decisions. Moreover, the level of frustration with decision-making had reached a climax at the time of the priority setting exercise in January 2001.

Process: Two "Decision Days" were held to meet three goals:

1. Identify which clinical service areas would be priorities

2. Identify areas of expense reduction and improved efficiency

3. Develop an operating plan based on these priorities

Each Decision Day required invitees to attend a full day session of decision-making and vote taking.

Decision Day #1 focused on the development of a three-year Clinical Services Plan. Members of the senior management team, to facilitate decision-making for what is determined to be a Clinical Service Priority, created a "decision tree". At the end of Decision Day #1, five clinical service priorities (CSP) were identified: Cancer, Cardiac, Musculoskeletal, Perinatal & Gynecology and Trauma.

In between the "Decision Days", a CSP workbook was created to assist the programs in prioritizing their work according to the five CSP.

Decision Day #2 focused on the development of the 2001/2002 Operating Plan. Seventy decisions were made primarily concerned with initiative to reduce expenses and create greater efficiency.

Limitations

First, the findings from this study may not be generalizable to other hospitals. However, generalizability is not the goal of qualitative research. Other hospitals may see themselves in this work and benefit from the lessons we have described. The description of the priority setting process itself can be helpful in any hospital wanting to set priorities in the context of strategic planning. Moreover, 'accountability for reasonableness' can be used to evaluate priority setting in a variety of healthcare settings. Second, we have not evaluated the consequences of these decisions. It will be important to study the subsequent budget cycles to evaluate the actual operational decisions that follow from each priority setting initiative. By continuing to study and analyse priority setting initiatives all healthcare organizations can learn and grow through improved priority setting. In order to make the most of the recommendations given, careful consideration should be given to how to implement, what the implementation will look like and what desirable outcomes are. Third, the participants may have been influenced by a social desirability bias — participants may have described what they thought the researcher wanted to hear rather than actual events. Describing priority setting is not the same as conducting actual priority setting.

Conclusion

This case study has provided an in-depth analysis of a priority setting process at a large urban teaching hospital and compared the lessons with those from a previous study. By focusing on the process of decision-making, capturing the lessons from these initiatives helps to contribute to an 'evidence base' for these important policy decision.

Acknowledgements

References

Altheide, D. L., Johnson, J. M. 1994, 'Criteria for Assessing Interpretive Validity in Qualitative Research', In *Handbook of Qualitative Research*, eds. Denzin, N. K., Lincoln, Y. S., Sage Publications, Inc., Thousand Oaks, pp. 485–99.

Cohen, J. 1994, 'Pluralism and Proceduralism', *Chicago-Kent Law Review*, vol. 69, pp. 589–618.

Daniels, N., and Sabin, J.E. 1997, 'Limits to health care: Fair procedures, democratic deliberation and the legitimacy problem for Insurers', *Philosophy and Public Affairs*, vol. 26, no. 4, pp. 303–502.

Deber, R., Wiktorowicz, M., Leatt, P., Champagne, F. 1994, 'Technology Acquisition in Canadian Hospitals: How is it done, and where is the information coming from?' *Healthcare Management FORUM*, vol. 7, no. 4, pp. 18–27.

Deber, R., Wiktorowicz, M., Leatt, P., Champagne, F. 1995, 'Technology Acquisition in Canadian Hospitals: How are We Doing?', *Health Care Management FORUM*, vol. 8, no. 2, pp. 23–8.

Downey S. 'Clinical Activity Target Setting Process: Planning and Priorities Council Decision Methodology.' University Health Network, 18 September 2001 (unpublished).

Foy R., So, J., Rous, E., Scarffe, J. H. 1999, 'Perspectives of commissioners and cancer specialists in prioritising new cancer drugs: impact of the evidence threshold', *British Medical Journal*, vol. 318, pp. 456–459.

Gibson, J. L. 2002, 'Philosophical Approaches to Health Care Priority Setting: Moral Obligations, Practical Realities', Doctoral Thesis, University of Toronto.

Ham, C. 1999, 'Tragic choices in health care: lessons from the Child B case', *British Medical Journal*, vol. 319, pp. 1258–61.

Ham, C., Coulter, A. 2001, 'Explicit and implicit rationing: taking responsibility and avoiding blame for health care choices', *Journal of Health Services Research and Policy*, vol. 6, pp. 163–9.

Ham, C., McIver, S. 2000, *Contested Decisions: Priority setting in the NHS*, King's Fund Publishing, London.

Hope, T., Hicks, N., Reynolds, D. J. M., Crisp, R., Griffiths, S. 1998, 'Rationing and the health authority', *British Medical Journal*, vol. 317, pp. 1067–1069.

Martin, D. K., Pater, J. L, Singer, P.A. 2001, 'Priority Setting Decisions for New Cancer Drugs: A Qualitative Study', *Lancet*, vol. 358, pp. 1676–81.

Martin D. K., Singer P. A. 2000, 'Priority setting and health technology assessment: beyond evidence-based medicine and cost-effectiveness analysis', In: *The Global Challenge of Health Care rationing*, eds. Coulter, A. and Ham, C., Open University Press, Philadelphia, pp. 135–145.

Martin, D. K., Singer, P. A. 2003, 'Priority setting for health technologies in Canada', In *Priority setting in health care. Institutions, information and 'accountability for reasonableness*, eds. Ham, C., and Roberts, G. King's Fund Publications: To be published May 2003.

Mays, N., Pope, C. 1995, 'Rigour and qualitative research', *British Medical Journal*, vol. 311, pp. 109–12.

Mays, N., Pope, C, 2000 'Qualitative research in health care: Assessing quality in qualitative research', *British Medical Journal* 2000, vol. 320, pp. 50–2.

McKneally, M. F., Dickens B., Meslin, E. M., Singer, P. A. 1997, 'Bioethics for Clinicians: Resource Allocation', *Canadian Medical Association Journal*, vol. 157, pp. 163–7.

Norheim, O. 2000, 'Procedures for Priority Setting and Mechanisms of Appeal in the Norwegian Health Care System, Presentation at the 3rd International Conference on Priorities', In *Health*, The Netherlands.

Rawls, J. 1993, *Political Liberalism*, Columbia University Press, New York.

Santiago-Sorrell, P., Shulman, K., Martin, D. K., and Singer, P. A., 2001. 'Priority Setting Case Study Final Report', Sunnybrook and Women's Health Science Centre and the University of Toronto Joint Centre for Bioethics. (Unpublished)

Singer, P. A., Martin, D. K., Giacomini, M., Purdy L. 2000, 'Priority Setting for New Technologies in Medicine: A Qualitative Case Study', *British Medical Journal*, vol. 321, pp. 1316–8.

Strauss A, Corbin J. 1998, 'Basics of Qualitative Research: Techniques and Procedures of Developing Grounded Theory'. Thousand Oaks, CA: Sage Publications Inc., pp. 55–242.

Strauss, A., Corbin, J. 1998, 'Basics of Qualitative Research: Techniques and Procedures of Developing Grounded Theory', Sage Publications Inc., Thousand Oaks, pp. 204–205.

Yin, R. K. 1994, *Case Study Research: Design and Methods*, Sage Publications, Inc., Thousand Oaks.

Beyond EBM: Critical Realism as the Foundation for Evidence-Based Public Health

Wendy McGuire

Introduction

Concerns with the high cost of health care, growing public scepticism about clinical effectiveness, persistent health inequalities and emerging global health threats are leading to a renewed focus on public health. The emergence of new infectious diseases that can be rapidly transmitted around the world has led to the intensification of traditional public health interventions, such as immunization, vaccination and global surveillance systems. At the same time, what many are calling the "new public health" emphasises a return to the social and political origins of public health by targeting the social determinants of health (Dean and Hunter 1996). In the past three decades, a consensus has emerged among public health researchers, policymakers and practitioners that health is the outcome of complex interactions between determinants at the level of the biological, the behavioral, the social and the environmental. This is reflected in a number of broad policy documents such as the Ottawa Charter for Health Promotion (WHO 1986), the Lalonde Report (Lalonde 1974) and the Black Report (Macintyre 1997).

Evidence-based medicine (EBM), a movement initiated to improve the efficacy and efficiency of clinical practice, has increasingly come to the fore of public health debates. In a very short period of time, the EBM movement has succeeded in establishing a well-defined hierarchy of evidence, with the meta-analysis of randomised controlled trials (RCT) at the pinnacle, and in promoting the use of the meta-analysis among practitioners and policymakers. Debates over the development of standards for evaluating public health research have centered on the applicability of the EBM model to the

field of public health. Unlike clinical medicine, public health has always been informed by a wide variety of disciplines and research paradigms spanning the natural and social sciences. Historic differences between how positivists/empirical realists and interpretivists/constructionists understand the nature of reality are embedded in current debates over public health evidence. Attempts to develop standards for the appraisal and synthesis of public health research have been bogged down by these differences.

Critical realism offers a reconciliatory position that may provide a way out of the current stalemate. Based on the philosophy of Roy Bhaskar (1998), critical realism is a relatively recent approach to understanding the nature of reality and causality in the natural and social sciences. There are three features of critical realism which distinguish it from the naïve empiricism of the natural sciences and the anti-realism of some strains of constructionism in the social sciences. First, it offers a deep concept of reality: the essential properties of natural and social phenomenon are not observable at the empirical level of reality. Therefore, science generally, and social science especially, must abandon a rigidly empiricist approach to uncovering the truth about things. Second, reality is hierarchically layered, from the cell to the body to the person to the social to the cultural, increasing in complexity at each ascending layer. Interactions within and between levels cause new and more complex formations to emerge that cannot be understood by understanding the components which produced them. Therefore, each layer must be seen as being analytically distinct. Thirdly, critical realism challenges positivist pattern-event regularity theory of causality which cannot account for interactions between the body, the self and society. While it may be applicable at lower levels of reality, it fails to account for the role of agency of individuals and groups who continually interpret and respond to their natural and social environments, co-creating the shape of the natural world and their social institutions.

Using critical realism as a framework for thinking about public health knowledge, I argue that the adoption of existing models of EBM, which are based on positivist methodologies and assumptions about causality, are insufficient and potentially harmful to human well-being. Yet constructionists/interpretivists who have put forward a critique of EBM have failed to develop a plausible alternative. In

this paper, I suggest that critical realism can provide this alternative and provide common ground for public health researchers to work together, bringing the full weight of scientific knowledge to bear on increasingly complex and global public health problems. In the first section of this paper, I will summarise the main ideas of critical realism. In the second section, I will provide an overview of the movement for the development of an evidence base for public health. In the third, I will discuss the limitations of current approaches. Finally, I will consider how critical realism can provide an alternative approach to the production and appraisal of public health knowledge and consider some of the implications for public health interventions.

Critical Realism: An Overview

The main argument put forward by critical realists is that the empirical methods of the natural sciences are not applicable to the social world, but that social research does not need to succumb to the relativism of postmodernism. Social reality <u>can</u> be known, even though the social world is complex and unpredictable, but critical realism offers a conception of the *real* that is very different from the empirical realism of the natural sciences. In a critical realist ontology, reality is layered. Bhaskar (1998) identified three layers, the empirical level where observations are made, the actual level, where events happen, and the deep level where the generative mechanisms exist, which have the power to produce events but cannot be directly observed. Furthermore, reality is stratified. Mechanisms, objects and events exist at different strata which are hierarchically organised, from the physical, chemical and biological to the psychological, behavioral and social. Each strata is distinct and separate, yet each interacts with the layer above and below to produce new mechanisms, objects and events. The ability of mechanisms to combine to create something new is what Bhaskar called *emergence* (Bhaskar 1998).

Figure 1 depicts these layers, from real to actual to empirical across the body, the self and society, providing examples at each level. The most significant aspect of Bhaskar's layered reality is that real mechanisms can exist as causal tendencies without necessarily being activated. Generative mechanisms, at the level of the <u>real</u>, can exist without manifesting themselves in an event at the level of the actual that can then be observed or measured. A mechanism at one level

may be triggered or blocked by other mechanisms at any other level. When a mechanism is activated, leading to emergence, the new mechanisms or objects that are generated cannot be understood by the components that produced it. For example, psychological thought processes cannot be fully understood by examining the neurological processes that make thought possible; group dynamics cannot be understood as the sum of the behavior of individual group members.

