FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Vivian Welch
AUTEUR DE LA THÈSE / AUTHOR OF THESIS

Ph.D. (Population Health)
GRADE / DEGREE

Institute of Population Health
FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

What is the Role of Systematic Reviews in Tackling Health Inequity?
TITRE DE LA THÈSE / TITLE OF THESIS

Peter Tugwell
DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

George Wells
CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

Kevin Brand

Elizabeth Kristjansson

Alan Shiell (University of Calgary)

Dawn Stacey

Gary W. Slater
Le Doyen de la Faculté des études supérieures et postdoctorales / Dean of the Faculty of Graduate and Postdoctoral Studies

[UNIVERSITY OF OTTAWA]

What is the role of systematic reviews in tackling health inequity?

Thesis for PhD in Population Health, University of Ottawa

Vivian A Welch
May 11, 2010

Supervisors: Peter Tugwell and George Wells

Library and Archives
Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 978-0-494-69116-8
Our file Notre référence
ISBN: 978-0-494-69116-8

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Table of Contents

Contents

Index of Figures:

Index of Tables:

Index of Appendices:

Contribution of authors

Paper 1:Contribution of authors:

Vivian Welch was responsible for the idea for this study, designed data collection tools and search strategies, screened studies, extracted data, analyzed data and wrote the first and subsequent versions of the paper. Peter Tugwell, Mark Petticrew, Betsy Kristjansson, George Wells provided comments on the protocol and study design, provided consultations regarding inclusion criteria and reviewed and approved the final version of this paper. Jessie McGowan reviewed the search strategy and provided comments on the paper. Joanne deMontigny, Maria Benkhalti, Erin Ueffing conducted screening and data extraction and provided comments on the final version of the paper.

Paper 2: contribution of authors:

Vivian Welch had the idea for this study, designed the data collection forms, extracted data, synthesized data in tables and wrote the first draft and final draft of this paper. Peter Tugwell, Mark Petticrew, Elizabeth Kristjansson, Janet Smylie, Kevin Brand, George Wells provided input on the protocol of this study and commented on the drafts of the paper. Elizabeth Kristjansson contributed statistical advice. Erin Ueffing, Bharbhoor Dhaliwal, Maria Benkhalti extracted data and verified double extraction, and provided comments on the final paper.

Paper 3: contribution of authors

Vivian Welch had the idea for this study, designed the interview question and sampling methods, conducted interviews and analyzed data and wrote the first draft and final draft of the paper. Janet Smylie and Elizabeth Kristjansson provided comments on the methods, and on previous drafts of this paper. Peter Tugwell and George Wells provided comments on the protocol for this study and the final manuscript. Bharbhoor Dhaliwal conducted interviews and analyzed data and provided comments on the final paper.

Paper 4: contribution of authors

Vivian Welch had the idea for this manuscript, developed the draft algorithm, recruited participants, analyzed answers and wrote the first and final versions of the manuscript. Peter

Tugwell contributed to the idea for the study, questions and framework as well as the final paper. George Wells contributed statistical advice. Janet Smylie contributed to the protocol for this study. Mark Petticrew contributed to the development of the questions and selection of systematic reviews for testing. Elizabeth Kristjansson and Kevin Brand contributed to the protocol for this study and provided comments on the final paper.

Acknowledgements

I am indebted to my children Skye, Nicholas, Noel and my husband, Jordi. This thesis could not have been completed without the support and encouragement of my supervisors, Peter and George, and thesis committee of Betsy, Janet and Kevin. I am grateful to Liz Lacasse for moral support and for helping with all the forms, signatures and copies throughout this thesis. Many thanks to Roseline Savage who provided ongoing support with the University regulations and forms. I am also grateful to my mother, father and sister, colleagues and friends who have supported me throughout this thesis. I am grateful to David Moher and Jennifer Tetzlaff for providing access to their cohort of systematic reviews from 2004. I am thankful to my professors and the doctoral program in Population Health at the University of Ottawa. I would like to thank the funders of this research: the University of Ottawa, the Canadian Institutes of Health Research, and the Government of Canada Lifelong Learning Program. I am grateful to Wendy Muckle for helping identify key informants and to the informants themselves for their participation. I am grateful to all the members of the Cochrane Collaboration who participated in the survey. I hesitated to list names here since I am sure to have forgotten some people without whom this thesis could not have been completed, so for all those people, I also give my heartfelt thanks.

List of abbreviations

CI: confidence intervals

CIHR: Canadian Institutes of Health Research

CINAHL: Cumulative Index to Nursing and Allied Health Literature

CSDH: Commission on Social Determinants of Health

EPPI-centre: Evidence for Policy and Practice Information and Coordinating Centre

EQUATOR network: Enhancing the Quality and Transparency of health Research

HIC: high income country

HIV: Human immunodeficiency virus

HTA: health technology assessment

JAMA: Journal of American Medical Association

LMIC: low and middle income country

MEDLINE: Medical Literature Analysis and Retrieval System Online

MeSH: Medical Subject Heading

NGO: non-governmental organization

NSAIDs: Non-steroidal anti-inflammatory drugs)

OR: odds ratio

PAIS International: Public affairs, public and social policies, international relations

PRECIS: **P**ragmatic-**E**xplanatory **C**ontinuum **I**ndicator **S**ummary)

PRESS: Peer Review of Electronic Search Strategies

PRISMA: Preferred reporting items for systematic reviews and meta-analyses

PROGRESS-Plus: Place of residence; Race/ethnicity/culture; Occupation; Gender; Religion; Education; Socioeconomic status; Social capital. The "Plus" includes additional factors across which inequalities in opportunity for health exists such as high income country (HIC) vs. low and middle income country (LMIC) setting, age, sexual orientation and disability.

PTSD: Post-traumatic stress disorder

RE-AIM (reach, effectiveness, adoption, implementation and maintenance)

SES: socioeconomic status

SR: systematic review

SUPPORT collaboration: SUPporting Policy Relevant reviews and Trials

WHO: World Health Organization

WHO-CHOICE: World Health Organization Choosing Interventions that are cost-effective

WHOSIS: World Health Organization Statistical Information Service

Abstract for thesis

Introduction:

Enhancing health equity remains of international political importance with endorsement from the World Health Assembly in 2009. The failure of systematic reviews to consider effects on health equity is described by decision-makers as a limitation to using systematic reviews as a basis for evidence-informed decisions. Hence, there is a need for guidance on the role of systematic reviews in assessing effects on health equity.

Methods:

Four studies were conducted to assess the role of systematic reviews in assessing effects on health equity. A Cochrane Collaboration methodology review and a methodology study assessed methods used in published systematic reviews to assess effects of interventions on health equity across ten categories defined by the acronym PROGRESS-Plus: Place of residence, Race/ethnicity, Occupation, Gender/sex, Religion, Education, Socioeconomic status, Social capital. Plus considers other factors associated with unequal opportunities for good health such as age, disability and developing country settings. A qualitative study assessed implementation factors that are associated with success of interventions in vulnerable populations and mapped these factors to the equity-effectiveness loop framework. An equity plausibility algorithm was developed and tested to predict the likelihood of effects of interventions on health equity.

Results:

Only 13% of published systematic reviews assess effects on health equity. Four methods were used to assess effects of interventions on health equity: 1) description of people in studies; 2) description of subgroup analyses; 3) analysis of differences; and 4) applicability assessment. Only 1 out of 20 methodological studies used an analytic method. Implementation factors that predict success of interventions on improving health of homeless people in Ottawa mapped well onto the equity effectiveness loop framework, suggesting this framework can be used to appraise and improve interventions to promote health equity. Testing of the equity plausibility algorithm developed based on these studies showed that 67% of respondents thought that differences in

relative effects of interventions were likely across sex and socioeconomic status, but there was little to no inter-rater agreement for these judgments.

Discussion

These studies show that systematic reviews lack consideration of effects of interventions on health equity. This dissertation makes recommendations to improve reporting and conduct of systematic reviews to improve the contribution of systematic reviews to the evidence-base on promoting health equity. Methodological research is needed to improve methods for assessing applicability of systematic reviews for populations across PROGRESS-Plus characteristics, by both those who conduct systematic reviews and those who use them as a basis for decision-making.

Introduction

This thesis aimed to develop and evaluate an equity plausibility algorithm to predict whether systematic reviews of interventions need to consider effects in disadvantaged populations. This thesis used mixed methods. Two quantitative studies were conducted to describe and compare assessment of effects on health equity in systematic reviews. A qualitative study was conducted to assess factors associated with success or failure of health and social care interventions in disadvantaged populations. Evidence from these three studies was used to develop an equity plausibility algorithm to predict whether interventions are likely to have effects on health equity.

Unfair and avoidable inequalities in health across socioeconomic, demographic and geographic factors have been termed *health inequities* [1]. Not only is the burden of illness and risk of most diseases higher for the disadvantaged, but effectiveness of interventions may be up to one third lower in disadvantaged populations due to a staircase effect of lower coverage/access, worse screening, poorer provider compliance and lower consumer adherence [2]. Health inequity exists within countries (e.g. between rich and poor); as well as between countries (e.g. between high income countries and low and middle income countries).

Systematic reviews represent an opportunity to identify what works to reduce health inequity [3]. Furthermore, there is increasing acceptance that systematic reviews are the best source of evidence of effectiveness as the basis for policy and practice since they reduce the chance of being misled, increase confidence in results, are an efficient use of time and are more easy to critically appraise and apply [4, 5]. However, systematic reviews rarely assess whether interventions have an impact on health equity. For example, only 1% of a random sample of Cochrane reviews considered potential differences in effectiveness of interventions across socioeconomic or demographic factors [6]. Furthermore, decision-makers have identified lack of evidence on health equity effects as a major barrier to using systematic reviews as an input to evidence-informed decision-making [7]. This information gap hinders the ability to make decisions about policies and practices to enhance health equity.

Systematic reviews can contribute to assessing effects on health equity in three ways: 1) Assess differences in relative effects in different populations defined by PROGRESS-Plus mediated through different mechanisms of action; 2) assess differences in absolute effects which are mediated by differences in baseline risk of outcomes in populations across PROGRESS-Plus; 3) assess differences in effects in the "real world" effects due to different access, diagnosis/screening, provider compliance or patient adherence.

Figure 1 shows these three ways that systematic reviews can contribute to our knowledge about effects of interventions on health equity. Firstly, differences in relative effects are uncommon. Two empirical studies of meta-analyses showed that differences in the relative risk ratio across differences in control group risk (as a proxy for differences in patient characteristics) occur in only 13-18% of published meta-analyses of health care interventions[8, 9]. In the example in Figure 1, increases in tobacco price are known to have a greater relative effect on poor people than rich people [10]. Secondly, since the absolute risk reduction is the product of the relative risk reduction and the baseline risk, greater absolute effects are expected for populations with a higher baseline risk of the outcome of interest. For example, people of low socioeconomic status have more comorbidities for cardiovascular disease such as obesity which increase their risk of stroke. Thus, lipid lowering drugs that seek to prevent stroke may have a larger absolute effect in low socioeconomic status populations (Figure 1) [11]. Thirdly, differences in effects may occur because of different uptake or utilization of interventions due to a staircase effect of factors which affect whether interventions reach the intended population (e.g. access, diagnosis, provider compliance and patient adherence[2, 12]). For example, the incidence of bicycle crashes with head injuries may be similar for both low and high socioeconomic status children. However, low socioeconomic status children are less likely to adopt helmets following a health promotion campaign thus reducing the real-world effectiveness of helmets[13] (Figure 1).

The Commission on Social Determinants of Health Public Health Knowledge Network proposed three upstream and two downstream pathways for promoting health equity [12]. The three upstream pathways are social stratification, differential exposure (e.g. to safe water) and differential vulnerability (e.g. poor people may be at higher risk due to clustering of other factors such as crowded housing and alcohol abuse) (see Figure 3). The two downstream factors are 1) reducing differential health outcomes, and 2) reducing differential consequence of health outcomes (see Figure 3). Both downstream pathways are mediated through action on health services characteristics which include availability, accessibility, acceptability, coverage due to patient adherence and provider compliance [12]. Interventions for both upstream and downstream pathways can be designed as universal or targeted programs. Universal programs are intended to reach the whole population from most disadvantaged to most advantaged, such as smoking bans [14]. Targeted programs are focused on disadvantaged populations, such as conditional cash transfers which pay poor people to attend school and regular health visits [15].

Systematic reviews and primary studies have the potential to assess effects on health equity by evaluating the effects of interventions focused on one of the five pathways above. However, both primary studies and systematic reviews are limited in their ability to detect differences in relative, absolute and real-world effects by whether their study samples include a diversity of participants across PROGRESS-Plus categories, and by whether they report the population characteristics. Primary studies and systematic reviews that do include different populations and describe their characteristics may be limited by their power to detect subgroup differences. Both systematic reviews and primary studies may also be limited by reporting or selection bias. For example, if effects of an intervention are negative in disadvantaged groups, they may be less likely to be reported. Primary studies may use restrictive inclusion criteria to maximize the signal to noise ratio and improve the likelihood of detecting differences [16]. Such inclusion criteria may result in a biased sample of relatively advantaged people who are most likely to benefit from the intervention of interest.

Purpose and objectives

The purpose of this thesis is to improve our understanding of the factors which influence whether upstream and downstream interventions have effects on health equity in one of three ways: 1) different relative effects; 2) different absolute effects; and 3) different real-world effects. Qualitative and quantitative methods were used to address the following four objectives:

- To review methods of assessing effects on health equity in systematic reviews of effectiveness;
- To describe the extent to which health equity effects are considered in systematic reviews, describe methods used, and assess the implications of their equity related findings for implications, practice and research.
- To identify factors associated with improved effectiveness of interventions in disadvantaged populations that affect health equity;

- To develop and validate an equity plausibility algorithm to predict the likelihood that an intervention will have different relative effects in disadvantaged populations

Research question

The overarching research question was: What factors predict whether an intervention assessed by a systematic review will have effects on health equity by differences in relative effects, absolute effects or real-world effectiveness?

This research question was answered by conducting four linked studies:

- A Cochrane methodology review to assess effects on health equity in groups of systematic reviews of interventions. A "*methodology review*" is a systematic review of methodology studies which study methods used in randomized trials, other health care evaluations or systematic reviews [17].
- A methodology study of a cohort of 300 systematic reviews to assess the extent to which published systematic reviews assess effects on health equity, how equity is assessed, and the influence of equity assessment on conclusions
- A qualitative study to assess implementation factors that influence the effectiveness of interventions in disadvantaged populations
- A health measurement scale development and validation study to assess the reliability and construct validity of an equity plausibility algorithm to predict whether interventions are likely to have different relative effects in disadvantaged populations, that will influence health equity

Rationale for the thesis

Decision-makers are increasingly under public pressure to consider the effects of programs and policies on health equity [3, 18]. Systematic reviews of effectiveness are a comprehensive analysis of all studies of a particular intervention, encompassing a range of populations, settings and implementation methods. As such, systematic reviews are increasingly being promoted as a comprehensive source of evidence for decision-making.

However, decision-makers are interested not only in what works, but also the costs involved in implementation, the potential risks or adverse effects and health equity, where health equity is defined as the distribution of health benefits across sociodemographic factors [19]. Lack of evidence for identifying health equity implications was highlighted by policy-makers as a major barrier to the use of systematic reviews as a basis for decision-making [7].

Systematic reviews rarely consider effects on health equity (only one out of 95 systematic reviews assessed differences in effectiveness across sociodemographic factors that could affect health equity), even if the data was available from primary studies [6]. Yet, unequal benefits or harms across different socioeconomic or demographic population groups could contribute to worsening health equity [20]. Systematic reviews can provide information relevant to health equity in different ways: 1) the gap approach to assess whether an intervention has differential effectiveness between most and least advantaged; 2) the gradient approach to assess differences in effects across the gradient of socioeconomic status in the population; or 3) Assessing the effectiveness of targeted interventions that aim to improve the health of disadvantaged populations (e.g. school-feeding for disadvantaged children). Systematic reviews can also be used to assess external validity; that is by assessing whether benefits, though studied in advantaged populations, can be extrapolated to disadvantaged populations and settings.

This thesis aimed to develop and validate an equity plausibility algorithm to assess the likelihood that an intervention assessed by a systematic review will have effects on health equity. This equity plausibility algorithm could be used by authors of systematic reviews to help decide whether to assess health equity effects within a systematic review; as well as by users of systematic reviews to make judgments about whether results of a systematic review can be extrapolated to disadvantaged populations and settings.

Literature review

Introduction

Health inequities are created and maintained through a complex interaction of different factors at multiple levels, illustrated by the conceptual model of the World Health Organization (WHO) Commission on Social Determinants of Health (CSDH) which shows the pathways which affect health equity including health system factors (quality, accessibility) as well as non health system factors, such as globalization and political climate (Figure 3) [21].

What is health equity?

Health equity has been defined by Margaret Whitehead as "differences in health which are not only unnecessary and avoidable but, in addition, are considered unfair and unjust"[22]. This definition has been chosen as the basis for this thesis since it is both widely accepted (e.g. by the Commission on Social Determinants of Health [23], the World Health Organization, the Global Health Equity Initiative [24], the Cochrane Health Equity Field [20] and the Cochrane Public Health Review Group [25]), and provides a framework of seven types of health differences which guides the judgment of the whether these health differences are unfair and unjust [22]. Whitehead proposes that health differences due to natural, biological variation, freely chosen actions and transient health advantage may be considered unavoidable and are unlikely to meet the criterion of unfair or unjust health differences. However, health differences due to behavior where choices are restricted (e.g. due to environment or social position), exposure to unhealthy living or working settings, inadequate access to essential services (e.g. health and social services), or consequences of ill health that causes sick people to suffer loss of income are likely to be judged as both avoidable and unfair differences in health.

Health equity refers to equity in health outcomes, not utilization or access to health care. Horizontal equity refers to equal health care for equal needs. Vertical equity refers to unequal health care for unequal needs. For example, greater resources may need to be devoted to people with severe, disabling disease than people with mild, chronic disease [26]. While horizontal and vertical equity influence equity in health outcomes, other factors such as social stratification, differential exposure and vulnerability are additional determinants of health inequalities.

How to promote health equity remains a subject of debate. The capabilities approach proposed by Amartya Sen states that all members of society deserve equal capabilities of being and achieving. Sen has argued that no list of minimum capabilities can be generated since capabilities depend on the context. Sen states that health equity includes fairness related to three main concepts: 1) capability to achieve good health and health achievement; 2) delivery of health care (e.g. non-discrimination) and 3) broader issues of social justice such as the distribution of social and financial resources and their implication on opportunity and ability to achieve good health [27]. In contrast, Martha Nussbaum has proposed a list of ten minimum capabilities, which include life and bodily health [28]. Another approach to promoting a just society, proposed by John Rawls, is that each member of society is entitled to a basket of “primary goods” which includes rights, liberties, income and wealth. Rawls explicitly excludes health from this basket of primary goods[29]. Despite these differences, the arguments of both Rawls and Sen have been used to suggest that ensuring health equity entails access to resources (primary goods) and opportunities (capabilities) needed for good health, particularly for the most disadvantaged[30].

How is health equity measured?

Assessment of health equity requires a judgment of "fairness" and avoidability. Because this judgment is difficult, Braveman suggested that health inequalities across sociodemographic categories across which most people agree disparities should not exist could be used as a proxy for health inequities [31]. The Commission on Social Determinants of Health proposed in their final report that any differences in health due to societal and structural conditions which determine where we live, work and play are avoidable and unjust [23]. For this thesis, the seven types of health differences suggested by Whitehead and whether health differences were considered causally linked to social structure have been used for judging whether differences in health are both avoidable and unfair.

Any measurement of health inequalities entails implicit value judgments [32]. To measure health inequalities between groups, one group must be chosen as the reference category, against which other groups are compared [33]. Inequalities in health status can be expressed in relative terms (e.g. as a percent of the reference group) or as an absolute difference. Conclusions based on relative and absolute measures of disparities can disagree [34],[32]. Assessment of the gradient from the lowest to highest category when there are more than two categories is needed to capture the effect on the middle-class, which is missed by only looking at the gap between lowest and highest (e.g. rich-poor gap).

Methods for measuring health inequalities have been summarized elsewhere [35], [36],[32]. The difference between two groups can be assessed with the rate ratio, rate difference, low-to-high ratio or shortfall (see Table 1 for examples of each). The difference between more than two groups, such as assessing the gradient across socioeconomic quintiles, can be assessed by the slope index of inequality, concentration index or index of dissimilarity. For each of the above-mentioned methods, a set of six issues need to be considered.

1) Measure of health status (e.g. whether to assess favourable or adverse events)
2) Population grouping (e.g. sex, ethnicity, education, income)

3) Reference group or norm to which other groups will be compared
4) Tradeoffs between absolute and relative measures
5) Weights attached to improvements in ill health (e.g. are improvements in the worst-off weighted more)
6) How to interpret inequalities in health across ordered categories

These issues and the consequences of different decisions are discussed by Anand et al. [36]. Decisions about what to measure and how to measure make a difference and reflect underlying value judgments[32].

What factors are associated with differences in effectiveness of interventions that affect health equity?

As acknowledged by the Commission on Social Determinants of Health, health equity exists across diverse factors, some of which interact with each other and with context and setting [23]. The acronym PROGRESS-Plus defines sociodemographic factors across which differences in effectiveness of interventions could be considered inequitable [37]. PROGRESS-Plus is an adaptation of an acronym defined by Tim Evans and Hilary Brown [38] to summarize the factors across which disadvantage might exist: Place of residence; Race/ethnicity/culture; Occupation; Gender; Religion; Education; Socioeconomic status; Social capital. The "Plus" was proposed to promote the inclusion of additional factors such as age, sexual orientation and disability, which may also result in reduced opportunity for equal care for equal need. For example, some have argued that discrimination against older people (ageism) results in inadequate treatment for depression in the elderly[39, 40]. Factors such as the political climate and health system characteristics are also implicated in the systematic patterns of health inequities across the socioeconomic gradient. PROGRESS-Plus has been adopted by the Cochrane Health Equity Field [41], the Cochrane Public Health Review Group [25], the UK-based EPPI-Centre [37] and the SUPPORT Collaboration [42]. PROGRESS-Plus has been adopted for this thesis to enhance consistency with these international groups dedicated to conducting and using systematic reviews for decisions related to enhancing health equity.

Social, cultural and biological factors may be associated with differences in effectiveness across PROGRESS-Plus factors. As proposed by Whitehead, differences which are associated with limited lifestyle choices due to societal structure or discrimination, exposure to unhealthy living or working conditions, inadequate access to essential services or differential consequences of ill health are most likely to meet the criteria for avoidability and unfairness. The process of implementation of interventions such as the degree and intensity of follow-up and compliance of providers may also mediate PROGRESS-Plus related differences in effectiveness. For example, pragmatic trials which allow greater flexibility in the selection and follow-up of patients as well as provider compliance and patient adherence are more able to assess differences in "real world" effectiveness in disadvantaged populations than explanatory trials. Pragmatic trials are defined as trials which assess effectiveness in real-world or usual clinical practice conditions; whereas explanatory trials are designed to assess efficacy in ideal, controlled conditions [43]. The PRECIS group (**P**ragmatic-**E**xplanatory **C**ontinuum **I**ndicator **S**ummary) identified 8 "spokes" across which pragmatic and explanatory trials differ: 1) Primary analysis; 2) Eligibility criteria; 3) Flexibility of the intervention; 4) Practitioner expertise; 5) Follow-up intensity; 6) Follow-up duration; 7) Participant compliance; 8) Practitioner adherence [44].

This thesis aims to identify the factors associated with health equity effects in systematic reviews, and use these factors to develop an equity plausibility algorithm to predict the likelihood of health equity effects for an intervention assessed by a systematic review.

How are effects on health equity measured? Targeting, gaps and gradients

Effectiveness of interventions may be lower for the disadvantaged, due to a staircase effect of worse coverage, diagnostic accuracy (e.g. delivery of health care), provider compliance (e.g. inverse care law) and consumer adherence (e.g. capability to achieve good health) [2]. Although these factors are most easily applied to clinical interventions, they can also be applied to public health interventions.

Targeted approaches to enhancing health equity involve interventions delivered only to disadvantaged populations. For example, two systematic reviews assessed all available evidence on targeted approaches to improving the impact of clinical health services for disadvantaged populations (e.g. low-income or ethnic minorities)[45, 46]. For example, intensive home visits were successful in promoting hypertension control in low-income populations [45]. Gap approaches focus on the differences between the worse-off and the most advantaged.

Assessing the effects on the gap and assessing effectiveness of targeted interventions fail to capture the effects of interventions on the gradient in health across levels of disadvantage [34]. A gradient approach recognizes health differences across the whole population and acknowledges that health inequities exist on a continuum along different axes such as socioeconomic status. A gradient approach is advocated by the Measurement and Evidence Knowledge Network of the WHO Commission on Social Determinants to address the causes of health inequities [47].

Why is equity important in context of systematic reviews?

Systematic reviews have tended to focus on average results, ignoring possible differences in effectiveness across socio-economically, geographically and demographically defined subpopulations, as well as factors such as differences in baseline risk as well as acceptability and affordability of the intervention that might affect effectiveness for these sub-populations [6, 48]. This lack of equity information in systematic reviews has been identified by decision-makers as a serious limitation to their use as a basis for evidence-based policy and practice [5, 7].

The methods for assessing how equity is considered in systematic reviews is assessed in this thesis by two types of study design: a methodology study and a methodology review (Figure 2Figure 1: Conceptual model for how systematic reviews can contribute to evidence-base on health equity

		Outcome			**Outcome**	
Low SES		**yes**	**no**	**High SES**	**yes**	**no**
	intervention	a	b	**intervention**	A2	B2
	control	c	d	**control**	C2	D2

Relative risk	**Low SES**	**High SES**
	(a/a+b)/(c/c+d)	(A2/A2+B2)/(C2/C2+D2)

Three ways that systematic reviews demonstrate effects on health equity, with three hypothetical examples:

1. **Difference in relative effects (e.g. greater effect of tobacco price increases on people with low SES)**

	Low SES	**High SES**
	Relative risk: 1.5	Relative risk: 1.1

2. **Same relative risk, difference in absolute effects (e.g. lipid lowering drugs may have greater effects in low SES if they have more comorbidities and a higher risk of stroke)**

Absolute risk reduction= relative risk * baseline event rate

	Low SES	**High SES**
Baseline 5 year risk of stroke:	5% of people (5 in 100)	1% of people (1 in 100)
Same relative risk (RR):	RR: 1.5	RR: 1.5
Absolute risk reduction:	7.5% (7.5 in 100)	1.5% (1.5 in 100)

3. **Differences in real-world effects due to access, diagnosis, provider compliance or patient adherence (e.g. bicycle helmets for preventing head injuries)**

	Low SES	**High SES**
Same relative risk in studies	RR: 1.5	RR: 1.5
Same baseline injury rate	5%	5%
Less uptake of helmets by low-income families	real-world RR: 1.0	RR:1.5

Figure 2). A methodology study of systematic reviews is an assessment of methods used in systematic reviews[17]. For example, Ogilvie and colleagues conducted a methodology study to assess whether six systematic reviews of community-based tobacco control interventions contained useful information on the effects of these interventions on reducing social inequalities in smoking [49]. A methodology review is a systematic review of methodology studies which study methods used in systematic reviews [17]. For example, a methodology review of the effect of including grey literature on the size of effect in meta-analyses of randomized trials included five methodology studies, each of which included 10 to 60 systematic reviews [50]. This thesis includes both a methodology study and a methodology review of assessing health equity in systematic reviews.

What guidance is available regarding how to assess health equity effects in systematic reviews?

Subgroup analysis is one method for assessing whether the effects of an intervention are different in different population groups defined by PROGRESS-Plus. Guidance on subgroup analysis in the Cochrane Handbook recommends that subgroup analysis in systematic reviews needs to be approached with caution to minimize spurious associations [51]. Seven criteria have been proposed for judging the believability of subgroup analyses: 1. clinically important difference? 2. statistically significant difference? 3. A priori hypothesis? 4. Subgroup analysis one of small number of hypotheses tested? 5. Difference suggested by comparisons within studies? 6. Difference consistent across studies? 7. Indirect evidence that supports hypothesized difference? [52] More recent discussion of the justification of subgroup analysis echoes these same principles, for both clinical trials [53] and meta-analyses [54].

Another way to assess the effects of interventions on health equity is to appraise the likelihood that interventions may have different effects in different populations defined by PROGRESS-Plus. Assessment of the likely effects in different populations can be assisted by checklists on applicability, transferability, generalizability, external validity and extrapolation (Table 2). There is some overlap in the literature with regard to the use of these terms. Generalizability has been defined as "*The extent to which a study's results provide a correct basis for generalisation beyond the setting of the study and the particular people studied*" by Glasziou [55].

Generalizability can be assessed across units (people and populations), treatments (variation in treatment delivery), occasions (timing of treatment) and settings (where interventions are delivered) [56]. The terms extrapolation and external validity have been used to refer to the same concept as generalizability, i.e. whether an interventions' effects can be used to predict effects in different populations from those studied [57]. Applicability has been used to describe individual particularization of treatment including an assessment of an individual's baseline risk of outcomes (both benefits and harms) and the likely impact of these on treatment effects for that individual (in absolute terms) or on the decision to treat or not to treat [53, 57, 58]. Applicability has also been used to refer to likely effects in subpopulations or communities, such as the Community Guide and the RE-AIM (reach, effectiveness, adoption, implementation and maintenance) frameworks which assess the representation of the target audience in the effectiveness studies [59, 60]. Transferability has been used to refer to the feasibility of implementing an intervention in a particular setting, context and population, considering things such as resources required, acceptability and relationship to other programs such as the EUROTHINE project which assessed transferability to the Netherlands setting[61] and Wang et al. [62]. The items in the checklists in table 2 focus on individual differences, treatment variations, timing of outcomes, and setting characteristics. All of these checklists were designed for use after a systematic review has been completed, not for predicting likelihood of differences in relative effects on health equity at the planning stage of a systematic review. If differences in relative effects are hypothesized at the outset of a systematic review, this has implications for the methods needed in the systematic review to assess and evaluate these differences.

Conceptual framework for health equity

Multiple frameworks have been proposed for population health and health equity, including the simplistic model of population health by Evans and Stoddart [63], the webs of causation models proposed by Krieger et al [64], and models which highlight individual and community-level factors [65]. This thesis used the conceptual framework described by Ted Schrecker and Ron Labonte, in their final report of the Globalization Knowledge Network to the WHO Commission on Social Determinants of Health (Figure 3)[21]. This framework was chosen because it is likely to gain wide acceptance following the publication of the final report of the World Health Organization

Commission on Social Determinants of Health [23], and also because it elucidates pathways for intervention. This framework provides an organizing framework for describing interventions which act on downstream factors such as health system characteristics, as well as upstream public health interventions (such as those designed to improve water quality and living conditions), and legal and policy interventions such as tax transfers and educational policy. This framework also encompasses the PROGRESS-plus factors with its explicit description of income, educational, gender, ethnicity and geospatial disparities as factors that lead to social exclusion, and which can be acted on by programs and policies to enhance capabilities. This framework is consistent with the Whitehead definition of health equity, which proposes that behaviours chosen in the context of restricted choices are avoidable and unfair.

This framework highlights the difference between determinants of health and determinants of health inequalities. Determinants of health are shown in this model as factors which influence health outcomes directly, including living conditions, employment conditions, food quality and water quality, as well as health system characteristics and consequences of illness. In contrast, determinants of health inequalities are those factors influenced by the societal structure and social stratification that cause differential vulnerability and exposure. Determinants of health inequalities have been called the "causes of the causes" [23], based on Geoffrey Rose's thesis that the causes of differences in individual health are different from the causes of differences between groups in health [66].

This conceptual framework suggests how systematic reviews can provide evidence related to reducing health inequity through the following four intervention points:

- 1) Increasing capabilities through education and income transfers (e.g. conditional cash transfers [15]),
- 2) Reducing differential exposure and vulnerability, (e.g. population level tobacco control [67]),
- 3) Reducing differential consequences of ill health (e.g. return to work interventions for breast cancer survivors [68]), and
- 4) Improving health system characteristics (e.g. systematic reviews relevant to human resources for health [69]).

These four intervention points and examples suggest that systematic reviews can answer questions about the effects on health inequity of clinical interventions related to health system characteristics, public health interventions and non-health interventions that aim to alter the social and built environment. The relative proportion of systematic reviews available for each of these intervention types is not known. In the methodology study described later in this thesis, only 9 out of 224 systematic reviews (4%) published in MEDLINE-indexed journals in November 2004 were health systems or public health interventions.

The Labonte and Schrecker conceptual framework was used in this thesis to inform the ways of assessing health equity effects (gap, gradient, targeting and judging applicability).

This thesis seeks to address two gaps in knowledge related to the role of systematic reviews in tackling health equity. First, this thesis seeks to identify the extent to which health equity effects are assessed in systematic reviews, including what methods are used and how health equity assessment informs the conclusions of the systematic reviews. Second, this thesis seeks to identify factors associated with effects on health equity in systematic reviews of effectiveness, and use these factors to develop an algorithm to predict whether interventions will have different effects on health of disadvantaged populations relative to more advantaged populations that are likely to influence health equity.

Paper 1: How are effects on health equity assessed in systematic reviews of interventions: a Cochrane Methodology Systematic Review

Authors: Vivian Welch, Peter Tugwell, Mark Petticrew, Betsy Kristjansson, Jessie McGowan, Joanne deMontigny, Maria Benkhalti, Erin Ueffing, George Wells

Word count: 3,684 words

Word count Abstract: 319 words

Contribution of authors:

Vivian Welch was responsible for the idea for this study, designed data collection tools and search strategies, screened studies, extracted data, analyzed data and wrote the first and subsequent versions of the paper. Peter Tugwell, Mark Petticrew, Betsy Kristjansson, George Wells provided comments on the protocol, provided consultations regarding inclusion criteria and reviewed and approved the final version of this paper. Jessie McGowan reviewed the search strategy and provided comments on the final paper. Joanne deMontigny, Maria Benkhalti, Erin Ueffing conducted screening and data extraction and provided comments on the final version of the paper.

Abstract

Background

Enhancing health equity has now achieved international political importance with endorsement from the World Health Assembly in 2009. The failure of systematic reviews to consider effects on health equity is cited by decision-makers as a limitation. Hence, there is a need for guidance on the advantages and disadvantages of how to assess effects on health equity in systematic reviews.

Objectives

To systematically review methods to assess effects on health equity in systematic reviews of effectiveness.

Study design

Cochrane methodology review of equity assessment in systematic reviews

Data sources

We searched the following databases up to Jan 2009: MEDLINE, PsychINFO, the Cochrane Methodology Register, CINAHL, Education Resources Information Center, Education Abstracts, Criminal Justice Abstracts, Index to Legal Periodicals, PAIS International, Social Services Abstracts, Sociological Abstracts, Digital Dissertations and the Health Technology Assessment Database, as well as using SCOPUS to identify citing articles of all included studies.

We included studies that assessed methods of measuring effects on health equity in a purposefully selected group of systematic reviews.

Study appraisal and synthesis methods

Risk of bias was appraised for included studies according to potential for bias in selection and detection, of systematic reviews using pre-determined questions. Data were extracted using a pre-tested form by two independent reviewers.

Results

Twenty methodological studies were included. Four methodological approaches to assessing effects on health equity were identified: 1) descriptive assessment of reporting and analysis in systematic reviews (all studies used a type of descriptive methods); 2) descriptive assessment of reporting and analysis in primary studies (4/20 studies); 3) analytic approaches (7/20 studies); and 4) applicability assessment (3/20 studies). Advantages and disadvantages of these methods are discussed.

Limitations

We did not assess the methods used or the availability of data in the individual systematic reviews included in these methodology studies.

Conclusions and implications

There is a need for methodological guidance, based on empirical evidence, on how to assess effects on health equity in systematic reviews. Subgroup analyses need to be justified and reported in sufficient detail that their credibility can be assessed by users.

Introduction

Description of the problem or issue

Health inequalities that are unfair and avoidable are classed as health inequities [22]. Health inequities persist, and are worsening, across almost all diseases, both within and between countries. For example, people living in the poorest countries have a life expectancy that is at least 40 years shorter than for people living in the richest countries. Within a single city (Nairobi, Kenya), the mortality rate of children younger than 5 years is 15 per 1000 in high-income areas and 254 per 1000 in the slums [70]. There is a moral obligation to address these health inequalities[23].

There is increasing acceptance that systematic reviews of the best available evidence are the most credible source of information on which to base policy and practice. This view was endorsed by a World Health Assembly resolution, which was based on the Mexico Ministerial Statement on Health Research [71]. A similar recommendation emerged from the "Role of Science in the Information Society" health conference[72] that was held as part of the World Summit of the Information Society in December 2003. The recommendation stressed the need for evidence to be delivered in a timely manner and in the right format. Systematic reviews are a useful basis for decision making because they reduce bias, increase the confidence in results, and are an efficient use of resources [73].

A recent study of policy maker perceptions found that policy makers increasingly consider systematic reviews as a useful source of knowledge to support decision making [74]. However, decision makers are interested not only in what works, but also the potential risks or adverse effects, costs in implementation and equity, defined as the distribution of benefit across sociodemographic factors [75]. The lack of evidence on the distribution of effects and impact on health equity has been highlighted by policy makers as a major barrier to the use of systematic reviews as a basis for decision making [76]. Unequal distribution of benefits or harms across different socioeconomic or demographic population groups could contribute to worsening health equity [77].

Health inequities have been defined as "differences in health which are not only unnecessary and avoidable but, in addition, are considered unfair and unjust" [78]. Assessing the effects of interventions on health equity is difficult because it requires a normative judgment on both the avoidability and the fairness of the distribution of effects [79]. Hence, assessments of the distribution of effects across potentially disadvantaged groups in both clinical trials and systematic reviews focus on health inequalities that can be measured [80, 81]. Health inequalities may exist across many dimensions; for example those defined by the acronym PROGRESS: Place of residence (urban/rural), Race/ethnicity, Occupation, Gender, Religion, Education, Socioeconomic status, and Social capital[41]. In addition, health inequalities may exist due to unequal opportunities across other factors such as age, disability, and sexual orientation, which has been termed PROGRESS-Plus [82].

Despite the demand for the assessment of the effects of interventions on equity by policy makers, these assessments are rare in systematic reviews. In one study only 1 out of 95 randomly sampled Cochrane Reviews assessed differences in effects across PROGRESS-Plus factors [6]. This was due to a lack of this information in the primary studies (only 10% reported differences in effects across PROGRESS-Plus factors), as well as a lack of assessment by the review authors[6].

Description of the methods being investigated

The different methods used to describe and assess health inequity in systematic reviews were investigated. Because health equity requires a normative judgment, this review focused on the assessment of health inequalities across PROGRESS-Plus factors. It also assessed whether these inequalities were described as unfair and unjust by the authors of the included studies.

There are a number of ways to measure health inequalities. For example, health inequalities can be expressed as the difference between the most and least advantaged groups in relative or absolute terms [35], or they can be expressed using indices such as the Gini index, concentration index [83], or benefit-incidence estimate [84]. The choice of method and comparator or reference group influences both the magnitude of the result and its interpretation [35]. See Table 3 for selected methods of assessing effects on health inequalities.

How these methods might work

Relative or absolute differences for health inequalities measured over time can demonstrate either an increase or decrease in health inequalities for the same data, because relative measures are affected by the underlying rate of the reference group. A detailed example of this can be found in Table C in Keppel 2005 [35]. Furthermore, economic measures of health inequalities, such as the Gini index, concentration index, and the benefit-incidence ratio, may be too complex to interpret and require too many data points to be useful in the context of systematic reviews[41]. This methodology review seeks to assess whether these methods have been used to assess health inequalities in empirical studies analyzing systematic reviews, and to explore the advantages and disadvantages of each method.