Figure 1. Determinants of Health and Layers of Reality

	Body	Self	Society
Empirical (observable)	Biological determinants Diagnosis of illness Treatment	Individual determinants (e.g. lifestyle) Experience and meaning of illness Coping	Social determinants (e.g. poverty) Disruption of social participation Health and social costs
Actual (objects/events)	Normal and pathological processes Signs and symptoms	Cognition Emotion Development Behavior	Political, economic, social welfare, health care systems Social behavior, norms, relations
Real (mechanisms)	Biological, physical, chemical, genetic mechanisms	Psychological, emotional, cognitive, spiritual mechanisms	Social, cultural, political, economic, religious mechanisms

Given these characteristics of reality, Bhaskar demonstrates how positivist assumptions about causality are inapplicable at the level of the social. Positivism locates causality in patterns or regularities of observable events. 'If A always follows B, then B causes A'. The reason for the pattern of events can sometimes by identified and observed through the manipulation of the object under study and the examination of which conditions are consistently linked to the object's presence. Bhaskar argues that the experiment is not possible at the level of the social due to the open nature of the social system and the complexity and unpredictability of the social world. At lower levels of reality, natural objects do not interpret and respond to the actions of the scientist or the meaning of the changes in their environment; they simply react. The effects of the experiment can, therefore, be

attributed to the variables manipulated in the experiment, not to the object's response to the experiment. However, in the realm of the social, human beings continually interpret and adapt to changes in our environment. We ascribe meaning to our world, which influences how we both define and produce social reality (Danermark et al. 2002). It is not sufficient to identify observable social patterns and regularities to locate causal mechanisms, as this disregards the role of meaning as a source of motivation for action.

Also, at higher levels of reality, Bhaskar argues that there are a greater number of generative mechanisms and the potential for emergence is much higher. Social reality is the outcome of complex interactions between multiple mechanisms within the social strata and between the biological, the psychological, behavioral and the social. Research at the level of the social will never be able to generate accurate predictions of social behavior. However, by understanding the essential nature of generative mechanisms and the factors which are likely to trigger or block their activation, causal tendencies can be identified.

> What causes something to happen has nothing to do with the number of times we have observed it happening. Explanation depends instead on identifying causal mechanisms and how they work, and discovering if they have been activated and under what conditions (Sayer 2000, p.14).

The identification of causal mechanisms in open social systems requires that theory and analysis be given a central role in the production of knowledge, particularly at the level of the social (Sayer 2000; Danermark et al. 2002). To generate explanatory knowledge of the determinants of health, public health research must be able to identify causal mechanisms at each distinct level of reality and target interventions to the appropriate level. This requires the continuation of distinct disciplinary knowledge production, as well as interdisciplinary collaboration and increased syntheses of knowledge across disciplines and research paradigms.

From EBM to Evidence-Based Public Health

In this section, I will review how public health researchers have responded to calls for the development of standards for the appraisal of public health research. The field of public health includes a broad range of activities which are carried out in a loosely coordinated fashion by different governmental and non-governmental agencies, in

developing and developed countries. Public health activities include health monitoring and surveillance, the prevention of chronic and communicable disease, health promotion, disaster preparedness and response, and the development of laws and policies to support health at the level of the population. In an on-going US study, the "Evidence-Based Practice for Public Health Project", investigators identified twenty knowledge domains for public health and linked each with the essential public health functions as identified by the US Centers for Communicable Disease Control (CDC) and the Pan American Health Organization (PAHO) (Martin and Simpson 2004). For example, these included epidemiology, biostatistics, social and behavioral sciences, public health nursing, laboratory sciences, and informatics. This study was funded by the CDC to assess the applicability of clinical EBM to public health and to provide resources to practitioners to increase the use of evidence in public health practice. Similar efforts are underway to develop a framework for a public health evidence base that can be used in Canada (Health Canada 2001) and the UK (Kelly et al. 2002; Swann et al. 2003).

Although a wide variety of study designs and theoretical perspectives have been used to examine the determinants of health and the implementation and evaluation of public health interventions, epidemiology is the dominant source of public health evidence. Public health researchers representing different research paradigms have taken clinical EBM as the starting point, either defending it or critiquing it as the basis for evidence-based public health. Unlike clinical medicine, which is based on a biomedical model of health and emphasises change at the level of the individual, public health is based on a social model and interventions can target the individual, the family, the community, or the nation. Settings for disease prevention and health promotion extend beyond the clinic or hospital to schools, homes, workplaces and the street. Interventions are often large-scale, multi-component, and difficult to evaluate. The sheer range of disciplines involved in public health, the dispersion of studies across journal types, and the absence of filters for retrieval have been identified as significant pragmatic obstacles to identifying potential evidence for public health (Jackson 2004).

Despite these differences between clinical medicine and public health, the EBM model dominates the debate. While natural and social scientists both agree that modifications are necessary, differences exist on the extent and nature of changes required. Epidemiologists largely

continue to defend the EBM hierarchy while advocating from some rejigging for public health (Rychetnik et al. 2002). In their article, "Criteria for evaluating evidence on public health interventions", Rychetnik and colleagues (2002) conclude that the RCT should be retained as the 'gold standard' for appraising studies of large-scale, complex public health interventions. However, they concede that RCTs are frequently impractical at the level of the population and that alternative methodologies should be considered. They also recognise the need for epidemiologists to develop new ways of theorising multiple causes of poor health and multiple dimensions of intervention effectiveness, including feasibility, cost-effectiveness and implementation.

Social scientists and transdisciplinary health researchers also call for methodological and theoretical pluralism and innovation. However, they tend to question the underlying assumptions of EBM and propose more substantial changes to the criteria for evaluating public health research (Dean and Hunter 1996; Popay and Williams 1996; Upshur et al. 2001). Health researchers in the social sciences, particularly those using qualitative methods, have long argued that the social context and meanings that people give to their experience of illness affects behavior, access to health services, and, ultimately, the effectiveness of health interventions (Popay and Williams 1996). An individual's objective social position, personal biography, and the personal meaning they ascribe to their status in the social hierarchy all come together to influence the biological processes that lead to illness and the lifestyles that are associated with poorer health. What underlying social relations produce the conditions for poor health? What social and interpersonal factors minimise the impact of social structures on health? How is social change facilitated? What public health interventions can target structural causes of poor health and support individual resilience?

According to social scientists, epidemiological tools and methods are inadequate for understanding determinants of health, or generating interventions, at the level of the social (Hayes et al. 1994). Epidemiology is well suited to measuring variations in the incidence and prevalence of illness across specific populations. Yet, the epidemiological population is conceived as the aggregation of individuals who share a particular trait that is linked to a specific disease or risk of disease. Traits that put individuals 'at risk' of negative health outcomes may be biological (such as genes), behavioral (such

as smoking) or social (such as socio-economic status). Policy solutions based on population trait-based evidence will target 'at risk' individuals rather than the social relations that produce risk. For example, epidemiology has established a relationship between race and hypertension. Race is conceptualised as a trait, rather than the outcome of a set of <u>social relations</u> which is based on the domination of one group of people by another. Educating racialised groups to change their fitness and dietary behavior will not take away the root cause of hypertension, which is located in historic social patterns of racial domination and exclusion. Changes in the type of disease racialised groups suffer from may change as a result of targeted interventions, but overall health will remain poorer than the dominant social group unless social relations themselves are transformed. Yet, epidemiological evidence most often targets change at the level of the individual, or 'downstream' factors, rather than 'upstream' factors in social and environmental conditions (i.e. policy and legislation) (Macintyre et al. 2001).

Social science research seeks to understand the interactions between broad social forces and individual and group perceptions and behavior that reproduce or transform social relations. Key public health problems, such as obesity, inactivity, smoking, addictions, and diet and persistent inequalities in health across socio-economic and racial groups have proven resistant to change at the level of the individual and require broad social interventions. In their critique of the use of EBM as the model for evidence-based public health, social scientists question core assumptions and argue that this model cannot provide the knowledge base needed to address public health challenges. In fact, public health research that meets the EBM criteria is potentially harmful. It can lead to misplaced responsibility onto individuals when the root causes of health inequalities are located in the social structure (Coburn and Poland 1996). It can be used to justify the reduction of spending on health care services without redirecting public money into other health or social programs that would contribute to improved health of the population (Labonte 1995; Poland et al. 1998; Eakin et al. 1996).

Limitations of Current Approaches to Evidence-Based Public Health

Despite significant differences between epidemiologists and social scientists, shared calls for theoretical and methodological pluralism

and the development of an interdisciplinary infrastructure for public health research suggest that there may be common ground for the development of an interdisciplinary evidence base for public health. However, pluralism can simply disguise institutional patterns of disciplinary dominance in public health. Epidemiology is the dominant public health discipline in terms of research funding dollars, leadership in public health organizations, and uptake of research into practice. Public health nursing and social and behavioral sciences take a distant second. Broadening the types of methodologies that are 'acceptable' as evidence in a new evidence-based public health will do little to increase their legitimacy and utilization if they remain at the bottom of an evidence hierarchy. An interdisciplinary public health infrastructure is just as likely to systematically perpetuate current patterns of dominance and marginalisation in public health research as to transform them.

In response to their marginalised position in public health and their own disciplinary evolution, social scientists often minimise the role of the biological, advocating for explanations solely at the level of the social. Epidemiologists, on the other hand, claim to be able to use their own set of tools to understand the social without drawing on the methodological and theoretical strengths of the social sciences. As a result, they fail to conceptualise the social as distinct from the aggregate of individuals. They also focus on the objective traits of populations, rather than the meaning that people ascribe to their social position, biography and health, failing to account for the role of meaning in shaping action. Neither of the current approaches to public health evidence has been able to address the urgent need to conceptualise and measure complex interactions across the biological, psycho/behavioral and social in order to fully grasp the causes and consequences of ill health that threaten the well-being and even survival of our species and societies.

The obstacles to bringing together existing public health research to inform complex public health problems are exemplified by the Cochrane Health Promotion and Public Health Field group. This field was established as a member of the Cochrane Collaborative in 1996 but reviewers have been unsuccessful in using the Cochrane systematic review methods to appraise and synthesise public health and health promotion research. In a background paper analysing the problems, it was found that none of the existing guidelines for conducting systematic reviews (the Cochrane Reviewers Manual, the

NHS CRD Report Number 4, the EPPI-Center Manual and the NHMRC Report) provided sufficient guidance for summarising interpretive research, integrating qualitative and quantitative research, evaluating the use of theory, or understanding the role of context in multi-component interventions (Jackson 2004). The possibility for these groups to be indefinitely stalled by differences in assumptions about the nature of reality, causality and methodology seems high. Without forward movement, it is likely that epidemiology will continue to dominate as the knowledge base for public health policy and practice while social scientists clamour for a greater voice or ignore the debate altogether. In the following section, I will argue that critical realism can move this project forward and improve the evidence base for public health decision-making.

A Critical Realist Alternative

Critical realism offers a way out of the current stalemate by re-conceptualising reality as deep and layered and by challenging the positivist pattern-event theory of causality which cannot account for complex and unpredictable interactions between the body, the self and society. Critical realism can provide a foundation for overcoming two key problems with current approaches to public health evidence. The first problem is the lack of sufficient attention to causal mechanisms at the level of the social. This is due to the exaggerated legitimacy awarded to a model of empirical research that is consistent with the assumptions underlying EBM. Second is the inability of current approaches to explore interactions between mechanisms at multiple layers of reality while keeping these layers analytically distinct. How can we bring together knowledge of the body and society without falling into the extremes of biological/social determinism or the unmitigated free will of the rational actor?

In this section, I review two examples of how critical realism can produce better explanations and interventions than either positivism or constructionism. Using Bhaskar's concept of emergence, Simon Williams (Williams 1999) proposes an alternative conceptualisation of disability to the two dominant schools of thought in the social sciences. In the first, the body itself is seen as a fabrication of discourse and the experience of physical sensation, including pain, is an effect of the meaning given by biomedical discourses on the body. In the second, disability theorists argue that disability is an outcome of social oppression, and the failure of society to provide barrier-free

access to full participation. Rather than argue that the biological body is irrelevant, due to its fabrication or manipulation by either social discourses or social structures, Williams argues that disability is an <u>emergent property</u>. Disability is located in the interplay between the physiological impairment, pre-existing social structures and attitudes towards disability, and the interactions between agents and structures to reproduce or transform social reality.