Why it is important to do this review

Despite the demand for health equity assessment in systematic reviews by policy makers and practitioners, there remains little empirical evidence on the advantages and disadvantages of the different methods available for assessing health inequalities in the context of systematic reviews of effectiveness.

Objectives

We aim to describe and assess the effects of using different methods to assess health inequalities in empirical research studies of systematic reviews of the effectiveness of interventions.

Methods

Criteria for considering studies for this review

Types of studies

We included empirical studies of a cohort (more than one) of systematic reviews that assessed the impact of interventions on health inequalities across one or more socioeconomic and demographic factors. An example of an eligible study is one that assessed the availability of information on health inequalities and the implications for policy in a cohort of all six Cochrane reviews of community-based tobacco control interventions [49].

As we aimed to assess methods for comparing health inequalities across different systematic reviews, we excluded individual systematic reviews assessing health inequalities. Furthermore, including individual systematic reviews might introduce bias because they are less likely to report health inequalities analyses when no substantive differences are found [85].

Types of data

We included data from published or unpublished empirical studies of a cohort of systematic reviews on the advantages and disadvantages of methods used to assess health inequalities. When possible, we examined the advantages and disadvantages of each of the methods used for assessing health inequalities in effects across sociodemographic factors defined by the acronym PROGRESS-Plus [82]. Within country inequalities were also considered by assessing differences in effects between high-income countries and low- or middle-income countries using the classification of the World Bank.

For the health inequalities to be judged inequitable, unfairness and avoidability (or remediability) need to be assessed. Therefore, we assessed whether the empirical studies of cohorts of systematic reviews include a judgment about the fairness and avoidability of health differences. If the studies make no judgment about health equity, we used the Whitehead criteria of avoidability and unfairness to make a judgment about whether health differences across these factors for the particular intervention and setting could be considered health inequities [78]. Judgments made using these criteria were documented, including whether sufficient information was available to make such a decision. For example, gender differences in incidence of osteoporosis or cervical cancer that are due to unavoidable underlying differences in biology would not meet the criteria for a health inequity [78]. We expected substantial heterogeneity in definitions of health inequity and inequalities. Therefore, we documented the variety of existing definitions to inform the development of universally accepted definitions.

Empirical studies of cohorts of systematic reviews were included if they focused on the following methods:

- Targeted approaches: evaluating effects (benefits or harms) in disadvantaged populations only.
- Gap approaches: evaluating differences in effects (benefits or harms) between the most and least advantaged groups (see Table 3).
- Gradient approaches: evaluating effects (benefits or harms) on the gradient from the most disadvantaged to the least disadvantaged groups (Table 3

Types of methods

We compared different methods used by these methods studies for assessing effects on health inequalities in terms of: the expertise required to implement the strategy; the availability of data; their usability; and whether and how judgments about health equity were made (e.g. judgments about fairness and avoidability of differences in benefits or harms). For each of these factors, we extracted descriptions by the authors of the methods studies. Data was not extracted from the systematic reviews included in the methods studies, nor the primary research included in those systematic reviews.

Types of outcome measures

Primary outcomes

Advantages and disadvantages of the methods used for assessing health inequalities, as described by the authors of the methodology studies. In the absence of authors' description, we judged the advantages and disadvantages from the perspective of the user of a systematic review. Data were not extracted from the systematic reviews included in the methods studies, nor the primary research on which these were based.

Whether the analyses of effects on health inequalities across PROGRESS-Plus factors meet the following criteria for credible subgroup analyses [86]:

- Clinically important difference.
- Statistically significant difference.
- A priori hypothesis.
- Subgroup analysis is one of a small number of hypotheses tested.
- Difference suggested by comparisons within (as opposed to between) primary studies of meta-analyses.
- Difference consistent across primary studies of meta-analyses.
- Indirect evidence that supports hypothesized difference.

Secondary outcomes

The following secondary outcomes were extracted from the methodology studies:

- Whether and how health inequity is defined and measured (e.g. whether proxy measures of disadvantage, such as nutritional status, are used).
- Information on the availability of data from primary trials or meta-analyses to conduct analyses across PROGRESS-Plus factors.
- What factors are associated with health inequalities (e.g. the types of primary studies included in the systematic reviews and implementation factors, such as the degree to which flexibility was allowed in the implementation).
- Implications for practice, policy, and research based on analysis of effects on health inequalities.

Search methods for identification of studies

The search strategy was developed by one author (VW) using a systematic scoping exercise to assess the effects of different MeSH terms and the use of limits on publication type (i.e. limited to meta-analyses or systematic reviews) and type of studies (i.e. intervention studies). The final search strategy does not include limitations on publication type as these were found to be too restrictive. An information scientist (JM) reviewed the search strategy, as recommended by the Peer Review of Electronic Search Strategies (PRESS) guidelines [87].

The search strategy was not limited by publication type or study design as there is no indexing term for studies that assess cohorts of systematic reviews. We included published and unpublished articles, as well as abstracts.

Electronic searches

We searched the following electronic databases:

- the Cochrane Methodology Register (to January 2009);
- MEDLINE (January 1950 to January 2009) using the Ovid interface;
- EMBASE (1980 to January 2009) using the Ovid interface;
- PsycINFO (1806 to January 2009) using the Ovid interface
- CINAHL (1998 to January 2009).

See

Appendix 1 for the MEDLINE search strategy. This search strategy was adapted for the other electronic databases.

To identify systematic reviews of social, legal, and educational interventions, we searched non-health literature databases using the Scholars Portal interface including:

- Education Resources Information Center (ERIC, 1965 to January 2009),
- Education Abstracts (1983 to January 2009),
- Criminal Justice Abstracts (1968 to January 2009),
- Index to Legal Periodicals (1994 to January 2009),
- PAIS International (public affairs, public and social policies, international relations - 1972 to January 2009),
- Social Services Abstracts (1979 to January 2009),
- Sociological Abstracts (1952 to January 2009), and
- Digital Dissertations (1997 to January 2009)
- Health Technology Assessment Database (available on the Cochrane Library) to January 2009.

Searching other resources

We also handsearched abstracts from the 2007, 2008 and 2009 Cochrane and Campbell Collaboration Colloquia if they were not yet indexed in the Cochrane Methodology Review Group Specialized Register.

We searched using SCOPUS on Sept 20, 2009 for any articles which cited the studies which met inclusion criteria for this systematic review. These articles identified by SCOPUS were screened for inclusion in this systematic review.

We searched the reference lists of included studies for other potentially relevant studies, and we contacted the authors of included studies to ask if they were aware of similar studies.

We also asked the editorial board members of the Cochrane Health Equity Field and Campbell Equity Methods Group whether they were aware of other potentially relevant studies.

Data collection and analysis

Selection of studies

Two reviewers (chosen from EU, JdM, MB, and VW) independently screened the titles and abstracts of all references retrieved by the search strategy to exclude those that are obviously irrelevant. The reviewers were not blinded to the authorship of the titles and abstracts because this is difficult to achieve and may not affect the screening process [88].

Potentially relevant articles were retrieved and screened independently by two reviewers (chosen from EU, JdM, MB, and VW) using an eligibility checklist. Disagreements were resolved by consensus in consultation with another reviewer (MP or PT). We documented all reasons for exclusion at both stages of screening for entry into a PRISMA (Preferred reporting items for systematic reviews and meta-analyses) flowchart [89].

Data extraction and management

Two reviewers (chosen from EU, JdM, MB, and VW) extracted data independently from the included methodology studies using a pre-tested data extraction form designed in an Excel spreadsheet, which was used to manage and summarize data. We compared the data extracted by both reviewers for each study. Disagreements were resolved by consensus. Another reviewer (MP or PT) mediated when consensus could not be reached.

Systematic reviews and primary research included in the methodology studies was not retrieved for further evaluation. Data was extracted only from the methodology studies.

We extracted data (see items in Appendix 2) on:

- how the sample of systematic reviews was selected;
- the characteristics of the systematic reviews (population, intervention, comparison, outcomes, study designs included, quality assessment, year of publication);
- characteristics of the interventions being studied (e.g. pharmacologic, health services);

- the method used to assess effects on health equity (how and whether equity is defined; which elements of PROGRESS-Plus were compared; whether other factors, such as the study design of primary studies, setting, or context, were assessed that might explain differences in effects across PROGRESS-Plus factors;
- how effects were compared (e.g. relative or absolute differences, or gradient approaches such as the Gini coefficient);
- the size of the difference in effects across different populations defined by PROGRESS-Plus.

We also assessed whether data on PROGRESS-Plus were available from the systematic reviews included in the studies.

Assessment of risk of bias in included studies

The Cochrane Handbook suggests assessing five categories of sources of systematic bias: 1) selection bias; 2) detection bias; 3) reporting bias; 4) attrition bias; and 5) performance bias [51]. We assessed selection bias (bias in how reviews are chosen) and detection bias (bias in how data is collected from reviews)[51]. We did not assess performance bias (how participants are exposed to intervention) or attrition bias (withdrawals) since these are related to exposure to an intervention which is not relevant for these methodology studies (Table 8.4a of the Cochrane Handbook [51]). We did not assess reporting bias (e.g. selective outcome reporting) since reporting bias in these studies is considered in the assessment of detection bias. For each of selection and detection bias, we assessed the transparency of the methods described by the authors and the potential for bias in the methods used to select the systematic reviews included in the cohort and in the methods used to extract and analyze data. In the context of empirical studies designed to assess health inequalities in cohorts of systematic reviews, selection and detection bias were defined as follows.

Selection bias: potential for bias in the selection of the systematic reviews to be included or excluded in the methodology study. We extracted details on the inclusion and exclusion criteria used to select systematic reviews for the methodology study.

Detection bias: potential for bias in the assessment of analytic methods and outcomes. We extracted information on how the details of the analysis of effects on health equity were extracted from the systematic reviews by the authors of the methodology study.

Measures of the effect of the methods

We conducted a comparative analysis of the methods used to assess effects on health inequalities by comparing the advantages and disadvantages of each of the methods. We extracted details on the availability of data from the systematic reviews and their included primary studies, as well as on the methods used to compare differences in disadvantaged populations to the overall pooled effect.

We also compared any subgroup analyses against the seven criteria for credible subgroup analyses described above [86]. We also assessed additional criteria for subgroup analyses of whether the statistical significance was tested with a subgroup-treatment interaction test and whether the trials stratified randomization across the subgroup factor of interest [54, 90].

Dealing with missing data

Authors of the included studies were contacted if insufficient information was available regarding sample generation, methods, and outcomes.

Assessment of heterogeneity

Results were not pooled. Results for each outcome (e.g. data availability, advantages, disadvantages, and credibility of subgroup analyses) were presented across each factor of PROGRESS-Plus for each included study.

Assessment of reporting biases

Reporting bias occurs when dissemination of research findings is influenced by the nature and direction of results [51]. We attempted to minimize the identification of only studies with positive results by using a comprehensive search strategy in diverse electronic databases, assessing relevant conference proceedings, reviewing citations, and contacting both the authors of eligible empirical studies and other experts. Positive studies, in the context of this review, include studies that are able to show statistically significant and substantive differences in effects across one or more PROGRESS-Plus categories.

Data synthesis

Results were synthesized in tables. Where data were available on subgroup analyses, we summarized the methods used to compare effects in different populations across PROGRESS-Plus categories. For continuous outcomes, we calculated mean differences and their 95% confidence intervals (CIs). We compared absolute differences in mean differences and differences expressed relative to the reference group. For dichotomous outcomes, we calculated the risk differences and risk ratios and their 95% CIs.

Subgroup analysis and investigation of heterogeneity

As this is a descriptive methodology review, the results were not pooled and subgroup analyses were not conducted.

Sensitivity analysis

As this is a descriptive methodology review, the results were not pooled and sensitivity analyses were not conducted.

Results

Study selection

6,164 potential articles were screened for inclusion (

Figure 4). Of these, 75 potentially eligible studies were retrieved in full text. Of these, 50 studies were excluded since they did not assess health inequalities in a cohort of systematic reviews or meta-analyses, five studies were excluded since they did not describe a focus on health equity [91-95]. These studies assessed health effects of interventions in special populations that could be classified as vulnerable across one or more PROGRESS-Plus factor (e.g. sexual offenders, elderly, children with chronic disease), but the authors of the study do not describe a focus on vulnerability or disadvantage (Table 4). One study [96] was a subset of another larger study [97].

Study characteristics

The 20 empirical studies of a cohort of systematic reviews were identified by electronic databases (n=16), searching SCOPUS for references to included studies (n=2) [69, 98], and contact with experts (n=3)[97, 99, 100] (Table 5). Two ongoing studies assessing equity aspects in health technology assessment reports were identified which may meet inclusion criteria (

Table 6). The methods used by these studies were: 1) Targeted approaches (n=9); 2) gap approaches (n=10) and gradient approach (n=1). Gender or sex was assessed in 14 out of 20 studies, socioeconomic status in 12 studies, race/ethnicity in six studies, age in six studies, LMIC in three studies, and two studies assessed all PROGRESS-Plus factors. The rationale for assessing effects on health inequalities in these studies was to better understand the mechanism of action of the intervention in five studies, to improve understanding of what works to reduce health inequalities in nine studies, to assess direct evidence on effectiveness in particular populations in seven studies, and to assess applicability and relevance of evidence for disadvantaged populations of settings in four studies. The number of meta-analyses or systematic reviews included in these studies ranged from 5 to 420 systematic reviews. Six out of 20 of these studies assessed cohorts of Cochrane Collaboration systematic reviews.

Definition of health equity

Equity was defined in three studies, as unfair and avoidable inequalities in health across socioeconomic strata [6, 100, 101].. None of the studies described making a judgment about the fairness or avoidability of differences in health. One study described using an "equity lens" [102] to assess whether systematic reviews could be used to answer questions about reducing health inequalities across SES, ethnicity or education. Three studies used the "SUPPORT equity checklist" [69, 103, 104] which assesses access to health care across LMIC, gender, age, ethnicity or SES (Appendix 3). The studies which assessed differences between men and women were inconsistent in their use of the terms gender and sex; for example, one study described assessment of differences between men and women as gender analysis [105], while another study described this assessment as gender and sex based analysis [106]. Four studies focused on assessing relevance of systematic reviews for decisions about health care in low and middle income countries (LMIC)[6, 69, 101, 107]. Two of these studies described differences in access to health care across geography and socioeconomic status in LMIC as inequitable [69, 103].

Risk of bias within studies

Risk of selection bias was low for all included studies since each empirical study of a cohort of systematic reviews used an explicit search method and inclusion criteria to identify relevant systematic reviews (Table 7). Detection bias was low for 11 of the included studies which reported explicit methods of data extraction, using forms and data verification. The other eight studies did not fully report methods for data extraction, and may be subject to a higher risk of missing relevant information.

Across studies, there is a low risk of selection bias since all of these empirical studies of cohorts of systematic reviews used a systematic search to identify studies that met predetermined criteria. Six out of 20 of these studies assessed cohorts of Cochrane systematic reviews which may be least likely to assess effects on health inequalities since they are most likely to assess efficacy questions where differences in effectiveness across PROGRESS-Plus factors are least likely to occur [101].

Methods identified to assess consideration of effects on health inequalities or health inequities

We identified four categories of methods used to assess whether systematic reviews considered effects of interventions on health equity: 1) descriptive assessment of systematic reviews; 2) descriptive assessment of primary studies included in the systematic reviews; 3) analytic approaches and 4) judgment of applicability to disadvantaged populations or settings (Table 8).

1) Descriptive assessment of systematic reviews

All 20 studies used at least one of the five descriptive approaches described below to assess whether their sample of SRs had considered effects of interventions on health equity.

1a) Mention of PROGRESS-Plus in introduction, objectives, discussion, implications

Only three methodological studies included in their objectives the assessment of explicit mention of PROGRESS-Plus in the introduction, objectives or discussion. This strategy provides information about whether SRs consider health equity in a broad sense, but provides no evidence on effects on health equity.

1b) Methods study assessed whether SRs describe populations in the primary studies across PROGRESS-Plus factors

Details on the populations included in the primary studies were available for 0% to 57% of SRs across PROGRESS-Plus factors (Table 8), with the most data available on sex distribution of the population (84/147 trials). The advantage of this approach is that information about the diversity of populations increases confidence in applying results across different populations and settings. The disadvantages are lack of data, and that description of populations does not assess differences in effects across these populations.

1c) Methods study assesses whether SR describes primary research as targeted at disadvantaged populations across PROGRESS-Plus

Ten methodology studies assessed whether systematic reviews described interventions as being evaluated in specific disadvantaged populations. Of these, six methodology studies selected SRs which focused only on disadvantaged populations (targeted). The disadvantaged populations targeted in these six methodology studies were elderly with mental health problems [39, 40], youth with disabilities [108], socially disadvantaged mothers [109], people in low and middle income countries [107] and low birth weight children [110]. These methodology studies described these five different populations as disadvantaged because of avoidable and unfair poorer health outcomes than other people due to lack of evidence, lack of guidelines or lack of resources to access and use preventive and curative interventions. Four methodology studies reported assessing whether the SRs described at least one study conducted in a disadvantaged population [6, 49, 101, 102]. While this descriptive method identifies whether interventions have been evaluated in disadvantaged populations, it does not assess the effects on health inequalities. Furthermore, it can be misleading since SRs with no studies in disadvantaged populations may still be relevant and applicable to disadvantaged populations.

1d) Methodology study assessed whether SRs have outcomes related to equity of access

Six methodology studies described whether SRs reported outcomes related to access to care or coverage of health services. Access to health care across disadvantaged groups (e.g. rural, low SES, LMIC, ethnicity) was reported in 18/173 SRs in these methodology studies. Access to health care is a determinant of both health and health inequalities. This strategy does not measure effects on health equity. Evidence on access to care may be affected by the eligibility criteria of the methodology studies. For example, one methodology study required that SRs contain information about access to care in LMICs by the focus of the review [103].

1e) Methodology study assesses whether SRs planned or conducted subgroup analyses across one or more PROGRESS-Plus factors

Nine methodology studies assessed whether subgroup analysis was conducted in groups of SRs. Outcomes were analyzed using subgroup analysis across one or more PROGRESS-Plus factor in only 22 out of 262 SRs assessed in these methodology studies (8%). For those that reported details of these subgroup analyses, subgroup differences were assessed across gender/sex (n=14), race/ethnicity (n=12) and socioeconomic status (n=1). Differences in effects across other factors of PROGRESS-Plus were not reported at the level of the SR in these methodology studies (LMIC, place of residence, occupation, religion, social capital). The advantage of this strategy is that subgroup analysis summarizes the data available in specific populations. However, these subgroup analyses are limited in their ability to detect differences due to statistical issues (e.g. post-hoc analyses, probability of finding a false association, lack of data in the primary studies, or lack of reporting stratified data in primary studies) [97]. Furthermore, subgroup analyses that were conducted were poorly reported (Table 9).

2) Descriptive assessment of primary studies included in the systematic reviews

2a) Methodology study assesses whether populations in primary studies are described according to PROGRESS-Plus:

Four methodology studies retrieved primary studies of included SRs to assess whether data was available from primary studies to conduct subgroup analyses in SRs. Population characteristics were reported in primary studies for sex most frequently (209/250 studies), followed by race, education, place of residence, socioeconomic status, occupation and social capital. This strategy has the advantage of assessing whether data is available in primary studies, thus assessing whether there is a risk of bias that PROGRESS-Plus characteristics are under-reported in systematic reviews [97, 101]. However, this strategy does not assess effects on health inequalities, and data may not be available from the primary studies stratified by PROGRESS-Plus characteristics.

2b) Methodology study assesses whether subgroup analyses conducted in primary studies:

Four of the methodology studies of systematic reviews assessed whether data was available from the primary studies on population characteristics across PROGRESS-Plus and whether outcomes were analyzed using subgroup analysis in the primary studies [6]. In the included primary studies, outcomes were reported separately for sex most commonly (from 13-36% of clinical trials), followed by SES (4 out of 103 trials in one study). Advantages of this approach are that more details are available regarding the methods of subgroup analyses by assessing information in the primary studies than in systematic reviews. Disadvantages of this approach are that it is time-consuming to locate and assess all primary studies [49, 97].

3) Analytic approaches

3a) Methodology study used regression analysis to assess association of PROGRESS-Plus factors with size of effect

Regression analysis was used by one methodology study of SRs on interventions to improve adherence [111]. Data was available for age (8 out of 12 SRs), sex (7 out of 12 SRs) and socioeconomic status (5 out of 12 SRs). Advantages of this approach are that it could be used to assess which PROGRESS-Plus factors are associated with effects on health equity and the dose-response of their effect. The disadvantage of this approach is that data may be unavailable (e.g. in this one study, one third of SRs lacked data to conduct this analysis).

3b) Methodology study compares effect size using an odds ratio, relative risk or risk difference between two groups across PROGRESS-Plus (e.g. men vs. women)

None of the 20 methodology studies reported a quantitative comparison of the difference between advantaged and disadvantaged populations or settings.

3c) Methodology studies assessed effects of interventions targeted at a specific population which is disadvantaged (e.g. older people with depression [39], youth with disabilities [109]).

Six methodology studies searched for systematic reviews of the effects of interventions targeted at populations which were described by the authors as disadvantaged by unequal opportunities for optimal health or high quality health care. These methodological studies identified a median of 11 SRs (range 5-23), and three studies reported clinically important and statistically significant effect sizes in these populations. The advantage of this approach is that evidence on effectiveness can be directly used to inform decisions about interventions aimed at specific disadvantaged populations (e.g. older people with depression) [39] and to identify gaps in the evidence-base. However, this approach may not be possible for some disadvantaged groups where systematic reviews or primary trials have not been conducted. Furthermore, this approach is limited by the methodological quality of the SRs and whether sufficient details about the process of implementation are reported to replicate the interventions. Also, the gap or gradient between these disadvantaged populations and others is not assessed, so the extent to which interventions generate health inequalities is not assessable [112].

4a) Methodology studies assess applicability to different populations across PROGRESS-Plus

Three methodology studies assessed the applicability and relevance of systematic reviews to improve health of people in LMIC [69, 103, 104]. These three methodology studies all used the SUPPORT Collaboration checklists (Table 8) for equity, applicability and scaling up to make judgments about whether the results from systematic reviews could be transferred to LMIC settings and could be expected to confer health benefits (details of SUPPORT checklists available in Appendix 3, and at: http://www.support-collaboration.org/summaries/methods.htm). Advantages of this approach are that it makes use of the best available evidence to make judgments that can be used to inform policy in LMIC. Disadvantages are that the judgment of applicability and equity are extremely challenging and requires content expertise, knowledge of LMIC settings and methodological knowledge [104]. Furthermore, this method does not assess the likely magnitude of effects and, since LMIC settings are extremely heterogeneous, the judgments required for this checklists need to be framed for specific settings.

Comparison against the "seven rules of when to believe a subgroup analysis"

For the nine methodology studies which reported subgroup analyses in SRs across a PROGRESS-Plus factor, we assessed whether these analyses met the Oxman and Guyatt seven credibility criteria of when to believe a subgroup analysis (Table 9) [86]. We also assessed two additional criteria suggested by Rothwell that subgroup analyses should be tested with a subgroup by treatment interaction and that randomization of trials should be stratified across the intended subgroup analyses [90]. The nine methodology studies provided insufficient data to assess seven out of nine criteria. Five studies provided a rationale to support the subgroup analyses, four studies described an *a priori* hypothesis, three studies reported statistical or clinically important differences, without details on the type of statistical test. None of these methodological studies described whether the differences assessed by subgroup analyses were due to differences in absolute effects (e.g. because of higher baseline risk in disadvantaged groups) or relative effects (e.g. because of different mechanisms of action).

Factors associated with differences in effects

None of the methodological studies described factors that might plausibly be associated with differences in effects across PROGRESS-Plus.

Discussion

Systematic reviews represent an opportunity for increasing the ability to detect subgroup differences because they include studies conducted in diverse settings and populations[113].. These systematic reviews can increase the confidence in their subgroup analyses by reporting the rationale and methods in sufficient detail [86, 90]. Measurement of effects on health inequalities is an active field of research, with over half of the included studies published in the last two years.

We identified four methods to assess effects on health equity in cohorts of systematic reviews: 1) describe populations in SRs; 2) describe populations in primary studies (e.g. randomized controlled trials or cohort studies); 3) analysis of different effects (benefit or harm); and 4) applicability assessment. However, the poor availability of data, both in primary studies and systematic reviews, for all of these approaches limits their usefulness.

The descriptive and analytic methods used in these methodology studies (described above) require data on outcomes stratified for specific populations across PROGRESS-Plus to assess effects in these populations. However, a lack of population-specific stratified outcome data does not mean that an intervention will not be effective in other populations (e.g. because primary studies have not been conducted in these populations or data has not been reported in the primary studies or the systematic reviews). For example, vaccination is expected to be effective in diverse populations, across a range of baseline risk and settings. For interventions tested in relatively advantaged populations, clinical epidemiology principles suggest that the relative risk reduction will remain the same across differences in baseline risk[114]. Thus, the absolute risk reduction is expected to be larger for populations with a higher baseline risk. For example, therapeutic drug monitoring was shown to be effective at improving adherence to antiretrovirals in clinical trials conducted exclusively in high-income countries. If the relative risk of 1.49 can be applied to low and middle income countries with higher HIV endemicity, a greater absolute effect may be achieved on population health [115].

None of these studies assessed what factors are associated with differences in effects on health equity. Identifying characteristics of interventions, population, comparison, setting, study design which are associated with effects on health equity could be used to inform *a priori* decisions to assess effects on health equity in systematic reviews and primary studies.

Descriptive and analytic approaches used by these methodology studies have the advantage of assessing whether an intervention has been tested in a specific disadvantaged population, which is appealing to practitioners and decision-makers deciding whether to implement an intervention in a specific population and setting. Analytic approaches have the advantage of providing an estimate of the magnitude of effect in either advantaged or disadvantaged populations, or both. However, we found few systematic reviews which conducted subgroup analyses, and none of them described the analyses in sufficient detail to assess the credibility of the findings, since they failed to report details on the seven Oxman and Guyatt credibility criteria [86]. Updated guidelines on subgroup analyses suggest also assessing four more items: 1) consideration of baseline characteristics; 2) independence of the subgroup effect (i.e. the subgroup effect is not confounded by association with another factor); 3) *a priori* specification of the direction of effect and 4) consistency across related outcomes [116].

None of the systematic reviews which reported effects on health inequalities described whether these different effects were due to differences in absolute or relative effects. Differences in absolute effects are expected in groups with a higher baseline risk of the outcome. For example, women from low and middle income countries have a higher rate of maternal mortality, and might achieve a larger benefit in absolute terms from interventions such as having a skilled attendant at the birth than women in high-income countries with a very low maternal mortality. Differences in relative effects suggest that the mechanism of action of an intervention is different. For example, the relative effect of increases in tobacco price is greater in low income populations [67].

Judgment of applicability of evidence to disadvantaged populations and settings makes use of available evidence to inform decisions. These methods have the potential to reduce needless replication of studies in different populations. However, applying these checklists is challenging and requires significant content, methodological and setting-specific expertise to judge whether there may be differences in absolute effects due to different prevalence of the condition or in relative effects due to differences in how the intervention is delivered or received (e.g. lack of follow-up could lead to more serious adverse events if early signs of toxicity are missed).

There is a lack of conceptual clarity regarding the definition of health equity. Only three out of 20 studies defined health equity explicitly. Use of the terms gender and sex in these studies conflicted with internationally accepted definitions, i.e. that sex refers to biological differences and gender refers to cultural and socially determined roles of males and females [117].

Six out of 20 studies involved collaboration of the Cochrane Health Equity Field[41], suggesting an increased interest in health equity within the Cochrane Collaboration. These studies analyzed cohorts of Cochrane reviews, which may be limited in their ability to detect subgroup differences since Cochrane reviews tend to contain fewer trials (median 8 studies) than other systematic reviews [118]. Furthermore, Cochrane reviews tend to assess efficacy questions where the effect size might be less likely to vary in different populations than for implementation questions which are more likely to be assessed by pragmatic trials [119]. None of the methodology studies assessed Campbell Collaboration systematic reviews, which focus on social, legal and educational interventions.

We used a rigorous and transparent process to identify and describe methods for assessing effects on health equity in systematic reviews, following up to date guidelines from the Cochrane Collaboration Handbook [120]. We used the PRISMA reporting guidelines to facilitate replicability[89]. There is a risk that we have missed some relevant studies since methodological studies of cohorts of systematic reviews are not well-indexed. We addressed this by using a comprehensive search strategy of both health and non-health databases, that imposed no limits on study design based on pilot-testing of the search strategy and review by a librarian scientist (JM) [87]. We also searched reference lists and used SCOPUS to identify citations of included studies. Four out of 20 of the included studies were published as abstracts [6, 106, 121] or reports [110] and one included study was identified by contact with experts [97]. Furthermore, two ongoing studies were identified as potentially eligible by contacting authors of included studies (

Table 6).

A limitation of this systematic review is that we did not include individual systematic reviews. We decided *a priori* that their inclusion could lead to bias since they may be less likely to report analyses of effects on health equity if none were found.

Another limitation of this review is that systematic reviews are dependent on the availability of data in primary studies. This systematic review did not assess whether data was available in primary studies nor the different biases which determine the representation and reporting of different populations and stratified analyses in primary research.

Authors Conclusions

Implication for systematic reviews and evaluations of health care

There is a need for improved reporting of subgroup analyses both in systematic reviews and primary studies to improve their credibility. These include the need for description of the rationale for subgroup analyses, assessment of clinical importance of subgroup differences, description of whether differences between groups are due to differences in absolute effects or relative effects.

Implications for methodological research

Methodological guidance, based on empirical data about the advantages and disadvantages of different approaches, is needed on how to assess effects on health equity in systematic reviews.

This systematic review identifies three areas for future research. Firstly, there is a need for methodological research to identify factors associated with differences in absolute and relative effects to improve our understanding of the rationale for exploring subgroup effects. Secondly, there is a need for methodological studies to assess the extent to which subgroup analyses can be used to assess intervention-generated inequalities. For example, individual patient data meta-analysis of individual level factors can be compared with study-level subgroup analyses to assess consistency of the findings across and within studies [122]. Thirdly, there is a need for methodological studies to assess differences in absolute and relative effects between advantaged and disadvantaged populations to provide guidance on the importance of clear reporting differences in both relative and absolute effects.

Funding

Vivian Welch received a Canadian Institutes of Health Research (CIHR) graduate scholarship for her doctoral studies. CIHR had no role in the decision to conduct or the conduct of this systematic review.

Paper 2: Do systematic reviews of effectiveness of interventions consider health equity?

Vivian Welch, Peter Tugwell, Mark Petticrew, Erin Ueffing, Bharbhoor Dhaliwal, Maria Benkhalti, Elizabeth Kristjansson, Janet Smylie, Kevin Brand, George Wells

Word count: 4104 words

Author contributions:

Vivian Welch had the idea for this study, designed the data collection forms, extracted data, synthesized data in tables and wrote the first draft and final draft of this paper. Peter Tugwell, Mark Petticrew, Elizabeth Kristjansson, Janet Smylie, Kevin Brand, George Wells provided input on the protocol of this study and commented on the drafts of the paper. Erin Ueffing, Bharbhoor Dhaliwal, Maria Benkhalti extracted data and verified double extraction, and provided comments on the final paper.

Acknowledgements: We are grateful to David Moher and Jennifer Tetzlaff for providing access to their cohort of systematic reviews.

Abstract: 244 words

Abstract

Background

Tackling health inequities both within and between countries remains high on the agenda of international organizations including the World Health Organization, as well as Regional, national and local governments. Systematic reviews can be a useful tool to assess effects on health equity because they include studies conducted in a variety of settings and populations.

Objective

This methodology study aims to describe the extent to which health equity effects are considered in systematic reviews, describe methods used, and assess the implications of their equity related findings for implications, practice and research.

Study design

Methodology study of equity assessment in systematic reviews

Methods

Two independent reviewers extracted information on the reporting and analysis of health equity effects in a cohort of 300 systematic reviews collected from all systematic reviews published in MEDLINE in November 2004, using a pre-tested data collection form. Any differences in data extraction were resolved by discussion.

Results

Of the 300 systematic reviews, 224 assessed the effectiveness of interventions on health outcomes. Of these 224 reviews, only 29 systematic reviews (13%) assessed effects of health, social and educational interventions on health inequalities. Of those which assessed effects on health inequalities, two thirds (18/29) used these differences in discussing applicability and implications for research, policy and practice.

Conclusion

Current methods for conducting and reporting systematic reviews do not take full advantage of the opportunity to assess effects on health equity. Development of methods for assessing effects on health inequalities is needed to improve the usefulness of systematic reviews for decisions about policy and practice related to reducing health inequities.

Introduction

Unfair and avoidable inequalities in health across socioeconomic, demographic and geographic factors have been defined as *health inequities* [78]. Not only is the burden of illness and risk of most diseases higher for vulnerable populations, but effectiveness of interventions may be up to one third lower in vulnerable populations due to a staircase effect of lower coverage/access, worse screening, poorer provider compliance and lower consumer adherence [77]. Health inequity exists for most diseases both within countries (e.g. between rich and poor); as well as between countries (e.g. between high income countries and low and middle income countries). For example, between countries, under five childhood mortality is less than 6 per 1000 births in industrialized countries compared to 160 per 1000 in sub-Saharan Africa in 2006 [123]. Within low and middle income countries, under-five mortality is a median of two times higher in the poorest people compared to the highest wealth quintile (source: WHOSIS data 2009).

Decision-makers are increasingly under public pressure to consider the effects of programs and policies on health equity [3, 18]. For example, a qualitative study found that policy-makers stated that tailoring of evidence to specific populations would facilitate the use of evidence in policy-making [124]. There is increasing acceptance that systematic reviews can inform policy and practice since they reduce the chance of being misled, increase confidence in results, are an efficient use of time and are more easy to critically appraise and apply [4, 5]. Systematic reviews also represent an opportunity to identify what works to reduce health inequity [3, 125]. However, decision-makers have identified lack of evidence on health equity effects as a major barrier to using systematic reviews as an input to evidence-informed decision-making [73, 76].

Health services and therapeutic interventions can have an effect on reducing health inequities. For example, integrated management of childhood illnesses reduced stunting in children in Tanzania, and reduced inequity between poorest and least poor children in stunting [126]. In the USA, stepped care and intensive follow-up decreased the socioeconomic gradient in mortality from hypertension [127]. This potential for health services and therapeutic interventions to reduce health inequalities needs to be documented in systematic reviews.

Systematic reviews can assess effects on health equity in four ways: 1) the gap approach to assess whether an intervention has differential effectiveness between most and least advantaged; 2) the gradient approach to assess differences in effects across levels of disadvantage (e.g. socioeconomic status) in the population; 3) Assessing the effectiveness of targeted interventions that aim to improve the health of disadvantaged populations (e.g. school-feeding for disadvantaged children [128]); and 4) assessing applicability of results to vulnerable populations or settings [69, 103]. For example, although tamoxifen for breast cancer has only been assessed in high income country settings, the results almost certainly apply to women in low and middle income countries because it is unlikely there are differences in response due to biologic factors, adherence, compliance due to the consistency of effect across diverse country settings (mostly high-income), diverse populations studied in the trials, and large number of trials and women [129].

Despite the demand for evidence on distributional effects from intended users of systematic reviews and the potential of systematic reviews to provide this evidence, systematic reviews rarely assess whether interventions have an impact on health equity. For example, only 1% of a random sample of Cochrane reviews considered differences in effectiveness of interventions across socioeconomic or demographic factors [6]. Other similar studies of groups of systematic reviews have found that there is a lack of reporting of socioeconomic characteristics of people included in the studies and a lack of consideration of the differences in effects across populations[100, 101]. None of these studies assessed whether the lack of consideration of effects on health inequalities was appropriate and justified.

Differences in effects of therapeutic or preventive interventions across vulnerable population groups can contribute to worsening health equity [41]. However, repetition of clinical research in each vulnerable population without a strong empirical rationale represents a needless waste of resources that could be directed towards enhancing health equity. Without assessing the effects of interventions in different populations in systematic reviews, we risk wasting resources on needless replication of results in vulnerable populations.

Objective

This methodology study aims to describe the extent to which health equity effects are considered in systematic reviews, describe methods used, and assess the implications of their equity related findings for implications, practice and research.

Methods

Definition of health equity

Health equity is defined using Margaret Whitehead's definition; i.e. unfair and avoidable health differences are considered health inequities [78]. The acronym PROGRESS-Plus defines sociodemographic factors across which differences in effectiveness of interventions could be considered inequitable [37]. PROGRESS-Plus is an adaptation of an acronym defined by Tim Evans and Hilary Brown [38] to summarize the factors across which disadvantage might exist: Place of residence; Race/ethnicity/culture; Occupation; Gender; Religion; Education; Socioeconomic status; Social capital. The "Plus" was proposed to promote the inclusion of additional factors across which inequalities in opportunity for health exists such as high income country (HIC) vs. low and middle income country (LMIC) setting, age, sexual orientation and disability.

Differences across PROGRESS-Plus were considered health equity differences if they were classified as inequitable or unfair by the authors of the systematic review. If the authors of the review provided no judgment on fairness or avoidability, the reviewers conducted this assessment using the criteria described by Whitehead [78].

Selection of studies:

A cohort of all systematic reviews indexed on MEDLINE in the month of November 2004 was used for this study. This cross-sectional cohort was assembled by Moher and colleagues for the purpose of describing reporting characteristics such as types of questions being asked, number of primary studies included and quality of reporting. The cohort was assembled using Montori's empirical search terms for high sensitivity (>98%) in retrieval of systematic reviews[130]. The search was limited to English-language due to resource implications of including non-English articles. The search strategy used was: (1) 200411$.ed; (2) limit 1 to English (3) 2 and

(Cochrane database of systematic reviews.jn. or search.tw. or metaanalysis metaanalysis.pt. or medline.tw. or systematic review.tw. or ((metaanalysis.mp,pt. or review.pt. or search$.tw.) and methods.ab.)) [118]. This cohort contains an intentional oversampling of the Cochrane Database of Systematic reviews since the Cochrane Database of Systematic reviews is released quarterly. Articles were included as systematic reviews in this cohort if "the authors stated objective was to summarize evidence from multiple studies and the article described explicit methods" [118].

This cohort of systematic reviews was chosen since it represented a large random sample of published systematic reviews and the reports were available.

For the purpose of this methodology study on health equity, systematic reviews were included if the stated purpose was to assess the effects of an intervention on health outcomes. This criterion resulted in excluding 76 systematic reviews which did not assess the effectiveness of an intervention on health outcomes. For example, systematic reviews of the association of risk factors with health outcomes were excluded since no intervention was assessed. Also, systematic reviews assessed effects of interventions on non-health outcomes (e.g. literacy) were excluded because no health outcomes were assessed.

Data extraction:

Two reviewers (two of BD, EU, VW, MB) independently extracted data on reporting and analysis of differences in effectiveness across PROGRESS-Plus factors, using a pre-tested data extraction form (see Appendix 4 for items). This data extraction form included 51 items on factors likely to be associated with effects on health equity such as characteristics of the population, intervention, comparison (setting), outcomes and study design (i.e. randomized controlled trials, observational studies or both), whether health inequalities or health inequities were described and how they were defined. The data extractors also judged the likelihood that differences in absolute or relative effects might occur based on the mechanism of action of the intervention across social, biological, cultural or environmental and the description of possible effect modifiers by the authors of the systematic review. This judgment was based on information from the background and methods section on the type of intervention, target population and the authors' proposed theory or mechanism of action.