Williams' (1999) critical realist analysis suggests that a balance must be struck between intervening medically to reduce disability and intervening socially to improve the conditions for barrier-free living. A wide variety of knowledges are needed to design these interventions: knowledge of the experience and meaning of living with disability and the variations across age, gender, sexual orientation, race and culture; understanding the impediments to access and participation, physically, psychologically and socially; developing new medical treatments to reduce impairment and the appearance of difference that can lead to stigma; and understanding how to increase tolerance for diversity and difference in institutional and social settings.

In the second example, critical realism is proposed as the basis for an alternative approach to risk assessment in child welfare (Houston 2001). Houston critiques the positivist position towards risk assessment for presenting risk as both neutral and calculable. This position, he argues, does not account for cultural variations in the construction of risk or the inherent uncertainty in determining which children are at risk and at how great a risk. Social constructionists have critiqued the individualist and rationalist bias of cognitive risk assessment, arguing that risk is constructed as a tool of social regulation to achieve goals that are historically and culturally situated. Individuals who are defined as being 'at risk' are subject to greater external and self-regulation than those who are not. The problem Houston identifies with the social construction position towards risk is that it fails to acknowledge the harm done to children as real.

According to Houston (2001), critical realism can be used to develop an alternative approach for assessing risk and developing interventions. Critical realism's deep view of reality suggests that causal mechanisms operating within the child, in her immediate environment or in the wider culture, <u>when activated</u>, produce or prevent harmful outcomes which affect the child's emotional, cognitive, social or physical development. Causal mechanisms are

triggered by day-to-day events, and by the child's own reflection and action in response to these events, producing patterned interactions over time. The social worker assesses the sources of harm by observing and forming hypotheses based on theories of child development and testing these theories in practice. Interventions are designed to activate or block protective or adverse mechanisms at multiple levels and to make individuals more aware of the social factors which shape their lives. This minimises self-blame thereby enabling clients to make personal changes and to challenge oppressive social structures. A critical realist approach to social work practice demands a high level of reflection by the social worker and client, to link everyday experience with broader social forces and to observe and be aware of the effects of minute changes. It requires that individuals live with the tension of knowing that many causes of suffering are outside of their control but that change is only ever possible when individuals instigate it.

These examples illustrate the value of critical realism in generating better explanations and better interventions than dominant approaches. It is widely accepted by researchers, government officials, medical and public health practitioners, and the general public that human health and well-being is fundamentally tied to the ways in which we organise our societies. Yet this common wisdom is not reflected in health research funding practices and standards for evaluating health research. EBM, if adopted as the basis for the evaluation of public health evidence, will continue to legitimise research that is incapable of generating knowledge of the social relations affecting health. Interventions that are not based on a solid understanding of deep causal mechanisms across levels of reality and the factors which trigger or block their activation are unlikely to be effective. Furthermore, they may be quite harmful by transferring excessive responsibility to individuals for conditions that are beyond their control.

Theoretical and methodological pluralism is not sufficient to overcome these deficiencies in the EBM hierarchy. A new standard is needed. A critical realist standard would give theory a central role in research, particularly at the level of the social. There would continue to be a strong role for epidemiological methods and knowledge. Epidemiology identifies patterns between social traits and health outcomes that may signify enduring causal relations. To strengthen the claims made about these relations, epidemiologists can make

much better use of social theory in developing social constructs, such as class, race and gender, and in generating explanations for epidemiological findings.

While the synthesis of public health research is necessary to bring the full weight of knowledge to complex public health problems, it will be more fruitful if there is a higher quality of research to begin with. This requires researchers to develop greater clarity in defining which level of reality is the unit of the analysis and to explore causal mechanisms at this level and the levels above and below it. Often researchers turn to the level above or below to explain their findings without having designed the study to adequately theorise or measure these levels (Hackman 2003). This is the danger that epidemiologists encounter when straying into the realm of the social without the adequate conceptual or methodological tools. By making clear distinctions between levels of reality, critical realism provides a means of evaluating whether knowledge claims are being made at the appropriate level, using methods appropriate to that level.

Conclusion

Bhaskar has described critical realism as a philosophical "under laborer and occasional midwife" rather than a substantive theoretical perspective: it clears away some of the rubbish or undergrowth which stand in the way of the search for useful knowledge (Sayer 2000, p.28). It makes room for the use of a variety of methodological strategies and places theory at the centre of the research endeavour. It encourages the researcher to think about the quality of the explanation and the analytic product of public health research rather than a checklist to see where a study fits in the EBM methodological hierarchy of evidence.

The greatest challenge in developing a multilevel, interdisciplinary evidence base for public health will be in bringing public health researchers together to develop a new standard. Institutional patterns of research funding, practice and publication constitute a key obstacle to improving the quality of public health research and changing the criteria for its appraisal. In clinical medicine, the rules for determining truth have become institutionalised and embodied in a set of orthodox beliefs about the production of knowledge (Dean 2004). When applied to social structures, and the interactions between the biological and the social, these rules function to (re) produce flawed visions of society and the social determinants of health and illness.

These rules cannot produce the type of knowledge needed to address entrenched and growing public health challenges. Neither will they lead to an understanding of the interactions between broad social and political forces, emerging public health threats, individual behavior and health outcomes. Critical realism offers an alternative, if enough stakeholders in public health are willing to transform an accepted way of thinking about health and illness into a new form of practice.

References

Bhaskar, R. 1998, *The Possibility of Naturalism*, Routledge, London.

Coburn, D. and Poland, B. 1996, 'The CIAR vision of the determinants of health: A critique ', *Canadian Journal of Public Health*, vol. 87, no. 308, p. 310.

Danermark, B., Ekstrom, M., Jakobsen, L., and Karlsson, J. 2002, *Explaining Society: Critical Realism in the Social Sciences*, Routledge, London.

Dean, K. 2004, 'The role of methods in maintaining orthodox health beliefs in health research', *Social Science and Medicine*, vol. 58, no. 4, pp. 675–685.

Dean, K. and Hunter, D.1996, 'New directions for health: Towards a knowledge base for public health action', *Social Science and Medicine*, vol. 42, no. 5, pp. 745–750.

Eakin, J., Robertson, A., Poland, B., Coburn, D., and Edwards, R. 1996, 'Towards a critical social science perspective on health promotion research', *Health Promot. Int.,* vol. 11, no. 2, pp. 157–165.

Hackman, J. 2003, 'Learning more by crossing levels: Evidence from airplanes, hospitals, and orchestras', *Journal of Organizational Behavior*, vol. 24, pp. 905–922.

Hayes, M., Foster, L., and Foster, H. 1994, *The Determinants of Health: A Critical Assessment*, University of Victoria Western Geographic Series #29, Victoria.

Health Canada 2001, 'The Population Health Template: Key Elements and Actions that Define a Population Health Approach', Health Canada, Ottawa.

Houston, S. 2001, 'Transcending the fissure in risk theory: Critical realism and child welfare', *Child and Family Social Work*, vol. 6, no.3, p. 219.

Jackson, N. 2004, 'Background Paper: Guidelines for Reviewers of Health Promotion and Public Health Interventions', The Cochrane Collaboration.

Kelly, M., Swann, C., Killoran, A., Naidoo, B., Barnett-Paige, E., and Morgan, A. 2002, 'Starter Paper: Methodological Problems in Constructing the Evidence Base in Public Health', UK Health Development Agency.

Labonte, R. 1995, 'Population health and health promotion: What do they have to say to each other?', *Canadian Journal of Public Health*, vol. 86, no. 3, pp. 165–168.

Lalonde, M. 1974, 'A New Perspective on the Health of Canadians (The Lalonde Report)', Ottawa.

Macintyre, S., Chalmers, I., Horton, R., and Smith, R. 2001, 'Using evidence to inform health policy: Case study', *BMJ*, vol. 322, pp. 222–225.

Macintyre, S. 1997, 'The black report and beyond what are the issues?', *Social Science and Medicine*, vol. 44, no. 6, pp. 723–745.

Martin, E. and Simpson, E. 2004, *Evidence-Based Practice for Public Health Project* [online]University of Massachusetts Medical School, [online], Available from: http://library.umassmed.edu/ebpph [Accessed 15 May 2005].

Poland, B., Coburn, D., Robertson, A., and Eakin, J. 1998, 'Wealth, equity and health care: A critique of a "Population Health" perspective on the determinants of health', *Social Science and Medicine*, vol 46, no. 7, pp. 785–798.

Popay, J. and Williams, G. 1996, 'Public health research and lay knowledge', *Social Science and Medicine*, vol. 42, no. 5, pp. 759–768.

Rychetnik, L., Frommer, M., Hawe, P., and Shiell, A. 2002, 'Criteria for evaluating evidence on public health interventions', *J Epidemiol Community Health*, vol. 56, pp. 119–127.

Sayer, A. 2000, *Realism and Social Science*, Sage, Thousand Oaks.

Swann, C., Falce, C., Morgan, A., Kelly, M., and Powell, G. 2003, 'HDA Evidence Base: Process and Quality Standards Manual for Evidence Briefings', UK Health Development Agency.

Upshur, R., VandenKerkoff, L., and Goel, V. 2001, 'Meaning and measurement: A new model of evidence in health care', *Journal of Evaluation in Clinical Practice*, vol. 7, pp. 91–96.

WHO. 1986, 'Ottawa Charter for Health Promotion', Rep. No. WHO/HPR/HEP/95.1, World Health Organization, Geneva.

Williams, S. 1999, 'Is anybody there? Critical realism, chronic illness and the disability debate,' *Sociology of Health and Illness*, vol. 21, no. 6, pp. 797–819.

Demanding Referral
in the Wake of Conscientious
Objection to Abortion

Carolyn McLeod

Many moral issues surrounding abortion are philosophically complex. Among these issues is whether physicians who conscientiously object to abortion should be required to refer patients to an abortion provider.[1] Usually, that is what is expected of physicians: they can refuse to accede to patients' requests for abortion so long as they refer patients to providers who perform abortions (Blustein 1993, Wicclair 2000, Dickens 2002). The fact that some conscientious objectors routinely violate the referral requirement, however, has sparked controversy.[2] Some people think such behaviour—refusing even to refer a woman for an abortion (!)—is abhorrent; I recently heard a pro-choice activist describe it as *"heinous."* Others think such behaviour is entirely appropriate. For example, Sean Murphy,[3] a leading advocate of conscience protection laws for health care providers, claims that the requirement to refer is unconscionable because it makes physicians complicit in the performance of acts which they find offensive. Murphy's position actually makes more sense to me than does that of the pro-choice activist (surely "heinous" is too strong!). Still, I don't think any of us should side with Murphy. Rather, we should maintain that even physicians who conscientiously object to referrals for abortion and to other relevantly similar practices are not heinous, but that nonetheless, they should be required to make the referrals. As physicians, they are morally obligated to do so.

Arguing conclusively that physicians have that obligation is tricky, however: too tricky for a paper of this length. So instead, I will simply

1. Here and throughout I am concerned with physicians who conscientiously object to abortion. For simplicity, therefore, I will simply talk about conscientious objectors, leaving it understood that the objectors I have in mind are ones who object to abortion.

2. Routine violations in Canada were a finding of a study done by CARAL (the Canadian Abortion Rights Action League) called "Protecting Abortion Rights in Canada" (2003). I imagine that such violations occur in the United States even more than in Canada, since opposition to abortion is stronger there than in Canada.

3. He is also the administrator of the Protection of Conscience Project, a non-profit international initiative that advocates for conscience protection laws for health care providers. His comments about referral appear at http://www.consciencelaws.org/Examining-Conscience-Issues/Ethical/Articles/Ethical12.html.

explain why the issue of referral in the wake of conscientious objection is complex. There are two main reasons why: 1) demanding referral can conflict with protecting conscience, but protecting conscience is important; 2) demanding referral limits conscience protection, but deciding what reasonable limits one can place on conscience protection is difficult.