Subgroup analysis across PROGRESS-Plus was assessed, using the definition in the Cochrane Handbook: "Subgroup analyses involve splitting all the participant data into subgroups, often so as to make comparisons between them. Subgroup analyses may be done for subsets of participants (such as males and females), or for subsets of studies (such as different geographical locations)" [51].

The data extractors assessed whether the authors of the systematic reviews described applicability issues. Discussion of applicability issues was defined as discussing magnitude of relative or absolute effects, acceptability or feasibility in different populations or settings. The data extractors assessed whether the authors of the systematic review described implications for practice or research across PROGRESS-Plus in the methods and discussion sections. The data extraction form assessed whether checklists for applicability were used, such as the Cochrane Handbook [51].

In addition to these factors, the data extractors judged whether the question of the systematic reviews was pragmatic or explanatory. We used the definition proposed by the PRECIS group that pragmatic systematic reviews are designed to determine the effects of an intervention under the usual conditions in which it will be applied," [44]. Explanatory systematic reviews are designed to answer a question about the effects of an intervention under ideal conditions [44]. The systematic reviews were classified as "pragmatic" if the inclusion criteria of the systematic review allowed flexibility in the comparator, the patient eligibility criteria and the delivery of the intervention (in terms of practitioner compliance and patient adherence). These factors were selected based on the ten factors proposed for PRECIS: 1) eligibility criteria, 2) intervention flexibility, 3) practitioner expertise in delivering the intervention, 4) flexibility of comparison (e.g. placebo, wait list, do nothing, alternatives), 5) practitioner expertise in giving comparison, 6) intensity of follow-up, 7) type of primary outcome, 8) intensity of measuring participant adherence, 9) intensity of measuring practitioner compliance, and 10) scope of analysis of primary outcome [44]. The extractors based this judgment on the description of the objectives and eligibility criteria for the systematic review.

Extraction by two independent extractors was compared and differences were resolved by discussion.

Analysis: Odds ratios were computed to assess the association of methods of systematic reviews to whether authors of systematic reviews hypothesized differences in effects of interventions across PROGRESS-Plus. Odds ratios and their 95% confidence intervals were computed for all dichotomous outcomes. The following is an example of the calculation for whether the systematic review included only randomized controlled trials (RCT).

	Difference across PROGRESS-Plus hypothesized	
Study designs are RCTs only	*Yes*	*No*
Yes	28	149
No	16	31
Totals	44	180

Odds ratio (OR) = (28/149)/(16/31)=0.36

Lower 95% CI= OR* $e^{-[1.96*SE(\log OR)]}$

Upper 95% CI= OR* $e^{[1.96*SE(\log OR)]}$

SE (log OR)= $\sqrt{\frac{1}{28}+\frac{1}{149}+\frac{1}{16}+\frac{1}{31}}$

The methods used by systematic reviews to assess differences in effects across PROGRESS-Plus were classified as using one of four methods (gap, gradient, targeted or applicability assessment). Subgroup analyses conducted across PROGRESS-Plus factors were assessed according to the seven 'credibility' criteria for subgroup analysis, proposed by Oxman and Guyatt [86]. These credibility criteria are intended to minimize the over-interpretation of spurious differences: 1. clinically important difference? 2. statistically significant difference? 3. A priori hypothesis? 4. Subgroup analysis one of small number of hypotheses tested? 5. Difference suggested by comparisons within studies? 6. Difference consistent across studies? 7. Indirect evidence that supports hypothesized difference?

Results

Selection of reviews

Of the cohort of 300 systematic reviews, 224 were classified as assessing the effects of an intervention on health outcomes. Of the 76 systematic reviews that were excluded, 16 assessed test characteristics of diagnostic methods, 21 conducted systematic reviews of research methods (e.g. quality assessment) and 39 assessed the association of patients' characteristics with outcomes.

Definition of health inequalities

No systematic reviews used the term "health equity" or described differences in health as unfair and avoidable. Nineteen systematic reviews out of 224 (8%) described health inequalities across PROGRESS-Plus factors that were hypothesized to affect the effectiveness of interventions. Of these, seven (37%) described differences in baseline prevalence of the disease across one or more PROGRESS-Plus factor, 11 (58%) described one or more PROGRESS-Plus factors as effect modifiers and one systematic review described disease status (severe learning disabilities that lead to disruptive and culturally unacceptable behaviour) as a cause for social exclusion and discrimination [131].

Reporting PROGRESS-Plus in systematic reviews

Characteristics of the populations across PROGRESS-Plus included in the primary studies of systematic reviews were reported in 153 out of 224 systematic reviews (68%). In these 224 systematic reviews, gender (49%), age (47%), place of residence (22%), and developing country setting (9%) were most frequently reported (Table 10), followed by race or ethnicity (4%), socioeconomic status (3%), occupation (1%), education (1%) and social capital (1%).

PROGRESS-Plus characteristics were also described in the theory of the intervention (8%), analysis (12%) and in applicability and implications (21%).

Pragmatic vs. explanatory questions

Only 12 out of 224 systematic reviews were classified as asking a pragmatic question, designed to assess effects of an intervention under usual practice conditions.

Evaluation of differences in effects in vulnerable populations

Of 224 systematic reviews, 29 (13%) assessed effects of interventions in vulnerable populations defined across PROGRESS-Plus, using subgroup analysis (n=15) or targeted analysis of vulnerable populations (n=14) (Table 11). Nine of these 29 SRs described health inequalities which were hypothesized to influence the effectiveness of interventions. Of the 29 SRs which assessed effects in vulnerable populations, 24 (11% of the total cohort of systematic reviews) found differences in effects of interventions on health or social outcomes for vulnerable populations. Of these 24 SRs, 14 assessed targeted interventions, two described different effects due to acceptability issues, two found differences in relative effects and six found differences in absolute effects because of differences in baseline risk. Only one of the systematic reviews reported conducting a statistical analysis to compare effects in different populations [132]. These 29 systematic reviews were less likely to be Cochrane reviews (odds ratio 0.28, 95% confidence interval 0.12 to 0.65); only 31% (9/29) were Cochrane systematic reviews compared to 57% of the total sample of systematic reviews (129/224).

Five systematic reviews were judged as not assessing effects on health equity because they evaluated differences in effects across age that were judged as unavoidable due to developmental differences. For example, subgroup analysis to compare weight gain for children compared to adults for Risperidone was not considered inequitable because it is likely due to biological or developmental differences in response to Risperidone that are unavoidable.

Fifteen systematic reviews judged differences in applicability of interventions across one or more PROGRESS-Plus factor (Table 12). The reasons for hypothesized differences in applicability were differences in baseline risk (e.g. higher rate of epilepsy in lower social classes [133]), hypothesized different mechanisms of action (e.g. tamoxifen for hepatocellular carcinoma may work differently in men vs. women because of estrogen effects[134]), contextual differences which might affect effectiveness (e.g. fluoride mouth-rinses are more effective for people with lower exposure to fluoride from other sources[135]), and potential confounding of effects (e.g. effects of vitamin C for asthma may be confounded by smoking status which is associated with socioeconomic status[136]).

None of these systematic reviews reported using a published checklist to assess applicability of interventions in different populations across PROGRESS-Plus factors.

Influence of considering health equity on conclusions

For those 29 systematic reviews which evaluated effects in vulnerable populations, 18 (62%) described implications for policy, practice or research (Table 11). Of these, four out of 29 SRs (14%) concluded the intervention was effective in a vulnerable population (due to no differences in effects) and one of 29 SRs (3%) concluded the intervention was more effective in vulnerable populations due to higher baseline risk. Four out of 29 SRs (14%) concluded that different or additional interventions were required to avoid inequalities in health outcomes of these interventions due to differences in effects (n=2) or different patient preferences or feasibility (n=2). Seven out of 29 SRs (24%) concluded that more research was needed on the effects of the intervention in the vulnerable population of interest.

Factors associated with measured or hypothesized differences in effects across PROGRESS-Plus

Factors statistically significantly associated with measured or hypothesized differences in effects across PROGRESS-Plus were differences in the theorized mechanism of action (OR 16.5, 95%CI: 5.0-54.4), description of health inequalities (OR 9.3, 95%CI 3.4-25.3), acceptability across PROGRESS-Plus (OR 14.2, 95%CI 6.4-31.9), pragmatic question for the systematic review (OR 6.6, 95%CI 2.0-22.0), inclusion of RCTs only (OR 0.36, 95% CI 0.18-0.75), description of the population across PROGRESS-Plus (OR 2.4; 95%CI 1.1, 5.5), and discussion of implications across PROGRESS-Plus (OR 14.1, 95%CI: 5.9-33.9) (Table 13). The judgment of likelihood of equity differences, as judged by the data extractors based on the background and methods, was also significantly associated with hypothesized or measured differences by the authors of the systematic reviews (OR 7.8, 95% CI 3.8-16.2).

Methods used to assess differences in effects across PROGRESS-Plus

Three methods were used to assess effects of interventions on health outcomes in populations across PROGRESS-Plus factors: 1) Subgroup analysis; 2). Targeted approach; and 3) Judgment of applicability. Subgroup analyses were classified into two types: 1) pooled, and 2) description of within-study differences. Seven systematic reviews compared pooled results from more than one study using statistical methods. Eight systematic reviews described differences in effects across one or more PROGRESS-Plus factor within individual studies, without combining data.

1. Subgroup analysis (n=15 systematic reviews)

Pooled results (n=7 systematic reviews)

Seven systematic reviews assessed differences across population characteristics defined by one or more PROGRESS-Plus factor by combining results of studies with different characteristics or of subpopulations within studies with different characteristics (Table 11). These seven systematic reviews assessed differences in relative effects (n=2) and absolute effects (n=5) (Table 14). Only one of these seven systematic reviews found a statistically significant difference in effects; between studies of educational interventions which reported the gender ratio and those that did not [137]. According to the authors of the systematic review, this result may be related to methodological flaws rather than true differences in effect size between boys and girls[137]. Five of these systematic reviews found a difference in effects across one or more PROGRESS-Plus factor. The systematic reviews met a median of three out of the seven credibility criteria (range 2-5) for subgroup analyses (Table 14).

Only two systematic reviews suggested implications for policy, practice or research based on the differential effects, suggesting the difference was clinically important [132, 138]. One systematic review found that fluoride mouth-rinses in populations with low background exposure to fluoride were statistically significantly beneficial, resulting in a preventive fraction of 29%. In contrast, the effect was not statistically significant in studies of children with exposure to toothpaste or fluoridated water (0-6% preventive fraction). The authors conclude that fluoride mouth-rinses may be useful in vulnerable populations with low exposure to fluoride (e.g. lower socioeconomic status)[132]. Another systematic review found that the benefit of pneumococcal vaccine was statistically significant in randomized trials only in populations from LMIC with a high incidence of pneumococcus (RR 0.66, 95%CI: 0.57 to 0.77), but no effect was found in elderly people in industrialized countries (RR 1.03, 95%CI: 0.86-1.25). This systematic review concluded that properly conducted randomized trials are needed to support the use of conjugate vaccines to prevent all cause pneumonia[138].

b) Description of within-study differences without pooling (n=8)

Eight systematic reviews described differences in effects across PROGRESS-Plus characteristics observed within individual studies (Table 11). These systematic reviews described the effects within individual studies across gender (7), socioeconomic status (3), education (1), occupation (1), religion (1) and age (2).

Of these eight systematic reviews, two systematic reviews reported differences in effects across PROGRESS-Plus. One systematic review reported that four studies found a greater relative effect of clonidine on smoking cessation in women, possibly due to differences in the success rate of women in the placebo groups [139]. One systematic review stated that adverse events of moxifloxacin for respiratory tract infection were associated with female sex and age, implying a gradient of differences in relative effects, but the data were not shown in the publication [140].

2. Targeted approach (n=14 SRs)

Of 14 systematic reviews which assessed the effects of interventions aimed at vulnerable populations across one or more of PROGRESS-Plus (Table 11), nine were non-pharmacologic interventions such as stress reduction, linguistically appropriate services, exercise training and psychological therapy.

Different needs due to stigma, discrimination or cultural acceptability were described as the reasons for targeting by ten systematic reviews. For example, one systematic review assessed the relative efficacy of buprenorphine and methadone for opioid dependence, with emphasis on the acceptability and feasibility issues of alternate day dosing for heroine-addicted drug users [141]. Higher baseline prevalence of disease in specific vulnerable populations was described by four systematic reviews as the reason for targeting. For example, one systematic review assessed the effects of household water treatment and storage interventions on diarrhea in low and middle income countries where prevalence of diarrhea is higher than industrialized countries [142].

3. Judgment of applicability (n=15 SRs)

Fifteen systematic reviews concluded that the interventions might have different effects in disadvantaged populations across one or more PROGRESS-Plus factor because of context, confounding, differences in baseline risk of different mechanisms of action(Table 12). Methods used to judge applicability were not described by any of these systematic reviews, Reasons for different judgments were related to context (n=11), mechanism of action (n=2), confounding (n=2) or different baseline risk (n=2).

Discussion

Systematic reviews have the potential to inform policy and practice decisions related to health equity since they include studies conducted in different populations and settings. The assessment of applicability and implications for vulnerable groups in systematic reviews needs to be improved. Although applicability issues in vulnerable populations were discussed by eleven systematic reviews, none of these systematic reviews used a published checklist to appraise applicability. Systematic reviews in this cohort described implications for vulnerable populations for 62% of the systematic reviews that were hypothesized by the authors to have differential effects for vulnerable populations.

None of the systematic reviews in this cohort used the term "health equity" or described differences in effects as unfair. This suggests that researchers are cautious about making a subjective assessment about fairness.

This methodology study found that 13% of systematic reviews published in MEDLINE assessed effects of interventions in vulnerable populations using three methods: 1) subgroup analysis, 2) targeted approaches, and 3) judgment of applicability. None of the systematic reviews in this cohort assessed a gradient of effects. Subgroup analyses need to be conducted based on a theoretical rationale for different effects [125]. The statistical and clinical importance of results as well as the consistency of those effects both within and between trials need to be reported [86]. Based on this sample of health care related systematic reviews, there is a need to improve the methods of conducting and reporting subgroup analyses as a tool for assessing differences in effects. Individual patient data systematic reviews represent an opportunity to assess differences in effects across socio-demographic population characteristics which vary within individual studies of systematic reviews [143].

Systematic reviews are an underexploited evidence-base of differences in effects across population characteristics. Of the 24% of systematic reviews which were judged by the extractors as likely to have different effects due to differences in biology, culture, social or environmental reasons, only half (26 out of 54 systematic reviews) assessed differences in effects across RPOGRESS-Plus. However, systematic reviews are highly dependent on the conduct and reporting of the primary studies that they review.

The assessment of effects in disadvantaged populations in primary studies may be subject to reporting and selection biases. Vulnerable populations may be excluded from primary research. For example, minority populations and women are under-represented in clinical trials and these demographic characteristics are under-reported[144-146]. Reporting and representation of race/ethnicity has increased over time from 1970 to 2009 [147]. Causes of underrepresentation of minorities and women in clinical trials are due to a number of factors that have evolved over time including a desire to protect vulnerable populations by research ethics boards[146], willingness to participate in clinical trials, and trial design features that seek to maximize the size of effect with specific inclusion criteria[16]. Studies conducted in vulnerable populations or with stratified analyses across PROGRESS-Plus factors may be subject to publication and outcome reporting bias [148].

Systematic reviews need to differentiate between differences in absolute effects that are due to baseline differences in risk of outcomes and differences in relative effects. This methodology study shows that differences in relative effects between the general population and vulnerable populations are rarely reported in systematic reviews of interventions published in MEDLINE. Only two out of 224 SRs (1%) compared differences in relative effects, and neither of these reported statistically significant effects. For those systematic reviews that did conduct subgroup analysis, the methods for reporting them met only three out of seven criteria for believing a subgroup analysis [86]. If vulnerable populations have higher baseline risk of outcomes, then they are likely to achieve a greater absolute response than more advantaged populations, if the same relative effect holds [11].

Interventions targeted at vulnerable populations have been criticized for not addressing upstream determinants of health inequalities such as education and income inequalities [149]. This methodology study suggests that targeted systematic review questions can both identify effective interventions (e.g. interventions to reduce inequalities in understanding of genetic testing in non-

English speakers) and identify research gaps (e.g. lack of effectiveness of antipsychotic medications for people with challenging behaviours due to learning disabilities).

Pragmatic systematic review questions were associated with consideration of effects in vulnerable populations. Pragmatic studies and questions allow greater flexibility in delivery, adherence, compliance and follow-up that are closer to real-life implementation than tightly controlled research studies [119]. Because of this flexibility, pragmatic systematic reviews may be able to better discern differences across PROGRESS-Plus because of a greater diversity of populations and settings in the included studies, which allow the assessment of applicability issues related to providing health and social care to harder to serve populations. For example, one pragmatic systematic review question included diverse study designs (case-control, cohort studies) of people taking any type or dose of non-steroidal anti-inflammatory drug [150]. The inclusion criteria of this systematic review on NSAIDs (Non-steroidal anti-inflammatory drugs) allowed the assessment of differences in hepatotoxicity across age and gender.

This methodology study was limited to systematic reviews from MEDLINE. Systematic reviews in non-medical databases may be more likely to assess differences in effects across socioeconomic status. For example, Campbell Collaboration reviews, which assess effects of legal, educational, social and justice interventions on health and other outcomes, may be more likely to assess differences across PROGRESS-Plus because they are focused on more upstream interventions[151]. For example, a Campbell systematic review of the effectiveness of school-based programs to reduce aggressive behavior included 73 studies, and found that children from low socioeconomic status families achieved greater benefits in reduction of aggressive behavior [152]. Campbell reviews are more likely to be indexed in education and social databases [153].

This study did not retrieve the individual studies included in each systematic review, thus we did not assess the availability of disaggregated data across PROGRESS-Plus in primary studies that could have been used for subgroup analysis. We found in another study that 10% of primary studies report disaggregated data across PROGRESS-Plus factors [101]. Individual patient data systematic reviews allow greater statistical power to assess differences across patient characteristics and does not depend on the adequate reporting in primary studies [143].

Systematic reviews published in journals may be restricted by editorial policies such as page limits. Therefore, the poor reporting of statistical methods and methods for judging applicability may be due to editorial policies which were not assessed in this study.

Conclusions

There is a need for greater consideration of effects across PROGRESS-Plus characteristics in systematic reviews of effectiveness of interventions. The distinction between relative and absolute differences in effects of interventions needs to be improved in systematic reviews. Reporting standards for subgroup analyses and agreement on how to assess applicability of systematic reviews are needed.

Paper 3: Factors associated with success of interventions for the homeless population in Ottawa

Welch VA, Smylie J, Kristjansson E, Tugwell P, Wells GA, Dhaliwall B

Word count: 3460 words

Abstract: 254 words

Author contributions:

Vivian Welch had the idea for this study, designed the interview question and sampling methods, conducted interviews and analyzed data and wrote the first draft and final draft of the paper. Janet Smylie and Elizabeth Kristjansson provided comments on the methods chosen, and on previous drafts of this paper. Peter Tugwell and George Wells provided comments on the protocol for this study and the final manuscript. Bharbhoor Dhaliwal conducted interviews and analyzed data and provided comments on the final paper.

Abstract

Introduction

Health inequity persists both between and within countries across socio-demographic factors which interact to influence opportunities for good health. Knowledge translation in the context of health equity implies improving the application of health and social services to reduce existing health inequities.

Objective:

To identify factors associated with improved effectiveness of interventions in vulnerable populations.

Study design: Qualitative study with convergent interviewing

Methods

The homeless population in Ottawa was selected as the vulnerable population of interest for this study because the agencies which provide services to homeless people can be identified.

Snowball sampling was used to recruit practitioners and managers based on maximum variation of opinion, until saturation was reached. The convergent interviewing method was used to identify factors considered important in determining the success of programs aimed to improve the health of homeless people in Ottawa. Verbatim transcribed interviews and notes taken during the interview were used to identify success factors.

Results

Saturation was reached after interviewing six practitioners and six managers. Most of the identified factors (15 out of 20) were related to improving accessibility and acceptability of services such as preventing discrimination from practitioners and reducing exposure to unsafe environments. All of these factors fit into the equity-effectiveness loop framework.

Conclusions

Success factors identified in this study could be used in the design and delivery of new programs and services to improve health of homeless people. These findings suggest that the equity effectiveness framework can be used to identify characteristics of health services which need to be modified to enhance their effectiveness in disadvantaged populations, thus enhancing health equity.

Introduction

Health inequity is defined as avoidable and unfair differences in health [78]. Health inequity persists for most diseases both within and between countries, despite over three decades of high-level political interest in tackling health inequalities within countries and internationally [23, 154]. PROGRESS-Plus (Place of residence, Race/ethnicity, Occupation, Gender/sex, Religion, Education, Socioeconomic status, Social capital, plus any additional factors such as disability or disease status across which differences in opportunity may exist) has been proposed as an acronym to capture factors across which health inequity may exist because of differences in equality of opportunity [37, 38].

In the last ten years, there has been a call for an improved evidence base on what works to decrease health inequalities [3, 47]. Systematic reviews of all available evidence have been proposed as a method for assembling the evidence on effectiveness of interventions which aim to improve health equity [73, 76]. However, only 1-10% of systematic reviews report effects of interventions on health inequalities [6, 101, 155]. As few as 10% of clinical trials report effects stratified across vulnerable populations [6, 101].

The WHO Commission on Social Determinants of Health proposed a classification of four types of upstream and downstream interventions that could act on promoting health equity (Figure 3): 1) Reducing social stratification (e.g. by education and income transfers such as conditional cash transfers [15]); 2) Reducing differential exposure and vulnerability, (e.g. population level tobacco control [67]); 3) Reducing differential consequences of ill health (e.g. return to work interventions for breast cancer survivors [68]); and 4) Improving health system characteristics (e.g. systematic reviews relevant to human resources for health [69]). Two previous studies on characteristics of systematic reviews has found little consideration of implementation factors and program design features of health interventions that could act on the latter three pathways to reduce inequities.

The equity-effectiveness loop is a framework which explicitly considers implementation and program design factors and their influence on health equity [77]. This framework has been tested with hypothetical examples [156] which highlight the gap in community effectiveness due to differences in the implementation factors of access, diagnostic accuracy, patient adherence, provider compliance (Figure 5) . However, these hypothetical case studies did not explore the modifiable factors which can be targeted in program design to improve effectiveness of interventions in vulnerable populations.

The aim of this qualitative study was to identify factors associated with improved effectiveness of interventions in vulnerable populations. The homeless population in Ottawa was selected for this study because the agencies which provide services to the homeless can be readily identified, in contrast to other vulnerable populations who could be served by a wider range of services and programs which are more difficult to identify and define (e.g. services for low income people are provided in a wide range of community health centres, hospital settings and community-based organizations).

Methods

Equity effectiveness framework

The equity effectiveness framework (Figure 5) was chosen as a basis for this study because it focuses on assessing the effects of implementation and program design features on health equity from needs assessment, program design, delivery through to monitoring and reassessment[77].

Qualitative research strategy

We selected convergent interviewing as the interview technique because it reaches saturation after fewer interviews than grounded theory by using constant comparative analysis, using negative case analysis to explore agreement and disagreements as well as exceptions to rules with each new interview [157, 158]. Furthermore, convergent interviewing is useful when the aim is to expand on existing theory [159]. This study aimed to explore modifiable factors for improving effectiveness of health and social interventions in homeless populations within each of the steps of the equity-effectiveness framework [77].

Convergent interviewing was developed in the field of market research, and has been applied to management and health areas [158]. Informants are selected based on maximum variation in their content expertise and other important characteristics using purposeful sampling. Interviews are open-ended and non-directive, allowing the participants to direct the conversation based on their experience. After each interview, key themes are identified. These themes are checked with subsequent informants for agreement or disagreement, and informants are asked to explain disagreements or agreements. For each theme, the interviewer seeks exceptions to proposed rules and seeks to explain agreements and disagreements with specific prompts, based on previous interviews [160].

The researcher in the convergent interviewing method is responsible for proposing a statement or question to start the conversation. This study applied a pragmatic epistemology [161-163]. Pragmatism implies that a research method or hypothesis is tested by the practical effects or workability in answering the question of interest [161]. Pragmatism supports a pluralist approach to research which supports mixed methods research [161]. This study used convergent interviewing to validate and support the use of an existing theory [161, 164], unlike other qualitative methods (e.g. grounded theory), where the data is used to build new theory [164]. This type of study recognizes that the interpretation of experiences depends on the researcher's own personal experiences [165].

The interviewers were responsible for providing prompts to continue the conversation as well as identify exceptions and request explanations about disagreements. The interviewer prompts provided may have biased the informants. In this study, one of the interviewers was the primary investigator (VW). This potential for bias was assessed by documenting the researcher`s viewpoint and documenting the prompts used.

Data collection procedures

Participants

We recruited practitioners and managers who provide health or social care services to the homeless in Ottawa. According to the methods described by Dick 1998 [160], maximum variation in opinion was sought using snowball sampling. The first informant was an experienced member of the Ottawa homeless service provider community with links to many of the health and social service agencies which provide services for the homeless in Ottawa. The first informant was asked to suggest two other practitioners or managers who might have different opinions. Each subsequent informant was asked to suggest people who might have different opinions than their own.

Practitioners were defined as medical doctors, nurses, community developers, community educators, social workers, and community outreach workers. Managers were defined as people who make decisions about which services will be provided, how, where and by whom.

This study was designed to identify what factors are important in designing and implementing successful health and social care and prevention programs for the homeless in Ottawa. This study focuses only on manager and practitioner viewpoints since they have experience designing, monitoring and implementing programs for the homeless in Ottawa. In the future, it might be important to hear from the homeless themselves, since there may be some factors that they could identify from their experience of living with inequitable differences in health and services that would be overlooked by practitioners and managers.

Interview methods

The interviews were conducted by two interviewers (VW, BD). The interviews were non-directive and in-depth, starting with the following opening statement:

> "Can you describe a particularly successful or unsuccessful health or social care program for homeless people in Ottawa; what was good or bad about it?"

Prompts were not used in the first interview. The interviewer asked questions to explore ideas described by the informant, and used active listening to encourage the informant to speak freely.

In subsequent interviews, two types of prompts were used: 1) active listening (e.g. "*They feel welcome*"), or 2) questions to seek opinion on an issue mentioned in a previous interview to assess concordance or discordance (e.g. "*some services have a focus on giving responsibilities to the clients and others have been described as based on a 'mercy model'. What is your opinion*?"

Each interview lasted about one hour. Interviews were conducted face to face in a setting which was familiar and comfortable for the informant, and free of distractions. This facilitated the development of a rapport [159]. The interviews were recorded and transcribed verbatim. The interviewers also took notes during the interview.

After each interview, themes were extracted in tabular format with illustrative quotes based on the notes. The two interviewers compared ideas and themes from their respective interviews.

The interviewers developed prompts for use in subsequent interviews to explore agreement, possible exceptions, identify and request explanations for disagreements [160].

Interviews proceeded until saturation, which was defined as the point where no new themes and no new exceptions or disagreements arose in two successive interviews. According to Driedger 2006, saturation requires approximately 6-8 interviews, which is consistent with other studies using convergent interviewing [159, 166]

Data analysis procedures

Data were analyzed by BD and VW after each interview using a table to summarize issues and informant opinions identified by note-taking. Verbatim quotes from the transcribed interviews were pasted into the table of issues identified by the note-taking. Details of explanations for disagreements and exceptions were summarized in the same table. The two interviewers compared factors identified and discussed quotes used to ensure they both agreed with the factor and understood the meaning.

Further analysis was conducted by VW to map the identified factors onto the six steps of the equity-effectiveness framework: 1) burden of illness; 2) equity effectiveness; 3) economic evaluation; 4) knowledge translation; 5) monitoring; and 6) reassessment. This framework was chosen since it focuses on health equity at each step of program design from assessing needs through to implementation (Figure 5).

Strategies for ensuring consistency and credibility

The procedures for data collection and analysis aimed to ensure credibility and consistency. Credibility is defined as confidence in the themes identified. We maximized credibility by comparing themes with each subsequent interview using prompts for exceptions and agreement, using in-built negative case analysis to explain disagreement and using an open interview question to allow the informant time to describe their own experience, thus minimizing interviewer bias.

Consistency across interviews was maximized by using an interview guide with the open-ended question, as well as an accompanying list of prompts which was expanded with each new interview.

Ethics approval

Ethical approval was obtained from the University of Ottawa Research Ethics Board (Ethics approval certificate, University of Ottawa Research Ethics Board #H02-09-11) (Appendix 5).

Results:

Saturation was reached after interviewing six managers and six practitioners.

Informants were diverse across age (approximately 30-60 years), sex (3 males, 9 females) and type of agency [hospital (1), clinic (3), shelters (4), non-governmental organizations (NGO) (2), government (2)]. Four out of 12 informants were from a faith-based organization. Clinical expertise included people with degrees in nursing (1), medicine (1), health education (3), social work (3), addictions counseling (2) and business (2). All informants had at least 10 years experience working with homeless people.

Themes

The coding identified 20 factors which were described as important in determining success of programs and services for homeless people in Ottawa (Table 15). These factors are:

1. Assess needs of individuals
2. Assess community needs
3. Provide a welcoming setting
4. Ensure services are feasible
5. Diagnose underlying problems
6. Provide training for staff
7. Enforce policies for safety
8. Train practitioners to be patient with clients
9. Ensure services are acceptable
10. Need to consider business case to provide right services and the right time for the right people

11. Build financial sustainability

12. Raise awareness and support from the public

13. Design programs based on best available evidence and theory

14. Provide individualized, multi-faceted treatment

15. Engage the homeless in designing and providing programs

16. Convergence of political agendas with agency priorities

17. Provide services in partnership with multiple agencies

18. Provide supportive environments for homeless to reduce exposure to unsafe behaviours

19. Agencies need leadership and vision

20. Improve performance by monitoring and evaluation

These factors were the same for managers and practitioners, with three exceptions. Managers described the need for a business case and for a window of opportunity as success factors. For example, a business model was described as necessary to obtain support from funders as well as the general tax-paying public. Furthermore, a business model helped to prevent duplication of services and reduce costs.

Two managers described convergence of political agendas with agency priorities as an important success factor. One manager stated that *"It was a time when homeless was considered to be a crisis. It was a time that it captured the attention of the public and the politicians."* Another stated that: "*then you have some opportunity with this funding that kind of comes out of the blue and then some luck of meeting people and so there is sometimes this convergence of things that occur and you are given the opportunity to go forward*".

Practitioners described patience with homeless clients as a success factor which was not identified by managers. One practitioner stated: *"You have to be very patient. They might miss appointments until they finally make it to one. It's like saying hard to serve which I don't like. I say harder to serve."* Practitioners described a cycle of substance abuse, antisocial behavior and homelessness as a process which needs to be accepted: "*they are going to do it over and over again*".

Exceptions, disagreement and explanation

Exceptions were identified for four factors: 1) empowering clients, 2) monitoring, 3) sustainability of staff and funding, and 4) unique needs of women.

Empowering clients to take responsibility of their own health, and providing them an opportunity for normal lives by participating in program delivery and household chores was described by several informants. For example, this was described as "*moving away from do-gooder, paternalistic model to a model of evidence-based practice with education, training, skill development and competencies*". However, some informants disagreed with this concept, explaining that clients need to be accepted "*where they are at*", stating that clients in crisis or with severe, unstable mental illness needed to be treated with a "*mercy model*". When probed further, this difference of opinion was explained by considering that there is a wide range in abilities of the homeless to participate in decision-making about their health and environment. Thus, the discrepancy of opinion represents a continuum of services to match clients' abilities to participate.

Monitoring and evaluation was described as a factor that contributed to success by allowing celebration of success with staff, clients and funders as well as constant improvement of services based on evaluation. For example, one manager stated: "*Data collection is a tool to improve performance*" Missing or inadequate evaluation and monitoring was described as a barrier to sustained enthusiasm and funding. There were no disagreements about the need for evaluation.

Sustainability of funding and staff was described as a success factor because consistent staff and reliable programs built trust in the homeless community. For example, one manager stated: "*...we need to do as an organization is to try as best as we can to put you into a position where you are financially stable, where you have long-term sustainability and that you are not subject to the control of any organization*". In contrast, two informants described a high staff turn-over of young graduates as a positive factor since the young graduates are enthusiastic and generate new ideas, which often led to improvements in services, based on the latest evidence.

There was disagreement about whether women need special or different services from men. Unique characteristics of women were acknowledged by over half of the informants. These needs included a greater vulnerability to physical and sexual abuse, a history of inadequate treatment or "*medicalization*" of trauma with drugs, and a greater orientation towards achieving personal goals than men. Despite these differences, three informants disagreed that women always need separate services. For example, women and men are included in the Managed Alcohol Program in Ottawa in the same building. The palliative services at the hospice are also open to both women and men. Thus the need for different services depends on the type of services and context.

Further analysis: Equity-effectiveness framework

These success factors were mapped to the six steps of the equity-effectiveness loop framework (Figure 5 and Table 16), with illustrative quotes. Factors clustered mainly into two steps: 1) the second step of efficacy-modifying factors which modify equity-effectiveness (i.e. diagnosis, access, adherence and compliance); and 2) the fourth step of knowledge translation and implementation.

Step 1: Burden of illness: This step assesses the burden of illness across the socioeconomic gradient. Informants identified the need to assess individual and community needs by PROGRESS-Plus factors (e.g. age, sex or ethnicity) as important for designing successful programs to meet those needs. At the individual level, this included assessing the multiple needs of each individual for health and social services as well as non-health needs, such as employment skills development and housing with appropriate support. At the community level, practitioners and managers stated that it is important to assess community statistics on the types of homeless populations who need services since these define the development of new programs (e.g. Inuit and seniors were described as two growing populations within the homeless community in Ottawa which were inadequately served by existing Ottawa services).

Step 2: Equity effectiveness: This step assesses what factors modify the real-world effectiveness of interventions across PROGRESS-Plus factors, including access, diagnosis, provider compliance and patient adherence. Factors related to improving acceptability, training for providers and appropriate diagnosis were described as important to improve effectiveness of interventions for homeless populations. The factors identified in this study contribute importantly to the application of the equity-effectiveness loop for a homeless population by identifying concepts related to discrimination (e.g. "stop judging") and social inclusion (eg "normalizing" and "opportunity") that were not explicitly stated in the equity-effectiveness loop. Furthermore, the factors identify the need for cultural sensitivity, building trust and empowering clients to take responsibility.

Step 3: Economic evaluation: This step assesses the outcomes that can be achieved for a specific cost. Informants described three factors related to economic evaluation: 1) the need for a business case; 2) ensuring sustainable funding and staff and 3) raising public awareness. As described above, the need for a business case was described by managers only, and entails justifying funding, minimizing duplication and increasing efficiency. Informants described a need to develop a sustainable funding base so that they were not reliant on pilot projects for funding, since short-term projects and short-term staffing compromise the trust of clients. Raising public awareness is classified as related to economic evaluation because the outcomes important to the public are related to showing that homeless programs are cost-saving; for example by reducing emergency room costs and police calls.

Step 4: Knowledge translation: This step entails applying knowledge about interventions to improve health outcomes across PROGRESS-Plus. Seven knowledge translation factors were described which ranged from identifying the problem to adapting knowledge, assessing barriers and tailoring interventions. In identifying the problem, informants stated that having a vision for change was important, as well as services based on evidence and theory such as harm reduction, youth engagement and non-violent communication. Partnership and service integration with other agencies was a major success factor described by all informants: '*Health brings a level of professionalism to the other agencies…we saw an advancement of the health culture across all agencies, professionalization of services of other agencies and learning about evidence based practice*". Environmental barriers were described as a negative influence such as the lack of appropriate supportive care (e.g. for addictions counseling) or exposure to negative influences (e.g. housing near place of drug trafficking or prostitution). Engagement of clients in designing and delivering services was described as a factor associated with success of programs; e.g. "*we let the clients start to speak and gave them some responsibility for the coordination of the program. We gave them some opportunity to speak on our behalf and they spoke much more eloquently then we could ever*".

Steps 5 and 6: Monitoring and reassessment: These two steps reinforce the need to monitor the effects of interventions on burden of illness and outcomes across PROGRESS-Plus and the need to modify services as required to have the intended effect on improving health of the disadvantaged. Both practitioners and managers described evaluation as a tool to improve service delivery, with a focus on outcomes.

Discussion

This study identified factors specific to assessing and overcoming barriers to improved effectiveness of interventions in the homeless population in Ottawa including identifying and minimizing discrimination from practitioners and the public, providing opportunities for normalization and providing supportive, safe environments. These factors mapped relatively well onto the equity-effectiveness framework, suggesting that this framework can be used to identify health system characteristics which need to be considered to maximize benefits for disadvantaged populations, thus enhancing health equity. The convergent interviewing method was a time-efficient method to identify factors related to the success and failure of health and social programs for homeless populations.

Mapping success factors to the equity-effectiveness framework illustrated that informants identified success factors mainly related to enhancing effectiveness (7 factors) and knowledge translation (7 factors). Mapping also identified the need for economic evaluation to focus on showing the cost-benefit for the general public and tax-payers in order to justify continued funding and public support, consistent with other studies on homelessness [167]. Thirdly, mapping identified relatively less consideration of burden of illness (2 factors), economic evaluation (3 factors), and monitoring and reassessment (1 factor) suggesting that informants either do not consider these steps important for success or do not think services for the homeless need adaptation relevant to these four steps. For example, the electronic health record system developed for the homeless in Ottawa was not described by any of the informants as important for success, despite being described as an important success factor in scientific articles and news stories [168]. Mapping factors to the equity-effectiveness framework was useful in identifying this lack of consideration of monitoring.

The strengths of this study are the maximum variation sampling strategy, mapping of results to an existing framework focused on enhancing health equity and minimal bias introduced from prompts by the investigator. The sampling strategy yielded a diverse sample across age, types of agencies and professionals' experience. Mapping the results onto the equity-effectiveness framework builds on this framework by identifying specific factors related to success such as public perception of homelessness. Thirdly, the convergent interviewing technique minimized the bias that can be introduced by directive or leading questions.

As with any qualitative research, the collection of data and its interpretation may be influenced by the researchers' own views and the interviewers' views since respondents may tell interviewers what they think they want to hear[165]. The risk of being influenced by only one person's experience and views was minimized by using two researchers to conduct interviews and analyze the data, by using a planning group to develop the open-ended question, and by developing prompts for subsequent interviews by two interviewers through discussion.

One limitation of this study is the reliance on agencies that are associated with improving health. No informants were interviewed from police services or employment sectors which are involved in non-health sector interventions that also have an effect on the health of the homeless and might consider different factors important in determining success. Secondly, equal importance was given to all factors identified since convergent interviewing does not seek to rank the importance of factors identified. Thirdly, the interpretation of results was subject to the investigators' bias since two of the authors of this paper published the equity-effectiveness framework. An effort was made to minimize this risk by the use of an open question and non-directive prompts.

Another limitation of this study is that the framework being tested was not used to develop prompts for interviews. If the framework had been used to develop prompts, the respondents might have provided more detail about some factors which were not well-described such as monitoring and reassessment.

In conclusion, this study builds on the existing framework of the equity-effectiveness loop by identifying that the factors considered most important by practitioners and managers to improving effectiveness of interventions in the homeless population in Ottawa are related to the knowledge translation and effectiveness steps of the framework. Convergent interviewing appears to be a useful tool for identifying factors influencing the degree of success of programs and services for the homeless population in Ottawa, with the exception of failing to identify the electronic health record as a useful monitoring tool. Further exploration of the convergent interviewing technique across programs and services designed to reduce health inequities in different populations will enhance our understanding of the utility of this tool in addressing the dearth of information regarding the effects of knowledge translation interventions designed to reduce health inequities.