When Referral and Conscience Protection Conflict

I will call "the common arrangement" that which allows conscientious objectors to refuse to perform abortions, but which requires them to refer patients to abortion providers. Many physicians see a fundamental problem with the common arrangement. The problem is this: while the arrangement is often thought to strike a balance between physician conscience and patient interests, if that were true, the referral requirement could not conflict with the protection of conscience, which it can do. In response one might ask why since physicians do not have to perform the "offensive" act themselves, their conscience isn't protected to some extent. Shouldn't physicians recognise referral as a *compromise* between their conscience and patient interests?

The issue is complicated. Referral can be a true compromise, but it can also be a false one, depending on the wider belief system or character of the provider. For example, referral is a true compromise when the provider believes that it is always worse to do wrong than to allow others to do wrong, which is what one does, from one's own perspective, when one makes a referral for a procedure that one finds offensive. Alternatively, the provider could be so epistemically humble when it comes to controversial moral issues that she can live with allowing others to do what she perceives to be wrong (Blustein 1993, pp. 309–11).[4]

Referral is a *false* compromise in the following sorts of cases. The provider believes that allowing others to do wrong is always as bad as doing the wrong himself. Alternatively, while he does not lack epistemic humility, he thinks that he should always err on the side of caution rather than risk committing what in his view is a serious moral wrong (e.g., murdering a fetus).

The likelihood that providers will see referral as a false compromise may be quite high, given the frequency with which people object to

---

4. There are other aspects of a provider's belief system or character that could make referral a true compromise. See Blustein (1993, pp. 309–11).

facilitating moral wrongs. Just consider whether you (reader) would object to the referral requirement if you had the same beliefs as a physician who fervently opposes abortion. From this perspective, you would surely say, with profound skepticism in your voice, "So I can refuse to commit a murder, but I have to tell someone who I know intends to commission a murder how she can go about doing that?" Many of us—no matter how epistemically humble—would refuse to help the woman.

Putting ourselves in the shoes of pro-life physicians may help us to better understand their concerns about conscience; but it may not convince us that those concerns should outweigh respect for patient autonomy. In other words, we may agree that referral can be a false compromise, in which case the provider receives no protection for conscience, while also agreeing that that is as it should be. Surely, where patient autonomy and provider conscience conflict, the former should win out. But is that *obviously* the case? Why should we care about protecting provider conscience anyway?

There are at least two important reasons why we should care. The first has to do with harm to providers: requiring that they violate their conscience whenever it conflicts with patient requests puts them, the providers, at serious risk of losing moral integrity through self-betrayal, which could profoundly compromise their psychological health and agency. The second reason concerns the good of patients. One might think that for patients' sakes, providers should follow the voice of their profession, not of their conscience; providers should simply abide by the standards of care that their profession lays out for them. But is it really in patients' best interests to have providers simply adhere to external standards? Don't we want providers also to care about their patients? I and a colleague have argued elsewhere that people with a conscience care about others, to some degree at least (McLeod & Bendik-Keymer 2002). They are morally connected to others, so that they perceive them not as objects to be manipulated, but as "subjects-of-a-life" who deserve some sympathy and respect (Regan 1983). If they did not believe that, they would be sociopaths, and a sociopath is a paradigm of someone who lacks a conscience. We should want providers to act on their conscience, therefore, and we should protect their ability to do so as part of our general desire that they care about their patients.

While we do, or should, hope that providers will care about their patients, we don't want them to care in any old way, however. For

example, we don't want them to care about whether their patients will burn in hell because of a blood transfusion or because of a patient's desire to have sex before marriage. Surely there are limits to the kind of conscience we're willing to protect in medicine.

Limiting Conscience Protection

I have said that protecting conscience in medicine is important; and yet, despite its importance, there should be limits to our willingness to accommodate conscientious objection. It is clear that the common arrangement effectively eliminates protection for some of these objections. I think that is as it should be; but explaining *why* that is as it should be is no mean feat. The difficulty lies in defending the relevant limit on conscience protection as a reasonable limit. Let me show how we run into this difficulty if we give what I take to be a common response to the refusal to refer.

Here is the common response: A physician's conscience *as a physician* should not preclude her from making referrals for abortion (or any referrals that the profession demands) because presumably, she freely chose to enter the profession and freely chooses to stay in it, and in doing so, "agreed [or agrees] to practice medicine according to the norms of the profession" (Blustein 1993, p. 312). Jeffrey Blustein considers this response in his work on conscientious objection, and claims that it "is open to an obvious reply. Physicians who deny a duty to refer are saying, in effect, that they do not share the medical profession's prevailing conception of itself" (Blustein 1993, p. 312). To me, this reply is not obvious, however, or is at least not obviously a good reply. Some physicians may indeed have a conception of medicine that differs from the prevailing one; yet the point of the common response is that they have agreed to practice in accordance with the prevailing conception.

Underlying the common response to the refusal to refer for abortion is a view about the proper scope of conscientious objection in medicine (hereafter *"The View about Scope"*): *physicians cannot make conscientious objections in their practices that violate established norms of the profession.* Central to The View about Scope is the idea that conscientious objection is permissible only if the practice being objected to is not an established professional norm. Notice too that The View about Scope restricts conscientious objection only within clinical practice; it does not prohibit physicians from rallying for change in other contexts, including those in which the real debate

about medical norms takes place (e.g., professional meetings). What is an appropriate form of conscientious objection in those other contexts is different from what is appropriate in clinical practice, for reasons that will become clearer below.

The View about Scope would reasonably account for why conscientious objectors should refer if it, 1) in fact supported the common arrangement, and 2) is morally justified. It is to these questions that I'll now turn.

First, does The View condone the common arrangement? The answer is "yes" because there are no established norms in medicine (in North America) saying that physicians have to perform abortions, but there are many norms suggesting that they have to refer for abortions.

Let me explain the point about referral first. Are there norms that demand referral? There must be if the common arrangement for conscientious objection to abortion is as I've said it is. The norms that support the referral part of this arrangement have to do with not abandoning patients, with respecting their autonomy, honouring their trust, and being beneficent towards them. Arguably, in medicine, there is an ethic of not abandoning patients, which is sometimes cited as the source of the referral requirement (Blustein 1993, p. 312). But one could also get this requirement out of the duty of physicians to respect the autonomy of their patients. To be autonomous, patients must have real options for pursuing their goals, and abortion is one option for women in unwanted pregnancies. Respecting the autonomy of these women must involve making them aware of this option and how they can access it. In other words, it must involve referral. A third relevant norm is that physician ought to honour the trust that patients place in them to act in their, the patients', best interests, and to interpret "best interests" not just from the provider's own perspective, but from the perspectives of the patient and of the profession as well. According to many patients and the profession, abortion should be an option for women in unwanted pregnancies. It follows that not referring can violate patient trust.[5] A final norm that supports referral is that physicians should act in their patients' best interests, as defined above. The profession tends to interpret beneficence toward women in unwanted pregnancies in terms of women having access to abortions.[6] Thus, according to the prevailing norm of beneficence, as well as those

5. In legal terms, it violates the fiduciary duty of the physician.

6. It does so, in all likelihood, because of its history of dealing with the horror of botched abortions when abortions were inaccessible.

of trust, patient autonomy, and not abandoning patients, physicians should do referrals for abortion.

One might wonder where these norms come from. How can we be certain that the prevailing norm of beneficence, for example, is as I've described it? I assume that we can refer, in part, to the codes and statements of medical associations. For example, the Society of Obstetricians and Gynaecologists of Canada endorses the International Planned Parenthood Federation Charter on Sexual and Reproductive Rights, which includes "the right to decide whether or when to have children."[7] This right suggests that beneficence would demand referral for (if not the performance of an) abortion.[8] For some norms, one could also look to the law and how it governs the medical profession. For example, fiduciary law and the law of disclosure surrounding informed consent (in Canada and the U.S.), respectively, support the above interpretations of norms about trust and autonomy. A further source for norms would be actual medical practice, and the norms, implicit or explicit, that prevail within it.

The results about norms from these different sources could conflict, of course, in which case, sorting out how to apply The View about Scope would involve having to determine, if possible, which source of norms has the greatest authority. If completing that task were impossible and conflicts were widespread, so that for many situations prevailing norms did not exist, surely the fault would lie with the profession, not with The View about Scope. No profession can legitimately call itself *a* profession if it does not have reasonably clear norms that define it.

To recapitulate, I've been suggesting that there are clear norms that favour referral for abortion. These norms imply that referrals are mandatory, according to medicine's prevailing conception of itself. While I think this conclusion is uncontroversial for many jurisdictions in North America, I recognize that it is controversial for others. For example, according to the College of Physicians and Surgeons of New Brunswick, and the law in some states (e.g.,

7. See http://www.sogc.org/sogcnet/sogc%5Fdocs/intl/chart%5Fe.shtml

8. To give another example about a different norm, the Canadian Medical Association (CMA), along with other Canadian health care associations including the Catholic Health Association of Canada, have agreed that physicians should not put patients at "risk of … abandonment." (See the "Joint Statement on Preventing and Resolving Ethical Conflicts Involving Health Care Providers and Persons Receiving Care," found on the CMA website at http://www.cma.ca.) Hence, not abandoning patients is a norm for these associations. Norms about abandonment, autonomy, and trust also exist for the American Medical Association and the American College of Obstetricians and Gynecologist. See the former's "Fundamental Elements of the Patient-Physician Relationship" (http://www.ama-assn.org/ama/pub/category/2498.html) and the latter's "Code of Professional Ethics" (http://www.acog.org/from_home/publications/ethics/).

Michigan and Illinois[9]), physicians who conscientiously object to standard medical procedures have no obligation to refer.

Whether The View about Scope supports the ability of physicians to refuse to perform abortions is more straightforward. The answer is "yes," because there is no medical norm that says physicians *must* perform abortions upon request. To be sure, if a woman had nowhere else to turn to get an abortion, a physician could violate norms of beneficence, trust, or autonomy if he turned the patient away. But if that were not the case—if another physician was available to do the abortion—the first physician's refusal would be permissible.

Let me now turn to the second question, namely whether The View about Scope is morally justified. The worry is this: even if The View about Scope is consistent with the common arrangement, it may lack independent credibility as a theory about when conscientious objection in medicine is permissible. In what remains, I'll consider this worry.

Limiting conscientious objection in medical practice to what does not violate established norms is appealing for a number of reasons. First, it helps to preserve the integrity of the profession. Second, it helps to maintain patient trust, since, as I have argued elsewhere, confusion about what norms someone will follow can seriously inhibit trust (McLeod 2002). Further, when what people medically require is crucial to their well-being, their trust can be shattered by conscientious objection. Third, the point about consent is worth repeating. Providers have agreed to follow the norms of their profession, if only by accepting the privileges that go along with membership in the profession. They do not need to be providers; surely, they have the ability to do other things with their lives.

But what if, in deciding to be a health care provider, their goal was to improve the profession from the inside out? While perhaps we would not value such behaviour if it aimed to make the profession more loving toward the "unborn," we would value it if it targeted areas of medicine that we thought were simply corrupt. It seems that The View about Scope would have us oppose such behaviour, however; and that is the most serious objection to The View about Scope: it is too conservative with respect to professional norms. The norms themselves might be corrupt. They might say, to give some stark examples, "Gas the Jews," or "Falsify the death certificates of torture

9. See Michigan's "Conscientious Objector Policy Act" and Illinois's "Health Care Right of Conscience Act," found on the websites for the legislatures of those states.

victims."[10] Shouldn't we allow physicians to object in their clinical practice to these sorts of norms, even if the norms are well-established?