Paper 4: Predicting plausibility of effects on health equity in systematic reviews: an equity plausibility algorithm

Welch V, Tugwell P, Wells G, Smylie J, Petticrew M, Kristjansson E, Brand K

Word count: 2685

Abstract: 300 words

Abstract

Introduction: Promoting health equity is a high priority on local, national and international policy agendas. The WHO Commission on Social Determinants of Health concluded we have a moral obligation to ensure fairness in the opportunity for good health. Promoting health equity entails assessing not only the distribution of health in the population, but also the distribution of effects of health care policies and programs across sociodemographic characteristics.

Objective: This study aimed to develop and evaluate an algorithm to predict whether an intervention is likely to have different relative effects in disadvantaged populations compared to advantaged populations.

Study design: Health measurement scale development and validation

Methods: An equity plausibility algorithm was developed using clinimetric methods with three questions based on literature review, key informant interviews and methodology studies. The three questions asked about possibility of differences in relative effects across sex or socioeconomic status due to: 1) different patient characteristics; 2) delivery of the intervention; and 3) different comparators. Thirty-five clinicians, methodologists and research users assessed the likelihood of differences across sex and socioeconomic status for ten systematic reviews with these questions. The consistency of responses across respondents was assessed using Fleiss' Kappa for multiple raters to assess agreement in answers to 60 questions (3 questions X 10 scenarios X 2 sociodemographic factors).

Results: There was little to no agreement for each of the three questions, with Fleiss kappa ranging from -0.001 to 0.199. Eight-eight percent of the 60 assessments of systematic reviews were judged by more than half of respondents as likely to have important differences in the magnitude of relative effects across gender and socioeconomic status.

Conclusions: This study found disappointing inter-rater reliability of questions for predicting the likelihood of different effectiveness of interventions in disadvantaged populations. Systematic reviews where differences in effects are plausible could be encouraged to use methods to explore differences in the distribution of health effects that may influence health equity.

Introduction

Systematic reviews represent an opportunity to identify what works to promote health equity because they include studies conducted in a diversity of settings and populations [3,67, 69]. There is increasing acceptance that systematic reviews are useful as a basis for evidence-informed policy and practice since they reduce the chance of being misled, increase confidence in results, are an efficient use of time and are more easy to critically appraise and apply [4, 73]. However, systematic reviews rarely assess whether interventions have an impact on health equity. For example, only 1% of a random sample of Cochrane reviews assessed differences in effectiveness of interventions across socioeconomic or demographic factors [6]. Failure to assess or consider effects on health equity in systematic reviews may lead to rejection of systematic reviews as a useful source of evidence for policy-makers who seek information on distribution of effects in the population [75, 76], or may lead to implementation of policies and programs which inadvertently increase health inequities [41, 125].

Promoting health equity remains high on the agenda of local, national and international policy agendas [23]. The World Health Assembly in 2009 issued statement 62.14, calling on all states to assess effects of policies on health equity and to address social determinants of health (http://apps.who.int/gb/ebwha/pdf_files/A62/A62_R14-en.pdf). In the UK, a strategic review has been launched on reducing health inequalities post-2010 (http://www.ucl.ac.uk/ghe.g./marmotreview).

Health inequity is defined as avoidable and unfair differences in health [78]. PROGRESS-Plus is an acronym which summarizes factors across which differences in health may be considered inequitable depending on the setting and context: Place of residence; Race/ethnicity/culture; Occupation; Gender; Religion; Education; Socioeconomic status; Social capital [38]. The "Plus" was proposed to capture additional factors that are associated with unequal opportunities for health, such as disability, disease status and age [37].

Health inequities are created and maintained through an interaction of factors at multiple levels, illustrated by the conceptual model of the World Health Organization (WHO) Commission on Social Determinants of Health (CSDH) which shows the pathways which affect health equity including health system factors (e.g. quality and accessibility) as well as non health system factors, such as social stratification (Figure 3) [169].

Differences in effects of interventions across PROGRESS-Plus factors influence health inequalities because they imply a different distribution of benefits and harms across sociodemographic factors. Differences in relative effects measured as differences in ratios (e.g. relative risk reduction) suggest different mechanisms of action. For example, the relative risk reduction of smoking in response to increases in tobacco price is highest for people with the lowest income [67]. Differences in absolute effects are likely due to differences in baseline risk. For example, immigrants and refugees have a lifetime risk of developing active tuberculosis of over 35%. Assuming a relative risk reduction with isoniazid treatment of 93%, treating all immigrants and refugees will reduce lifetime cases of active tuberculosis by 33 in 100 people. In contrast, the typical Canadian-born population has a low risk of developing latent tuberculosis of less than 5 in 100 people, thus the absolute reduction in number of cases with treatment would be only 2 per 100 people [170].

Systematic reviews can assess whether an intervention has an effect on health equity in four ways: 1) the gap approach to assess the difference in effects between advantaged and disadvantaged populations (e.g. by assessing effect modification using subgroup analyses or meta-regression [171]); 2) the gradient approach to assess a continuum of differences in effects across degrees of social disadvantage; 3) assessing the effects of targeted interventions that aim to improve the health of disadvantaged populations (e.g. school-feeding for disadvantaged children [128]); 4) assessing transferability or applicability of results of systematic reviews to different settings (e.g. low-income countries) [69, 103, 104] or disadvantaged populations, such as low-income, single mothers [98].

This study aimed to develop and evaluate an algorithm to assess the likelihood of differences in relative effects of interventions in disadvantaged populations (across PROGRESS-Plus categories) compared to advantaged populations.

Methods

The equity plausibility algorithm was developed using nine steps based on Feinstein's *Clinimetrics* [172] and Streiner and Norman's book on constructing a health measurement scale (Table 17) [173].

Ethics approval

This research study was approved by the University of Ottawa Research Ethics Board (ethics approval certificate #H02-09-11b and #H02-09-11c) (Appendix 7 for consent form and ethics approval certificate).

Purpose

The purpose of the algorithm was defined as assessing the likelihood of differences in relative effects of an intervention across population characteristics defined by PROGRESS-Plus.

Item generation

Items were generated using four methods. First, existing checklists for applicability, transferability and external validity were assessed for factors related to judging likely differences in relative or absolute effects (Table 2). Second, factors associated with subgroup analyses across PROGRESS-Plus were assessed in a systematic review of methods for assessing effects on health equity (Welch paper #1, systematic review of methods)[174]. Third, factors associated with statistically significant or clinically important subgroup analyses across PROGRESS-Plus factors were assessed in a group of 224 systematic reviews (Welch, paper #2, methodology study). Fourth, practitioners and managers were interviewed using convergent interviewing [158] to identify factors associated with success or failure of program implementation in a vulnerable population (Welch, paper #3, qualitative study).

Item reduction, questionnaire format, scaling, face validity

One author (VW) developed a draft algorithm using the items in Table 18. Items were phrased using wording from previous checklists where possible. Dichotomous yes/no categories were chosen for the responses for two reasons: 1) a "don't know" category was considered unhelpful in determining likelihood; and 2) the ability of respondents to discriminate more finely than yes/no was uncertain. The draft algorithm was discussed with two other authors (PT and GW). Face validity refers to the clinical sensibility of a scale and its component items [172]. The face validity was tested by asking four clinician methodologists with experience in clinical epidemiology, systematic reviews and health equity to review the items and judge their ability to measure the concept of interest, defined as assessing the likelihood of differences in relative or absolute effects across PROGRESS-Plus characteristics. These four clinician methodologists were also asked whether the questions were clear. The questions of the algorithm were modified after this face validity testing.

Consistency

Inter-rater reliability of the equity plausibility algorithm was assessed in a field study by recruiting methodologists, clinicians and users of systematic reviews to apply the algorithm to a sample of 10 systematic reviews. Thirty-five methodologists, clinicians and users of systematic reviews were recruited from members of Cochrane Collaboration entities. Respondents were asked to judge the likelihood of differences for two PROGRESS-Plus items: 1) female vs. male and 2) low socioeconomic status vs. high socioeconomic status.

Raters were given a summary of the 10 systematic reviews inclusion criteria and methods for the population, intervention, comparison, outcomes and study designs included (columns 3-7 of Appendix 8). Study designs were given to raters since the study design may affect the ability to discern differences in effects. Raters were not given the results of the systematic reviews.

The equity plausibility algorithm was presented to raters as a list of three questions on one page, with a checkbox requiring Yes or No answers. Raters were also given examples of how each of the factors might create important differences in the magnitude of relative effects, based on examples from the updated Journal of American Medical Association (JAMA) User's Guide on applying results to individual patients [58]. An example of the survey is provided in Appendix 9.

Inter-rater reliability was assessed using Fleiss Kappa for multiple raters for 60 assessments (10 systematic reviews X 3 questions X 2 PROGRESS factors) [175].

Ten systematic reviews were selected for the consistency testing based on the following criteria: 1) proven effective and cost-effective interventions identified by the WHO-CHOICE (: World

Health Organization CHoosing Interventions that are Cost-Effective) initiative [176]; 2) representation of different types of intervention using the categories defined by the Disease Control Priorities Project of: i) Population-based primary prevention; ii) Personal interventions and iii) Policy instruments [177]; 3) included in the top ten causes of burden of disease in the world projected for 2030 [178]; 4) availability of a systematic review with greater than five included studies including a diversity of settings and populations. The systematic reviews are described in Appendix 8.

Construct validity

Construct validity of a health measurement scale refers to whether the scale is appropriate for its intended purpose. When there is no gold standard, construct validity can be assessed by comparing results of a scale to results of other similar scales thought to measure the same or similar attribute or construct. The ultimate test of the construct validity of this equity plausibility algorithm is whether it predicts which systematic reviews have found differences in effects by sex and socioeconomic status. The effects on sex and socioeconomic status were extracted from the 10 systematic reviews (column 2 of Appendix 8). However, since the demonstration of differential effects across sex and socioeconomic status are dependent on the availability of data from the primary studies, as well as the methods of the systematic review, the judgment by the authors of the systematic review about applicability and generalizability was taken as the criterion with which the equity plausibility ratings were compared. Therefore, the construct validity of the equity plausibility algorithm was assessed by comparing whether the raters' assessment of likelihood of differences in effects was supported by the discussion of applicability and generalizability by the authors of the systematic reviews across PROGRESS-Plus. Only those comparisons where greater than 80% of raters agreed on a rating were compared with systematic review conclusions about applicability and generalizability. This cut-off of 80% was chosen based on use of this criterion for consensus recommendations in rheumatology [179].

Results

Item generation

A list of 50 potential items was generated (Table 18). A draft equity plausibility algorithm was developed by three authors (VW,GW,PT) that consisted of four yes/no items: 1) differences in implementation factors across PROGRESS-Plus (e.g. differences in resources); 2) likelihood of variations in the delivery of the intervention across PROGRESS-Plus (e.g. because of poor acceptability, inappropriate literacy level); 3) different mechanism of action of the intervention across PROGRESS-Plus; and 4) differences in expected absolute effects because of higher risk or prevalence across PROGRESS-Plus.

Item reduction, questionnaire format, scaling, face validity

Consultation with four content experts resulted in adding an item about possible differences in the comparator across PROGRESS-Plus that might affect the effectiveness of the intervention. The first three questions were combined into two questions: 1) differences in mechanism of action because of patient characteristics (e.g. biology, social, cultural, environment interacting with the intervention) and 2) whether the intervention is delivered in the same way across PROGRESS-Plus, taking into account possible provider effects, geography etc. The question about absolute effects was removed, since it did not fit with the focus of this checklist to identify likelihood of differences in relative effects across PROGRESS-Plus. The revised equity plausibility algorithm consisted of three questions (Table 19).

Inter-rater Consistency

Ten systematic reviews were selected for the field test (Appendix 8): 1) mass media to promote HIV testing [180]; 2) Population level tobacco control [102]; 3) psychological therapy for post-traumatic stress disorder [181]; 4) first-line anti-hypertensive drugs for people with hypertension [182]; 5) surgical interventions for age-related cataract [183]; 6) vaccines for measles, mumps and rubella [184]; 7) antidepressants vs. placebo in primary care [185]; 8) artemisinin-based combination therapy for uncomplicated malaria [186]; 9) primary safety belt laws [187]; and 10) handwashing for the prevention of diarrhea episodes [188].

Of 43 people contacted, 35 filled out the questionnaire (81%). The 35 respondents represented a mix of users, methodologists and clinicians with a median of 7 years of experience using or conducting systematic reviews and diverse clinical experience (e.g. public health, musculoskeletal, dermatology, cancer, infectious disease) (Table 20).

67% of the ratings across all questions and systematic reviews endorsed the likelihood of important differences in effects for sex and socioeconomic status. There was little to no inter-rater agreement, as assessed by Fleiss' kappa for multiple raters, across all questions for both sex and socioeconomic status, except for the first question about patient difference across sex which showed slight agreement (k=0.199) (Table 21).

The reason for the answers was most commonly based on theory (n=10 people) and personal experience (n=11). Empirical data was described by 3 people and 4 people admitted to using "guesses" (Table 22). Respondents stated that other information would have been useful to complete the task, including more information about the interventions and outcomes, clarity on the comparator question and details on the size of difference considered important. Comments on the task reflected endorsement of the importance of considering differences across PROGRESS-Plus in the design of systematic reviews (n=10 people).

Construct validity

Fourteen assessments out of 60 (23%) attained greater than 80% agreement between raters (bolded text in Table 21). Six of these assessments were consistent with the judgments of applicability described by the authors of the systematic reviews (indicated with a "Y" in superscript). For example, differences in effects due to patient characteristics (question 1) was judged likely for population tobacco control across SES, which is supported by the conclusions of the systematic review that state that price control is more effective for low income populations [67]. Similarly, 92% of raters assessed that there were not likely to be differences in effects of measles, mumps and rubella vaccine due to different patient characteristics across sex (question 1), which agreed with the conclusions of the systematic review. Eight of the assessments which reached greater than 80% agreement between raters were not consistent with the judgments of the authors of the systematic reviews (indicated with an "x" in superscript). For example, 91%

of raters judged that there were likely differences in relative effects due to differences in delivery of PTSD intervention for low-income compared to high income populations. However, the systematic review of PTSD does not discuss any differences in implementation or delivery across SES. Similarly, 100% of raters judged that surgery of cataracts would have different implementation/delivery in low-income people, but this was not discussed or mentioned in the systematic review of cataract surgery [183].

Discussion

The equity plausibility algorithm of three items showed little to no inter-rater agreement for any item across sex and socioeconomic status. There was a lack of agreement between the raters' assessments and the conclusions of authors of the systematic reviews about applicability. The findings of this study suggest that prediction of the likelihood of different effects may require more detailed consideration of how raters make these judgments and the information required to make these ratings.

The ability of systematic reviews to assess effects in disadvantaged populations is dependent on the data availability in the primary studies. For example, primary studies may be underpowered to conduct subgroup analyses. Primary research may not include disadvantaged populations across PROGRESS-Plus or not report PROGRESS-Plus characteristics. Publication of primary research conducted in disadvantaged populations may be subject to publication or reporting biases.

Raters predicted likely differences in relative effects for 67% of the assessments. This finding is in conflict with empirical studies of meta-analyses which show that the relative risk reduction is constant across different patient characteristics for 82 to 87% of health care treatments [8, 9]. Furthermore, less than 11% of systematic reviews of effectiveness report statistically significant differences across PROGRESS-Plus factors (Welch study#2, methodology study). Raters in this study may have considered implementation issues in their judgments such as whether the intervention would be feasible or affordable in different contexts which could explain the lack of agreement with the conclusions of the systematic reviews.

The frequency of endorsing the likelihood of important differences in effects of 67% and the raters' comments about subgroup effects are in conflict with the widely accepted guidance on the need for caution in conducting and interpreting subgroup analyses [51, 53, 54]. Even if differences are found across populations defined by PROGRESS-Plus, these differences may be due to other factors that are confounded with population characteristics such as undetected co-interventions and differential follow-up [113].

Lack of content expertise and a desire for more information from the primary studies was cited by one third of respondents as a limitation in making the judgments about equity plausibility. These findings suggest that there is a need for more consistent assessment of applicability in systematic reviews themselves. This is supported by interviews with policy-makers who describe the need for consideration of local applicability in systematic reviews [189]. The advantage of promoting assessment of applicability in systematic reviews is that authors have content expertise and access to the primary studies to make the judgments.

This study followed established steps in developing a checklist of item generation, pilot-testing and assessed consistency and construct validity in a field study. The field study selected systematic reviews based on predetermined criteria to maximize diversity of types of interventions (personal and population level) and disease areas with sufficient data in the systematic reviews to make judgments about applicability.

A limitation of this study is that the rationale for affirmative responses was assessed by self-report on a questionnaire. Investigation of the rationale of judgments using in-depth discussion or think out loud protocols could elucidate reasons for lack of agreement. For example, it is possible that raters confused relative and absolute effects or that they made judgments about implementation broader than the systematic review questions. Think out loud protocols could also assess how respondents justified their expectation of different effects, for example, using their knowledge about population health and social determinants of health. Another limitation is that the examples provided in the questionnaire guide may have biased respondents to answer affirmatively. The examples provided focused on differences across sex or SES which might be important (e.g. surgeon skill in low-volume settings).

Implications for research

This study has shown that these three questions cannot be used to reliably judge the likelihood of differences in relative effects across PROGRESS-Plus. The finding that 67% of raters judged important differences likely across sex and socioeconomic status reinforces the need to develop, evaluate and promote methods for considering and reporting effects on health equity in systematic reviews. The three methodological areas identified by this study are: 1) methods for judging applicability which differentiate between relative and absolute effects, and achieve acceptable inter-rater reliability [189]; and 2) methods for conducting and reporting subgroup analyses across patient and setting characteristics that increase their credibility (e.g. meta-regression of individual patient data [122, 190]); and 3) methods to promote the assessment of applicability and generalizability by authors of systematic reviews. The Cochrane and Campbell Equity Methods Group aims to act as a forum to stimulate methodological research on these topics.

General discussion and recommendations

This thesis has increased our knowledge about the role of systematic reviews in tackling health inequity. Firstly, this thesis has described the extent to which health equity effects are assessed in systematic reviews, what methods are used and how assessing effects on health equity is used in the conclusions of the systematic reviews. Secondly, this thesis identified factors associated with effects of interventions on health equity and used these factors to develop and test the validity of an algorithm to predict whether interventions will have different effects on health of disadvantaged populations that are likely to influence health equity.

This dissertation assessed what factors are associated with effects of interventions on health equity, by assessing these factors in systematic reviews and methodology studies of systematic reviews, as well as by interviewing managers and practitioners with experience implementing interventions for the homeless. The results of these studies were used to develop and test three questions to judge the plausibility of an intervention having effects on health equity.

The first paper, a Cochrane methodology systematic review, found 20 methodology studies of groups of systematic reviews which assessed effects on health equity. Half of these studies were published in the last two years (2008 and 2009), suggesting an increased interested on assessing effects of interventions on health equity. Of these 20 studies, only one study used an analytic approach to assess effectiveness between advantaged and disadvantaged populations. Of the nine methodological studies which described subgroup analyses, none provided sufficient detail to assess the credibility of subgroup analyses using the 1992 Oxman and Guyatt criteria [52] and Rothwell criteria [53]. There was a lack of description of whether differences in effects on health equity were due to differences in relative or absolute effects. Assessment of likely effects on health equity was described as challenging because of the need for content and methodological expertise by three studies which applied an equity checklist to systematic reviews to assess likely effects in low and middle income countries [69, 103, 104].

The second paper, a methodology study of a cohort of all systematic reviews indexed in MEDLINE in November 2004, found that 13% (29 out of 224) of systematic reviews of effectiveness assessed effects of interventions in disadvantaged populations. Of these, 15 systematic reviews reported effects in disadvantaged populations (7 used statistical approaches to subgroup analysis and 8 described subgroup analyses conducted in the primary studies). These subgroup analyses were reported in insufficient detail to judge their credibility; systematic reviews reported a median of three out of seven of the Oxman and Guyatt credibility criteria for subgroup analyses [52]. Another 14 systematic reviews were targeted at a vulnerable population, thus providing direct evidence on effects in a vulnerable population, but no data on whether these interventions affect the gap in health between advantaged and disadvantaged populations.

The third paper, a qualitative study of implementation factors associated with success of programs designed for the homeless in Ottawa identified 20 factors associated with successful programs and these mapped well onto the equity-effectiveness loop framework [20], suggesting the equity effectiveness framework's utility in identifying factors associated with effects of interventions on health equity.

The fourth study, which developed and tested the inter-rater reliability of an equity plausibility algorithm, found low inter-rater agreement in predicting the likelihood of interventions having different effects across sex or socioeconomic status. The finding that 67% of raters judged important differences likely across sex and socioeconomic status reinforces the need to develop, evaluate and promote methods for considering and reporting effects on health equity in systematic reviews.

Strengths and limitations

The four papers in this dissertation followed accepted guidelines to enhance their reproducibility and credibility. These strategies included double extraction of data using a pre-tested extraction form for the methodology study, peer-review of the search strategy and publication of the search terms [191], and assessment of risk of bias of included studies [51]. The qualitative study used constant comparative analysis with each subsequent informant to increase credibility [159]. The equity plausibility algorithm followed accepted steps for developing a checklist, including item generation, item reduction, pilot-testing and assessment of consistency and construct validity [172].

This dissertation focused on the methods of systematic reviews, and did not assess the methods used by primary research studies to assess effects on health equity. Two methodology studies found that assessment of effects across PROGRESS-Plus factors in primary studies is mainly used to control for confounders [6], and that analysis of effects in different populations is more frequent in primary studies than in systematic reviews [101]. This dissertation did not assess characteristics of primary studies, such as whether they were sufficiently powered to detect subgroup analyses.

Because the sample of systematic reviews for the methodology study on systematic reviews was based on all systematic reviews published in MEDLINE in November 2004, the findings may not be generalizable to systematic reviews of non-health interventions (e.g. Campbell Collaboration reviews) since these may be less likely to be indexed in MEDLINE.

The Schrecker and Labonte framework was used to identify how systematic reviews can provide evidence on reducing health equity. For example, the importance of social, financial, legal and educational interventions in reducing health inequities suggested by this framework guided the choice of databases for the Cochrane methodology review. Interpretation of the second study of this thesis in light of this conceptual model highlights the lack of consideration of legal, financial and educational interventions since 96% of the included systematic reviews were focused on clinical interventions. The link between how health services are delivered (e.g. accessibility, quality) and health equity proposed by this framework were explored in the third study using qualitative methods.

Implications for authors of systematic reviews

State the rationale for or against subgroup analyses, based on how the intervention is expected to work

The rationale for how an intervention is expected to work provides the basis for justifying subgroup analyses across different populations or settings. The fourth paper in this dissertation found that 67% of assessments using the equity plausibility algorithm questions predicted differences in relative effects across sex or socioeconomic status (Table 21). These findings suggest that users of systematic reviews are likely to expect different effects in disadvantaged populations.

Report subgroup analyses, if conducted, with sufficient detail to assess credibility using Oxman and Guyatt criteria and additional criteria as necessary

If subgroup analyses are conducted, they need to be reported with sufficient detail about the rationale, *a priori* hypotheses, both statistical and clinical significance, support from other indirect evidence and from both within and between study comparisons, to judge their credibility. Without sufficient detail, users of systematic reviews cannot judge the importance of these subgroup effects or assess their credibility.

Distinguish between relative and absolute effects

Greater improvement is expected with most interventions for vulnerable populations since the baseline risk of adverse health outcomes is higher for most diseases in vulnerable populations [11]. This dissertation found that subgroup analyses reported in systematic reviews did not describe whether differential effects in different populations were due to differences in absolute or relative effects. A lower relative risk reduction in vulnerable populations suggests the potential for intervention generated inequalities due to differences in the mechanism of action [112]. Whereas, a higher absolute effect suggests that an intervention has the potential to reduce health inequalities in a population.

Discuss applicability to vulnerable populations using a published checklist

Assessment of applicability has the advantage of reducing needless duplication of studies in different populations. All three methodology studies that judged likely equity effects using applicability assessment were published in the last year, suggesting an increased interest in using applicability assessment to judge equity effects. The fourth study in this dissertation showed little to no agreement between multiple raters about the likelihood of different effects across sex or socioeconomic status. Authors of systematic reviews have content and methodological expertise as well as access to the primary studies which may facilitate the judgment of applicability.

Describe populations and settings of studies to facilitate judgments of applicability

Details of the populations included in the systematic review facilitate judgments of applicability. The second study found that the sociodemographic details of the included studies were reported by 68% of systematic reviews, but most of these reported sex and country setting. Socioeconomic status was reported in only 6 out of 224 systematic reviews. Improved reporting of sociodemographic details facilitates judgments about applicability to different populations and settings [58].

Consider comparing effects in pragmatic vs. explanatory trials

Pragmatic trials are more similar to real-world conditions, by allowing flexibility in the intensity of follow-up, more variability in patients (e.g. more co-morbidities), practitioner compliance and patient adherence and other factors [44]. These studies may be more likely to find smaller differences in relative effects, relative to the control group because the intervention is not delivered to the patients as well as it could be or because patient co-morbidities confound the effects. Pragmatic trials are also more likely to include people from vulnerable populations, and may be more relevant to decision-makers who must consider the distribution of effects in a population.

Recommendations for future research

Test whether systematic reviews with pragmatic questions are more likely to have credible subgroup effects

Despite finding only 12 systematic reviews with a pragmatic question (i.e. that allowed flexibility in patients eligibility, adherence, practitioner compliance, follow-up, practitioner skill) in the cohort of 224 systematic reviews, pragmatic questions were statistically significantly more likely to assess effects on health equity (odds ratio 6.6, 95% confidence interval 2.0-22.0). A methodological study of systematic reviews with a pragmatic focus would be useful to confirm this finding.

Assessment of subgroup analyses in primary studies

Lack of conduct and reporting of subgroup analyses in systematic reviews may be due to lack of data in primary studies. Furthermore, primary studies may be underpowered to detect subgroup differences. Methodological assessment of reporting and assessment of effects on health equity in primary studies is needed to support guidelines on assessing health equity for reporting and conducting primary research.

Compare methods of assessing subgroup analyses- IPD vs. within study

Subgroup analyses reported in the systematic reviews assessed as part of the second study were limited to describing within-study comparisons or analyses of study-level differences. Assessment of differences in effect across patient characteristics which vary within a study can be more accurately assessed with individual patient data, and has shown that injury prevention strategies are not likely to increase health inequalities [122, 192]. Methodological studies comparing the feasibility and credibility of different approaches to subgroup analyses are needed to inform guidance on subgroup analyses.

Develop and test standardized applicability checklists for systematic review authors

There is a need for methodological guidance on assessing applicability and generalizability in systematic reviews. It is unknown the extent to which applicability is assessed in systematic reviews at the moment, and how these assessments are made. This dissertation found that only 30/224 systematic reviews discussed applicability to vulnerable populations, and none of these reported using a published checklist for applicability. Without an assessment of whether the results apply to different populations by the authors, the user of the systematic review must make this judgment. Evidence from this dissertation suggests that users, clinicians and methodologists are unreliable in making this judgment about applicability.

There is a need to: i) assess the extent to which applicability is assessed in systematic reviews and how the judgments are made; ii) assess how users of systematic reviews make judgments about applicability when there is no such assessment in the systematic review; iii) develop methods for assessing applicability with clearly defined items that achieve good inter-rater agreement for raters with content and methodological expertise; iv) assess acceptability to users (decision-makers, managers and consumers) of systematic reviews with applicability assessments to those without.

Update reporting guidelines for systematic reviews on subgroup analyses and applicability

The EQUATOR network has shown that reporting guidelines improve the conduct of clinical trials and systematic reviews [193]. The Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) contain two items related to subgroup analyses (item 16 and 23 which state that results of subgroup analyses should be reported if done, and should specify if these were specified *a priori*) and one item on interpretation of results (item 26 which states that a general interpretation of results should be given but does not describe applicability) [194, 195]. Updating of the PRISMA reporting guidelines on subgroup analyses could improve the ability of users of systematic reviews to judge the credibility of subgroup analyses by promoting the reporting of all of the Oxman and Guyatt criteria (e.g. clinical and statistical importance), as well as promote the distinction between relative and absolute effects, and consider other criteria for subgroup analyses [53]. There is no item in PRISMA which relates to assessment of applicability or generalizability.

Dissemination and diffusion of results

The results of this thesis are relevant to organizations and individuals that produce systematic reviews, organizations that fund the production of systematic reviews, and users of systematic reviews. Results will be published in traditional journals that are likely to reach these audiences, such as the British Medical Journal, Public Library of Science (PLoS) Medicine and the Journal of Clinical Epidemiology. Results will also be presented at the joint Cochrane and Campbell Colloquium in 2010 to disseminate to authors and users of Cochrane and Campbell systematic reviews.

Results of this thesis will be further disseminated by including them in a forthcoming chapter on health equity in the Cochrane Handbook guidelines for the conduct of systematic reviews. The results will also be discussed with authors of the PRISMA guidelines to design and conduct methodological studies to support the updating of PRISMA guidelines on subgroup analyses and applicability [194].

Contribution to Population Health

One of the goals of population health is to assess the impact of interventions and policies on reducing health inequalities. As described in the introduction, systematic reviews can contribute to the evidence-base on reducing health inequalities in four ways by providing evidence on the effects of: 1) reducing social stratification (e.g. by education and income transfers), 2) reducing differential exposure and vulnerability, 3) reducing differential consequences of ill health, and 4) improving health system characteristics. To inform the evidence-base on reducing health inequalities, systematic reviews must also assess the distribution of effects in disadvantaged populations across PROGRESS-Plus using one of four methods: gap (comparing effects in two groups), gradient (assessing effects across the socioeconomic gradient), targeted (assessing effects in disadvantaged populations) or judging applicability of interventions to disadvantaged populations that may or may not have been included in the trials.

This dissertation contributes to the field of population health by describing and critically appraising methods used in systematic reviews to assess the effects of interventions on health inequalities. This dissertation also developed and assessed the validity of an equity plausibility algorithm for predicting the likelihood of differences in effects across PROGRESS-Plus factors. This dissertation proposes six recommendations for authors of systematic reviews on how to improve systematic reviews as a basis for making decisions that aim to reduce health inequalities. These recommendations focus on improved reporting of the proposed mechanism of action of the intervention, reporting sociodemographic characteristics of participants in primary studies, reporting subgroup analyses in sufficient detail that their credibility can be appraised and assessing applicability using published checklists.

Figure 1: Conceptual model for how systematic reviews can contribute to evidence-base on health equity

		Outcome			Outcome	
Low SES		**yes**	**no**	**High SES**	**yes**	**no**
	intervention	a	b	**intervention**	A2	B2
	control	c	d	**control**	C2	D2

Relative risk	**Low SES**	**High SES**
	(a/a+b)/(c/c+d)	(A2/A2+B2)/(C2/C2+D2)

Three ways that systematic reviews demonstrate effects on health equity, with three hypothetical examples:

4. **Difference in relative effects (e.g. greater effect of tobacco price increases on people with low SES)**

	Low SES	**High SES**
	Relative risk: 1.5	Relative risk: 1.1

5. **Same relative risk, difference in absolute effects (e.g. lipid lowering drugs may have greater effects in low SES if they have more comorbidities and a higher risk of stroke)**

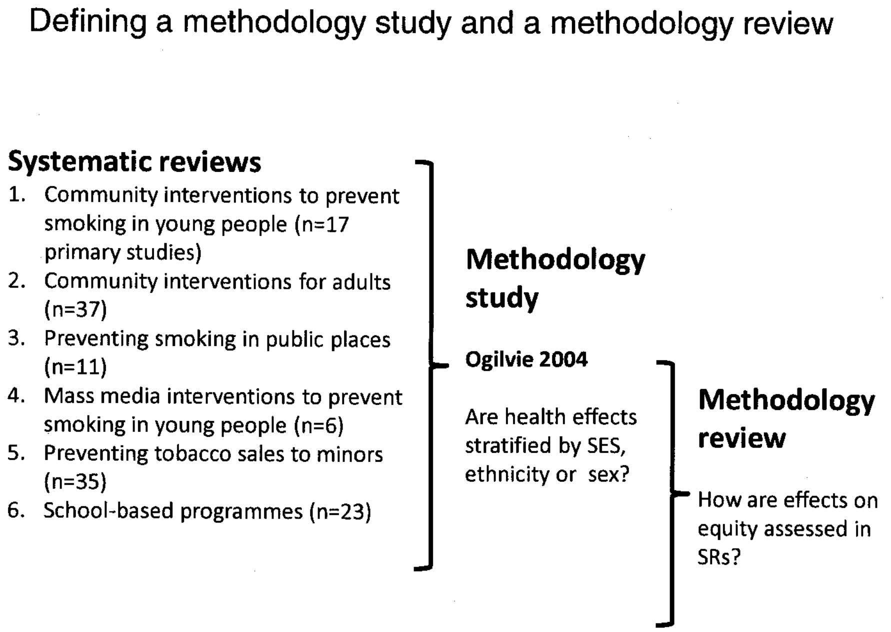
Absolute risk reduction= relative risk * baseline event rate

	Low SES	**High SES**
Baseline 5 year risk of stroke:	5% of people (5 in 100)	1% of people (1 in 100)
Same relative risk (RR):	RR: 1.5	RR: 1.5
Absolute risk reduction:	7.5% (7.5 in 100)	1.5% (1.5 in 100)

6. **Differences in real-world effects due to access, diagnosis, provider compliance or patient adherence (e.g. bicycle helmets for preventing head injuries)**

	Low SES	**High SES**
Same relative risk in studies	RR: 1.5	RR: 1.5
Same baseline injury rate	5%	5%
Less uptake of helmets by low-income families	real-world RR: 1.0	RR:1.5

Figure 2: Example of a methodology study on community interventions to prevent smoking (Ogilvie 2004) [49]



Defining a methodology study and a methodology review

Systematic reviews

1. Community interventions to prevent smoking in young people (n=17 primary studies)
2. Community interventions for adults (n=37)
3. Preventing smoking in public places (n=11)
4. Mass media interventions to prevent smoking in young people (n=6)
5. Preventing tobacco sales to minors (n=35)
6. School-based programmes (n=23)

Methodology study

Ogilvie 2004

Are health effects stratified by SES, ethnicity or sex?

Methodology review

How are effects on equity assessed in SRs?