I think this objection is serious. Indeed, I think it calls for an amendment to The View about Scope. Here's one amendment that seems appropriate, for reasons I'll describe momentarily: revise The View about Scope so that it says physicians cannot make conscientious objections that violate established norms of the profession, *except when the norms require them to lay hands on someone not for that person's benefit, but for the benefit of others*. Call this simply "The Amended View." The exception within it comes from a principle defended by Judith Jarvis Thomson in "A Defense of Abortion": she says that "one has a right to refuse to lay hands on people, even where it would be just and fair to do so, even where justice seems to require that somebody do so" (1971, p. 54). Justice or fairness are the relevant norms because Thomson is referring to situations in which someone has to lay hands on another out of fairness to someone else (e.g. lay hands on Jones, who has taken Smith's coat, and return the coat to Smith). Similarly, I think that where "justice," as defined by prevailing norms of medicine, says that physicians must do harm to patients for the sake of others, physicians can refuse to do so.

The amendment is worthy of support for a number of reasons. 1) It accommodates many of our intuitions about norms to which we tend to think conscientious objection should occur in medicine, including discriminatory norms (e.g. gas the Jews, sterilize welfare moms) and norms about torture. 2) The amendment is consistent with the tradition of conscientious objection to war; in the past, the state permitted objectors to wars that it thought were just because people should not have to lay hands on people for others' sakes. 3) Something like the amendment underlies the common arrangement for conscientious objection to abortion, which allows physicians to refuse to lay hands on fetuses for women's sakes. 4) In the face of corrupt practices—gassing the Jews, sterilizing welfare moms—the amendment would allow physicians to object without having to refer, because in such situations there is no patient to refer (the patients themselves are not asking to be killed or sterilized!).

But unfortunately, The Amended View will not do. For what about situations where the norms of medicine require physicians to harm

10. The second may have been the norm for military medical personnel at Abu Ghraib. Some of them did falsify death certificates of the Iraqi prisoners who were tortured there (Miles 2004; Lifton 2004).

patients for the sake of others by *omitting* treatment for the patient (e.g. omitting infertility treatment for a lesbian woman, omitting treatment for a torture victim)? The principle about not laying on hands does not include harmful omissions. *To* include them, however—to say, that is, that physicians can conscientiously object to norms that require them to make omissions they deem to be harmful—could put the arrangement of objecting-then-referring for abortion in jeopardy. Integral to that arrangement is the understanding that physicians will omit from doing whatever is in their power to prevent a woman from having an abortion. But that is an omission to which some physicians who are opposed to abortion would strongly object. They would want to do whatever they could to save fetuses, rather than simply refer women on for abortion.

The upshot is that further amendment to The View about Scope is in order; but it's doubtful that any simple amendment would work. While one might have thought that determining whether physicians who conscientiously object to abortion ought to refer for abortion would be straightforward, as I have tried to show, it is not. I certainly hold out hope for an answer; but I can't pretend that the issue is easy. I don't think any of us can.

References

Blustein, J. 1993, 'Doing What the Patient Orders: Maintaining Integrity in the Doctor-Patient Relationship', *Bioethics*, vol. 7, no. 4, pp. 289–314.

Dickens, B. 2002, 'Informed Consent', In *Canadian Health Law and Policy*. Eds. Downie, J., Caulfield, T., and Flood, C., Butterworths, Markham.

Lifton, R. J. 2004, 'Doctors and Torture', *The New England Journal of Medicine*, vol. 351, no. 5, pp. 415–416.

McLeod, C. 2002, *Self-Trust and Reproductive Autonomy*, The MIT Press, Cambridge.

_, and Bendik-Keymer, J., Unpublished manuscript. *Protecting Legitimate Appeals to Conscience by Health Care Providers*.

Miles, S. H. 2004, 'Abu Ghraib: its legacy for military medicine', *The Lancet*, vol. 364, August 21, pp. 725–729.

Regan, T. 1983, *The Case for Animal Rights*, University of California Press, Berkeley.

Thomson, J. J. 1971, 'A Defense of Abortion', *Philosophy and Public Affairs*, vol. 1, no. 1, pp. 47–66.

Wicclair, M. 2000, 'Conscientious Objection in Medicine', *Bioethics*, vol. 14, no. 3, pp. 205–227.

The Marginalisation of
Public Health
in Ontario

Mary Powell

This paper is organized to fulfil three objectives: The first is to introduce evidence that public health in Ontario has been marginalised. The second is to consider three theoretical issues that are fundamental to the discussion. In the third section, I explain why marginalisation has occurred and make the argument that public health has been in dismal shape in Ontario for most of the past 125 years.

The term, public health, is used here to refer to preventive measures taken to safeguard the public from communicable diseases and unsanitary conditions. In general, the focus of public health is not on the health of the individual *per se*, but on the potential public impact of a person's health (for example, a case of smallpox could endanger the health of the entire community, while someone's appendicitis could not) or on the essentially social factors that influence people's health (for instance, tuberculosis is more likely to spread in overcrowded and unsanitary housing).

Evidence to support the argument that public health has been marginalised is drawn from three contemporary events: first is the 1997 decision by the Harris Government to abandon a policy (in place since 1940) of providing provincial operating grants to local health authorities and instead to make municipalities bear 100% of the costs. Further evidence of marginalisation is that, despite modern expectations of clean water, the Walkerton water supply was contaminated with bacteria in May of 2000, leading to seven deaths and hundreds of cases of illness, some of them severe and life-threatening. Third, the pattern of weaknesses in public health revealed when the SARS epidemic hit Toronto in February 2003 is further confirmation that public health has been marginalised. The effect of cutbacks and downsizing within the provincial Ministry of Health, the

consistent underfunding of public health, and patterns of institutional interaction that hindered communication among health units, between health units and hospitals, and between the field and the Ministry of Health all made dealing with the epidemic more difficult.

One of the theoretical issues often raised in discussion of public health is the prevention-treatment conundrum. In every sector, from child abuse to cigarette smoking to criminal behaviour, societies understand the need to spend on *treatment* in a way that they never really grasp the need to spend on *prevention*. Certainly there is a strong intellectual commitment to prevention but, at the level of social spending, prevention is consistently underfunded or disconnected so that it is less effective (McKinley 1998). It has been demonstrated repeatedly that trying to integrate prevention with treatment leads inevitably to the leak of prevention funds into treatment urgencies. The philosophical issues involved here are beyond the scope of this paper, but they remain essential. We must be able to reconceptualise prevention so that governments can move beyond the unproductive prevention-treatment dichotomy.

A second theoretical question is how to approach the study of public health policy and structures in Ontario over the course of almost 125 years. To grasp the pattern of development, it is helpful to examine the main institutions in their historical context, an approach referred to as historical institutionalism. An example drawn from the creation of the Provincial Board of Health may illustrate the point. It has been argued that institutions "emerge from a collision between people with new ideas and an already existing world with people who have other ideas already institutionalised…." (Kloppenberg 1995, 125). Medical men and leading physicians worked for years to convince the provincial government to act on the need for sanitary measures, which they understood clearly to be imperative in Ontario. The Mowat government resisted for almost 10 years, at times giving no reason, and at other times stating that it would cost too much. When the medical men did convince Mowat, the Premier, to act on their scientific convictions, Mowat's action was to introduce a bill into the legislature that would give institutional form to scientific conviction through the establishment of the Provincial Board of Health (PBH). During Mowat's first meeting with Dr Peter Bryce (secretary of the newly appointed PBH), the Premier made it clear that the Cabinet was not necessarily committing itself to the 'sanitary crusade' but rather was giving authority to the PBH to act on *their* convictions. Mowat

said to Bryce: "we have passed this health legislation but have … little knowledge of what there is to do or of its extent … its success will wholly depend upon your energies" (Bryce 1921, 1). The PBH as an institution derived its legitimacy from the Ontario Legislature, and was the first permanent institution in Ontario history to have authority (albeit limited) to act to on matters of public health. Throughout this paper, reference will be made to the rise, transformation, decline, weakness, and conflicts among historical institutions.

The third theoretical issue relates to the question of common professional training. Although many civic leaders became involved in campaigns for clean water supplies and control of infectious diseases, the majority of those involved and in leadership positions were physicians. By 1912 the University of Toronto offered a one-year course of study leading to a Diploma in Public Health; most physicians, intending to work full-time as an MOH or for the province, pursued this advanced training. When District Officers of Health were hired, all seven (later eight) were physicians who were given special training in public health by the PBH. It later became standard practice (if not yet a legislative requirement) for any physician working full-time in public health to have post-graduate training in the field.

Public health physicians' views on public health were profoundly similar largely because of the common bonds of their professional education. One former official observed, referring to the uniformity of services in Ontario that, "…[public health] physicians all went to the [University of Toronto] School of Hygiene, took their DPH, had the same lectures … and that's how we got uniformity."(Ontario Standing Committee on Social Development 1982 (Sept 8), p. 27). Department officials had a growing group of allies among the full-time public health physicians (MD, DPH) working as MOHs in city health departments and county health units, whose work situations were different but who shared common professional bonds.

The impact of common training was two-fold. First, all full-time MOHs and public-health physicians in the Department of Health had a common perception of public health problems, of how to address them, of what programs were effective, and how they should be organized. This common culture allowed the two groups (local and provincial) to agree on policies, procedures, structures, and patterns

of funding. Second, while there is evidence that bureaucrats in general developed ties from years of working together (Breton and Wintrobe, 1982), the trust among public health physicians was deepened by their common professional training. The strength of professional bonds created harmony, smoothed over conflicts, strengthened cooperation, and made institutions work well together. Even when the institutions within which they worked had less power or authority than they needed for some of their plans or, for example, local authorities were reluctant to join larger, more economic health units, the provincial-local network of public health physicians felt free to innovate, unfettered by usual rules, and found all sorts of ways to encourage local authorities to reconsider, not the least of which was their home-made, back-of-the-napkin system of provincial grants to health units. When public health physicians later had less and less authority within the Department/Ministry of Health, and when health programs important to MOHs in the field were fragmented among other ministries, the level of contact between provincial public health physicians and MOHs declined markedly. Without contact with their provincial colleagues and without the common cause the two levels had shared for so long, MOHs faced institutional conflict, disharmony, and inadequate information in their attempts to deal with other provincial ministries.

We turn now to the question of why public health has been marginalised. The first point that needs to be made is that public health is a matter of provincial jurisdiction. The federal government took responsibility for quarantine stations, like the station at Grosse Île in the St. Lawrence, but refused to take broader measures for public health on the grounds that it might interfere with relations with the provinces (Brouse 1875).

Provincial jurisdiction over public health meant that the province could act itself or could delegate responsibility to local governments, over which the province had complete legal control, if not complete political control. By the pre-Confederation Baldwin Act governing municipalities, every municipality had the authority to establish a local board of health, and leading cities such as Toronto, Hamilton, and Windsor had already done so. But the rest of the municipalities were largely without any kind of public health authority, a subject of great concern for physicians at the medical schools and their colleagues who saw unsanitary conditions breeding disease and outbreaks of smallpox and typhoid fever spreading unchecked in rural Ontario.

After years of campaigning, this group of physicians and civic leaders convinced Oliver Mowat's government to create the Provincial Board of Health in 1882. The legislation establishing the PBH was both its guideline and its boundary. In terms of membership, the act required only that four board members be physicians, but in the PBH's 45-year history, 33 of the 34 appointees were physicians (the single exception was an engineer, appointed late in 1882 to replace a departing member). All appointments were by Order-in-Council, so it is clear that successive Cabinets regarded physicians as the appropriate specialists to direct public health in Ontario. At the same time, the PBH was at arm's length from the Cabinet. The originating board had hoped to have a Cabinet minister on the board itself, or at least to have a Minister of Health at the Cabinet table. The plea for a Minister of Health was renewed in 1905, but to no avail. What this meant to the PBH was that no minister took responsibility for the board or for public health. When the Provincial Board of Health sent memoranda asking for increased powers so that it could be more effective, no one at Cabinet supported the PBH case, and their requests were occasionally ignored and often rejected.

From the outset the Provincial Board of Health was a weak institution, not only because it had so little influence with Cabinet, but also because its powers over local boards of health were purely advisory. As the first PBH discovered once it was in operation, few municipalities had even appointed local boards of health. The Mowat government acknowledged the PBH's point about impossibility of working without local boards of health and passed the *Public Health Act of 1884*, which required all municipalities to appoint a local board of health. The line of argument pursued by the PBH was that local boards of health, whose only duty was to care for local health, would surely become champions of the cause (just as the PBH was), and would stand up for public health even in the face of local opposition.