Figure 3: Globalization and social determinants of health [21]

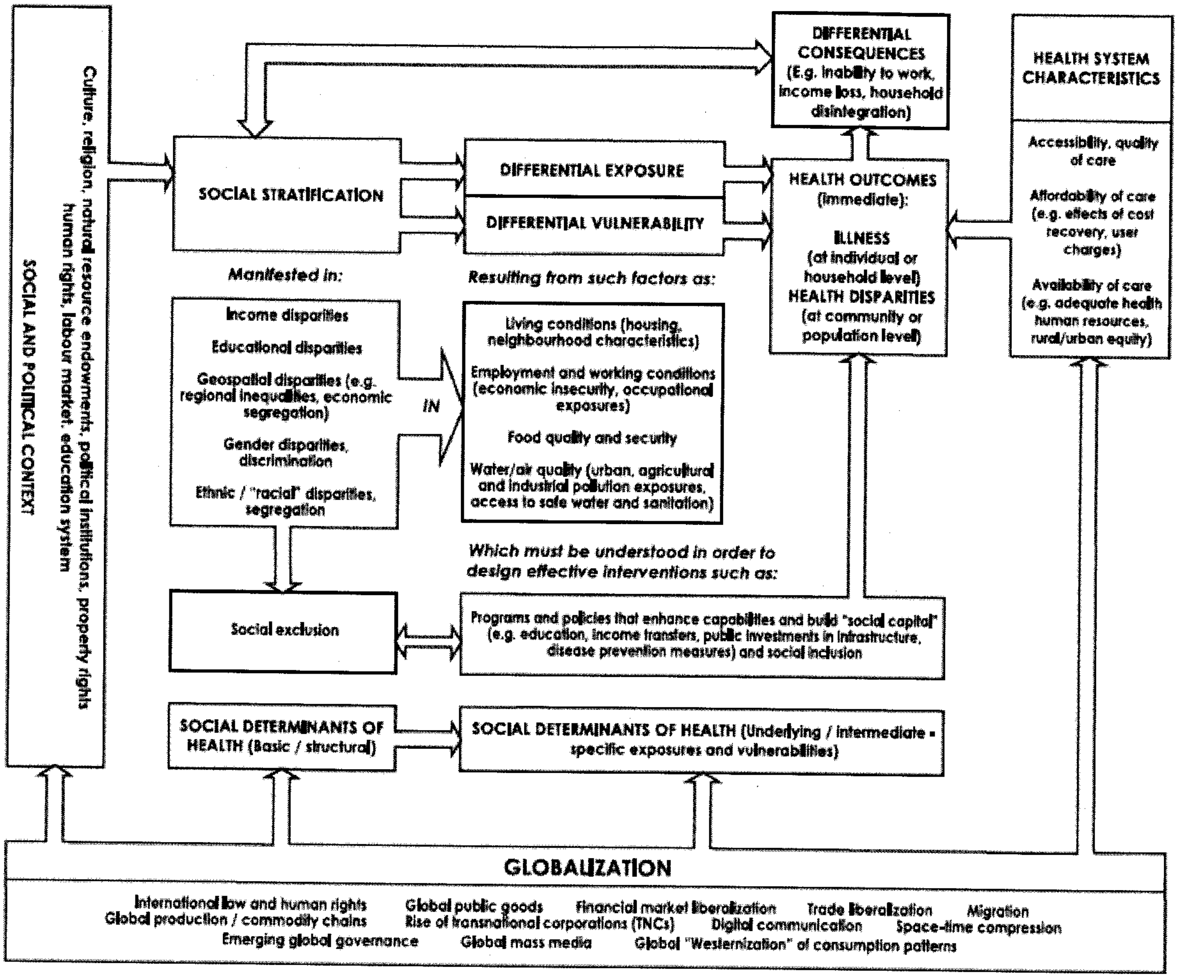
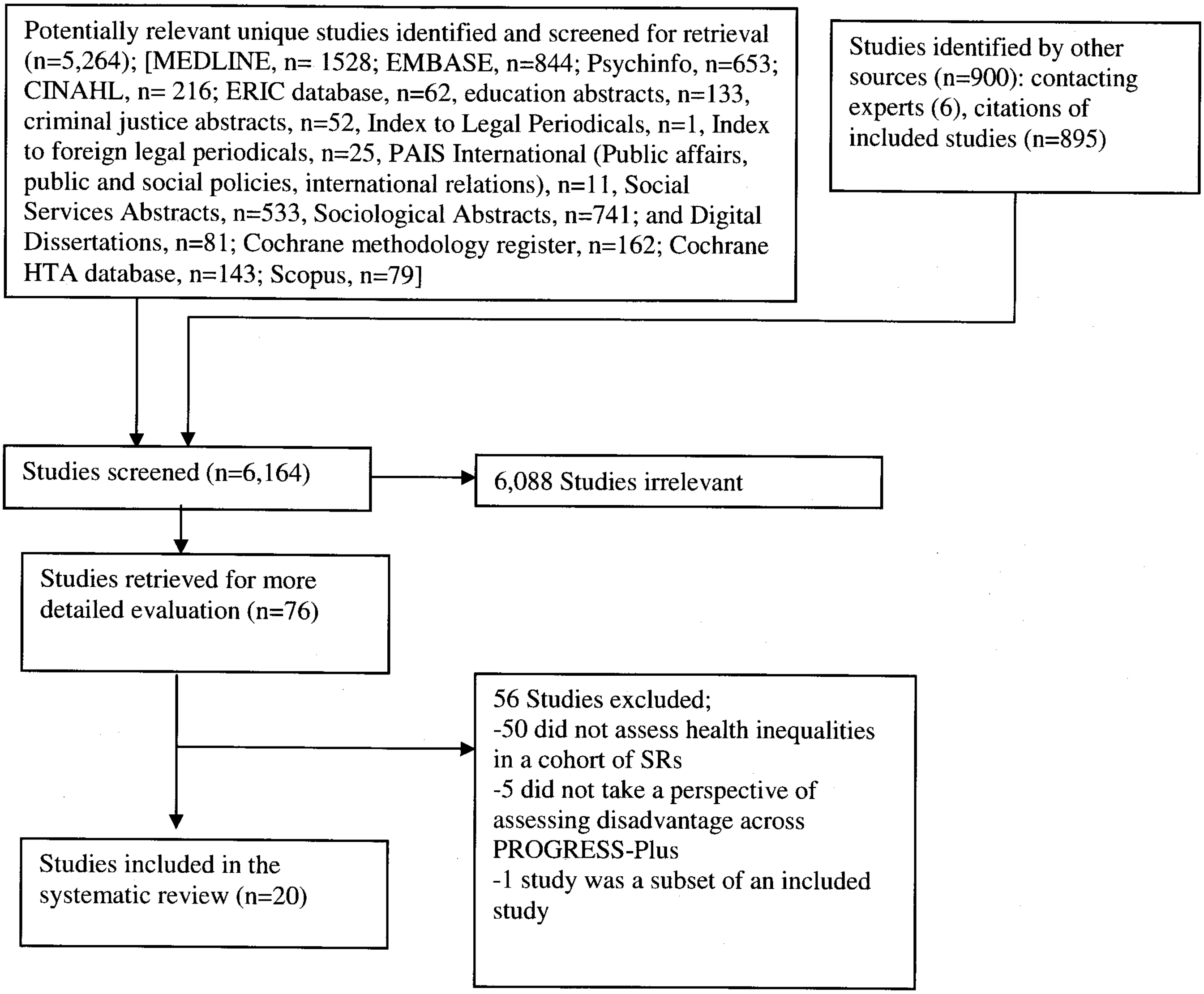


Figure 4: Flow of methodology studies included in Cochrane methodology review



Potentially relevant unique studies identified and screened for retrieval (n=5,264); [MEDLINE, n= 1528; EMBASE, n=844; Psychinfo, n=653; CINAHL, n= 216; ERIC database, n=62, education abstracts, n=133, criminal justice abstracts, n=52, Index to Legal Periodicals, n=1, Index to foreign legal periodicals, n=25, PAIS International (Public affairs, public and social policies, international relations), n=11, Social Services Abstracts, n=533, Sociological Abstracts, n=741; and Digital Dissertations, n=81; Cochrane methodology register, n=162; Cochrane HTA database, n=143; Scopus, n=79]

Studies identified by other sources (n=900): contacting experts (6), citations of included studies (n=895)

Studies screened (n=6,164)

6,088 Studies irrelevant

Studies retrieved for more detailed evaluation (n=76)

56 Studies excluded;
-50 did not assess health inequalities in a cohort of SRs
-5 did not take a perspective of assessing disadvantage across PROGRESS-Plus
-1 study was a subset of an included study

Studies included in the systematic review (n=20)

Note: PROGRESS: Place of residence (urban/rural), Race/ethnicity, Occupation, Gender, Religion, Education, Socioeconomic status and Social Capital; SR: systematic review; ERIC- Education Resources Information Center

Figure 5: Equity-effectiveness loop framework

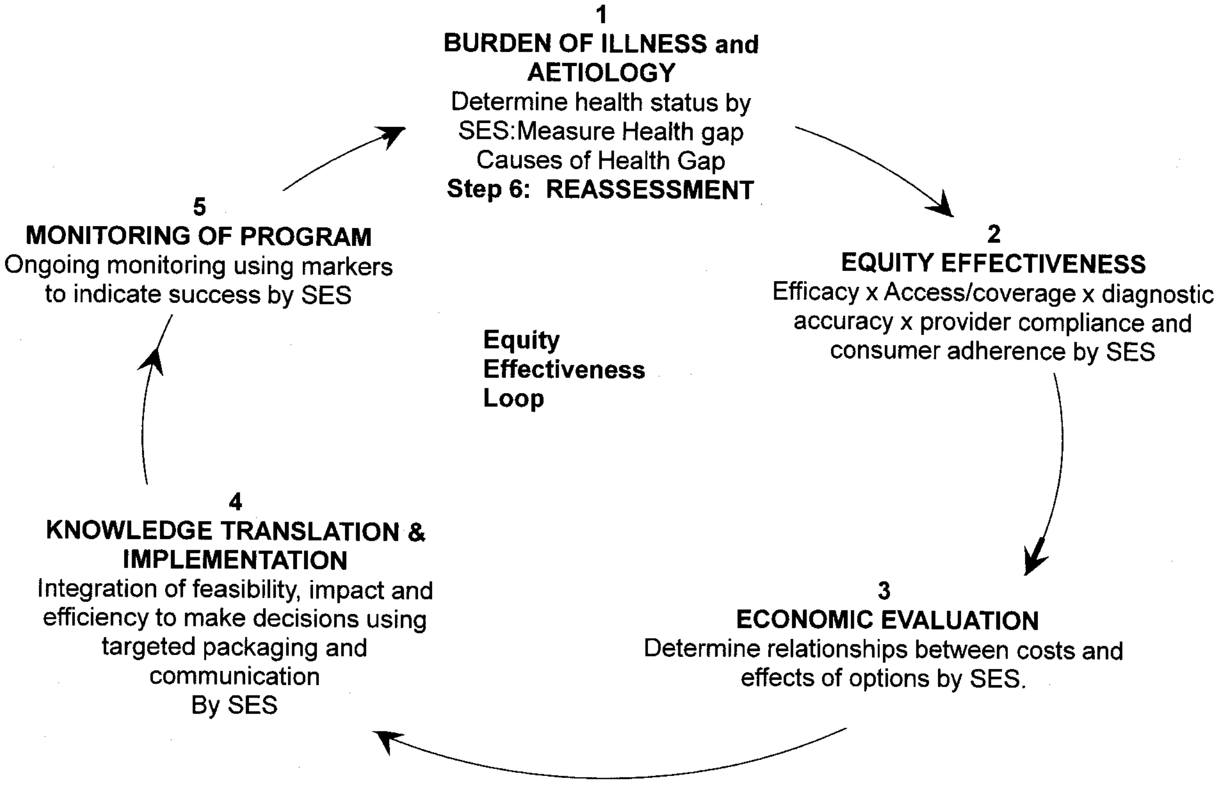


Table 1: Methods of measuring health inequalities

Method	Calculation	Example: infant mortality rate per 1000	Result
Rate ratio	Lowest/reference group	Low SES: 36 per 1000	2.4
Low to high ratio	Lowest/highest	High SES: 10 per 1000	3.6
Rate difference	Lowest-highest	Reference group= population average: 15 per 1000	26 deaths per 1000
Shortfall (from norms)	Lowest-norm		21 deaths per 1000

SES: socioeconomic status

Table 2: Applicability and transferability checklists

Authors	Applicability questions
Dans 1998, User's Guides to the Medical Literature [196]	Biologic (1) Are there pathophysiologic differences in the illness under study that may lead to a diminished treatment response? (2) Are there patient differences that may diminish the treatment response? Social and economic (3) Are there important differences in patient compliance that may diminish the treatment response? (4) Are there important differences in provider compliance that may diminish the treatment response? Epidemiologic (5) Do my patients have comorbid conditions that significantly alter the potential benefits and risks of the treatment? (6) Are there important differences in untreated patients' risk of adverse outcomes that might alter the efficiency of treatment?
Glasziou 1998 [197]	Is my patient **SO** different that results do not apply? Is treatment feasible in my setting? What are likely benefits and harms? How will my patients' values influence the decision?
Glasziou 2000 [55]	**Applicability defined as**: applying results from studies to individual What are predicted absolute risk reductions? Do the benefits outweigh the harms- absolute + net benefits and costs, strength of preferences **Transferability defined as**: applying results to wider population- infer/predict /project benefits and harms variations in treatment effect (e.g. subgroups analyses) how does treatment effect vary with baseline risk?
Briss 2000 Community guide to preventive services [198]	Define target populations and setting Assess whether target population included in studies Assess whether population in included studies is representative of target population Judgment about whether intervention works better or worse
Rothwell 2005 [90]	Setting of trial Selection of patients Characteristics of patients Differences between trials and routine clinical care Outcome measures and follow-up Adverse effects of treatment
Wang 2005 [199]	If implemented, will effectiveness be the same? baseline prevalence characteristics of population capacity to implement Whether the intervention could be implemented: political environment social acceptability

Authors	**Applicability questions**
	cultural adaptability Resource implications Educational level Organizational structure
Green and Glasgow 2006 [200]	Reach and representativeness Program or policy implementation and adaptation Outcomes for decision making (e.g. moderator effects, costs) Maintenance and institutionalization
AHRQ methods guide 2007	Population Intervention intensity and quality Comparator (and dosing) Outcomes Timing of follow-up
CONSORT 2001 [201]	External validity, also called generalizability or applicability, is the extent to which the results of a study can be generalized to other circumstances (181). Patient characteristics (e.g. age, sex, severity of disease, and comorbid conditions?) Intervention (e.g. drugs within a class of similar drugs, different dosage, timing and route of administration) Comparison (e.g. different concomitant therapies) Setting (e.g. primary, secondary, and tertiary levels of care) Outcomes (e.g. effect on related outcomes that were not assessed in the trial, and the importance of length of follow-up and duration of treatment)
Cochrane Handbook 2008 [171]	Biologic variation Variation in context and culture Variation in adherence Variation in values and preferences
Lavis [202]	Could it work? structural elements Will it work (or what would it take to make it work)? E.g. power dynamics, other priorities/realities Is it worth it? Balance of benefits and harms worth the incremental costs?
Dans 2007 [203]	Do the public health recommendations in the guidelines address a priority problem for disadvantaged populations? Is there a reason to anticipate different effects of intervention in disadvantaged and privileged populations? Are the effects of the intervention valued differently by disadvantaged compared with privileged populations? Is specific attention given to minimizing barriers to implementation in disadvantaged populations? Do plans for assessing the impact of the recommendations include disadvantaged populations?
PRISMA 2008 [195]	*"Item 24: Summary of evidence* Summarise the main findings, including the strength of evidence for each main outcome; consider their relevance to key groups (such as healthcare providers, users, and policy makers)… Although there is no standard way to assess applicability simultaneously to different audiences, some systems do exist."

Authors	**Applicability questions**
EUROTHINE – judging transferability to the Netherlands setting [204]	Description of results of intervention study -method, internal validity, -observed effects in lower SES groups Potential relevance of study to the Netherlands -potential for effect on health inequalities, -relationship to other programs, -evidence needed for evaluation Implementation of intervention in the Netherlands - experiences with implementation, -possibilities to adjust intervention to the Dutch situation Effectiveness in Dutch context -mechanisms that explain observed effect, -differences between Netherlands and study population in relevant factors, -observed variations in effectiveness according to population, place, period -estimation of effectiveness in the Netherlands
Dans 2008, Updated JAMA Users Guides [58]	A. Can I apply the results to my patients? 1. Have biologic factors that might modify the treatment response been excluded? (across sex, comorbidities, race, age and pathology) 2. Can the patients comply with treatment requirements? 3. Can the health care providers comply with treatment requirements? B. Are the benefits worth the risks and costs?

Table 3: Selected methods of assessing effects on health inequalities

Method	Calculation
Targeted approach	Evaluation of effect size in the disadvantaged population only (e.g. Cochrane Review on community animal health services for improving household wealth and health status of low income farmers by [205].
Relative difference (gap approach)	(advantaged - disadvantaged)/advantaged
Absolute difference (gap approach)	advantaged – disadvantaged
Gradient-approach regression	Regression-based index of relative effect across incremental categories of disadvantage.
Gradient-concentration index	Twice the area between the concentration curve and the line of equality (45 degrees line), defined with reference to the concentration curve, which graphs health status on the y-axis against categories of disadvantage on the x-axis (World Bank).
Gradient or gap-benefit incidence	Computes the distribution of public expenditure across different PROGRESS-PLUS groups according to actual utilization of services.
Gradient approach - Gini index	Measure of inequality of income distribution, defined as the area between the line of equality and the Lorenz curve, with categories of PROGRESS-Plus on the x-axis and percentage of total income on the y-axis [206].

Table 4: Characteristics of Methodology Studies excluded from the Cochrane methodology review

Study	Reason for exclusion
Gulmezoglu 1997 [91]	No focus on health equity. Assessed effects of interventions to prevent impaired fetal growth.
Barlow 2004 [92]	No focus on health equity. Included children with chronic disease, but did not make any judgment that these children were disadvantaged or that these interventions could affect health equity.
Espinosa 2007 [93]	No focus on health equity, clinical guidelines on elderly that included 8 SRs
Craig 2003 [94]	No focus on health equity. Assessment of effects of treatments for sexual offenders from SRs.
Gaes 1999 [95]	No focus on health equity. Effectiveness of correctional rehabilitation from SRs.

Table 5: Characteristics of methodology studies included in Cochrane methodology review

Author, year	PROGRESS-Plus dimension	Rationale for assessing PROGRESS-Plus dimension	Gap, gradient or targeted	Number SRs or MAs included	Equity considered and definition	Is judgment of equity made, if so, how?
Nasser 2007 [121]	Place of residence, SES, LMIC	different prevalence in LMIC; cost-effectiveness more important in LMIC, challenges in implementation may be different in LMIC than HIC	Gap	420 Cochrane SRs	not defined nor discussed	Not done
Tudiver 2008 [106]	Gender and sex based analysis	Research shows sex and gender are relevant in cardiovascular disease risk factors, but quality of evidence remains weak for many interventions for women	Gap	38 Cochrane SRs	not defined	not done
Morrison 2004 [111]	Gender, SES and age	SES, gender and age assessed as potential effect modifiers for interventions aimed at increasing adherence	Gradient	12 SRs, MAs or quantitative overviews	not defined	not done
Tugwell 2008 [101]	PROGRESS	Need for evidence on what works to reduce inequalities across socioeconomic strata	Gap	14 Cochrane SRs	Health equity defined as: 1) effective in people who are disadvantaged or; 2) at least as effective in low SES as in high SES	fairness and avoidability referenced, but not judged by authors
Ogilvie 2004 [207]	"any socio-demographic variable", not further defined, but later mentions age, sex and SES	Reducing social inequalities in a political priority- is evidence available that tobacco control policies will help achieve this?	Gap	6 Cochrane SRs	not defined, review intends to assess "social distribution of intervention effects", "social inclusiveness or targeting of interventions	no judgment on fairness, all inequalities across social dimensions are assessed
Johnson 2003 [208]	Gender	Imperative that the practice of medicine be based on high-quality evidence, including evidence on women. CVD is number one killer of women, yet clinical trials performed predominantly in men	Gap	31 Cochrane SRs	not defined, if study assessed results broken down by gender, then considered gender-based analysis	gender-based analysis if: 1) analysis of results broken down by gender; or 2) textual comparisons of results in men and women

Author, year	PROGRESS-Plus dimension	Rationale for assessing PROGRESS-Plus dimension	Gap, gradient or targeted	Number SRs or MAs included	Equity considered and definition	Is judgment of equity made, if so, how?
Tsikata 2003 [6]	PROGRESS-Plus	Need for evidence on what works to reduce inequalities across socioeconomic strata	Gap	95 Cochrane SRs	Health inequities defined as avoidable and unfair inequalities in health, across SES	All inequalities across PROGRESS were assessed
Adamek 2008 [39]	Plus- Age	Many older people suffer needlessly from debilitating mental health and addiction disorders because they are misdiagnosed due to discrimination based on age	Targeted	16	Equity not defined	Misdiagnosis of older people with depression is described as "needless"
Bartels 2003 [40]	Plus- age and mental health	Older adults with mental illness receive poorer quality of care (relative to younger people with mental illness and older people without mental illness) due to likelihood of more adverse effects and smaller magnitude of benefit for older adults with mental illness because of cognition, physiological and social functioning changes	Targeted	23	Equity not defined	Unmet need for age-appropriate evidence-based practice geriatric mental health care
Althabe 2008 [104]	Plus- LMIC	Need to appraise applicability and relevance of quality improvement strategies in LMIC to reduce health inequalities in maternal and child health	Targeted	23	Equity not defined, but authors used the SUPPORT "equity checklist"	Authors judged equality, defined as equitable access to the strategies for the most disadvantaged health providers, and for the most vulnerable women and children
Viswanathan 2008 [209]	Race, ethnicity	Persistent disparity in health outcomes by race suggests a role for further research on interventions that narrow the gap	Gap	64	Equity not defined.	Disparities in health outcomes remain after adjusting for medical risk
Stewart 2006 [108]	Plus- youth with disabilities	Youth with disabilities do not have same outcomes as peers for health, achievement and employment.	Targeted	5	Equity not defined	Authors describe worse outcomes in youth with disabilities than their peers
D'Souza 2004 [109]	SES	Poor health outcomes are concentrated in socially disadvantaged mothers	Targeted	5	Equity not defined	poor health concentrated in low SES women

Author, year	PROGRESS-Plus dimension	Rationale for assessing PROGRESS-Plus dimension	Gap, gradient or targeted	Number SRs or MAs included	Equity considered and definition	Is judgment of equity made, if so, how?
Ball 2002 [110], Ball 2004 [98]	SES, ethnicity, occupation	to assess which interventions have the greatest positive impact on prevention of low birth weight which is concentrated in families living in poverty and deprivation	Targeted	19	Equity not defined	Health inequalities described as deeply entrenched and generational.
Main 2008 [102]	SES	Reducing social inequalities in smoking and its health consequences is a public health priority	Gap	19	Authors mention applying an "equity lens"	"equity lens" to assess whether systematic reviews can be used to answer questions about reducing health inequalities across SES, ethnicity education
Browne 2004 [210]	SES, ethnicity, gender, age	Policy initiatives need to be developed to ensure sufficient funding and promote delivery of effective programs to appropriate children at risk of poor mental health	Targeted	23	Health equity not defined	Absence of resources and opportunities described as risk factor for mental health problems
Lewin 2008 [103]	Plus-LMIC	To reach Alma Ata goal of health for all, need to assess which health systems interventions are applicable and relevant in LMIC	Targeted	20	Health equity not defined, mention "equitable" distribution of health services	Inadequate delivery of cost-effective interventions is both unfair and avoidable
Odierna 2009 [100]	SES, ethnicity	Health policies need to be evaluated to ensure that they are not unintentionally increasing health inequalities	Gap	32	Health equity defined as eliminating avoidable inequalities, particularly those that result from injustice or social exclusion.	Health equity requires improving health outcomes for all sectors of the population, not benefitting only the advantaged
Bambra 2009 [97], and Bambra 2009 [97]	SES, gender, ethnicity, age	to identify evidence on interventions to reduce health inequalities by acting on social determinants of health	Gap	30	Health equity not defined	No judgment by authors
Chopra 2009 [69]	LMIC	Quality health care depends on sufficient health workers to deliver the care (e.g. in remote areas), policy makers need evidence on effects of policy options on equitable health care	Targeted	28	Health equity not defined: lack of health workers in remote/rural areas described as inequitable	inequitable distribution of health workers limits quality health care

Note: PROGRESS: Place of residence (urban/rural), Race/ethnicity, Occupation, Gender, Religion, Education, Socioeconomic status and Social Capital; SR: Systematic Review; LMIC: Low and Middle Income Countries; HIC: High income countries; MA: meta-analysis; SES: socioeconomic status

Table 6: Potentially eligible ongoing methodology studies

Title	Source
1) Comparative Effectiveness Methods – Clinical Heterogeneity	**http://effectivehealthcare.ahrq.gov/healthInfo.cfm?infotype=rr&ProcessID=93**
2) Addressing equity in HTA production, an analysis of equity aspects within current HTA domains and possibilities for creating equitable assessments- Dimitra Panteli, Reinhard Busse, Technische Universitat Berlin, 2009	**http://www.htacademy.de/htacademy/Panteli_Project_Proposal.pdf**

Table 7: Quality assessment of methodology studies include in the Cochrane methodology review

Author, year	Objective	Selection bias	Detection bias
Nasser 2007 [121]	To identify Cochrane reviews that are relevant to developing countries; and to determine how they tackled the developing country setting	Low risk of bias; all SRs that mentioned LMIC in title, abstract or text were included	High risk of bias; methods for data extraction not described
Tudiver 2008 [106]	Determine whether and how 'sex' and 'gender' are addressed in a sample of Cochrane systematic reviews in cardiovascular disease.	Low risk of bias: random sample of 1/3 of reviews from Cochrane heart, hypertension and peripheral vascular disease review groups	Low risk of bias: data extracted using pre-tested form by 1 research assistant
Morrison 2004 [111]	To compile quantitative reviews of studies of adherence; to critique this literature and summarize current knowledge of adherence	Low risk of bias; Any SR, MA or quantitative overview that assessed adherence with prescribed medications; aimed at patients; more than 1 included study	Unclear risk of bias, data extraction not described
Tugwell 2008 [101]	To apply the "equity lens" to Cochrane reviews of rheumatoid arthritis	Low risk of bias. All systematic reviews of the musculoskeletal review group published since Issue 1, 2003	Low risk of bias; double extraction by 2 reviewers with pre-tested form
Ogilvie 2004 [207]	To assess the potential contribution of evidence from existing systematic reviews of effectiveness to answering the question: what works in reducing social inequalities in smoking?	Low risk of bias; all completed reviews of the effectiveness of community-base tobacco control interventions	Unclear risk of bias; Data extraction methods not described but data to be collected were described
Johnson 2003 [208]	To assess whether Cochrane systematic reviews on cardiovascular disease handled gender differences and whether the data pertaining to treatment of CVD is applicable to the clinical care of women	Low risk of bias; Included all completed systematic reviews in Cochrane Heart group, Hypertension group, peripheral vascular diseases group.	Unclear risk of bias; Method of extraction not described but details of data to be extracted were provided
Tsikata 2003 [6]	To determine whether Cochrane reviews report and analyze the data needed to assess the effectiveness of interventions at reducing health inequities	Low risk of bias; Random sample of 10% of systematic reviews published between issue 1, 2000 to issue 2, 2003 in the Cochrane Library; stratified by review group (n=42; 7 Cochrane review groups excluded because <5 reviews)	Low risk of bias. Data extraction was done by 2 reviewers, using a pre-tested form; discrepancies resolved by discussion
Adamek 2008 [39]	To review the effectiveness of interventions for older people with mental health and addiction disorders, who are unfairly under-diagnosed and under-treated	Low risk of bias: Comprehensive search for reviews focusing on psychosocial or psychological interventions for people 50 years of age or older	Not clear, methods for extracting details from SRs were not described

Author, year	Objective	Selection bias	Detection bias
Althabe 2008 [104]	To systematically analyse the results of systematic reviews of strategies for improving the quality of care, where these strategies are relevant to maternal and child health (MCH) in developing countries.	Low risk of bias: 2 independent reviewers assessed SRs against explicit inclusion criteria	Low risk of bias: 2 independent reviewers extracted data using explicit extraction form
Viswanathan 2008 [209]	To summarize maternal health research priorities, map these priorities to existing reviews, identify gaps that can be addressed with systematic reviews, including racial disparities	Low risk of bias, search of medline using MESH delivery, obstetric, systematic reviews, relevant to MCH	Not clear, authors state "each study was reviewed to establish nature of intervention, primary outcome and subanalysis of racial disparities"
Stewart 2006 [108]	1. What factors help or hinder the process of transition to adulthood for youth with disabilities? 2. What service delivery methods have been used?	Low risk of bias: pre-specified inclusion criteria to identify systematic reviews	Low risk of bias: extracted details using Critical Appraisal Skills Program
D'Souza 2004 [109]	To review evidence on improving perinatal outcomes for disadvantaged women	Low risk of bias: systematic search in 8 electronic databases for systematic reviews and studies on 10 different subgroups of disadvantaged women	Not clear: methods for data extraction not described
Ball 2004 [98]**, Ball 2002** [110]	To review effectiveness of interventions to prevent low birth weight with focus on relevance to socially disadvantaged women	Not clear- inclusion criteria not clear – article described as "review of reviews"- "papers related to low birth weight"	High risk of bias- no description of how articles were selected
Main 2008 [102]	To review effectiveness of population-level tobacco control interventions to reduce social inequalities	Low risk of bias: pre-defined inclusion criteria to identify all SRs with details on sociodemographic characteristics of participants	Low risk of bias: 2 independent reviewers screened abstracts, extracted data and assessed quality
Browne 2004 [210]	To review effectiveness of interventions to improve mental health for children	Low risk of bias: all systematic reviews identified by systematic search using predefined inclusion criteria: mental health promotion initiatives for children	Low risk of bias: used critical appraisal tool to extract data
Lewin 2008 [103]	To summarize evidence from systematic reviews that have the potential to improve delivery of cost-effective interventions in primary health care in LMIC	Low risk of bias: systematic search with prespecified inclusion criteria,	Low risk of bias: 2 independent reviewers screened for inclusion and extracted data and assessed quality using forms
Odierna 2009 [100]	To assess racial/ethnic, gender, and socioeconomic status (SES) concordance between Medicaid populations and studies synthesized in Drug Effectiveness Review Project (DERP) systematic reviews	Low risk of bias: all drug effectiveness reviews from 2004-2007 were included	Low risk of bias: one reviewer extracted data, and this was verified by a second reviewer
Bambra 2009 [97] **and Bambra 2009** [97]	To identify evidence on interventions to reduce health inequalities by acting on social determinants of health	Low risk of bias: systematic search of electronic databases + handsearching	Low risk of bias: 2 independent reviewers screened titles and extracted specific data

Author, year	Objective	Selection bias	Detection bias
Chopra 2008 [69]	To assess effects of policy options on equitable distribution of health workers in LMIC	Low risk of bias: systematic search of electronic databases with inclusion criteria + handsearching	Low risk of bias: used structured forms, 2 reviewers extracted data
Bartels 2003 [40]	To assess geriatric-specific evidence-base for mental health care	Low risk of bias: Systematic search of three electronic databases, with specific inclusion criteria (geriatric specific guidelines, evidence reviews and meta-analyses)	High risk of bias: No description of how data was extracted or by who

Note: SR: Systematic Reviews; MAs: meta-analyses, MCH- maternal and child health, LMIC: low and middle income countries

Table 8: Methods used in methodology studies to assess whether health equity was considered in systematic reviews

Methods used to assess health equity effects	Which studies used this method? (author year)	Data availability	Advantages	Disadvantages
1a.Descriptive –SRs mention PROGRESS-Plus in introduction, objectives, discussion, implications	Tudiver 2008 [99], Nasser 2007 [121], Lewin 2008 [103],	Gender (9/38 SRs), LMIC (6/20 SRs),	Indicates whether authors of systematic reviews have considered health equity	Does not assess effects on health equity
1b.Descriptive-SRs describe population across PROGRESS-Plus factors	Nasser 2007 [121], Tudiver 2008 [106], Tugwell 2008 [101], Ogilvie 2004 [207], Tsikata 2003 [6], Lewin 2008 [103],	2 studies did not report data availability [121, 207]; For the other 3 studies, PROGRESS-Plus data was available for: Place of residence (5/95 SRs); race/ethnicity (7/95 SRs); occupation (1/95 SRs); gender (84/147 SRs); religion (1/95 SRs); education (0/95 SRs); SES (4/95 SRs); social capital (0/95 SRs), LMIC (13/58 SRs reported >1 study in LMIC)	Provides direct data on whether different populations included in SRs which is useful for judging applicability	Does not analyze influence of population characteristics or setting on effects on health inequalities Data available for gender in 57% of SRs, others are available in less than 25% of SRs
1c.Descriptive-SR describes if intervention is given only to disadvantaged populations across PROGRESS-Plus	Nasser 2007 [121], Ogilvie 2004 [207], Tsikata 2003 [6]; Main 2008 [102], Adamek 2008 [39], Stewart 2006 [108], D'Souza [109], Ball 2004 [98], Browne 2004 [210]; Tugwell 2008 [101]; Bartels 2003 [40]	Data not reported for 2 studies [207], [121]; 17/114 SRs described interventions aimed at people defined by race, gender, low SES or age [6, 102]; six methodology studies selected only SRs that focused on disadvantaged groups across PROGRESS; 21/109 SRs included studies conducted in LMICs,	Assesses if interventions have been tested in specific disadvantaged populations	Does not assess effects of intervention Can be misleading since SRs with no studies conducted in disadvantaged populations may still be relevant and applicable

Methods used to assess health equity effects	Which studies used this method? (author year)	Data availability	Advantages	Disadvantages
1d.Descriptive-Outcomes of SR related to equity of access	Tsikata 2003 [6], Nasser 2007 [121], Althabe 2008 [104], Lewin 2008[103], Bambra 2009 [97], Chopra 2008 [69]	Equity of access measured in 18/173 SRs. Data not reported by one study [121].	Provides data on access to health care, a determinant of health inequalities	Data on access to care does not measure effects on health inequalities Measuring access to health care is dependent on the question and availability of data depends on selection criteria of methodology review
1e.Descriptive-describe if SRs conduct or plan subgroup analyses across PROGRESS-Plus	Tugwell 2008 [101]; Ogilvie 2004 [207]; Johnson 2003 [208]; Tsikata 2003 [6], Viswanathan [209], Main 2008 [102], Lewin 2008 [103], Odierna 2009 [100], Bambra 2009 [97]	Analysis by PROGRESS-Plus subgroup in 22/198 SRs; Place of residence; 0; Race/ethnicity (12/262); Occupation (0); Gender 14/198; Religion (0); Education (0); SES (1/198); 6/49 SRs assess differences across SES, gender or race.	Subgroup analysis provides direct data needed to answer whether the intervention works the same or differently in populations of interest	Lack of data: data available by PROGRESS-Plus subgroups of interest in 10% of SRs (28/247 had data)
2a.Descriptive – assess if primary studies describe population across PROGRESS-Plus	Tugwell 2008 [101]; Tsikata 2003 [6], Johnson 2003[105], Ogilvie 2004 [49] for 1 SR	Place of residence (26/263), race (42/263), occupation (24/250), gender (209/250), religion (0), education (42/263), SES (25/263), Social capital (24/250)	Provides evidence on whether sufficient evidence is available from primary studies to conduct subgroup analyses in SRs	Data may not be available stratified by PROGRESS-Plus factors in the primary studies
2b.Descriptive-assess if primary studies stratified analyses by PROGRESS-Plus	Tugwell 2008 [101]; Ogilvie 2004 [207], Johnson 2003 [208], Tsikata 2003 [6]	11 of 147 primary studies stratified by one or more PROGRESS-Plus [101]; 64/258 assessed gender/sex [208]; 10/76 and 5/14 stratified by sex [207]; 7/103 stratified by education, gender or SES [6]	Identifies whether subgroup analyses across PROGRESS-Plus are available in primary studies and the direction and magnitude of effects in different populations	Time-consuming to assess all primary studies of included SRs Does not rule out the possibility of spurious statistical significance

Methods used to assess health equity effects	**Which studies used this method? (author year)**	**Data availability**	**Advantages**	**Disadvantages**
3a.Analytic: Association	Morrison 2004 [111]	Age in 8/12 SRs; Sex in 7/12 SRs; SES in 5/12 SRs	Indicates whether PROGRESS-Plus factors are associated with different relative effects Could be used to assess gradients of effect modification according to different levels of PROGRESS-Plus (e.g. poverty)	Data unavailable for 33% of SRs (4/12)
3b. Analytic: relative comparison of effect size in two groups using an odds ratio	None			
3c. Analytic: assess effects in a disadvantaged population	Adamek 2008 [39], Stewart 2006 [108], D'Souza [109], Ball 2004 [98], Browne 2004 [210]; Bartels 2003[40]	Identified median of 11 SRs with targeted evidence (range 5-23); three studies reported medium to large effect sizes of interventions targeted at depression in older adults [39], youth with disabilities [109] and mental health promotion in children [210]. Three studies did not report effect sizes	Directly applicable for decisions about interventions in these specific disadvantaged populations Identifies evidence gaps	Lack of data in some disadvantaged populations limits the use of this approach for other populations and settings Low methodological quality of SRs may limit applicability Lack of data on process of implementation

Methods used to assess health equity effects	**Which studies used this method? (author year)**	**Data availability**	**Advantages**	**Disadvantages**
4a.Applicability: assess likely impact on disadvantaged populations using checklists for applicability and equity	Althabe 2008 [104], Lewin 2008 [103], Chopra 2008 [69]	8/20 SRs were considered most transferable to LMIC setting [103], 1 study only included SRs if they were deemed applicable in LMIC settings [104], 1 study assessed applicability to LMIC settings using the SUPPORT checklist	Useful summary for policy-makers about likely relevance in LMIC settings Standardized format makes judgments explicit and transparent Does not require replication of studies in different populations and settings Not subject to statistical power issues of subgroup analyses	Does not assess the magnitude of effect in different populations Requires content and methodological expertise to make equity and applicability judgments Low availability of data to make judgments [104], [103], [69]

Note: SES: Socioeconomic status; PROGRESS: PROGRESS: Place of residence (urban/rural), Race/ethnicity, Occupation, Gender, Religion, Education, Socioeconomic status and Social Capital

Table 9: Comparison of subgroup analyses described in methodology studies against credibility criteria for subgroup analyses

	Johnson 2003[105]	**Tsikata 2003**[6]	**Ogilvie 2004**[207]	**Odierna 2009**[100]	**Lewin 2008**[103]	**Tugwell 2008**[101]	**Viswanathan 2008**[209]	**Main 2008**[102]	**Bambra 2009**[97]
Clinically important difference	Not described	Not described	Not described	Not described	Yes, differences in effect that could affect health equity in 4/20 SRs	No SRs (0/14) conducted subgroup analyses	No SRs (0/64) conducted subgroup analysis across race	Can't tell- 3/19 SRs assessed effects on health inequalities	Can't tell- 8/30 SRs assessed effects on health inequalities
Statistically significant difference	Not described, state "potential difference" in 3 out of 31 systematic reviews	Yes, in 1/95SRs	Not described	Yes, 5/16 SRs reported statistically significant difference in effects across gender or race	Not described	No data	No data	Not described	Not described
A priori hypothesis	Not described	Yes	Yes	Not described	Not described	No data	No data	yes	yes
One of a small number of hypotheses tested	Not described	Yes	Not described	Not described	Not described	No data	No data	Not described	Not described
Differences suggested by within study comparisons	Not described	Yes	Not described	Not described	Not described	No data	No data	Not described	Not described
Difference consistent across studies	Not described	NA- only 1 study	Not described	Not described	Not described	No data	No data	Not described	Not described
Indirect evidence to support hypothesis	Yes, evidence that cardiovascular risk factors, presentation, treatment and treatment outcomes vary	yes- economic rationale why transport incentives would work better for poorer people	Yes, smoking is associated with social disadvantage	Not described	Not described	No data	No data	Yes	Yes

	Johnson 2003 [105]	**Tsikata 2003** [6]	**Ogilvie 2004** [207]	**Odierna 2009** [100]	**Lewin 2008** [103]	**Tugwell 2008** [101]	**Viswanathan 2008** [209]	**Main 2008** [102]	**Bambra 2009** [97]
	between men and women								
Statistical subgroup by treatment interaction	Not described	Not described	Not described	Not described	Not described	No data	No data	Not described	Not described
Primary studies stratified randomization by subgroup of interest	Not described	Not described	Not described	Not described	Not described	No data	No data	Not described	Not described

Table 10: PROGRESS-Plus factors described in cohort of 224 systematic reviews published in MEDLINE in November 2004

	Population described by PROGRESS -Plus 68% (153/224)	Theory considers PROGRESS-Plus 8% (18/224)	Analysis of differences across PROGRESS-Plus (using qualitative, quantitative or targeted methods) 12% (29/224)	Applicability or implications consider PROGRESS-Plus 21% (49/224)
Overall	153	18	29	49
Place	49	2	0	1
Race/ethnicity	10	4	2	9
Occupation	2	0	0	1
Gender	109	3	9	11
Religion	0	0	1	0
Education	2	2	1	3
Socioeconomic status	6	3	7	7
Social capital	2	2	4	5
LMIC	21	3	3	15
Disability	3	3	3	13
Age	105	2	2	1

PROGRESS-Plus: Place of residence (including urban/rural); Race/ethnicity/culture, Occupation, Gender, Religion, Education, Socioeconomic status, Social capital; Plus includes age, disability and low and middle income country (LMIC)

Note: PROGRESS-Plus items add up to more than the overall because some systematic reviews assessed more than one PROGRESS-Plus factor.