Neither of these assumptions withstood the test of reality, because members of the local board were not convinced that what the PBH wanted was worth championing, and they were certainly not willing to defend it against their own friends and neighbours. The two most important institutions did not achieve their principal goals. The PBH tried but was constrained by its own weaknesses and by the recalcitrance of local boards of health. Local boards of health failed by choice; their mandate was to regulate local sanitary conditions and, to a great extent, they chose not to do so. When the PBH proposed

stronger legislative measures, a senior member of cabinet reminded Provincial Board members of "the propriety of hastening slowly, of educating the people until the time comes when ... [PBH] requirements will be received as reasonable …" (Ontario Provincial Board of Health 1887, p. xl).

It was also up to the council to hire a medical health officer (MHO). The key to understanding the role of the MHO is that the physicians who were appointed by the local councils were (except in the largest cities) physicians in private practice, whose livelihoods depended on their ability to attract and retain paying patients. Actions such as closing down tanneries or imposing quarantine were contrary to their economic self-interest, and municipal councils relied on this conflict of interest. Just as members were often appointed to the local board of health on the understanding that they would do as little as possible, so too were local physicians appointed as MHOs to spend little and interfere less. Their pay was intended to remind them of this agreement: in 1897, when 56% of the Ontario population lived in rural townships, the average annual salary for township MHOs who were paid (many were not) was $17.49. This contrasted sharply with the $3000 annual salary for the full-time MHO in Toronto (Provincial Board of Health 1897, p.9).

As early as 1891, the Provincial Board of Health knew that the system of local boards of health and medical health officers was a failure. At that time Dr Peter Bryce, Secretary of the PBH, had proposed county health units as the appropriate alternative, arguing that they would contain enough people to warrant hiring a full-time medical health officer who could be properly trained and given security of tenure. If a medical officer with a full-time salary could fulfil his duty without fear or favour, he could enforce sanitary bylaws, vaccinate and quarantine when necessary, and undertake other measures to safeguard the public health. This county health system was adopted in Ontario in 1940, when the first of what would be 46 units was established. But in the 19th and early 20th century, opposition to local boards of health and to county health units came from the local townspeople, tradesmen, and farmers, because of their persistent lack of belief in public health measures and an even more ingrained unwillingness to spend local funds on anything not regarded as utterly essential.

Another attempt at securing effective local public health came in 1912, when a new public health act made appointment of a medical

officer of health (MOH) compulsory. The act also required that the MOH be paid a reasonable salary, that he could only be fired for cause, and only then if the PBH approved. In an effort to improve public health knowledge among MOHs, they were required to attend an annual meeting of Health Officers in Toronto, where the PBH would provide educational sessions on the most important issues of the day. For the first time, District Officers of Health (DOH) (at first, seven, later eight) working for the PBH, were hired as field staff to encourage and back up the local MOHs (Ironically, the 1912 Act required that the counties being supervised by a DOH collectively pay his salary; this provision brought nothing but complaint, resistance, non-compliance, and threats to vote against the Government. It was dropped in 1918). Despite the 1912 changes, the structural weaknesses of the previous system remained: local municipalities did not want their local boards of health to spend money, and the vast majority of local MOHs were physicians in private practice, who still faced the fundamental conflict between public duty and private livelihood.

There was intermittent improvement in the state of public health in many municipalities, but it depended on the health of the local economy and on the connection local people could see between public health expenditure and their health and economic well-being. Rural Ontarians were a sceptical lot.

With the election of the Farmer-Labour government in 1919, and the appointment of a Minister of Health and Labour to argue for his portfolio, the provincial government was willing to spend more on health. By 1920, the staff of the Provincial Board of Health had increased to 50, from 28 the year before, and was organized into a standard divisional structure (Division of Maternal and Child Welfare, Division of Sanitary Engineering, etc).

In 1924, the Department of Health was created as part of a government-wide trend to departmentalization, a decision that stemmed largely from the need for the political executive—the Cabinet – to exercise greater control over rapidly growing government services. This growth was indicated by the increase in government expenditures, which rose from $5.4 million in 1905 to $57.9 million in 1930, and the growth of the provincial public service, which also increased tenfold, from 704 in 1904 to 7,760 in 1932 (Schindeler 1969).

For public health, this was an important point of institutional transition. The department was a new organisational form, headed by

a deputy minister with direct access to the Minister of Health, access that the PBH had never had. Furthermore, with an enlarged staff complement, the department was also (in terms of its potential) a new institution; almost half its employees had been hired since 1920. Dr William Bell, a paediatrician from the Division of Maternal and Child Welfare was appointed first deputy minister in 1925, signalling both an end for the PBH and a new focus for the Department on what it could do itself in public health, rather than serving as a guardian to local authorities. Because the PBH continued to exist, bitterness and acrimony arose between Board members and the new deputy minister. Superficially the point of argument was the PBH's championship of local public health and its disapproval of the new department's lack of concern; but in a more visceral way the PBH and their supporters were angry and dismayed at having served so long and then being cast aside, with no recognition of their expertise or years of service, without even courtesy appointments in the new Department. McCullough was kept on in the Department but was assigned no duties. In 1927, to bring the unhappy situation to an end, the government passed an act abolishing the Provincial Board of Health; this was roundly condemned in the newspaper by the old guard of the public health community.

Spending on health increased during these years, but most of the funds went to programs offered by the Department of Health itself: travelling exhibits, pamphlets, treatment of patients with venereal disease, and cancer control. While the eight DOHs still worked in the field, little other attention was given to supporting local public health.

In 1930, the Royal Commission on Public Welfare (the Ross Commission) recommended a reorganisation of responsibilities for health. In particular, the Commission wanted responsibility for inspection of, and administration of grants to, general hospitals assigned to the Department of Health. The Ferguson Government accepted this recommendation, but went beyond it to assign responsibility for psychiatric hospitals (then known as Ontario Hospitals) to the Department of Health.

The magnitude of the change is apparent in the budgets and staff complements involved. The original Department of Health had a budget of about $700,000 and a staff of fewer than 100. The Hospital Inspection and Grants Branch administered grants of almost $900,000 to general hospitals and had a staff of between 25 and 30. Ontario

Hospitals had operating expenditures of $4.2 million and employed 4000 people.

Because of the size of the organisation, it was decided that there should be two deputy ministers. Dr William Bell continued as deputy minister for the health side of the Department, and Dr BT McGhie, a former superintendent of a psychiatric hospital, became deputy minister for Hospitals (general hospitals inspections and grants and Ontario Hospitals). In a general shakeup under Hepburn in 1935, this unusual arrangement was ended and the Department reverted to the normal practice of one deputy minister, with McGhie continuing as deputy minister. However, in recognition of the longstanding lack of leadership in public health, Dr John Phair (DPH), then Director of Maternal and Child Welfare and Director of Public Health Nursing, was named Chief Medical Officer of Health, a position that had not existed earlier, though McCullough had had the title Chief Officer of Health.

Regardless of the loss of its own deputy, hospitals remained organisationally a separate division. It is important to note, for the future that, from 1935, despite being under the control of the same deputy minister, hospitals in Ontario formed a separate division in the Ontario Department of Health, a division organisationally distinct from public health services offered directly or in support of local public health. This separation was complete to the point that the Hospitals Division produced its own annual report. While they may have nominally been in the same department, health and hospitals were institutionally distinct and had few connections, common pathways, or methods of coordination.

During the Depression and the War, expenditures were cut back and many positions went unfilled or were held for staff in the armed services. When the Liberal Government of Mitch Hepburn was elected in 1934, one of the government's first moves was to fire all civil servants hired since 1933 and many others, including all eight DOHs. Innovation, to the extent it occurred, was financed by outside sources. A health survey of four counties in eastern Ontario was paid for by the Canadian and Ontario Dental Councils, the Ontario Medical Association and others, partly as a relief program and partly as a health survey. When this project was nearly completion, the provincial government agreed to apply to the Rockefeller Foundation for money to undertake a full-time county health unit. The Eastern Ontario

Health Unit operated from January 1ˢᵗ 1935 to the 31ˢᵗ of December 1939, offering full-time public health services under the direction of a full-time MOH and eight full-time public health nurses. Toward the end of trial period, Dr James Munro, a physician, local reeve, and strong supporter of the health unit, undertook a campaign to convince local decision makers that they should continue with the health unit. Munro met with the local MPP, with the Minister of Health, with local councillors and local businessmen. When the vote came, the United Counties of Stormont, Dundas, and Glengarry voted to carry on with the health unit, but wanted a grant of 50% of the costs from the province. It was only at the eleventh hour that the Premier (such a momentous decision was not left to the Minister of Health) informed the Counties' Council that the Government would meet half the cost for one year. The first county health unit in Ontario, with 50% provincial funding, went into operation on January 1ˢᵗ 1940, with a full-time staff offering full-time services.

It was clear to John Phair, the CMOH, that county health units could be developed all over the province once the war was over, and he began to prepare for the peacetime expansion. First, he set up a Division of Public Health Administration and hired two new public health physicians, one to be his assistant and one to head up the new Division. Then in August 1945, with the death of McGhie, Phair himself was appointed deputy minister. By 1950, Phair and his colleagues had helped to create 26 county health units around the province; in each case, this meant the abolition of local boards of health with part-time MOHs and the absorption of the population into the new county health unit that offered full-time services under the direction of a full-time MOH and staff.

Three institutionalised practices, each supporting the other, made the county health unit system not only a possibility but a reality. The most obvious and visible institutionalised practice was the provincial conditional grant. The province paid 50% of the costs for the county, town, village, and township participants (55% if the county was really poor), and subsidised cities that joined the county unit on a sliding scale inverse to population, so that cities over 100,000 population received only a 15% grant. One of the remarkable features of this grant system is that it was institutionalised, in the sense that the Department of Health and all participants in all the health units recognised its existence and its rules, but at the same time it was completely

informal, authorized from time to time by a memorandum from the Minister and structured according to the wisdom of the public health physicians and their sense of what would be appropriate or necessary to get a county health unit set up. In the early years, the usual practice was for the province to pay the grants quarterly in arrears (later, monthly in arrears) and it was a measure of both the institutionalisation and the informality of these arrangements that health units borrowed their operating funds from local banks, even though there was no legislative authority for them to do so.

The strong provincial-local network of public health physicians was the second institutionalised practice. Two developments made this possible. First, throughout the period from 1945 to 1965, public health physicians served as deputy ministers of health (Phair to 1957, Gordon Brown 1958 to 1965). This gave public health a high priority and authoritative leadership. The deputy ministers had strong subordinates, also public health physicians, administering public health programs and the health unit program in particular, which had risen to 38 units by 1965. Second, not only were there strong public health physicians at the provincial level, but in the field, public health physicians served as MOHs in the 15 city health departments and in the 38 county health units, giving the provincial physicians a strong, like-minded group with whom to plan public health programs.

Meetings between provincial officials and local health unit staff became quarterly events. Because there were so many full-time MOHs, public health nurses, public health inspectors, and public health dentists, regular meetings became very productive. In the early years, the information flow was from the provincial experts to those in the health units, but as the health units gained experience, the information flow balanced out and then shifted so that local health officials were making demands on the department for better materials to use with the public, more technical information relating to health hazards, and for a higher level of expert advice. This meant that, when provincial policy advisors in public health argued their case for new programs, they had strong evidentiary support from the local side of the institutionalised network.

The third institutionalised practice, in Ontario as in other provinces, was the departmentalized cabinet. Ministers exhibited what was referred to as portfolio loyalty, trying to ensure that their departments got a fair share (or a greater share) of resources and won any

interdepartmental conflicts. At the cabinet table during this period, the Minister of Health was expected to speak for his Department and his cabinet colleagues trusted that he knew what was best, just as he trusted the Minister of Education or the Attorney General to know what was best in their domains. Between 1945 and 1969, there were three Ministers of Health: R T Kelley (Jan 1945 – Aug 1950); Mackinnon Phillips (Aug 1950 – Dec 1958); Dr Mathew Dymond (Dec 1958 – Aug 1969). Each of these ministers served long enough to get to know his departmental staff well. Kelley had to be convinced about the county health unit and it is evident that he became a strong supporter. Mackinnon Phillips presided over the growth of county health units to the point that more than 70% of the population of Ontario had access to full-time public health services. Mathew Dymond, himself a physician, was also comfortable with his departmental staff and a strong supporter of their plans.