Table 11: Systematic reviews (n=29) which assessed differences across PROGRESS-Plus by: 1) subgroup analysis; 2) targeted; or 3) description of individual studies

Reference	Intervention	Population	Method	Health inequalities hypothesized to affect intervention effects in background	PROGRESS-Plus Factor	Finding	Implications related to vulnerable populations
McCord 2004 [86][211]	Maxillofacial prosthodontic care for oral cancer	People with oral cancer	Subgroup analysis-Descriptive	No	Gender, SES	Gender, SES had no significant impact on outcomes;	Research needed to assess statistical power of effect modifiers (e.g. age, marital status, smoking status)
Green 2001 [273][212]	Genetic screening for pregnant women	Pregnant women	Subgroup analysis-Descriptive	No	SES	Videos may be more effective than leaflets for those who cannot read.	Research is needed on interventions for low SES to reduce inequalities in understanding and anxiety
Pachler 2004 [314][213]	Permanent colostomy after rectal resection for cancer	People with rectal resection for cancer	Subgroup analysis-Descriptive	Yes	Education, Gender, Occupation, Religion and SES	One study included mainly patients with low social class, low income level and poor education and it is argued that patients in poor areas may have problems managing their stoma	Research needed: Inadequate reporting of social class in studies needs to be addressed
Nowak 2004 [399][134]	Tamoxifen for hepatocellular carcinoma	People with hepatocellular carcinoma	Subgroup analysis-Descriptive	No	Gender	Describe influence of gender in 2 trials: 1 trial no difference, 1 trial found benefit for men but not women	Not done
Gourlay 2004 [412][214]	Clonidine for smoking cessation	Smokers	Subgroup analysis-Descriptive	No	Gender	4 studies showed greater effect of clonidine for women	Not done
Ball 2004 [708][140]	Moxifloxacin for respiratory tract infections in adults	Respiratory infection	Subgroup analysis-Descriptive	No	Gender, Plus-age	In individual studies, predisposition to adverse effects was associated with greater age and female sex	Effective for women and men, and for elderly
Rubenstein 2004 [781][150]	Non-steroidal anti-inflammatory drugs	People taking NSAIDs	Subgroup analysis-Descriptive	No	Gender, Plus-Age	Older patients at higher risk for liver injury; 2 studies found no effect of gender, 1 study found men at increased risk	Effective for women and men

Reference	Intervention	Population	Method	Health inequalities hypothesized to affect intervention effects in background	PROGRESS-Plus Factor	Finding	Implications related to vulnerable populations
Mowatt 2004 [841][215]	Myocardial perfusion scintography for diagnosis and management of angina and myocardial infarction	People with angina or myocardial infarction	Subgroup analysis-Descriptive	Yes	Gender	6 studies which assessed gender concluded that SPECT provides important independent prediction of survival in both men and women	Effective for women and men (no difference)
Swanson 2000 [43][216]	Interventions for children with learning disabilities	Students (children or adults) with learning disabilities	Subgroup analysis-pooled results	No	Gender	No difference between males and females	Effective for both males and females (no difference in effects)
Twetman 2004 [101] [132]	Fluoride mouthrinses	Non-selected populations of all ages	Subgroup analysis-pooled results	Yes	SES (background fluoride exposure as a proxy for SES)	0.95 per year difference in DMF; 23% preventive fraction difference	More effective for low SES due to higher baseline risk of caries
Swanson 1999 [102] [137]	Reading interventions	Children with learning disabilities	Subgroup analysis-pooled results	No	Gender	Effect size was 0.48 greater for studies which did not report the gender ratio compared to those that did	Implications not discussed
Wendel-Vos 2004 [254] [217]	Exercise training to prevent stroke	General	Subgroup analysis-pooled results	No	Gender	Relative risk of stroke 0.54 for men, 0.76 for women, difference = 0.22 (p=0.07)	No implications related to gender
Ramakrishnan 2004 [553] [218]	Micronutrient interventions	Children	Subgroup analysis-pooled results	No	SES (measured by proxy of baseline nutritional status- height and weight for age)	Multi-micronutrient supplementation has positive effect on child growth	Need research on food-based approaches that are sustainable and feasible in resource poor settings
Matthews 2004 [705] [219]	Peg interferon Alfa-2a	Chronic hepatitis C virus infection	Subgroup analysis - pooled results	Yes	Race/ethnicity	Blacks have lower sustained viral response than white (15% vs 35% in one study and 26% vs 39% in another study)	More research is needed on blacks
Conaty 2004 [782] [138]	Pneumococcal polysaccharide vaccines to prevent pneumonia	Adults	Subgroup analysis-pooled results	No	Adults in LMIC vs. elderly in HIC	RCTs in LMIC where prevalence is high show benefit (RR 0.66, 95% CI: 0.57-0.77). RCTs in elderly show no effect (RR 1.03, 95% CI: 0.86-1.25)	Properly conducted randomized trials are needed to support the use of conjugate vaccines in the elderly

Reference	Intervention	Population	Method	Health inequalities hypothesized to affect intervention effects in background	PROGRESS-Plus Factor	Finding	Implications related to vulnerable populations
Parker 2004 [52] [220]	Hip protectors	Elderly	Targeted	No	Plus- frail elderly	Evidence of effectiveness of hip protectors in institutional settings with high background incidence of hip fracture	Effective for frail elderly
Heyn 2004 [188] [221]	Exercise training to improve function and cognition in elderly with cognitive impairment	Elderly people (>65 yrs) with cognitive impairment	Targeted	Yes	Plus- elderly with cognitive impairment	Exercise training improves fitness, physical function and cognitive function in elderly with dementia	exercise beneficial for elderly with cognitive impairment
Kenworthy 2004 [298] [222]	Psychological interventions for those at risk of offending	People who have sexually offended or are at risk of offending, , population described as vulnerable	Targeted	No	Plus- sexual offenders	Psychological interventions can be studies in RCTs. Limited evidence of their benefits for this population	ethics of providing this unproven treatment to vulnerable people outside of a trial is debatable
Ejere 2004 [348] [223]	Face-washing promotion for preventing active trachoma	Trachoma endemic communities (i.e. Africa, Asia, Middle East)	Targeted	No	Plus- LMIC endemic countries	Face-washing combined with antibiotics reduces severe trachoma	Not done
Mattick 2004 [374] [224]	Buprenorphine maintenance vs. placebo or methadone maintenance for opioid dependence	Opioid dependence	Targeted	No	Plus- Heroine-addicted drug users	Buprenoprphine is effective, but not as effective as methadone	Different intervention may be preferred for some vulnerable populations due to feasibility : Buprenorphine may have advantage when alternate day dosing is desirable due to feasibility issues
Dinh-Zarr 2004 [382] [225]	Interventions for preventing injuries in problem drinkers	Problem drinkers	Targeted	No	Plus- problem drinkers	Interventions reduce injuries and their antecedents	Not done
Brylewski 2004 [407] [131]	Antipsychotic medication for challenging behaviour in people with learning disability	People with learning disability	Targeted	Yes	Plus- disease status of learning disability described as "culturally abnormal behaviours"	No evidence on whether these medications help or harm adults with learning disability and challenging behaviours	Research needed: Policy: managers should insist on good quality research, clinicians should use judgment and clinical experience

Reference	Intervention	Population	Method	Health inequalities hypothesized to affect intervention effects in background	PROGRESS-Plus Factor	Finding	Implications related to vulnerable populations
Grossman 2004 [417] [226]	Mindfulness-based stress reduction for chronic illness	Chronically ill	Targeted	No	Plus- disability of chronic illness	Stress reduction beneficial on mental and physical health outcomes with effect size of 0.5	Not done
Carter-Pokras 2004 [521] [227]	Providing linguistically appropriate services	People with limited English proficiency	Targeted	Yes	Race/ethnicity - language, disadvantaged non-English speakers	Linguistic interpretation results in higher adherence and better health status	Different interventions for non-English speakers: procedures needed to identify patients who need linguistic interpreters
Gundry 2004 [611] [228]	Household water treatment and storage interventions to reduce diarrhea	People in LMIC	Targeted	Yes	Plus- LMIC	Household treatment reduces cholera but not diarrhea	Need to assess factors that affect post-source water quality (e.g. including hygiene education, household water treatment methods)
Lees 2004 [709] [229]	Therapeutic communities	Mainly disadvantaged prison inmates or drug abusers	Targeted	No	Plus- drug abusers or prison inmates	Confirmed effectiveness of therapeutic communities on success (e.g. reduction in criminal behavior)	Not done
Henry 2004 [717] [230]	Treatment interventions to improve insight in psychosis	Psychosis	Targeted	No	Plus- Disease status of schizophrenia	Possible benefits of psycho-education	Not done
Hesse 2004 [860] [231]	Psychosocial treatment + antidepressants to promote abstinence from drugs and alcohol	Alcohol or drug dependent people	Targeted	No	Plus- alcohol or drug dependency	No support for claim that psychosocial interventions can enhance effectiveness of antidepressants	Not done
Coon 2004 [926] [232]	Complementary and alternative therapies for hepatitis c	Hepatitis c virus infection	Targeted	No	Plus- disease status of HCV results in discrimination due to association with injection drug users and homeless	No evidence for the use of complementary therapies in subgroups who have been denied interferon-alpha therapy in the past (e.g. HIV co-infection or psychiatric conditions)	Not done

Table 12: Systematic reviews which considered PROGRESS-Plus in applicability, context, or baseline risk

Reference	Intervention description	Population	Why is applicability different?	Implications based on judgment of applicability
Ahmed 2004 [49] [233]	Supportive care	Patients with gastrointestinal cancer	Mechanism of action	Supportive care needs to consider cultural needs (including language, diet issues and literacy)
Benson 2004 [59] [135]	Fluoride mouthrinse	People with orthodontic devices	Context	Mouthrinses more effective when there is less use of fluoride toothpaste and non-fluoridated water supplies (proxy for low SES)
Tan 2004 [60] [234]	Air vs. oxygen for resuscitation of infants	Infants at birth	Context	Results may not apply in high income countries where more equipment and care available; 3 out of 5 studies were in low and middle income countries
Hoenig 2004 [113] [235]	Rehabilitation interventions in geriatrics	Geriatric	Context	Cognitive impairment affects effectiveness of rehab interventions in geriatrics
De Silva 2004 [117] [236]	Human milk to prevent infection rates in infants	Infants at birth	Context	Cohort studies showed bias in favour of human milk groups in terms of higher maternal sociodemographic variables, e.g. Greater avoidance of alcohol, smoking, and illegal drug taking and better antenatal care during the prenatal period
He 2004 [294] [237]	Salt reduction for blood pressure reduction	General	Context	Depends on background level of salt in diet- e.g. in developed countries, processed food contributes 75-80% of salt in diet
Adab 2004 [300] [133]	Common antiepileptic drugs	Pregnant women	Baseline risk	Epilepsy is more common in lower social classes
Turner 2004 [312] [238]	Community-based interventions for the prevention of burns and scalds in children	Children	Baseline risk	Higher baseline incidence in the control communities was reflective of lower socioeconomic characteristics in these areas.
Moore 2004 [325] [239]	Dietary advice	Type 2 diabetes mellitus in adults	Context	Adoption of affluent, westernized diets is affecting diabetes in developing countries
Nannini 2003 [339] [240]	Combined corticosteroid and longacting beta-agonist in one inhaler	COPD	Confounder	smoking may inhibit expression of corticosteroids and may be higher in low SES; smoking status was not well reported in all studies
Stade 2004 [351] [241]	Vaginal chlorhexidine during labour to prevent early-onset neonatal group B streptococcal infection	Women during labour	Context	in poorer areas of world with less skilled personnel and resources, this treatment may be of more benefit since intrapartum chemoprophylaxis limited availability
Rees 2004 [360] [242]	Exercise-based rehabilitation for heart failure	Heart failure	Context	Most of the studies have been conducted in men, more research needed to confirm effects in elderly and women
Kulier 2004 [390] [243]	Minilaparotomy and endoscopic techniques for tubal sterilisation	Fertile couples	Context	in resource poor settings (LMIC), cost of laparoscopy and surgeon expertise may be barriers
Zupan 2004 [398] [244]	topical umbilical cord care at birth	Newborn	Context	In low- and middle- income countries, neonates have a much higher risk of infection resulting in serious illness or death. The cord probably remains an important portal for bacteria, as demonstrated by neonatal tetanus.
Nowak 2004 [399] [134]	Tamoxifen for hepatocellular carcinoma	People with hepatocellular carcinoma	Mechanism of action	While tamoxifen has been used in both women and men with hepatocellular cancer, its putative mode of action raises question of whether effects differ between women and men due to estrogen effects
Ram 2004 [400] [136]	Vitamin C supplementation for asthma	Asthma	Confounder	smoking status may confound effect and is associated with lower SES, only reported in 2/8 studies
Scott 2004 [874] [245]	Angioplasty	Patients with ST segment elevation myocardial infarction	Context	thrombolysis is more effective than angioplasty when time since onset is longer, which may be related to rural place of residence (e.g. in rural Australia where time to transport is > 3 hours)

Table 13: Factors associated with differences in effectiveness in systematic reviews across PROGRESS-PLUS (measured or hypothesized differences)

	Difference across PROGRESS-PLUS measured or hypothesized by authors (44 systematic reviews)	No difference across PROGRESS-Plus or not assessed or discussed by authors (180 systematic reviews)	Odds ratio (95% confidence interval)
Cochrane review	22 (50%)	107 (59%)	0.7 (0.4 to 1.3)
Theory considers differences in effectiveness across PROGRESS-Plus	12 (27%)	4 (2%)	16.5 (5.0-54.4)
Health inequalities described	12 (27%)	7 (4%)	9.3 (3.4-25.3)
Non-pharmacological interventions	25 (57%)	78 (43%)	1.7 (0.9-3.3)
Judgment by extractors that differences may exist across PROGRESS-Plus	26 (59%)	28 (16%)	7.8 (3.8 to 16.2)
Acceptability discussed by author of SR across PROGRESS-Plus	24 (55%)	14 (8%)	14.2 (6.4-31.9)
Pragmatic question for systematic review	7 (16%)	5 (3%)	6.6 (2.0-22.0)
Study designs are RCTs only	28 (64%)	149 (83%)	0.36 (0.18-0.75)
Description of population across PROGRESS-Plus	36 (82%)	117 (65%)	2.4 (1.1-5.5)
Subgroup analysis planned across PROGRESS-Plus	6 (14%)	10 (6%)	2.7 (0.9 -7.8)
Subgroup analysis conducted across PROGRESS + factor	7 (15.9%)	0 (0%)	Not estimable
Implications discuss PROGRESS-Plus	20 (46%)	10 (6%)	14.1 (5.9-33.9)
Quantitative analysis	33 (75%)	142 (79%)	0.8 (0.4-1.7)
Process evaluation	13 (30%)	28 (16%)	2.3 (1.1-4.9)
Outcomes relevant to disadvantaged populations	1 (2.3%)	0 (0%)	Not estimable
Intervention culturally acceptable?	1 (1%)	5 (11%)	0.8 (0.1-7.1)
Context likely to be different across PROGRESS-Plus	26 (59%)	53 (29%)	3.5 (1.8-6.8)
Qualitative evaluation of differences across PROGRESS-Plus	8 (18.2%)	1 (0.6%)	39.8 (4.8-328)

Table 14: Characteristics of subgroup analyses in systematic reviews which combined data from different studies across PROGRESS-Plus factors, according to Oxman and Guyatt credibility criteria for subgroup analyses

	PROGRESS-Plus factor	Method to compare difference	Size of difference in effects	95% CI of difference in effects	Statistics used to compare groups	1. clinically important difference?	2. statistically significant difference?	3. a priori hypothesis	4. one of small number of hypotheses?	5. differences suggested by within study comparesons-	6. difference consistent across studies?	7. indirect evidence to support hypothesis?
217	Gender	Relative	0.22 difference in relative risk	N.d.	N.d.	N	N	Y	Y	Y	N.d.	N
138	LMIC	Relative	0.37 difference in relative risk	N.d	N.d	Y	N.d.	Y	Y	N	N	Y
219	Ethnicity	Absolute	13-20% difference in sustained viral response	Not done	Not done	N.d.	N.d.	Y	Y	Y	Y	Y
132	SES	Absolute	23% difference in preventive fraction	Not done	Chi square	Y	N.d.	Y	Y	N	N	Y
137	Gender	Absolute	0.48 difference in effect size between studies which report gender ratio and those that do not	N.d.	N.d.	N	Y	Y	N	N.d.	N.d.	N.d.
218	SES	Absolute	Effect size larger for those with low baseline nutritional status	N.d	N.d	N	N	Y	Y	Y	N	N.d.
216	Gender	Absolute	No difference	Not done	N.d.	N	N	Y	N	N	N	Y

*LMIC- low and middle income countries, SES: socioeconomic status, n.d.- not described

Table 15: Factors identified for optimizing benefit of programs and services for homeless people in Ottawa

Success Factors	Exemplar quotes
1.Assess needs of individuals	“has to be based on what are the felt needs of the individuals so when you see people are struggling with substance abuse issues and there is nothing there for them, that is a felt need” “we met with the clientele before we even started the program…asking them if a certain program was built, what would you be looking for, what would you stay in?”
2. Assess community needs	“a lot of different committees gathering information” “we hold community consultations… we did key informant interviews”
3.Provide a welcoming setting	“having a youth friendly environment where you see themselves represented, where there is not a lot of rules. It is pretty low threshold access… you can come after school hours… They can take a nap if they want to.”
4. Ensure services are feasible	Standardized assessment tools … difficult for the homeless community. It takes a couple of hours … It is hard to do with the homeless or transient population because you have to track them down.”
5. Diagnose underlying problems	“sometimes they don’t even have a definitive diagnosis when they come to us” “sometimes the wrong treatment for the problem” “stop judging”
6.Provide training for staff	“increasing capacity for staff as well through training because you can’t develop services if you don’t train staff to be able to recognize the issues and how do they support clients”
7. Enforce policies for safety	“policy for example on boundaries between clients and staff… have a policy and are following it for major incidents and conflicts of interests”
8. Need to be patient with clients	*“You have to be very patient. They might miss appointments until they finally make it to one.”*
9. Ensure services are acceptable	“We are going to have to match from the culture in order to communicate and assist”
10.Need to consider business case to provide right services and the right time for the right people	“one of my fortes is business so imagine starting an organization with very good people coming from almost like a religious background … you have to be able to understand who you are and who you are partnering with and don’t duplicate services. You really have to investigate the services that we are providing, are they the right services, the right time for the right people and does it make business sense”
11.Build financial sustainability	“we need to do as an organization is to try as best as we can to put you into a position where you are financially stable, where you have long-term sustainability and that you are not subject to the control of any organization”

Success Factors	Exemplar quotes
12.Raise awareness and support from the public	“we are actually saving the city hundreds of thousands of dollars over a period of time. … We are going to be having meetings. …at community centres in the area where local residents can come. … You want to be as open and transparent as possible but you also don’t want people to twist and convey what you are actually trying to do … You are going to tell them what they need to know…to the average citizen, there is that fear factor” When you are looking at outcomes and that is when the program gets a little harder because if I am explaining the program to a city official or a politician, they are going to say, what is your goal, what is the outcome of the program. You can break it down into different ways so if I am speaking to a city official, I would say, one of our outcomes is a decrease in city disruption
13.Design programs based on best available evidence and theory	“we essentially did a lot of homework… what worked well in the managed alcohol program …harm reduction based programs”
14.Provide individualized, multi-faceted treatment	“You are always trying to balance … It is very individualized.” “we have the addictions services … social workers … help people with employment from writing resumes to teaching computer skills … housing workers … full-time teacher … kitchen training program … anger management. They do whatever needs to be done. What do you need? Where are you at? … Why are there stumbling blocks?”
15.Engage the homeless in designing and providing programs	“they are starting to build a community now of women who are part of the homeless community…They support each other”
16.Convergence of political agendas with agency priorities	“First and foremost about good ideas but there also has to be opportunity and luck; That was luck. We got this federal funding… so there is sometimes this convergence of things that occur and you are given the opportunity to go forward”
17.Provide services in partnership with multiple agencies	“we went through a change management process with all of the agencies to resolve our cultural differences (eg social services, judicial, housing, health)”
18.Provide supportive environments for homeless to reduce exposure to unsafe behaviours	“We really put a big emphasis on supportive housing and getting people out of shelters into some form of supportive housing where they can actually have some sort of normal life”
19.Agencies need leadership and vision	“you also have to have a very strong vision of where you want to go”
20.Improve performance by monitoring and evaluation	“Outcome focused, evidence-based practice and adapt to the reality of the population and service providers abilities; Data collection is a tool to improve performance”

Table 16: Success factors mapped onto the equity effectiveness framework steps

EQUITY EFFECTIVENESS LOOP	FACTORS IDENTIFIED
1.Burden of illness and aetiology	1.Assess needs of individuals
	2. Assess community needs
2.Equity effectiveness	
Access/ coverage	3.Provide a welcoming setting
	4. Ensure services are feasible
Diagnostic accuracy	5. Diagnose underlying problems
Provider compliance	6.Provide training for staff
	7. Enforce policies for safety
	8. Need to be patient with clients
Patient adherence	9. Ensure services are acceptable
3.Economic evaluation	10.Need to consider business case to provide right services and the right time for the right people
	11.Build financial sustainability
	12.Raise awareness and support from the public
4.Knowledge translation and implementation	13.Design programs based on best available evidence and theory
	14.Provide individualized, multi-faceted treatment
	15.Engage the homeless in designing and providing programs
	16.Convergence of political agendas with agency priorities
	17.Provide services in partnership with multiple agencies
	18.Provide supportive environments for homeless to reduce exposure to unsafe behaviours
	19.Agencies need leadership and vision
5.Monitoring of programme	20.Improve performance by monitoring and evaluation
6.Reassessment	20. Improve performance by monitoring and evaluation

Table 17: Steps in generating a health measurement scale, from Feinstein *Clinimetrics* and Streiner and Norman *Health Measurement Scales*

Step	Description
1. Item generation	Patient or key informant interviews and systematic review of similar scales
2. Questionnaire format	Borrow from existing questionnaires, if possible
3. Questionnaire scaling	Scale needs to match the observer's ability to discriminate [172]
4. Pre-testing	Assess clarity of questions, assess redundancy
5. Item reduction	Use combination of clinician and patient expertise and experience, the purpose of the index, and psychometric methods
6. Internal consistency or homogeneity	Internal consistency is a measure of whether all of the items on a scale or index are measuring the same property.
7. Consistency or reliability	Assess whether the same result will be found if the index is repeated by the same method or observer (intra-observer variability) or another observer (inter-observer variability) for stable subjects
8. Validity	Assess whether the index measures what we think it is measuring [246]
9. Feasibility	Assesses acceptability to raters and subjects (e.g. how long does it take to use? What experience does the rater need to use it?)

Table 18: Items generated for the development of the equity plausibility algorithm, from literature review and the first three studies of this thesis

Source	*Potential items*
Literature review of existing checklists	1. Are there pathophysiologic differences in the illness under study that may lead to a diminished treatment response?
	2. Are there patient differences that may diminish the treatment response?
	3. Are there important differences in patient compliance that may diminish the treatment response?
	4. Are there important differences in provider compliance that may diminish the treatment response?
	5. Do my patients have comorbid conditions that significantly alter the potential benefits and risks of the treatment?
	6. Are there important differences in untreated patients' risk of adverse outcomes that might alter the efficiency of treatment?
	7. Is my patient **SO** different that results do not apply?
	8. Is treatment feasible in my setting?
	9. Do the benefits outweigh the harms?
	10. How will my patients' values influence the decision?
	11. What are predicted absolute risk reductions?
	12. Are there variations in treatment effect (e.g. subgroups analyses)
	13. How does treatment effect vary with baseline risk?
	14. If implemented, will effectiveness be the same? (taking into account baseline prevalence, characteristics of population, capacity to implement)
	15. Can the intervention be implemented (political environment, social acceptability, cultural adaptability, resource implications, educational level, organizational structure)
	16. Is there a reason to anticipate different effects of intervention in disadvantaged and privileged populations?
Cochrane Methodology review	17. No factors identified as associated with considering health equity
Methodology study of equity assessment in SRs	18. Theory considers differences in effectiveness across PROGRESS-Plus
	19. Health inequalities described
	20. Judgment by extractors that differences may exist across PROGRESS-Plus
	21. Acceptability discussed by author of SR across PROGRESS-Plus
	22. Pragmatic question for systematic review
	23. Study designs are observational
	24. Subgroup analysis conducted across PROGRESS + factor
	25. Implications discuss PROGRESS-Plus
	26. Process evaluation
	27. Outcomes relevant to disadvantaged populations
	28. Description of population across PROGRESS-Plus
	29. Context likely to be different across PROGRESS-Plus
	30. Qualitative evaluation of differences across PROGRESS-Plus

Source	*Potential items*
Qualitative study of success factors for homeless people in Ottawa	31. Assess needs of individuals 32. Assess community needs 33. Provide a welcoming setting 34. Ensure services are feasible 35. Diagnose underlying problems 36. Provide training for staff 37. Enforce policies for safety 38. Need to be patient with clients 39. Ensure services are acceptable 40. Need to consider business case to provide right services and the right time for the right people 41. Build financial sustainability 42. Raise awareness and support from the public 43. Design programs based on best available evidence and theory 44. Provide individualized, multi-faceted treatment 45. Engage the homeless in designing and providing programs 46. Convergence of political agendas with agency priorities 47. Provide services in partnership with multiple agencies 48. Provide supportive environments for homeless to reduce exposure to unsafe behaviours 49. Agencies need leadership and vision 50. Improve performance by monitoring and evaluation

Table 19: Equity plausibility algorithm questions

Question 1: Are there differences in patient/community/ population characteristics (e.g. underlying pathophysiology, comorbidities, patient attitudes, etc.) that are likely to create important differences in the magnitude of relative effect of the intervention versus the control for the outcome of interest?
Question 2: Are there differences in the way that the intervention is delivered (e.g. provider compliance, provider skill, technical resources, availability of drugs/treatments) that are likely to create important differences in the magnitude of the relative effect of the intervention versus the control for the outcome of interest?
Question 3: Are there differences in the comparator across patient, community or population that are likely to create important differences in magnitude of relative effects?

Table 20: Characteristics of 35 raters who assessed equity plausibility

Experience with systematic reviews	User-6 Methodologist-15 Clinician-12
Years of experience	Median: 7 Range : 2-15
Area of research/expertise	Public health-8 Musculoskeletal-7 Dermatology-1 Child health-1 Methods-9 Family medicine-5 Infectious disease-2 Reproductive health-1 Cancer-1

Table 21: Agreement of equity plausibility ratings between raters for each question and PROGRESS factor, across 10 systematic reviews

	Question 1: Patient differences Agreement between raters (Proportion of yes answers)		***Question 2: Delivery of intervention Agreement between raters (Proportion of yes answers)***		***Question 3: Comparator Agreement between raters (Proportion of yes answers)***	
Systematic review	Sex	SES	Sex	SES	Sex	SES
Mass media for HIV testing	**0.91^{x} (96%)**	**1^{x} (100%)**	0.56 (70%)	0.70 (83%)	0.49 (57%)	0.64 (78%)
Population tobacco control	0.56 (70%)	**0.83^{y} (91%)**	0.48 (48%)	0.60 (74%)	0.48 (48%)	0.76 (87%)
Psychological therapy for PTSD	0.70 (83%)	0.64 (78%)	0.70 (83%)	**0.83^{x} (91%)**	0.48 (52%)	0.49 (57%)
First line antihypertensives	0.53 (65%)	0.53 (65%)	0.48 (48%)	0.70 (83%)	0.49 (43%)	0.48 (52%)
Surgery for age-related cataract	0.52 (67%)	0.59 (75%)	0.52 (67%)	**1.00^{y} (100%)**	0.52 (67%)	0.70^{y} (83%)
Vaccines for measles, mumps and rubella in children	**0.83^{y} (8%)**	0.45 (50%)	0.70 (17%)	0.59 (75%)	0.59 (25%)	0.59 (75%)
Antidepressants versus placebo for depression in primary care	**0.83^{x} (92%)**	**0.83^{x} (92%)**	0.52 (67%)	**0.83^{x} (92%)**	0.45 (50%)	0.52 (67%)
Artemisinin-based combination therapy for treating uncomplicated malaria	0.52 (33%)	0.59 (75%)	0.52 (33%)	**0.83^{x} (92%)**	0.59 (25%)	0.45 (50%)
Primary safety belt laws	0.70 (83%)	0.47 (58%)	0.52 (67%)	1.00^{x} (100%)	0.52 (33%)	0.47 (58%)
Hand washing for preventing diarrhoea	0.52 (67%)	0.70 (83%)	0.45 (50%)	**0.83^{y} (92%)**	0.45 (50%)	**0.83^{y} (92%)**
Fleiss Kappa	0.199	-0.001	0.068	0.105	0.005	0.04

Notes: ^{x} Denotes judgments with >80% agreement that important differences in magnitude of effect are likely that disagree with the systematic reviews; ^{y} Denotes judgments with greater than 80% agreement that important differences in magnitude of effect are likely that agree with the systematic reviews. PTSD: Post-traumatic stress disorder; SES: socioeconomic status

Table 22: Comments and reactions to making equity plausibility judgments

Reason for answer	theory-10; personal experience- 11; empirical data-3 guesses-4;
Other information needed	more intervention specific information and data how big are the differences being sought? how does comparator overlap with intervention delivery was information available from trials? consider including community cluster trials Need information on how intervention was delivered
General comments	Important to consider these issues in design of SR-5 Difficult- 4; Interesting to consider these issues-5; Subjective-6; Why only gender and SES- need to consider other factors (e.g. sexual orientation)- 2 Country differences are important- e.g. if universal drug coverage is available-1 Accessibility of drugs is less of an issue-1 Ask questions about heterogeneity-1 Intervention delivery is important for understanding Easy-1

Appendix 1: MEDLINE search strategy for Cochrane Methodology Review

The search terms for MEDLINE are listed below, along with the number of hits obtained from searching MEDLINE from January 1960 to January week 2, 2009. The search strategy will be adapted for the other electronic databases.

1 exp Meta-Analysis as Topic/ 8694
2 systematic review.tw. 13239
3 meta-analys$.tw. 21728
4 meta-epidemiolog$.tw. 10
5 exp "Review Literature as Topic"/ 3979
6 (Cochrane adj2 review).tw. 967
7 OR/1-6 39977
8 (gender-based OR gender-related OR gender differences OR gender factors).mp. 12856
9 ((sex OR gender) adj2 (analysis OR specific OR difference? OR factor? OR inequit$ OR disparit$ OR inequalit$)).mp. 196205
10 exp sex factors/ 164145
11 exp geriatrics/ 23034
12 ((ethnic$ OR race OR racial OR religio$ OR cultur$ OR minorit$ OR refugee OR indigenous OR aboriginal) adj3 (analysis OR difference$ OR specific OR disparit$ OR inequalit$ OR inequit$)).tw. 26985
13 exp homosexuality/ 16581
14 exp disabled persons/ 35507
15 ((poverty OR low-income OR socioeconomic$ OR social) adj2 (analysis OR disadvantage$ OR specific OR difference? OR factor? OR inequalit$ OR depriv$ OR inequit$ OR disparit$)).mp. 87592
16 exp Educational Status/ 26647
17 exp Socioeconomic Factors/ 248934
18 ((discriminat$ OR social exclu$ OR social inclu$) adj3 (religion OR culture OR race OR racial OR aboriginal OR indigenous OR ethnic$)).tw. 475
19 ((urban OR rural OR inner-city OR slum) adj2 (difference$ OR specific OR analysis OR inequit$ OR disparit$ OR inequalit$)).tw. 1464
20 ((resource-poor OR (low-income adj countr$) OR (middle income adj countr$) OR africa OR developing countr$ OR south america OR china OR asia OR latin america) adj2 (relevance OR analysis OR specific OR difference OR applicab$ OR inequit$ OR disparit$ OR inequalit$)).tw. 541
21 OR/8-20 522320
22 7 AND 21 1672

Appendix 2: Cochrane Methodology Review Data Extraction Items

Ref ID
Author
Year
PROGRESS dimension
Definition equity (by author)
How is judgment of equity made? Ie fairness and avoidability?
Proxy measures used for PROGRESS-Plus? (e.g. nutritional status)
Reason/rationale for assessing equity
Number meta-analyses
Quality: Selection bias: how was sample of systematic reviews selected? Is there likelihood of selection bias?
Quality: Attrition bias: potential for bias in the exclusion of systematic reviews from analysis. Were any systematic reviews were excluded and why.
Quality: Detection bias: potential for bias in the assessment of analytic methods and outcomes in cohorts of systematic reviews. How did studies extract details of analysis of effects on health equity.
outcomes (benefits, harms, costs)
Quant measure of gaps/gradients
Statistical methods used (e.g. meta-regression, subgroup analysis...)
Methods of comparing gap (relative, absolute, gradient, risk difference)
Describe whether PROGRESS+ is mentioned in SR- in introduction, methods, discussion
Describe whether SRs describe population across PROGRESS+
Describe whether SRs include studies of targeted interventions aimed at disadvantaged
Describe outcomes related to equity of coverage or access
Describe whether primary studies included in the SRs stratify analyses by PROGRESS
Describe whether subgroup anlayses were planned or conducted across PROGRESS in the SRs
Subgroup analysis described in sufficient detail to answer 7 questions
Analytic- assess association of PROGRESS+ factor with effect size
Analytic: compare effect size between two groups using odds ratio, risk difference, relative risk
Analytic- assess likely impact on disadvantaged populations using checklists for applicability
Effect size
Standard error
95% CI
Expertise required to assess equity effects (as described by author, or paste in methods)
Availability of data to assess equity gap (as described by author)
Useability for end-user? (judgment by extractor or paste author's description
Advantages of method chosen to assess gap, as described by author?
Disdavantages of method chosen to assess gap, as described by author?
Clinically important difference?
Statistically significant difference?
A priori hypothesis
Subgroup analysis is one of small number of hypotheses tested?
differences suggested by within study comparisons
Difference consistent across studies?

Indirect evidence to support hypothesis?
Implications for policy, practice, research based on equity, equality analysis?
How was this study found? (searching databases, handsearching etc…
Factors associated with equity differences (e.g. study design, implementation adherence, compliance)
Limitations as described by author
Strengths as described by author

Appendix 3: SUPPORT Collaboration checklists to assess equity, applicability and scaling up

Available from: http://www.support-collaboration.org/summaries/methods.htm

APPLICABILITY

Consider differences in:

- structural elements of health systems (such that an intervention could not work in the same way)
- on-the-ground realities and constraints (that might substantially alter the potential benefits of the intervention)
- baseline conditions (different absolute effects, even if the relative effectiveness was the same)
- perspectives and influences of health system stakeholders (such that the intervention may not be accepted or taken up in the same way)

EQUITY

- Are there plausible reasons for anticipating differences in the relative effectiveness of the intervention in disadvantaged settings within the country?
- Are there likely to be different baseline conditions within the country, so that the problem would be more or less important in disadvantaged settings within the country?
- Are there likely to be different baseline conditions in disadvantaged settings within the country, so that the absolute effectiveness would be different?
- Are there important considerations that should be given to implementing the intervention to ensure that inequities are not increased and that they are reduced

SCALING UP

- What are the most important economic consequences?
- What information is there about the total resource implications of expanding coverage and sustaining an intervention?
- Is there important uncertainty about medium to long-term economic consequences?
- Is there important uncertainty about the applicability of any reported economic consequences?

Appendix 4: Data extraction items of methodology study

REFID
effectiveness SR (y/n)
Intervention description
Population
Population inclusion criteria: general population/specific population/targeted to disadvantaged
Targeted at disadvantaged? If yes, which PROGRESS+ factor
Intervention culturally acceptable? (y/n/not described, describe how it was assessed and for which populations)
Intervention: social/education/public health/pharmacological/surgery/complementary medicine/health systems reorganizing
Intervention: Theory described for how intervention expected to work (y/n)
Intervention: Theory consider different populations across PROGRESS+ (y/n; if yes, which factors)
Intervention: Theory of intervention (as described by SR)
Theory of intervention consider differences in effectiveness across PROGRESS+ (y/n, if yes, which factors?)
Intervention: extractors judgment: Is there biological, cultural, social reason why intervention would work differently in different populations, explain
Intervention: setting; does intervention require specific setting to work or be offered? E.g. resource-poor settings in LMIC
Question of systematic review: pragmatic (effectiveness) or explanatory?
Comparison: control/alternative interventions
Outcomes described as relevant to disadvantaged across PROGRESS+ (y/n, if yes, which ones)
Study design (RCT, Obs, both, other)
Number of studies/participants
Population characteristics described across PROGRESS + (y/n, if yes, which ones?
Quantitative synthesis (y/n)
Process evaluation (y/n)
Health Equity mentioned?
health inequalities/disparities mentioned?
Equity definition
Process evaluation consider differences across PROGRESS+? Which factors?
Qualitative evaluation of differences across PROGRESS+ (y/n), if yes, which technique, which factor(s)?
Subgroup analysis across one or more PROGRESS-Plus done? Which factor/factors of PROGRESS+
Applicability of results consider differences across PROGRESS + (y/n, if yes, which ones?)
Intervention dependent on context/setting (y/n, if yes, why)
Subgroup analysis across PROGRESS+: methods (relative, absolute, gradient, other)
1. clinically important difference?
2. statistically significant difference?
3. a priori hypothesis
4. subgroup analysis is one of small number of hypotheses tested?
5. differences suggested by within study comparisons
6. difference consistent across studies?
7. indirect evidence to support hypothesis?
Quantitative measure of gaps/gradients [gaps/gradient/targeted]

Statistical methods used (e.g. meta-regression, subgroup analysis…)
Methods of comparing gap (relative, absolute, gradient, risk difference)
Effect size (of gap/difference)
Std error (of gap/difference)
95% CI of gap/difference
Applicability of results considered across PROGRESS+? Which factors?
Implications for policy, practice, research based on equity or equality analysis?
Acceptability of intervention for disadvantaged
Quality of intervention assessed in relation to effect size
Contamination, confounding across PROGRESS+ categories
Organizational levels targeted by intervention
Community engagement (are members of disadvantaged communities involved in designing the intervention?)

Appendix 5: Consent Form and ethics approval for qualitative study (Ethics approval certificate, University of Ottawa Research Ethics Board #H02-09-11)

Name of Research Project:
What factors are associated with health equity effects of interventions? Qualitative study

I, ____, agree to participate in the research being conducted as part of the doctoral thesis of Vivian Welch, PhD candidate in Population Health at the University of Ottawa. I agree to be interviewed by Vivian Welch who is under the supervision of the Co-supervisors Peter Tugwell and George Wells, and by a thesis committee consisting of Betsy Kristjansson, Janet Smylie and Kevin Brand.

I understand that this study is gathering information from practitioners, decision-makers and systematic review authors about what factors related to the population, intervention, setting and outcomes might lead to different effects between more advantaged and less advantaged people. My participation will consist of an interview of 45 minutes to 1 ½ hours, during which notes will be taken and it will be audio-taped with my consent. There is no risk associated with the research. The interview will take place at a time and location convenient for me. My participation is voluntary. The potential benefits of this research are improved generalizability of research results to disadvantaged populations.

I have read the detailed information sheet provided by the researchers and am aware of the purpose of the research study and the time involved and any inconveniences.

I understand that I can withdraw consent and stop participating in the study at any time and for any reason, before or during an interview, and I will not be penalized because I have stopped. I also understand that the interview will be audio-taped. I can ask to have the audio-tape stopped at any time. If I stop the audio-taping, my interview will not be included in the data analysis.

I accept to be tape-recorded yes/no_________

I have read the information sheet and received assurance from the researchers that the information I will share will remain strictly confidential, the data will be kept in a secure manner, and my name will not be identified in the study.

I have been given the opportunity to ask questions about the interviews and the questions that I have asked have been adequately answered. I understand that I may also ask questions in the future.

If I wish, I may contact the Protocol Officer for Ethics in Research, 550 Cumberland Street, Room 160, (613) 562-5387 or ethics@uottawa.ca for more information about my rights as a research participant.

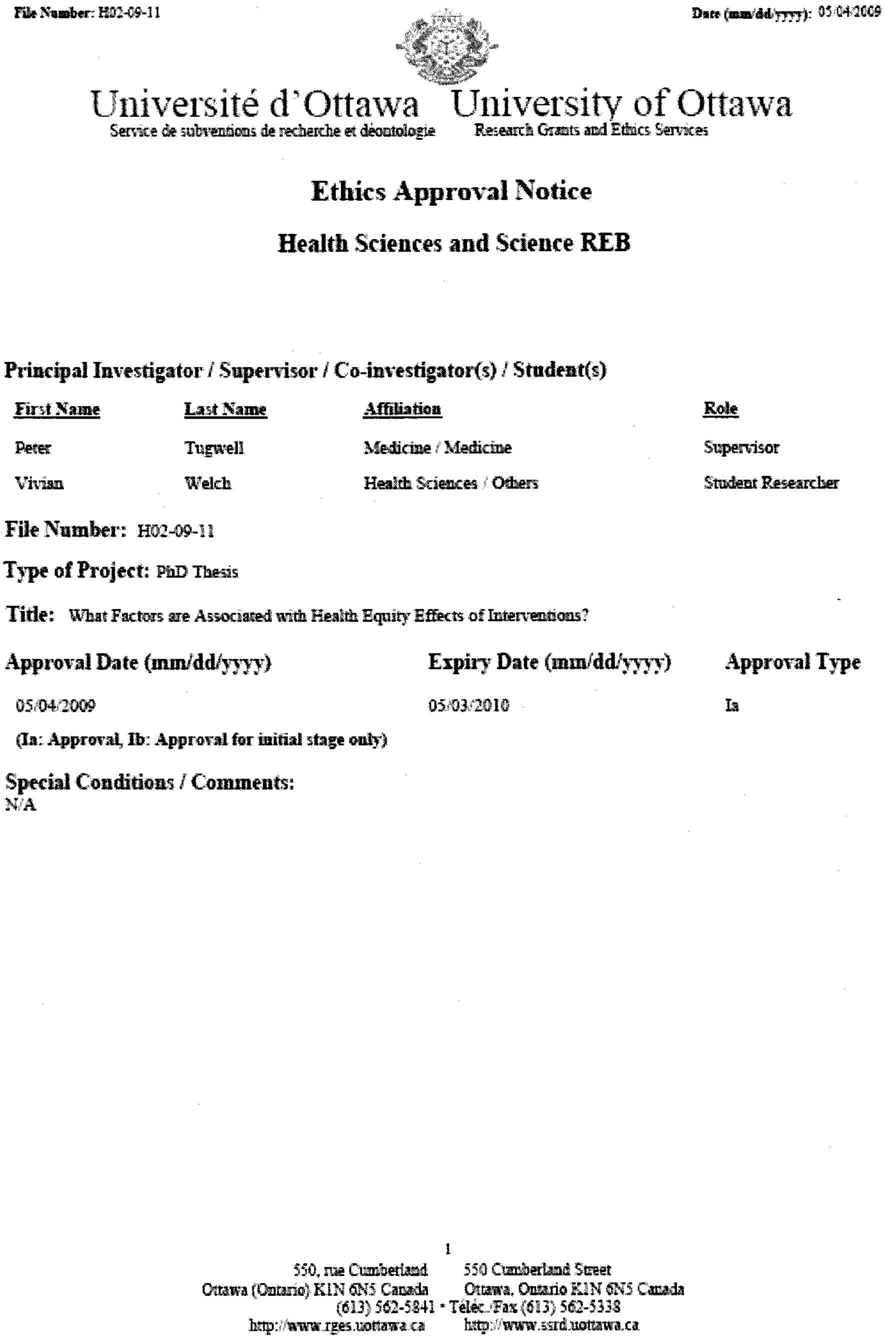
There are two copies of this consent form, one of which I may keep.

If I have any questions about the conduct of the research project, I may contact:

Peter Tugwell, supervisor at (613) 562-5800 ext. 1945 or at elacasse@uottawa.ca

Interviewer's Signature: ________________________________ Date: ____________

Participant's Signature: ________________________________ Date: ____________



File Number: H02-09-11

Date (mm/dd/yyyy): 05/04/2009

Université d'Ottawa University of Ottawa
Service de subventions de recherche et déontologie Research Grants and Ethics Services

Ethics Approval Notice

Health Sciences and Science REB

Principal Investigator / Supervisor / Co-investigator(s) / Student(s)

First Name	Last Name	Affiliation	Role
Peter	Tugwell	Medicine / Medicine	Supervisor
Vivian	Welch	Health Sciences / Others	Student Researcher

File Number: H02-09-11

Type of Project: PhD Thesis

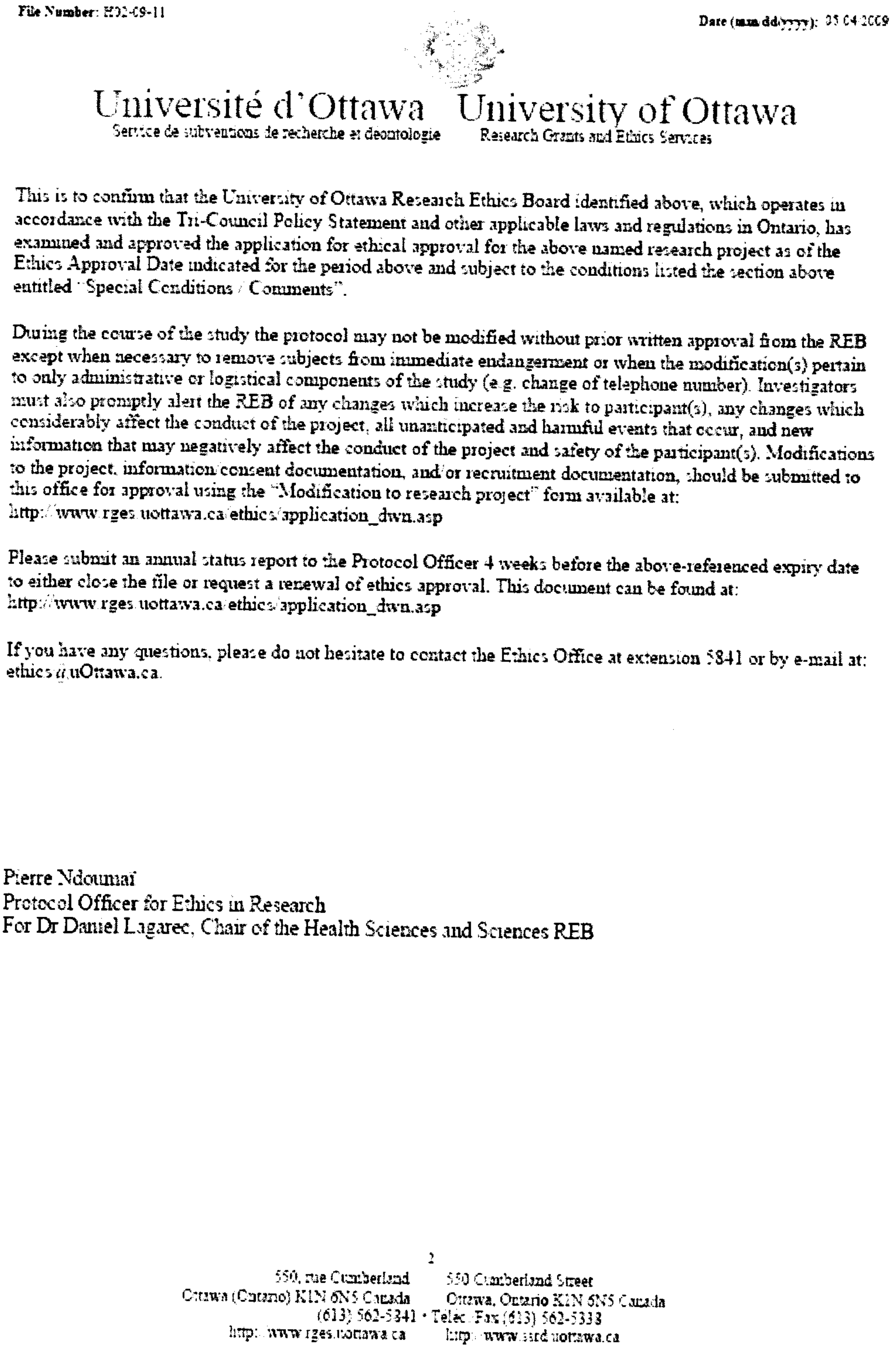
Title: What Factors are Associated with Health Equity Effects of Interventions?

Approval Date (mm/dd/yyyy)	Expiry Date (mm/dd/yyyy)	Approval Type
05/04/2009	05/03/2010	Ia

(Ia: Approval, Ib: Approval for initial stage only)

Special Conditions / Comments:
N/A

550, rue Cumberland 550 Cumberland Street
Ottawa (Ontario) K1N 6N5 Canada Ottawa, Ontario K1N 6N5 Canada
(613) 562-5841 • Téléc./Fax (613) 562-5338
http://www.rges.uottawa.ca http://www.ssrd.uottawa.ca



File Number: H02-09-11

Date (mm/dd/yyyy): 05/04/2009

Université d'Ottawa University of Ottawa
Service de subventions de recherche et deontologie Research Grants and Ethics Services

This is to confirm that the University of Ottawa Research Ethics Board identified above, which operates in accordance with the Tri-Council Policy Statement and other applicable laws and regulations in Ontario, has examined and approved the application for ethical approval for the above named research project as of the Ethics Approval Date indicated for the period above and subject to the conditions listed the section above entitled "Special Conditions / Comments".

During the course of the study the protocol may not be modified without prior written approval from the REB except when necessary to remove subjects from immediate endangerment or when the modification(s) pertain to only administrative or logistical components of the study (e.g. change of telephone number). Investigators must also promptly alert the REB of any changes which increase the risk to participant(s), any changes which considerably affect the conduct of the project, all unanticipated and harmful events that occur, and new information that may negatively affect the conduct of the project and safety of the participant(s). Modifications to the project, information/consent documentation, and/or recruitment documentation, should be submitted to this office for approval using the "Modification to research project" form available at:
http://www.rges.uottawa.ca/ethics/application_dwn.asp

Please submit an annual status report to the Protocol Officer 4 weeks before the above-referenced expiry date to either close the file or request a renewal of ethics approval. This document can be found at:
http://www.rges.uottawa.ca/ethics/application_dwn.asp

If you have any questions, please do not hesitate to contact the Ethics Office at extension 5841 or by e-mail at: ethics@uOttawa.ca.

Pierre Ndoumaï
Protocol Officer for Ethics in Research
For Dr Daniel Lagarec, Chair of the Health Sciences and Sciences REB

550, rue Cumberland 550 Cumberland Street
Ottawa (Ontario) K1N 6N5 Canada Ottawa, Ontario K1N 6N5 Canada
(613) 562-5341 • Telec./Fax (613) 562-5338
http://www.rges.uottawa.ca http://www.sard.uottawa.ca

Appendix 6: Interview Guide for convergent interviews

GENERAL QUESTION –

I'm interested in your experiences as a practitioner/program planner designing and delivering health and social care programs for homeless populations

What types of homeless populations have you worked with and what has your role been (e.g. designing and implementing programs, evaluating research)

- Can you describe a particularly successful or unsuccessful program, what was good or bad about it?
- What could have made the program work better for disadvantaged people?
- What do you think could have been done differently?
- Throughout your experiences, what have you found to be the most important factors in determining the success or failure of interventions in homeless populations
- Is there something that we have not covered that you think is important to take into account?
- Can you suggest 1-2 other people that we could talk to that might not necessarily share your views?

Appendix 7: Consent form and ethics approval for equity plausibility study (ethics approval certificate #H02-09-11c)

PARTICIPANT CONSENT FOR RESEARCH PROJECT

Name of Research Project:
What factors are associated with health equity effects of interventions? Field study

I, ______________________________, agree to participate in the research being conducted as part of the doctoral thesis of Vivian Welch, PhD candidate in Population Health at the University of Ottawa. I agree to fill out the equity plausibility checklist for 15 systematic reviews (introduction and methods only) developed by Vivian Welch who is under the supervision of the Co-supervisors Peter Tugwell and George Wells, and by a thesis committee consisting of Betsy Kristjansson, Janet Smylie and Kevin Brand.

I have read the detailed information sheet provided by the researchers and am aware of the purpose of the research study and the time involved and any inconveniences.

I understand that I can withdraw consent and stop participating in the study at any time and for any reason, before or during an interview, and I will not be penalized because I have stopped.

I have read the information sheet and received assurance from the researchers that the information I will share will remain strictly confidential, the data will be kept in a secure manner, and my name will not be identified in the study.

I have been given the opportunity to ask questions about the interviews and the questions that I have asked have been adequately answered. I understand that I may also ask questions in the future.

If I wish, I may contact the Protocol Officer for Ethics in Research, 550 Cumberland Street, Room 160, (613) 562-5387 or ethics@uottawa.ca for more information about my rights as a research participant.

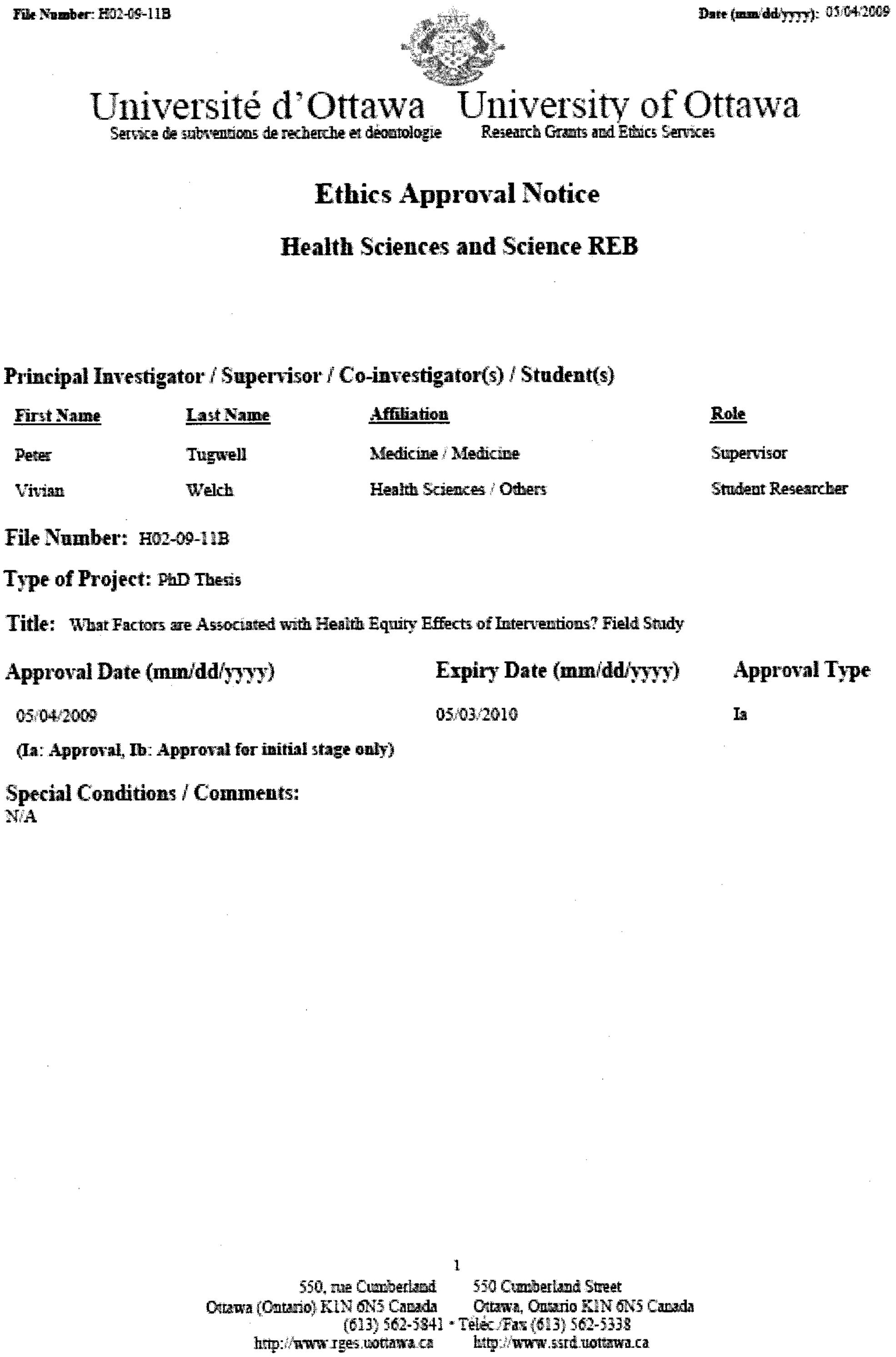
There are two copies of this consent form, one of which I may keep.

If I have any questions about the conduct of the research project, I may contact:

Peter Tugwell, supervisor at (613) 562-5800 ext. 1945 or at elacasse@uottawa.ca

Interviewer's Signature: ______________________ Date: __________

Participant's Signature: ______________________ Date: __________



File Number: H02-09-11B

Date (mm/dd/yyyy): 05/04/2009

Université d'Ottawa University of Ottawa
Service de subventions de recherche et déontologie Research Grants and Ethics Services

Ethics Approval Notice

Health Sciences and Science REB

Principal Investigator / Supervisor / Co-investigator(s) / Student(s)

First Name	Last Name	Affiliation	Role
Peter	Tugwell	Medicine / Medicine	Supervisor
Vivian	Welch	Health Sciences / Others	Student Researcher

File Number: H02-09-11B

Type of Project: PhD Thesis

Title: What Factors are Associated with Health Equity Effects of Interventions? Field Study

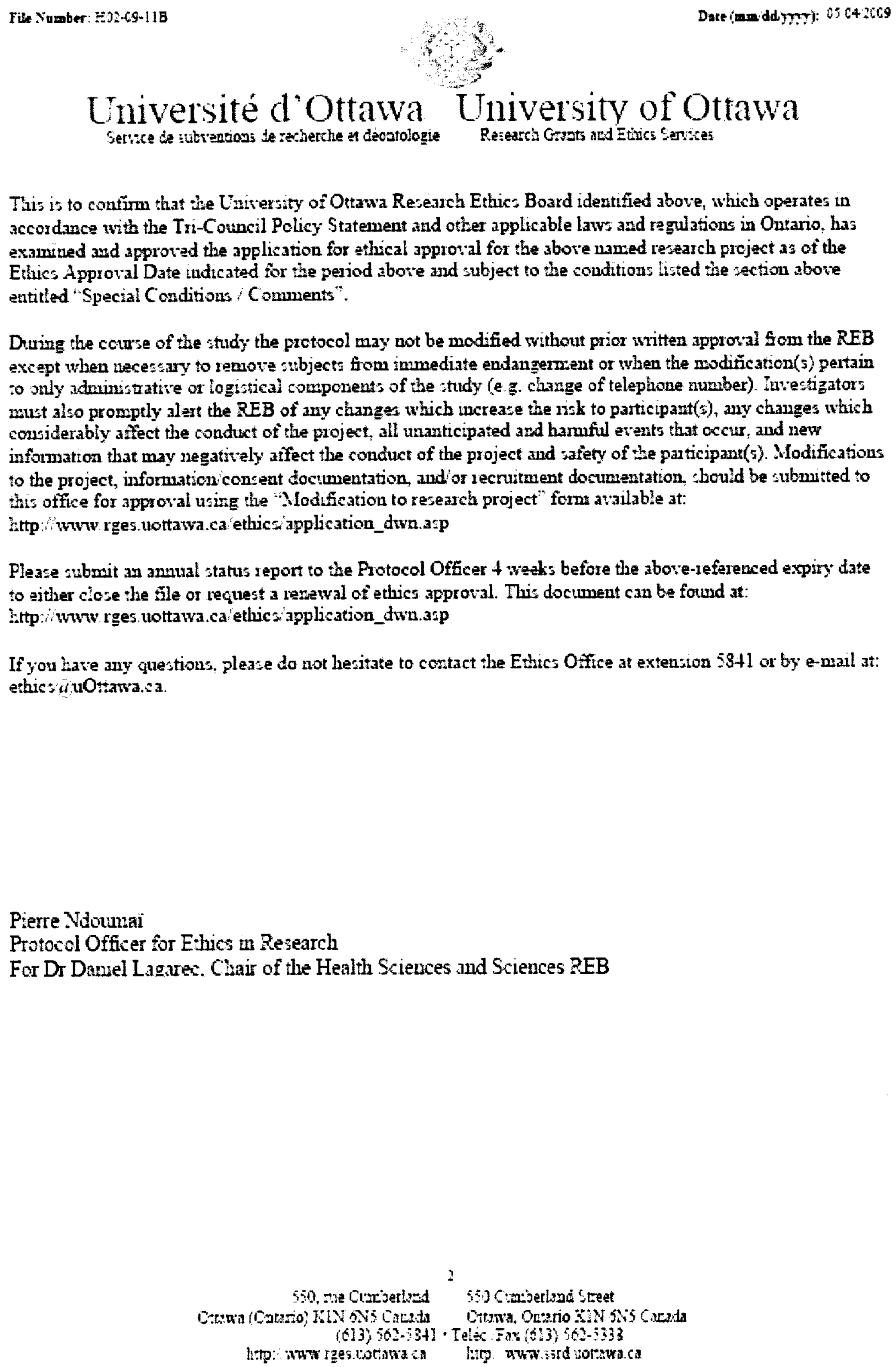
Approval Date (mm/dd/yyyy)	Expiry Date (mm/dd/yyyy)	Approval Type
05/04/2009	05/03/2010	Ia

(Ia: Approval, Ib: Approval for initial stage only)

Special Conditions / Comments:
N/A

1

550, rue Cumberland 550 Cumberland Street
Ottawa (Ontario) K1N 6N5 Canada Ottawa, Ontario K1N 6N5 Canada
(613) 562-5841 • Téléc./Fax (613) 562-5338
http://www.rges.uottawa.ca http://www.ssrd.uottawa.ca



File Number: H02-09-11B

Date (mm/dd/yyyy): 05/04/2009

Université d'Ottawa University of Ottawa
Service de subventions de recherche et déontologie Research Grants and Ethics Services

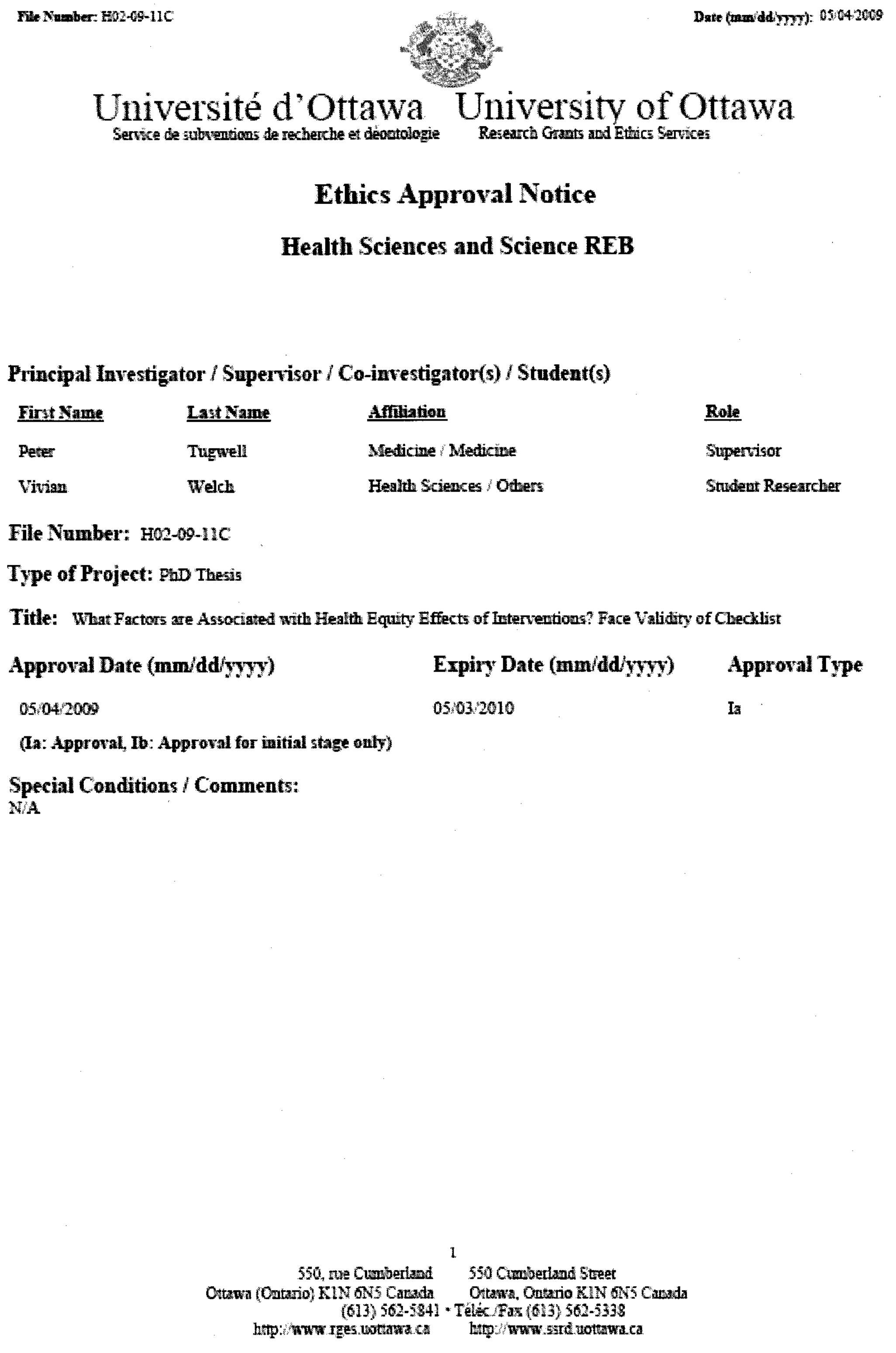
This is to confirm that the University of Ottawa Research Ethics Board identified above, which operates in accordance with the Tri-Council Policy Statement and other applicable laws and regulations in Ontario, has examined and approved the application for ethical approval for the above named research project as of the Ethics Approval Date indicated for the period above and subject to the conditions listed the section above entitled "Special Conditions / Comments".

During the course of the study the protocol may not be modified without prior written approval from the REB except when necessary to remove subjects from immediate endangerment or when the modification(s) pertain to only administrative or logistical components of the study (e.g. change of telephone number). Investigators must also promptly alert the REB of any changes which increase the risk to participant(s), any changes which considerably affect the conduct of the project, all unanticipated and harmful events that occur, and new information that may negatively affect the conduct of the project and safety of the participant(s). Modifications to the project, information/consent documentation, and/or recruitment documentation, should be submitted to this office for approval using the "Modification to research project" form available at: http://www.rges.uottawa.ca/ethics/application_dwn.asp

Please submit an annual status report to the Protocol Officer 4 weeks before the above-referenced expiry date to either close the file or request a renewal of ethics approval. This document can be found at: http://www.rges.uottawa.ca/ethics/application_dwn.asp

If you have any questions, please do not hesitate to contact the Ethics Office at extension 5841 or by e-mail at: ethics@uOttawa.ca.

Pierre Ndoumaï
Protocol Officer for Ethics in Research
For Dr Daniel Lagarec, Chair of the Health Sciences and Sciences REB



File Number: H02-09-11C

Date (mm/dd/yyyy): 05/04/2009

Université d'Ottawa University of Ottawa

Service de subventions de recherche et déontologie Research Grants and Ethics Services

Ethics Approval Notice

Health Sciences and Science REB

Principal Investigator / Supervisor / Co-investigator(s) / Student(s)

First Name	Last Name	Affiliation	Role
Peter	Tugwell	Medicine / Medicine	Supervisor
Vivian	Welch	Health Sciences / Others	Student Researcher

File Number: H02-09-11C

Type of Project: PhD Thesis

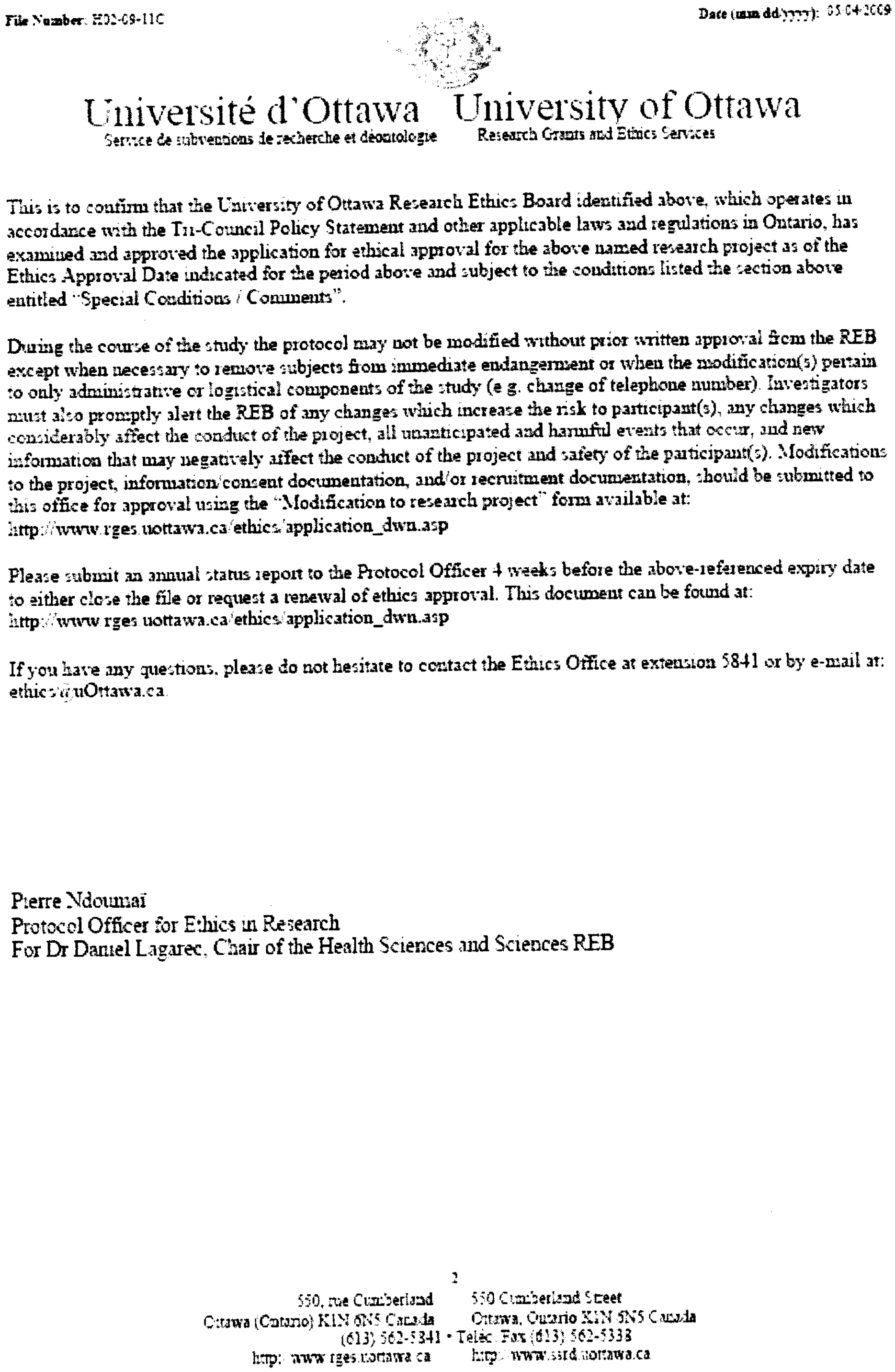
Title: What Factors are Associated with Health Equity Effects of Interventions? Face Validity of Checklist

Approval Date (mm/dd/yyyy)	Expiry Date (mm/dd/yyyy)	Approval Type
05/04/2009	05/03/2010	Ia

(Ia: Approval, Ib: Approval for initial stage only)

Special Conditions / Comments:
N/A

550, rue Cumberland 550 Cumberland Street
Ottawa (Ontario) K1N 6N5 Canada Ottawa, Ontario K1N 6N5 Canada
(613) 562-5841 • Téléc./Fax (613) 562-5338
http://www.rges.uottawa.ca http://www.ssrd.uottawa.ca



File Number: H02-09-11C

Date (mm/dd/yyyy): 05/04/2009

Université d'Ottawa University of Ottawa

Service de subventions de recherche et déontologie Research Grants and Ethics Services

This is to confirm that the University of Ottawa Research Ethics Board identified above, which operates in accordance with the Tri-Council Policy Statement and other applicable laws and regulations in Ontario, has examined and approved the application for ethical approval for the above named research project as of the Ethics Approval Date indicated for the period above and subject to the conditions listed the section above entitled "Special Conditions / Comments".

During the course of the study the protocol may not be modified without prior written approval from the REB except when necessary to remove subjects from immediate endangerment or when the modification(s) pertain to only administrative or logistical components of the study (e.g. change of telephone number). Investigators must also promptly alert the REB of any changes which increase the risk to participant(s), any changes which considerably affect the conduct of the project, all unanticipated and harmful events that occur, and new information that may negatively affect the conduct of the project and safety of the participant(s). Modifications to the project, information/consent documentation, and/or recruitment documentation, should be submitted to this office for approval using the "Modification to research project" form available at: http://www.rges.uottawa.ca/ethics/application_dwn.asp

Please submit an annual status report to the Protocol Officer 4 weeks before the above-referenced expiry date to either close the file or request a renewal of ethics approval. This document can be found at: http://www.rges.uottawa.ca/ethics/application_dwn.asp

If you have any questions, please do not hesitate to contact the Ethics Office at extension 5841 or by e-mail at: ethics@uOttawa.ca

Pierre Ndoumaï
Protocol Officer for Ethics in Research
For Dr Daniel Lagarec, Chair of the Health Sciences and Sciences REB

550, rue Cumberland 550 Cumberland Street
Ottawa (Ontario) K1N 6N5 Canada Ottawa, Ontario K1N 6N5 Canada
(613) 562-5841 • Téléc. Fax (613) 562-5338
http://www.rges.uottawa.ca http://www.ssrd.uottawa.ca

Appendix 8: Characteristics of systematic reviews chosen for testing the equity plausibility algorithm

	Difference across sex or SES? (likely or confirmed)	Population	Intervention	Comparison	Outcome	Eligible study designs	Studies included	Applicability judged by authors	Implementation	Setting	Results
Mass media to promote HIV testing [180]	Sex: Not described SES: No- "applicable in non-literate"	General public, target groups (e.g. sex workers, drug users, bisexual, pregnant women, adolescents)	Specific or general mass media campaigns, targeted at population or specific groups that aim to increase counseling and testing, including radio, TV, print, film, billboards, folk media.	Control group or pre-intervention levels, In relation to relative increase of HIV testing over time	HIV testing by the general population or specific target populations, HIV seroprevalence	RCT, CCT, interrupted time series	2 RCTs, 3 CCT, 9 ITS	"Radio and television interventions can be used in literate and non-literate communities; therefore they are applicable to LMIC"	Quality of media campaigns was difficult to assess- e.g. duration, intensity, adherence to social marketing principles	Did not find any studies in high risk groups or epidemic countries, or comparisons of different methods, or assessment of cost-effectiveness	Effective in short term (slope change 5.5 SD units (95% CI 2.4-8.6)
Population tobacco control interventions [67]	Sex: No difference found SES: Yes, for price control	People smoking, at risk of smoking, exposed to ETS or general population,	Population level tobacco control (remove subsidies on production, restricting trade, restrict advertising, health warning labels, restricting smoking in public places)	Usual exposure to smoking cessation and prevention	Changes in smoking behavior, sales for people with different demographic or socioeconomic characteristics	Any design	84 studies	Increasing the price of tobacco products is more effective in low-income populations. Smoking restrictions in workplaces and public places are not more effective among more advantaged groups	Studies often do not often describe contextual factors or co-interventions which may affect results	Over half of studies conducted in the USA and six in the UK. Priority to assess effects in other country settings and contexts	Most compelling evidence on favouring the less well-off is increased price of tobacco

	Difference across sex or SES? (likely or confirmed)	Population	Intervention	Comparison	Outcome	Eligible study designs	Studies included	Applicability judged by authors	Implementation	Setting	Results
Psychological treatment for post-traumatic stress disorder (PTSD)[181]	Sex: Yes, works better in women SES: not described	Adults suffering from traumatic stress symptoms for three months or more	Trauma-focused or group cognitive behavioural therapy, stress management, supportive therapies, psychodynamic therapies	Control, placebo, waiting list, usual care, alternative psychological therapies, Usual care allowed	Severity of clinician-rated or self-reported traumatic stress symptoms	RCTs	33 RCTs	Studies including only females, all of whom had been assaulted, produced more positive results than the overall results.	Diversity of clinical interventions	Issues with the control group since it is difficult to blind participants and therapists and to have a "placebo" for psychological therapy	Improvement in clinician assessed PTSD symptoms immediately after treatment (large effect size as assessed with SMD)
First-line drugs for hypertension[182]	SEX: No, females represented 45% of population SES: Not described	Adult with baseline resting BP => 140 mm Hg systolic or diastolic BP =>90 mm Hg. >70% patients high BP	First-line anti-hypertensive therapy: thiazide diuretics, beta blockers, calcium channel blockers, angiotensin converting enzyme (ACE) inhibitors, angiotensin II receptor antagonists or alpha adrenergic blockers	Must be placebo, or an untreated control. Other drugs allowed if taken by <50% of patients	Mortality, stroke, coronary heart disease, cardiovascular events	RCT >1 year	24 RCTs, 58,040 patients included	Females represented 45% of population, 10/24 trials reported ethnicity of 0-80% african-american), 72% of people studied were primary-prevention	72% of participants were primary prevention. Results for secondary prevention were less robust	Most participants were recruited from Western industrialized countries (15% USA, 66% Europe, 7% Australia)	Low dose thiazides decrease coronary heart disease (RR 0.72, 95% CI 0.61 to 0.84); but high dose thiazides did not (RR 1.01, 95% CI 0.85 to 1.20)
Surgical interventions for age-related cataract[183]	Patient differences: not discussed Delivery: expensive machines and highly skilled surgeons needed for most effective method; difficult in LMIC Comparator: higher prevalence of uncontrolled cataracts in LMIC	people with age-related cataract	different surgical interventions for age-related cataract	Different types of surgery or waitlists	Visual acuity	RCTs	17 RCTs, 9627 people	Need more studies in developing countries where access to expensive machines, volume of surgeries and skill of surgeons may be lower	Phacoemulsification requires expensive machine (20,000 pounds), highly skilled surgeons Barriers to surgery in LMIC	8 studies in Europe, 2 in the far east and 6 in India, 1 in Africa Higher prevalence and less control in LMIC	Pharmacoemulsification gives the best outcomes, but will only be accessible in poorer countries if the cost decreases

	Difference across sex or SES? (likely or confirmed)	Population	Intervention	Comparison	Outcome	Eligible study designs	Studies included	Applicability judged by authors	Implementation	Setting	Results
Vaccines for measles, mumps and rubella in children [247]	Patient characteristics: not described, but implications for practice state that effectiveness demonstrated world-wide Delivery: not described, but implications for practice state that effectiveness demonstrated world-wide Comparator: not described	healthy individuals up to 15 years of age	Vaccination with any combined MMR vaccine given independently, in any dose, preparation or time schedule	Do-nothing or placebo.	Clinical cases: measles, mumps or rubella	All comparative prospective or retrospective studies	5 RCTs, 1 CCT, 14 cohort studies, 5 case-control studies, 3 time series trials, 1 cross-over, 1 ecological trial, 1 self-controlled case series trial	External validity of included studies was low	inadequate description of populations, response rates, vaccine content and exposure	Setting of studies was not described in SR	Limited evidence of safety of MMR, lack of field studies on effectiveness of MMR
Antidepressants versus placebo for depression in primary care [185]	Patient characteristics: not discussed, Delivery: not discussed, primary care practitioners may be more likely to change therapy if adverse effects occur Comparator: not discussed	patients (under the age of 65 years) with depression in primary care	tricyclic antidepressants (TCAs) or selective serotonin reuptake inhibitors (SSRIs	Placebo	Depression symptoms	RCTs	14 RCTs	Results apply to patients with major depressive disorder and heterogeneous depression; NNT of 6-16 are comparable to other treatments in primary care	All studies gave a clear description of the treatments and concurrent therapies	9/14 studies had “representative samples” (ie, outpatient, general practice settings)	Benefit of TCA and SSRI (NNT of 6-16), consistent across subgroup and sensitivity analyses for quality, setting and type of depression

	Difference across sex or SES? (likely or confirmed)	**Population**	**Intervention**	**Comparison**	**Outcome**	**Eligible study designs**	**Studies included**	**Applicability judged by authors**	**Implementation**	**Setting**	**Results**
Artemisinin-based combination therapy for treating uncomplicated malaria [186]	Patient characteristics: endemicity and resistance affect effectiveness Delivery: not discussed Comparator: not discussed	People with uncomplicated *P. falciparum*malaria	Artemisinin-based Combination Therapy (ACT)	Other antimalarials	Treatment failure	RCTs	50 RCTs	Interpret summary statistics with caution since there are changing patterns of resistance which vary with place and time	Very young children and pregnant women were excluded in these trials	31 RCTs in Africa, 17 n Asia, in south America, 1 in Oceania; resistance and endemicity are important	All five ACTS met the WHO criteria for less than 10% failure rate
Primary safety belt laws to reduce fatal and nonfatal injuries in vehicle crashes [187]	Patient characteristics: more effective for lower use groups (e.g. African-American, Hispanic, male) Delivery: no difference in enforcement across race/ethnicity Comparator: prevalence of wearing seatbelt is lower for males, rural people, African-American and Hispanic	General population	Primary enforcement laws allow a police officer to stop a motorist solely for not wearing a safety belt	No safety belt laws	Safety belt use, crash-related morbidity and mortality	Any evaluations of safety belt laws	13 studies of primary vs. secondary enforcement	Seatbelt use increased more in African-Americans and Hispanics than whites; also higher uptake by rural people and males	Studies show that enforcement is similar across race and ethnicity; public opposition may be a barrier to implementation	Diverse settings and target groups;	Fatalities decreased by 8% with primary compared to secondary laws

	Difference across sex or SES? (likely or confirmed)	**Population**	**Intervention**	**Comparison**	**Outcome**	**Eligible study designs**	**Studies included**	**Applicability judged by authors**	**Implementation**	**Setting**	**Results**
Hand washing for preventing diarrhea [188]	Patient characteristics: not described (though analyses were age and gender adjusted) Delivery: need long-term studies of effectiveness in LMIC (less monitoring) Comparator: availability of water and materials in LMIC	General population in institutions (e.g. day-care centres), households, or communities	interventions to promote hand washing e.g., small group discussions, multimedia communication campaigns (tv,radio, leaflets, comic books, songs, slide shows, use of T-shirts and badges, pictorial stories, dramas, and games.)	No intervention to promote handwashing	Diarrhoeal episodes in children and adults	RCTs	14 RCTs	Uncertain if effectiveness will be maintained if scaled up to larger population with less monitoring and over longer time period Results cannot be generalized to all ages (all children were less than 15 and most less than 7 yr years)	Handwashing may require infrastructure, cultural, behavioural changes and resources; assessed hand washing method, style, materials provided, water availability	8 in high-income countries, 5 in LMIC, 1 in HIV/AIDS patients	39% reduction in diarrhea episodes in children in HIC, 32% reduction in LMIC

Appendix 9: Sample equity plausibility algorithm survey provided to raters

Task: Please think about the question described in PICO format above. Do you think there are likely to be important differences in the magnitude of relative effects across: 1)women vs. men or 2)poorer vs. richer people (low socioeconomic status- SES)?