This confluence of institutional practices and historical events made the expansion of public health services in Ontario possible. It was a time of relative prosperity in Ontario; all of these developments took place within the placid stability of continuing Conservative majority governments. The cabinet was a so-called departmentalized cabinet, in which decisions were largely left to individual ministers to make in conjunction with their departmental officials; their decisions were then taken to cabinet for ratification. In the Department of Health, the senior officials were public health physicians, as were many of their subordinates. Their views on public health were profoundly similar in part because of the common bonds of their professional education. Department officials had a growing group of allies among the full-time public health physicians working as MOHs in city health departments and county health units, whose work situations were different but who shared common professional bonds. The factors that facilitated the expansion of public health services in the twenty years after WWII, however we view them, are unlikely to come together again: prosperity, stable one-party rule, a departmentalized cabinet that gave individual ministers substantial autonomy, public health professionals as senior departmental officials and participants in strong provincial-local professional bonds, and the freedom to institutionalise a provincial conditional grant that officials tailored to suit the needs they saw.

While the Conservatives continued to win elections until 1985, changes began to occur for public health as early as 1966. Gordon

Brown, a public health physician was replaced as deputy minister by Dr Kenneth Charron, a physician but not a public health specialist. One of Charron's first acts was to restructure the department into four main divisions, one of which was the Public Health Division. He appointed Dr Gordon Martin, formerly MOH for North York, to be its Executive Director, reporting directly to the deputy minister. The Public Health Division was divided into four branches, one of which was the Local Health Services Branch, responsible for liaison with local health authorities through a central office and a group of regional medical officers.

In 1967 Dr Mathew Dymond, still Minister of Health, introduced major changes to the *Public Health Act*: first, to require that all municipalities participate in a county or district health unit so that the whole population of Ontario would have access to full-time public health services; second, to provide for the creation of larger health units to be known as district health units, funded by the province at 75%; and third, to regularise the system of grants to municipalities participating in county health units and give grants for the first time to city health units. Full-time public health services throughout the province were required by law for three reasons: first, with universal accessibility to hospital services, there would certainly be criticism if the population did not have comparable access to public health services; second, for public health reasons, the department wanted to be able to reach all MOHs in the event of a crisis, such as polio epidemics (which had occurred in 1955, 1958, and 1960) and not spend hours or days trying to reach the last 50 or 60 part-time MOHs. The third reason was that after twenty years of voluntary compliance, sweetened by provincial funding, not very many citizens or local boards or part-time MOHs were left to protest against the legislation. In 1966, 248 local boards of health employed 155 part-time MOHs; by 1967, only 90 local boards were left.

The new district health unit (DHU) policy was well received, because all participants in the unit received the 75% grant. By June 1969, 20 of the 29 district health units proposed by the Department of Health were underway, and a further 4 (which each formed the nucleus of a DHU even if not all the partners had yet joined) were also receiving the 75% grant. The case of Hamilton illustrates the power of the grant. In 1966, the city of Hamilton spent $954,000 on public health. In 1968, Hamilton was a founding member of the Hamilton-Wentworth

District Health Unit. By 1972, the DHU's budget was $2.3million; the province paid 75% and the balance was shared between the City of Hamilton and the County of Wentworth. Hamilton's share was $350,000, just under 37% of what it had spent on public health in 1966.

Other grant systems were revised at the same time, so that any municipality participating in a county health unit received a 50% grant (where before counties, towns, villages, and townships got 50%, but cities got between 33% to 15% depending on their population). City health departments, which until 1967 had received no provincial operating grants at all, received 25% grants.

To some extent, the departmental reorganisation, the creation of district health units and the regularization of grants were all part of an attempt to set the department house in order for the major changes that were to come, and in which public health physicians would play a very little part.

To summarise the major changes, it is possible to identify three separate elements, though in practice they were not separable. One immense change that occurred at the highest executive level in Ontario, in most other provincial governments, and certainly at the federal level, was an effort to restructure decision-making, to reduce departmentalised decision-making and to empower the cabinet as a whole to bring their collective view to the main issues of the day. The goal of cabinet restructuring was to emphasize "shared knowledge, collegial decision-making, and the formulation of government-wide priorities and objectives"(Dupré 1985, p.4).

A mechanism to accomplish this, apart from creating new structures like Management Board, was to take a new approach to the careers of minister and deputy ministers. Rather than having them serve for long periods in a single post, there was thought to be advantage in moving them from department to department, to give them a broader sense of government concerns and to eliminate any vestiges of departmental loyalty that had been a strong tradition in the 1950s and early 1960s. The contrast between stability and frequent change is evident in health. If we take the period from the appointment of John Phair as deputy minister in February 1945 as the starting point, and take the end point as the February cabinet shuffle before the May1985 defeat of the Conservative government, we have a 41-year period to deal with. The changeover point from departmentalized decision-making

to frequent turnover begins with the appointment of Tom Wells as Minister of Health in August 1969, replacing Dr Mathew Dymond who had served as Minister of Health for eleven years. The stable period, therefore, runs from February 1945 to August 1969, and the period of frequent changes from August 1969 to February 1985.

Table 1. Average Length of Service of Ministers and Deputy Ministers of Health in Ontario in Stable and Turnover Periods

	Number of Ministers	Average Length of Service in in months	Number of Deputy Ministers	Average Length of Service in months
Stable Period February 1945 – August 1969	3	86	2	125
Turnover Period August 1969 – February 1985	7	26.4	6	31.3

This table clearly illustrates that the stable period is characterised by a small number of long-serving ministers and deputies; this in part was what enabled them to develop their close ties and common positions on key questions in public health. By contrast, the turnover period shows twice as many ministers and deputies, and that they stayed in their posts for comparatively short times. A further complication is that ministers and deputies did not rotate together. For example, Dr Charron, who was deputy to Dr Gordon Brown, the last deputy in the stable period, was also deputy to the first three ministers of the changeover period. Some ministers, like Dennis Timbrell, who was Minister of Health for five years, had three deputy ministers during that period. In a similar way, some deputies, like Graham Scott, who was a deputy minister in Health for only a little over two years, served three Ministers of Health during his brief tenure.

Closely related to collective decision-making and government-wide priorities was a concern for expenditure. Initially, the question was not as much about deficits and uncontrollable spending as it was that the government had inadequate knowledge of and control over the funds it was spending. Of course the central concern was hospital costs, because they were so large, and Ontario also operated its own system

of psychiatric hospitals, where costs were rising and hospitals practices came under increasing scrutiny, especially for the children housed at the enormous Orillia hospital. This concern about spending was both a concern about expenditure itself: was the money being spent wisely, in a timely manner, according to proper procedures? It was also a concern about the accountability aspect of the process: were spenders documenting their expenditures, were they able to account for all the money spent, did they understand the requirements of an audit?

In order to address these concerns, the Department of Health was urged to bring in financial managers to help them put appropriate systems in place, and throughout the government there was an interest in bringing in staff with greater managerial expertise, because the Government of Ontario was committed to "the fullest adoption and implementation of the new management style" (Szablowski 1975, 116). Indicative of this trend was that physicians were no longer appointed as deputy ministers in the Department of Health. When Dr Ken Charron left his position as deputy minister, he was replaced by Stanley Martin, a non-physician but former head of the Ontario Hospital Services Commission with strong managerial and financial expertise. Since 1972, qualifications for deputy ministers in the Ministry of Health have not included medical training.

Another major change required by the new system was the shift from department to ministries, which involved not simply a change in title, but meant bringing together directly under the Minister, not only the department but also all the agencies, commissions and other structures for which s/he might be responsible. In the case of the Ministry of Health, this meant bringing together the existing Department and the Ontario Hospital Services Commission as well as the Ontario Hospital Insurance Registration Board. A necessary part of this process was to re-examine the groupings of responsibilities within each ministry, epecially given the changing social thinking about developmentally delayed (referred to as mentally retarded) children. Responsibility for them was taken from the Ministry of Health and transferred to the Ministry of Community and Social Services.

Other programs were transferred out the Ministry of Health, partly because it was believed that relatively self-contained programs might get more attention in a Ministry not dominated by a mammoth program like health insurance. The very first programs were transferred in 1957 when the Ontario Water Resources Commission

was set up, but the pace of transfer was rapid during the period 1969 to 1981, in response not only to government reorganisation but to intense partisan conflict over pollution and the environment. In 1969 responsibility for air pollution control and waste management was transferred to the Department of Energy and Resources Management (which became the Department — later Ministry — of the Environment). In 1971, responsibility for pesticides control and private sewage disposal systems was also transferred to the new Ministry of the Environment. The major change in occupational health, following the report of the Ham Commission on the Health and Safety of Mine Workers, was that the entire Occupational Health Protection Branch was transferred from the Ministry of Health to the Occupational Health and Safety Division of the Ministry of Labour in December 1976. Finally, in 1981, the Public Health Engineering staff was transferred from the Public Health Branch of the Ministry of Health to the Ministry of the Environment.

During the same period, 1973 to 1982, the Ministry of Health was reorganised five times, partly reflected its own search to decide what its role would be, and partly reflecting the efforts of a deputy or a minister to put its own stamp on the department. Several observations can be made. First, public health fell through the hierarchy of the organisational chart: from being one of four divisions in 1966, with an executive director reporting to the deputy minister, to a single branch in 1982, reporting to an assistant deputy minister. No one above the branch director had qualifications in public health. Second, members of the public health staff were under considerable pressure about their own positions because of the frequent reorganisations and because their usefulness to MOHs in the field had declined sharply.

The provincial-local public health relationship, founded on common professional bonds and a common commitment to public health, came under great strain beginning in 1969. With rising government costs because of 75% DHU grants, the Minister, Dr Mathew Dymond, was under pressure from cabinet to control costs, and wrote to local boards of health that controlling MOHs' salaries was probably the best starting point. Salaries were the costliest item and the MOH had the highest salary. Dr Gordon Martin, Executive Director of the Public Health Division, also wrote to local boards of health several months later and proposed using the salary of a Grade VI physician in the Department of Health as the basis of the MOH salary. The response of the MOHs to the "Martin letter" was immediate and outraged. Not

only did they believe that being tied to a Grade VI physician would give them too low a salary and make MOH recruitment more difficult, they were enormously offended that Martin had written directly to local boards of health and not to them. They regarded his action as a "gratuitous insult" and concluded that Martin's conduct "was unilateral and secretive and lacked courtesy" (Society of Medical Officers of Health of Ontario 1969). Local MOHs met immediately following the Martin letter, and formed their own association, the Society of Medical Officers of Health of Ontario (SMOHO) to deal with the provincial department, because they regarded the department as no longer trustworthy. Other professional groups followed the MOHs' lead and by 1975 a total of 10 professional public health associations had been formed in Ontario to safeguard their interests in dealing with the province.

But other problems also undermined the provincial-local relationship in public health. At the provincial level, both reorganisations and program transfers left public health uncertain and less able to be the source of all advice for MOHs that they had once been. For the first time, they had different interests and neither level really cared about the other; for example, reorganisations meant little to MOHs, but were a source of uncertainty and confusion for provincial staff. On the other hand, when Bert Lawrence was Minister of Health for a brief period, he suggested making local public health a provincial responsibility and doing away with local boards of health. This was not a crisis for provincial officials, but was very disconcerting for MOHs who were the executive officers of the local boards. When Lawrence left, he was followed as Minister by Dr Richard Potter, who gave an interview stating that "boards of health and their medical officers in Ontario will be phased out ... because better use can be made of the money involved" (*Toronto Sta*r 13 May 1972, p. 10).

When departmental public health officials were not coping with new deputies and new ministers, and when local MOHs were not trying to educate ministers about the need for public health, the provincial-local relationship continued to decline. The main reason was that with program transfer, the expertise that MOHs needed concerning water pollution, waste water, clean water, sewage treatment, private sewage systems, health and safety conditions, industrial accidents, and industrial pollutants was no longer available from the Ministry of Health. Nor was the Ministry in a position to function as an

interministerial moderator, despite some early attempts to set up tri-ministerial agreements among Health, Environment, and Labour. Even when the MOHs sat in interministerial committees, there seemed little concern for health issues. They pointed to the example of an interministerial committee dealing with, " ... terminal disinfection of effluent from water pollution control plants". In SMOHO's view, it appeared "... that concerns related to the health of fish, cost of chorine, potential for chlorine damage to the environment, have been of equal or greater concern than the potential health consequences of various enterically transmissible human illnesses" (Society of Medical Officers of Health of Ontario 1981–84). In the end, the MOHs themselves were left trying to get information out of other ministries and they often did not succeed. This meant that their capacity to safeguard the public health was reduced because of program transfer and because of the unwillingness of other ministries to recognise the importance of the MOH's job and cooperate accordingly.

Provincial-local relations in public health also deteriorated as more stringent financial controls were put in place. In place of the informality of earlier years, local health authorities had to conform to province-wide reporting requirements. Both the substantive decisions about what constituted a "shareable expense" and the procedural issues about when expenditure guidelines would be issued, and when budgets would be approved, created problems between the two levels. By 1975 officials from the Ministry's Financial Controls Branch were sent out to help health units deal with budgeting and accounting; in a way, this made the problem more complex because these experts did not view the issues from a public health perspective. The clash of professional cultures led at time for MOHs to seek support from public health physicians in the Ministry, but they had little influence in the Ministry, and were caught in the middle between serving the new regime of cabinet-directed cost control and offering sympathy and moral support but nothing substantive to the local level.

It was in this atmosphere of centralised cabinet decision-making and stringent cost control that a new public health bill was introduced. A draft act had been in the works since 1968, but it had never made it to first reading. A renewed process based on extensive consultation with the public health community began in July 1976; joint ministry-local health committees were established to prepared the principles of

the bill; the Ministry commissioned several studies of public health; the existing legislation was reviewed; committees were set up to prepare the core programs; meetings were held in nine cities to discuss the core programs; a discussion paper was published; the assistant deputy minister toured the province to discuss the act; a conference was held in 1981 to get feedback from the local public health community; and the draft act was circulated for comment. By the time the bill was introduced in July 1982, it was widely understood in the public health community; it was also referred to the Standing Committee on Social Development for a detailed review, and the committee met 23 times to discuss the bill.

The 1983 Health Protection and Promotion Act required that all local health authorities provide core public health services in seven areas: community sanitation, control of communicable diseases, preventive dentistry, family health, home care services, nutrition, and public-health education. For the public health professionals, the requirements were not an issue, either because they had already met them or because the new act would compel their board or municipality to provide what had been lacking. Among local boards of health, there was more concern, especially about where the money would come from to pay for these services. Some argued that if the services were required by the province, then the province should pay. However, the grant system remained the same; 75% for DHUs, which now included all units outside of Toronto, and 40% for the Toronto municipalities. The act also maintained the MOH as chief executive of the local board, and required that s/he inform him/herself about occupational and environmental health matters.

As funding had become more difficult, the issue of core programs became more controversial. Not only were core programs to be offered but the province specified exactly what was required for a service to be acceptable in a series of documents "Mandatory Health Program and Services Guidelines;" in a half-in-jest way, these were sometimes regarded by the province as mandatory, and by local authorities as guidelines. One of the enduring issues was and remains the problem that some local authorities cannot fill key staff positions, including MOH, while in other cases the local board of health has actively resisted provincial pressure to hire a fully qualified MOH, preferring to have a local physician work part time for the authority. The uniformity, long sought in local public health, remains an elusive goal.

On the provincial side, public health physicians, who would once have been in constant contact with their local counterparts, had been replaced by legislative requirements and mandatory guidelines. There were too few of them to ensure that the requirements and guidelines were actually followed, but they tried to keep the public health flag flying even if it was deep in the shadows of Medicare.

The Harris government, elected in 1995, was determined to untangle the mess of provincial and local responsibilities, and to that end appointed a panel under the chairmanship of David Crombie called "Who Does What?" Crombie set up a number of sub-panels, including one on social services, which included public health. The sub-panels did not issue reports as such, but sent letters setting out their recommendations. The sub-panel letter on social services, co-signed by Crombie and Grant Hopcroft, recommended that the provincial government "fully fund all boards of health to deliver mandatory programs" (Crombie and Hopcroft 1996, p.2). Because Harris intended to take over education spending, he needed to find expenditures that he could download onto the municipalities so that the whole exercise would be "revenue neutral". It was presumably for this reason, and not anything having to do with public health itself, that the Harris government completely rejected the Crombie/Hopcroft recommendation and, instead, withdrew provincial funding from public health, leaving municipalities to bear 100% of the cost. Three points can be inferred from this decision: 1) public health was not very important and if it ended up as collateral damage, so be it; 2) the government had virtually no knowledge of the history of public health funding nor interest in it; and 3) whatever public health was, municipalities could handle it. It is not known if the government was aware of the importance of universality and access to public health, or whether the government recognized that some localities might not able to offer services because of different local funding capabilities.

Funding pressures continued to have a major impact on policy-making. Under the Rae government's cost containment strategy, the Ministry of the Environment started charging municipalities for water quality tests in 1993. In response, some local authorities took their lab work to the private sector. The Ministry of Health, which had thirteen labs compared to the MOE's four, did not charge any fees but notified municipalities that it would no longer be offering routine lab tests after September 1996. The Ministry of Health was concerned about the

quality of private labs that would be doing all the routine drinking water tests, because they were not subject to the same standards and inspections as clinical labs were. At the Ministry of the Environment, senior officials were considering privatizing water testing as a cost saving measure, but their more gradual timeline to make the transition was dramatically shortened when the Harris government was elected and wanted privatization implemented immediately. Among the consequences was that private labs were never instructed to notify the Ministry of the Environment or the Medical Officer of Health of adverse water quality reports; such notification had been automatic when testing had been done by government ministries. During the early Harris years, the Ministry of the Environment was severely cutback, and had too few field employees to undertake regular inspections of small waterworks, like the system of wells in Walkerton, where operator incompetence, lab failure to pass on adverse water quality reports, inadequate inspection and monitoring all played a part in the *E. coli* contamination in May, 2000 (O'Connor 2002).

The Walkerton crisis, and especially the O'Connor report's revelations about how many close calls there may have been elsewhere, highlighted the events of the 1990s, but the underlying pattern of fragmentation of responsibility at the provincial level began in the late 1960s. This fragmentation complicated the job of the MOH and made it more difficult for him or her to fulfil or even remain appropriately aware of health responsibilities that had been shifted, provincially, to other ministries.

By the time the SARS epidemic hit Toronto in the spring of 2003 several significant and unfortunate statements about public health should be obvious. First, public health at the provincial level was very weak, both institutionally and in the relatively few staff members left. They had been relegated to the status of poor cousin, given few resources to be of use to local health authorities, and were too often in the role of provincial enforcer than local advisor or supporter. Planning for an epidemic, investing in appropriate data systems, and building links with the federal and other provincial governments—all of which might have made a significant difference—had not taken place (Campbell 2004). But it is essential to recognise that weak institutions are not able to attract resources. Related to the question of resources but separate from it is that weak institutions are very rarely able to innovate.

Second, local health authorities were under severe financial pressure. While the 100% municipal responsibility for public health could not be sustained, the Harris government restored funding only to the 50% level, leaving the vast majority of health authorities dealing with only half of their pre-downloading provincial grant. It was impossible for many to maintain their programs and staff complement, and they were stretched and stressed by the financial pressures.

Although professional organisations of public health professionals operated throughout the province, they were focused on individual professions (public health nurses, public health dentists, medical officers of health) and did not constitute a network of local health authorities that would have been so helpful during the SARS epidemic. The province-wide groups that did exist, like the Association of Local Public Health Authorities representing board members, were not those required in the SARS epidemic. Coordination among health units, particularly those in contiguous areas like the Greater Toronto Area, had not been established.

Third, institutionally, hospitals have been separate from public health since their alleged integration in 1930; they have few structural connections with public health, though the MOH's responsibility for infection control has sometimes provided a link in smaller centres. In some hospitals good relations exist between public health staff and hospital infection control experts, but this has tended to occur in response to local initiatives (e.g. the Capital Health Alliance in Ottawa) rather than any process of deliberate provincial planning.

Many other problems have been identified from downloading, Walkerton, and the SARS case, but the issues raised here relate to the longstanding dismal state of public health. Despite the serendipitous twenty years after the Second World War, the general pattern is one of weak and unsupported provincial institutions struggling to encourage weak and reluctant local institutions (pre-1940). The provincial institutions found it somewhat easier to encourage full-time institutions (post-1945) when financial inducements were offered. But the decline of public health within the Ministry of Health and the fragmentation of public health responsibilities at the provincial level meant a return to a weak and unsupported provincial role. As a consequence, MOHs were expected to coordinate locally what was fragmented provincially. While MOHs were prepared to take on this role, they had very limited success in doing so, because of cutbacks in

field staff at other ministries and a persistent view among those ministries that MOHs were 'children' of the Ministry of Health and should be dealing only with the Ministry of Health. The work of local health authorities, underfunded and understaffed, was made more difficult by provincial weakness.

References

Breton, A. and Wintrobe, R. 1982, *The Logic of Bureaucratic Conduct*, Cambridge University Press, Cambridge.

Brouse, W. H. 1875, 'Sanitary Legislation and the Ontario Medical Council', *The Sanitary Journal*, vol. 1 no. 8, 1 August.

Bryce, P. H. 1921, 'The Beginnings of Public Health in Ontario', 15 January; Archives of Ontario, RG 62, First Series A-1, vol. 45/46.

Campbell, A. 2004, *Interim Report of the SARS Commission.*

Crombie, D. and Hopcroft, G. 1996, Third Letter to Al Leach, Minister of Municipal Affairs and Housing from the Social Services Sup-Panel of Who Does What, 20 December.

Dupré, J. S. 1985, 'Reflections on the Workability of Executive Federalism', In *Intergovernmental Relations* vol. 63 of Research Studies of the Royal Commission on the Economic Union and Development Prospects for Canada.

Immergut, E. M. 1998, 'The Theoretical Core of the New Institutionalism', *Politics & Society,* vol. 26, no. 1, pp. 5–34.

Kloppenberg, J. T. 1995, 'Institutionalism, Rational Choice and Historical Analysis', *Polity* vol. 28, pp. 125–128.

McKinley, J. B. 1998, 'Paradigmatic obstacles to improving the health of populations: Implications for health policy', *Salud Pública de México*, vol. 40, no. 4, pp. 369–379.

O'Connor, D. 2002, *Report of the Walkerton Inquiry*, Part I.

Ontario, Department of Public Health 1934 *Annual Report*.

Ontario, Provincial Board of Health 1887 Sixth *Annual Repor*t.

Ontario, Provincial Board of Health 1897 Scrapbook; Archives of Ontario, RG 62, Series B-4, Vol 464, Item 109, 77, Envelope 4.

Ontario Standing Committee on Social Development 1982 'Bill 138, Health Protection Act,' 8 September, S–33.

Schindeler, F. F. 1969, *Responsible Government in Ontario*, Toronto.

Scott, R. W. 1987, 'The Adolescence of Institutional Theory', *Administrative Science Quarterly*, vol. 32, pp. 493–511.

Skrowronek, S. 1995, 'Order and Change', *Polity*, vol. 28, pp. 91–96.

Society of Medical Officers of Health of Ontario 1969, 'Minutes of Inaugural Meeting', 25 June.

Society of Medical Officers of Health of Ontario 1981–84, 'Minutes', untitled brief.

Steinmo, S, Thelen, K and Longstreth, F 1992, *Structuring Politics: Historical Institutionalism in Comparative Perspective*, Cambridge University Press, Cambridge.

Szablowski, G. 1975, 'Policy-Making and Cabinet: Recent Organizational Engineering at Queen's Park,' In D. C. MacDonald, *Government and Politics of Ontario*, Toronto.

Toronto Star (1972) 13 May.