About you: Are you practitioner, consumer, statistician: please describe? ____________

How many years experience with Cochrane and evidence-based medicine?___________

What is your area of research/expertise?_____________________

Question	**Surgical interventions for age-related cataract**	**Vaccines for measles, mumps and rubella in children**	**Antidepressants versus placebo for depression in primary care**	**Artemisinin-based combination therapy for treating uncomplicated malaria**	**Primary safety belt laws to reduce fatal and nonfatal injuries in vehicle crashes**	**Hand washing for preventing diarrhoea**
1.Are there differences in **patient/community/ population characteristics** (e.g. underlying pathophysiology, comorbidities, patient attitudes, etc.) that are likely to create **important differences** in the magnitude of relative effect of the intervention versus the control for the outcome of interest?	Gender/Sex	Gender/Sex	Gender/Sex	Gender/Sex	Gender/Sex	Gender/Sex
	Y/N	**Y/N**	**Y/N**	**Y/N**	**Y/N**	**Y/N**
[e.g. <u>pathophysiology</u>: aspirin has greater relative effect on stroke prevention in women than men; e.g. <u>co-morbidities</u>: poor people have higher co-morbidities than richer people; e.g. <u>patient attitudes:</u> women may have different preferences for type of care than men]	Low SES	Low SES	Low SES	Low SES	Low SES	Low SES
	Y/N	**Y/N**	**Y/N**	**Y/N**	**Y/N**	**Y/N**

Question	**Surgical interventions for age-related cataract**	**Vaccines for measles, mumps and rubella in children**	**Antidepressants versus placebo for depression in primary care**	**Artemisinin-based combination therapy for treating uncomplicated malaria**	**Primary safety belt laws to reduce fatal and nonfatal injuries in vehicle crashes**	**Hand washing for preventing diarrhoea**
2.Are there differences in the way that the **intervention is delivered** (e.g. provider compliance, provider skill, technical resources, availability of treatments) that are likely to create **important differences** in the magnitude of the relative effect of the intervention versus the control for the outcome of interest?	Gender/sex	Gender/sex	Gender/sex	Gender/sex	Gender/sex	Gender/sex
	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N
	Low SES	Low SES	Low SES	Low SES	Low SES	Low SES
e.g. surgeon skill in hospitals with less surgeries may be lower than in clinical trials	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N
3.Are there differences in the **comparator** across patient, community or population that are likely to create **important differences** in magnitude of relative effects?	Gender/sex	Gender/sex	Gender/sex	Gender/sex	Gender/sex	Gender/sex
	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N
	Low SES	Low SES	Low SES	Low SES	Low SES	Low SES
e.g. poor people may live in an environment with poorer water quality	Y/N	Y/N	Y/N	Y/N	Y/N	Y/N
4.What is your reason for your answer? Theory, personal experience, other data? Please feel free to use more space and to separate gender/sex and SES	Gender: SES:					

Reference List

1. Diderichsen,F., Evans,T., & Whitehead,M. The Social Basis of Disparities in Health. in *Challenging Inequities in Health. From Ethics to Action* (eds. Evans,T., Whitehead,M., Diderichsen,F., Bhuiya,A. & Wirth,M.) 13-23 (Oxford University Press, New York, 2001).

2. Tugwell,P., de,S.D., Hawker,G., & Robinson,V. Applying clinical epidemiological methods to health equity: the equity effectiveness loop. *BMJ* **332**, 358-361 (2006).

3. Mackenbach,J.P. Tackling inequalities in health: the need for building a systematic evidence base. *Journal of Epidemiology & Community Health* **57**, 162 (2003).

4. World Health Assembly. Resolution 5834 on the Ministerial Summit on Health Research. 5834. 2005.
Ref Type: Bill/Resolution

5. Lavis,J.N., Davies,H.T.O., & Gruen,R.L. Working within and beyond the Cochrane Collaboration to make systematic reviews more useful to healthcare managers and policy makers. *Healthcare Policy* **1**, 21-33 (2006).

6. Tsikata,S. *et al.* Do Cochrane systematic reviews contain useful information about health equity? [abstract]. *XI. Cochrane Colloquium. : Evidence, Health Care and Culture; 2003. Oct 26 31; Barcelona. , Spain.*77 (2003).

7. Petticrew,M., Whitehead,M., Macintye,S.J., Graham,H., & Egan,M. Evidence for public health policy on inequalities: 1: the reality according to policymakers. *J EPIDEMIOL COMMUNITY HEALTH* **58**, 811-816 (2004).

8. Schmid,C.H., Lau,J., McIntosh,M.W., & Cappelleri,J.C. An empirical study of the effect of the control rate as a predictor of treatment efficacy in meta-analysis of clinical trials. *Statistics in Medicine* **17**, 1923-1942 (1998).

9. Furukawa,T.A., Guyatt,G.H., & Griffith,L.E. Can we individualize the 'number needed to treat'? An empirical study of summary effect measures in meta-analyses. Int J Epidemiol 2002;31:72-76. *Int. J. Epidemiol.* **31**, 72-76 (2002).

10. Ogilvie,D. *et al.* The harvest plot: a method for synthesising evidence about the differential effects of interventions. [Review] [20 refs]. *BMC Medical Research Methodology* **8**, 8 (2008).

11. Anderson,G.M. *et al.* Both clinical epidemiology and population health perspectives can define the role of health care in reducing health disparities. *Journal of Clinical Epidemiology* **58**, 757-762 (2005).

12. Commission on Social Determinants of Health Priority Public Health Conditions Knowledge Network. Scoping paper: Priority Public Health Conditions. 2007. Geneva, World Health Organization.
Ref Type: Report

13. Parkin,P.C. *et al.* Evaluation of a promotional strategy to increase bicycle helmet use by children. *Pediatrics* **91**, 772-777 (1993).

14. Lewis,G.H., Osborne,D.C., & Brown,A.C. Partial smoking ban would worsen health inequalities. *British Medical Journal* **332**, 362 (2006).

15. Lagarde,M., Haines,A., & Palmer,N. The impact of conditional cash transfers on health outcomes and use of health services in low and middle income countries. *Cochrane Database of Systematic Reviews* **4**, (2009).

16. Sackett,D.L. Why randomized controlled trials fail but needn't: 2. Failure to employ physiological statistics, or the only formula a clinician-trialist is ever likely to need (or understand!). *CMAJ Canadian Medical Association Journal* **165**, 1226-1237 (2001).

17. The Editorial Team Cochrane Methodology Review Group Cochrane Methodology Review Group. About The Cochrane Collaboration (Cochrane Review Groups (CRGs)). *Cochrane Database of Systematic Reviews* **Issue 2, Art. No.: METHOD.**, (2005).

18. Marmot,M. & Commission on Social Determinants of Health Achieving health equity: from root causes to fair outcomes. *Lancet* **370**, 1153-1163 (2007).

19. Lavis,J.N. *et al.* Towards systematic reviews that inform health care management and policy-making. *J Health Serv Res Policy* **10**, 35-48 (2005).

20. Tugwell,P., Petticrew,M., Robinson,V., Kristjansson,E., & Maxwell,L. Cochrane and Campbell Collaborations, and health equity. *Lancet* **367**, 1128-1130 (2006).

21. Labonte,R. & Schrecker,T. Globalization and social determinants of health: Introduction and methodological background (Part 1 of 3). *Global Health* **3**, (2007).

22. Whitehead,M. The Concepts and Principles of Equity and Health. *INT J HEALTH SERV* **22**, 429-445 (1992).

23. Marmot,M. *et al.* Closing the gap in a generation: health equity through action on the social determinants of health. *Lancet* **372**, 1661-1669 (2008).

24. Evans,T., Whitehead,M., Diderichsen,F., Bhuiya,A., & Wirth,M. *Challenging Inequities in Health: From Ethics to Action*(Oxford University Press, London, 2002).

25. Waters,E. *et al.* Evidence synthesis, upstream determinants and health inequalities: the role of a proposed new Cochrane Public Health Review Group. *Eur J Public Health* **18**, 221-223 (2008).

26. Macinko,J. & Starfield,B. Annotated bibliography on equity in health, 1980-2001. *International journal for equity in health* **1**, 1 (2002).

27. Sen,A. Why health equity? *Health Econ* **11**, 659-666 (2002).

28. Nussbaum,M.C. Capabilities as Fundamental Entitlements: Sen and Social Justice. *Feminist Economics* **9**, 33-59 (2003).

29. Rawls,J. *A Theory of Justice*(Harvard University Press, Cambridge, MA, 1971).

30. Graham,H. *Unequal Lives: Health and Socioeconomic Inequalities*(Open University Press, London, 2007).

31. Braveman,P. Health disparities and health equity: concepts and measurement. *Annu Rev Public Health* **27**, 167-194 (2006).

32. Harper,S., King,N.B., Meersman,S.C.R.M.E., Breen,N., & Lynch,J. Implicit value judgments in the measurement of health inequalities. *Milbank Quarterly* **88**, 4-29 (2010).

33. Graham,H. & Kelly,M.P. Health inequalities: concepts, frameworks and policy. 2004. London, UK.
Ref Type: Report

34. Keppel,K., Pamuk,E., Lynch,J., & et al. Methodological issues in measuring health disparities. National Center for Health Statistics. Vital Health Stat Series 2, 141. 2005. National Centre for Health Statistics.
Ref Type: Report

35. Keppel,K., Pamuk,E., Lynch,J., & et al. Methodological issues in measuring health disparities. National Center for Health Statistics. Vital Health Stat Series 2, 141. 2005. National Centre for Health Statistics.
Ref Type: Report

36. Anand,S., Diderichsen,F., Evans,T., Shkolnikov,V.M., & Wirth,M. Measuring disparities in health: methods and indicators in *Challenging Inequities in Health: From Ethics to Action* (eds. Evans,T., Whitehead,M., Diderichsen,F., Bhuiya,A. & Wirth,M.) 49-66 (Oxford University Press, London, 2002).

37. Kavanagh,J. *et al.* Health promotion, public health and the health of young people: a systematic map of inequalities research [abstract]. *XV Cochrane Colloquium; 2007 Oct 23-27; Sao Paulo, Brazil.* 115-116 (2007).

38. Evans,T. & Brown,H. Road traffic crashes: operationalizing equity in the context of health sector reform. *Injury Control & Safety Promotion* **10**, 11-12 (2003).

39. Adamek,M.E., Slater,G.Y., Adamek,M.E., & Slater,G.Y. Depression and anxiety. *J GERONTOL SOC WORK* **50 Suppl 1**, 153-189 (2008).

40. Bartels,S.J.D. Evidence-based practices in geriatric mental health care: An overview of systematic reviews and meta-analyses. *Psychiatric Clinics of North America* **26**, Dec (2003).

41. Tugwell,P., Petticrew,M., Robinson,V., Kristjansson,E., & Maxwell,L. Cochrane and Campbell Collaborations, and health equity. *Lancet* **367**, 1128-1130 (2006).

42. Oxman,A., Lavis,J., Lewin,S., & Fretheim,A. SUPPORT Tools for evidence-informed health Policymaking (STP) 10: Taking equity into consideration when assessing the findings of a systematic review. *Health Research Policy and Systems* **7**, S10 (2009).

43. Roland,M. & Torgerson,D.J. What are pragmatic trials? *BMJ* **316**, 285-286 (1998).

44. Thorpe,K.E. *et al.* A pragmatic-explanatory continuum indicator summary (PRECIS): a tool to help trial designers. *Journal of Clinical Epidemiology* **62**, 464-475 (2009).

45. Arblaster,L. *et al.* A systematic review of the effectiveness of health service interventions aimed at reducing inequalities in health. *Journal of Health Services & Research Policy* **1**, 93-103 (1996).

46. Gunning-Schepers,L. & Gepkens,A. Reviews of interventions to reduce social inequalities in health: research and policy implications. *Health Education Journal* **55**, 226-238 (1996).

47. Kelly,M.P. *et al.* The social determinants of health: Developing an evidence base for political action. 1-165. 2007. Geneva, World Health Organisation.
Ref Type: Report

48. Ogilvie,D. & Petticrew,M. Reducing social inequalities in smoking: can evidence inform policy? A pilot study. *Tob Control* **13**, 129-131 (2004).

49. Ogilvie,D. & Petticrew,M. Reducing social inequalities in smoking: can evidence inform policy? A pilot study. [Review] [30 refs]. *Tobacco Control* **13**, 129-131 (2004).

50. Hopewell,S., McDonald,S., Clarke,M.J., & Egger,M. Grey literature in meta-analyses of randomized trials of health care interventions. *Cochrane Database of Systematic Reviews* **Issue 2. Art. No.: MR000010. DOI: 10.1002/14651858.MR000010.pub3.**, (2007).

51. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.2 [updated September 2009]*(The Cochrane Collaboration,2009).

52. Oxman,A.D. & Guyatt,G.H. A consumer's guide to subgroup analyses. *ANN INTERN MED* **116**, 78-84 (1992).

53. Rothwell,P.M. Treating individuals 2. Subgroup analysis in randomised controlled trials: importance, indications, and interpretation. *Lancet* **365**, 176-186 (2005).

54. Thompson,S.G. & Higgins,J.P. Treating individuals 4: can meta-analysis help target interventions at individuals most likely to benefit? *Lancet* **365**, 341-346 (2005).

55. National Health and Medical Research Council. *How to use the evidence: assessment and application of scientific evidence.* 2000. Canberra, Australia, National Health and Medical Research Council.
Ref Type: Report

56. Green,L.W. & Glasgow,R.E. Evaluating the relevance, generalization, and applicability of research: issues in external validation and translation methodology. [Review] [91 refs]. *Evaluation & the Health Professions* **29**, 126-153 (2006).

57. National Health and Medical Research Council. How to use the evidence: assessment and application of scientific evidence. Handbook series on preparing clinical practice guidelines. 2000. Sydney, Australia, National Health and Medical Research Council.
Ref Type: Report

58. Dans,A.L., Dans,L.F., & Guyatt,G. Applying results to individual patients in *Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice* (eds. Guyatt,G., Rennie,D., Meade,M.O. & Cook,D.J.) 273-289 (The McGraw-Hill Companies, Inc., New York, USA, 2008).

59. Briss,P.A. *et al.* Developing an evidence-based Guide to Community Preventive Services--methods. The Task Force on Community Preventive Services. *American Journal of Preventive Medicine* **18**, Suppl-43 (2000).

60. Green,L.W. & Glasgow,R.E. Evaluating the relevance, generalization, and applicability of research: issues in external validation and translation methodology. [Review] [91 refs]. *Evaluation & the Health Professions* **29**, 126-153 (2006).

61. Tackling health inequalities in Europe: an integrated approach: EUROTHINE: Final report. 2007. Rotterdam, Erasmus University Medical Center.
Ref Type: Report

62. Wang,S., Moss,J.R., & Hiller,J.E. Applicability and transferability of interventions in evidence-based public health. *Health Promotion International* **21**, 76-83 (2006).

63. Evans,R.G. & Stoddart,G.L. Producing health, consuming health care. Soc Sci Med. 1990;31:1347-1363. *Social Science & Medicine* **31**, 1347-1363 (1990).

64. Krieger,N. Epidemiology and the web of causation: has anyone seen the spider? *Social Science & Medicine* **39**, 887-903 (1994).

65. Starfield,B. Pathways of influence on equity in health. *Social Science & Medicine* **64**, 1355-1362 (2007).

66. Rose,G. Sick individuals and sick populations. *Int. J. Epidemiol.* **14**, 32-38 (1985).

67. Thomas,S. *et al.* Population tobacco control interventions and their effects on social inequalities in smoking: systematic review. [Review] [109 refs]. *Tobacco Control* **17**, 230-237 (2008).

68. Hoving,J.L., Broekhuizen,M.L., & Frings-Dresen.M.H. Return to work of breast cancer survivors: a systematic review of intervention studies. *BMC Cancer* **9**, 117 (2009).

69. Chopra,M., Munro,S., Lavis,J.N., Vist,G., & Bennett,S. Effects of policy options for human resources for health: an analysis of systematic reviews. *Lancet* **371**, 668-774 (2008).

70. World Health Organization. The world health report 2008 : primary health care now more than ever. 2008. Geneva, Switzerland, World Health Organization.
Ref Type: Report

71. 58th World Health Assembly Resolution. World Health Organization . 2008. 4-17-2006.
Ref Type: Electronic Citation

72. Role of Science in the Information Society. Role of Science in Information Society . 2008. 5-9-2005.
Ref Type: Electronic Citation

73. Lavis,J.N., Davies,H.T.O., & Gruen,R.L. Working within and beyond the Cochrane Collaboration to make systematic reviews more useful to healthcare managers and policy makers. *Healthcare Policy* **1**, 21-33 (2006).

74. Pope,C., Mays,N., & Popay,J. Informing policy making and management in healthcare: The place for synthesis. *Healthcare Policy* **1**, 43-48 (2006).

75. Lavis,J.N. *et al.* Towards systematic reviews that inform health care management and policy-making. *J Health Serv Res Policy* **10**, 35-48 (2005).

76. Petticrew,M., Whitehead,M., Macintye,S.J., Graham,H., & Egan,M. Evidence for public health policy on inequalities: 1: the reality according to policymakers. *J Epidemiol Community Health* **58**, 811-816 (2004).

77. Tugwell,P., de,S.D., Hawker,G., & Robinson,V. Applying clinical epidemiological methods to health equity: the equity effectiveness loop. *BMJ* **332**, 358-361 (2006).

78. Whitehead,M. The Concepts and Principles of Equity and Health. *International Journal of Health Services* **22**, 429-445 (1992).

79. Kawachi,I. Social capital and community effects on population and individual health. *Socioeconomic Status and Health in Industrial Nations* **896**, 120-130 (1999).

80. Arblaster,L. *et al.* A systematic review of the effectiveness of health service interventions aimed at reducing inequalities in health. [Review] [70 refs]. *Journal of Health Services & Research Policy* **1**, 93-103 (1996).

81. Gepkens,A. & Gunning-Schepers,L.J. Interventions to reduce socioeconomic health differences: A review of the international literature. *European Journal of Public Health* **6**, 218-226 (1996).

82. Oliver,S. *et al.* Health promotion, inequalities and young people's health: a systematic review of research. 2008. London, EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.
Ref Type: Report

83. Koolman,X. & van Doorslaer,E. On the interpretation of a concentration index of inequality. *Health Econ* **13**, 649-656 (2004).

84. Wagstaff,A. & Waters,H. How were the Reaching the Poor Studies Done? in *Reaching the Poor with health, nutrition and population services: What works, what doesn't and why.* (ed. Gwatkin DR,W.A.Y.AS.) 27-46 (The World Bank, Washington, 2005).

85. Chan,A.W., Hrobjartsson,A., Haahr,M.T., Gotzsche,P.C., & Altman,D.G. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. *JAMA* **291**, 2457-2465 (2004).

86. Oxman,A.D. & Guyatt,G.H. A consumer's guide to subgroup analyses. *Ann Intern Med* **116**, 78-84 (1992).

87. Sampson,M., McGowan,J., Lefebvre,C., Moher,D., & Grimshaw,J. PRESS: peer review of electronic search strategies. 2008. Ottawa, Canadian Agency of Drugs and Technology for Health (CADTH).
Ref Type: Report

88. Berlin,J.A. Does blinding of readers affect the results of meta-analyses? University of Pennsylvania Meta-analysis Blinding Study Group. *Lancet* **350**, 185-186 (1997).

89. Moher,D., Liberati,A., Tetzlaff,J., Altman,D.G., & PRISMA Group Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* **339**, (2009).

90. Rothwell,P.M. Treating individuals 2. Subgroup analysis in randomised controlled trials: importance, indications, and interpretation. *Lancet* **365**, 176-186 (2005).

91. Gulmezoglu,M., de,O.M., & Villar,J. Effectiveness of interventions to prevent or treat impaired fetal growth. *Obstetrical & Gynecological Survey* **52**, 139-149 (1997).

92. Barlow,J.H., Ellard,D.R., Barlow,J.H., & Ellard,D.R. Psycho-educational interventions for children with chronic disease, parents and siblings: an overview of the research evidence base. [Review] [29 refs]. *CHILD CARE HEALTH DEV* **30**, 637-645 (2004).

93. Espinosa-Aguilar,A. *et al.* Clinical guideline for diagnosis and treatment of depression in elderly. [Spanish]. [References]. *Salud Mental* **30**, Nov-Dec (2007).

94. Craig,L.A., Browne,K.D., & Stringer,I. Treatment and sexual offense recidivism. *Trauma & Abuse: A Review Journal* **4**, 89-95 (2003).

95. Gaes,G.G., Flanagan,T.J., & Motiuk,L.L. Adult correctional treatment. *Michael Tonry and Joan Petersilia (Ed). (1999). Prisons (pp. 361-426). Chicago*(1999).

96. Bambra,C. *et al.* Working for health? Evidence from systematic reviews on the effects on health and health inequalities of organisational changes to the psychosocial work environment. *PREV MED* **48**, 454-461 (2009).

97. Bambra,C. *et al.* Tackling the wider social determinants of health and health inequalities: evidence from systematic reviews. *J EPIDEMIOL COMMUNITY HEALTH* **Aug 19, (epub ahead of print)**, (2009).

98. Ball,L. Low birth weight: exploring an enigma of failure. *British Journal of Midwifery* **12**, 374-379 (2004).

99. Tudiver,S., deBoscoe,M., Runnels,V., & Doull,M. Context Matters: Applying Sex and Gender-based Analysis to Cochrane Reviews. 6th Annual Canadian Cochrane Symposium. 2008. 3-6-2008.
Ref Type: Conference Proceeding

100. Odierna,D.H. & Bero,L.A. Systematic reviews reveal unrepresentative evidence for the development
of drug formularies for poor and nonwhite populations. *Journal of Clinical Epidemiology* **Apr 16. [Epub ahead of print]Click here to read**, (2009).

101. Tugwell,P. *et al.* Is health equity considered in systematic reviews of the Cochrane Musculoskeletal Group? *Arthritis Care and Research* **59**, 1603-1610 (2008).

102. Main,C. *et al.* Population tobacco control interventions and their effects on social inequalities in smoking: placing an equity lens on existing systematic reviews. *BMC Public Health* **8**, (2008).

103. Lewin,S. *et al.* Supporting the delivery of cost-effective interventions in primary health-care systems in low-income and middle-income countries: an overview of systematic reviews. *Lancet* **372**, 928-939 (2008).

104. Althabe,F. *et al.* Strategies for improving the quality of health care in maternal and child health in low- and middle-income countries: an overview of systematic reviews. *Paediatric and Perinatal Epidemiology* **22 Suppl 1**, 42-60 (2008).

105. Johnson,S.M., Karvonen,C.A., Phelps,C.L., Nader,S., & Sanborn,B.M. Assessment of analysis by gender in the Cochrane reviews as related to treatment of cardiovascular disease. [Review] [48 refs]. *Journal of Women's Health* **12**, 449-457 (2003).

106. Tudiver,S., deBoscoe,M., Runnels,V., & Doull,M. Context Matters: Applying Sex and Gender-based Analysis to Cochrane Reviews. 6th Annual Canadian Cochrane Symposium. 2008. 3-6-2008.
Ref Type: Conference Proceeding

107. Nasser,M. *et al.* Guidelines for preparing Cochrane Reviews relevant to Developing Countries. Cochrane Colloquia XV. 2007.
Ref Type: Conference Proceeding

108. Stewart,D. *et al.* A critical appraisal of literature reviews about the transition to adulthood for youth with disabilities. *Physical & Occupational Therapy in Pediatrics* **26**, 5-24 (2006).

109. D'Souza,L. & Garcia,J. Improving services for disadvantaged childbearing women. *CHILD CARE HEALTH DEV* **30**, 599-611 (2004).

110. Ball,L. & Kirkham,M. Low birth weight in Sheffield. A review of interventions and their effectiveness. 1-57. 2002. Sheffield, UK, University of Sheffield, Women's Informed Childbearing and Health Research Group.
Ref Type: Report

111. Morrison,A. & Wertheimer,A.I. Compilation of quantitative overviews of studies of adherence. *Drug Information Journal. 38(2)()(pp 197-210), 2004. Date of Publication: 2004.*197-210 (2004).

112. Adams,J. & White,M. When the population approach to prevention puts the health of individuals at risk. *Int. J. Epidemiol.* **34**, 40-43 (2005).

113. Glasziou,P.P. & Sanders,S.L. Investigating causes of heterogeneity in systematic reviews. *Statistics in Medicine* **21**, 1503-1511 (2002).

114. Anderson,G.M. *et al.* Both clinical epidemiology and population health perspectives can define the role of health care in reducing health disparities. *J Clin Epidemiol* **58**, 757-762 (2005).

115. Kredo,T., Van der Walt,J.S., Siegfried,N., & Cohen,K. Kredo T, Van der Walt JS, Siegfried N, Cohen K. Therapeutic drug monitoring of antiretrovirals for people with HIV. *COCHRANE DATABASE SYST REV*Art. No.: CD007268. DOI: 10.1002/14651858.CD007268.pub2. (2009).

116. Sun,X., Briel,M., Walter,S.D., & Guyatt,G.H. Is a subgroup effect believable? Updating criteria to evaluate the credibility of subgroup analyses. *British Medical Journal* **340**, (2010).

117. Spitzer,D.L. Gender and Sex-Based Analysis in Health Research: A Guide for CIHR Researchers and Reviewers. 2008. Ottawa, Canadian Institutes of Health Research. 4-20-2008.
Ref Type: Report

118. Moher,D., Tetzlaff,J., Tricco,A.C., Sampson,M., & Altman,D.G. Epidemiology and reporting characteristics of systematic reviews. *PLoS Med* **4**, e78 (2007).

119. Thorpe,K.E. *et al.* A pragmatic-explanatory continuum indicator summary (PRECIS): a tool to help trial designers. *Journal of Clinical Epidemiology* **62**, 464-475 (2009).

120. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.0 [updated February 2008]*(The Cochrane Collaboration,2008).

121. Nasser,M. *et al.* Guidelines for preparing Cochrane Reviews relevant to Developing Countries. Cochrane Colloquia XV. 2007.
Ref Type: Conference Proceeding

122. Sutton,A.J., Kendrick,D., & Coupland,C.A.C. Meta-analysis of individual- and aggregate-level data. *Statistics in Medicine* **27**, 651-669 (2008).

123. UNICEF. Progress for Children: a child survival report card. 2009. UNICEF, Geneva.
Ref Type: Report

124. Jewell,C.J. & Bero,L.A. "Developing good taste in evidence": facilitators of and hindrances to evidence-informed health policymaking in state government. *Milbank Quarterly* **86**, 177-208 (2008).

125. Petticrew,M. *et al.* Better evidence about wicked issues in tackling health inequities. *J Public Health* **31**, 453-456 (2009).

126. Masanja,H., Schellenberg,J.A., de Savigny,D., Mshinda,H., & Victora,C.G. Impact of Integrated Management of Childhood Illness on inequalities in child health in rural Tanzania. *Health Policy and Planning* **20**, i77-i84 (2005).

127. Hypertension detection and follow-up program cooperative group Educational Level and 5-Year All-Cause Mortality in the Hypertension Detection and Follow-up Program. *Hypertension* **9**, 641-646 (1987).

128. Kristjansson,E.A. *et al.* School feeding for improving the physical and psychosocial health of disadvantaged elementary school children. [Review] [126 refs]. *COCHRANE DATABASE SYST REV*CD004676 (2007).

129. Early Breast Cancer Trialists' Collaborative Group (EBCTCG) Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. *Lancet* **365**, 1687-1717 (2005).

130. Montori,V.M., Wilczynski,N.L., Morgan,D., Haynes,R.B., & Hedges team Optimal search strategies for retrieving systematic reviews from Medline: analytical survey. *British Medical Journal* **330**, 1162-1163 (2005).

131. Brylewski,J. & Duggan,L. Antipsychotic medication for challenging behaviour in people with learning disability. [Review] [71 refs][Update of Cochrane Database Syst Rev. 2001;(3):CD000377; PMID: 11686959]. *Cochrane Database of Systematic Reviews*CD000377 (2004).

132. Twetman,S. *et al.* Caries-preventive effect of sodium fluoride mouthrinses: a systematic review of controlled clinical trials. [Review] [79 refs]. *Acta Odontologica Scandinavica* **62**, 223-230 (2004).

133. Adab,N., Tudur,S.C., Vinten,J., Williamson,P., & Winterbottom,J. Common antiepileptic drugs in pregnancy in women with epilepsy. [Review] [103 refs]. *Cochrane Database of Systematic Reviews*CD004848 (2004).

134. Nowak,A., Findlay,M., Culjak,G., & Stockler,M. Tamoxifen for hepatocellular carcinoma. [Review] [34 refs]. *Cochrane Database of Systematic Reviews*CD001024 (2004).

135. Benson,P.E. *et al.* Fluorides for the prevention of white spots on teeth during fixed brace treatment. [Review] [197 refs]. *Cochrane Database of Systematic Reviews*CD003809 (2004).

136. Ram,F.S., Rowe,B.H., & Kaur,B. Vitamin C supplementation for asthma. [Review] [39 refs][Update in Cochrane Database Syst Rev. 2009;(1):CD000993; PMID: 19160185].[Update of Cochrane Database Syst Rev. 2001;(4):CD000993; PMID: 11687089]. *Cochrane Database of Systematic Reviews*CD000993 (2004).

137. Swanson,H.L. Reading research for students with LD: a meta-analysis of intervention outcomes. *Journal of learning disabilities* **32**, 504-532 (1999).

138. Conaty,S., Watson,L., Dinnes,J., & Waugh,N. The effectiveness of pneumococcal polysaccharide vaccines in adults: a systematic review of observational studies and comparison with results from randomised controlled trials. *Vaccine* **22**, 3214-3224 (2004).

139. Gourlay,S.G., Stead,L.F., & Benowitz,N.L. Clonidine for smoking cessation. *Cochrane Database of Systematic Reviews* **Issue 3. Art. No.: CD000058.pub2. DOI: 10.1002/14651858.CD000058.pub2.**, (2004).

140. Ball,P., Stahlman,R., Kubin,R., Choudhri,S., & Owens,R. Safety profile of oral and intravenous moxifloxacin: cumulative data from clinical trials and postmarketing studies. *Clin Ther* **26**, 940-950 (2004).

141. Mattick,R.P., Kimber,J., Breen,C., & Davoli,M. Buprenorphine maintenance versus placebo or methadone maintenance for opioid dependence. *Cochrane Database of Systematic Reviews* **Issue 2. Art. No.: CD002207.pub2. DOI: 10.1002/14651858.CD002207. pub2.**, (2003).

142. Gundry,S., Wright,J., & Conroy,R. A systematic review of the health outcomes related to household water quality in developing countries. [Review] [51 refs]. *Journal of Water & Health* **2**, 1-13 (2004).

143. Sutton,A.J.K. Meta-analysis of individual- and aggregate-level data. *Statistics in Medicine* **27**, 28 (2008).

144. Rochon,P.A. *et al.* The inclusion of minority groups in clinical trials: problems of under representation and under reporting of data. *Accountability in Research* **11**, 215-223 (2004).

145. Geller,S.E., Adams,M.G., & carnes,M. Adherence to federal guidelines for reporting of sex and race/ethnicity in clinical trials. *Journal of Women's Health* **15**, 1123-1131 (2006).

146. Bartlett,C. *et al.* The causes and effects of socio-demographic exclusions from clinical trials. [Review] [289 refs]. *Health Technology Assessment (Winchester, England)* **9**, iii-iiv (2005).

147. Berger,J.S. *et al.* Reporting and representation of race/ethnicity in published randomized trials. *American Heart Journal* **158**, 742-747 (2009).

148. Sterne,J.A.C., Egger,M., Moher,D., & Cochrane Bias Methods Group Addressing reporting biases in *Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.2 [updated September 2009]* (eds. Higgins,J.P.T. & Green,S.) (The Cochrane Collaboration, 2009).

149. Lantz,P.M., Lichtenstein,R.L., & Pollack,H.A. Health policy approaches to population health: the limits of medicalization. *Health Affairs* **26**, 1253-1257 (2007).

150. Rubenstein,J.H. & Laine,L. Systematic review: the hepatotoxicity of non-steroidal anti-inflammatory drugs. *Aliment Pharmacol Ther* **20**, 373-380 (2004).

151. Shadish,W. & Myers,D. Campbell Collaboration Research Design Policy Brief. http://www.campbellcollaboration.org/artman2/uploads/1/Research_Design_Policy_Brief.pdf , 1-25. 2004. Campbell Collaboration.
Ref Type: Electronic Citation

152. Wilson,S.J. & Lipsey,M. The effects of school-based social information processing interventions on aggressive behavior, part I: Universal Programs. *Campbell Systematic Reviews* **5**, (2006).

153. Welsh,B.C. & Farrington,D.P. Evidence-Based Crime Prevention: Conclusions and Directions for a Safer Society. *Canadian Journal of Crimnology and Criminal Justice* **47**, 337-354 (2005).

154. Lalonde,M. A new perspective on the health of Canadians. A working document. 1974. Ottawa, Government of Canada.
Ref Type: Report

155. Welch,V. *et al.* P35: How effects on health equity are assessed in systematic reviews of interventions: a Cochrane methodology systematic review. Vivian Welch, Joanne Demontigny, Erin Ueffing, Maria Benghalti, Jessie McGowan, Mark Petticrew,Elizabeth Kristjansson, George Wells, Peter Tugwell. 17th Cochrane Colloquium. 2009. Cochrane Collaboration. 11-10-2009.
Ref Type: Conference Proceeding

156. Welch,V., Tugwell,P., & Morris,E.B. The equity-effectiveness loop as a tool for evaluating population health interventions. *Revista de Salud Publica* **10**, Suppl-96 (2008).

157. Williams,W. & Lewis,D. Convergent interviewing: a tool for strategic investigation. *Strategic Change* **14**, 219-229 (2005).

158. Driedger,S.M. *et al.* Finding common ground in team-based qualitative research using the convergent interviewing method. *Qualitative Health Research* **16**, 1145-1157 (2006).

159. Rao,S. & Perry,C. Convergent interviewing to build a theory in under-researched areas: principles and an example investigation of internet usage in inter-frim relationships. *Qualitative Market Research* **6**, 236-247 (2003).

160. Dick,B. Convergent interviewing: a technique for qualitative data collection. 1998.
Ref Type: Unpublished Work

161. Scott,P.J. & Briggs,J.S. A Pragmatist Argument for Mixed Methodology in Medical Informatics. *Journal of Mixed Methods Research* **3**, 223-241 (2009).

162. Johnson,R.B. & Onwuegbuzie,A.J. Mixed Methods Research: A Research Paradigm Whose Time Has Come. *Educational Researcher* **33**, 14-26 (/10).

163. Cornish,F. & Gillespie,A. A pragmatists approach to the problem of knowledge in health psychology. *Journal of Health Psychology* **14**, 800-809 (2009).

164. Charmaz,K. *Constructing Grounded Theory: A Practical Guide through Qualitative Analysis*(Sage Publications Ltd,2006).

165. Creswell,J. *Qualitative Inquiry and Research Design: Choosing Among Five Approaches*(Sage Pubn Inc,2007).

166. Bohle,P., Quinlan,M., Kennedy,D., & Williamson,A. Working hours, work-life conflict and health in precarious and "permanent" employment. *Rev Saude Publica* **38 (Suppl)**, 19-25 (2004).

167. Turnbull,J., Muckle,W., & Masters,C. Homelessness and health. *CMAJ Canadian Medical Association Journal* **177**, 1065-1066 (2007).

168. Podymow,T., Turnbull,J., Tadic,V., & Muckle,W. Shelter-based convalescence for homeless adults. *Canadian Journal of Public Health* **97**, 379-383 (2006).

169. Ostlin,P. *et al.* Priorities for research on equity and health: Implications for global and national priority setting and the role of WHO to take the health equity research agenda forward. 2009. Geneva, World Health Organization. 1-11-0009.
Ref Type: Report

170. Greenaway,C. *et al.* Tuberculosis: Evidence review for newly arriving immigrants. Canadian Medical Association Journal . 2010.
Ref Type: In Press

171. *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.0.0 [updated February 2008]. eds. Higgins, J. P. T. & Green, S. www.cochrane.org . 2008. The Cochrane Collaboration. 2-29-2008.
Ref Type: Electronic Citation

172. Feinstein,A.R. *Clinimetrics*(Yale University Press, New Haven, 1987).

173. Feinstein,A.R. *Feinstein AR. Clinical Epidemiology: The Architecture of Clinical Research. Philadelphia: Saunders, 1985: 39-52*(Saunders, Philadelphia, 1985).

174. Welch,V. *et al.* How effects on health equity are assessed in systematic reviews of interventions (Protocol). *Cochrane Database of Systematic Reviews*Art. No.: MR000028 (2009).

175. Fleiss,J.L. Measuring nominal scale agreement among many raters. *Psychological Bulletin* **76**, 378-382 (1971).

176. Evans,D.B., Lim,S.S., Adam,T., Edejer,T.T., & WHO Choosing Interventions that are Cost Effective (CHOICE) Millennium Development Goals Team Evaluation of current strategies and future priorities for improving health in developing countries. *BMJ* **331**, 1457-1461 (2005).

177. Jamison,D.T. *et al. Priorities in Health*(Oxford University Press, New York, 2006).

178. Mathers,C.D. & Loncar,D. Projections of global mortality and burden of disease from 2002 to 2030. *PLoS Med* **3**, 3442 (2006).

179. Tugwell,P. *et al.* OMERACT: an international initiative to improve outcome measurement in rheumatology. *Trials* **8**, (2007).

180. Vidanapathirana,J., Abramson,M.J., Forbes,A., & Fairley,C. Mass media interventions for promoting HIV testing. *Cochrane Database of Systematic Reviews*(2005).

181. Bisson,J. & Andrew.M. Psychological treatment of post-traumatic stress disorder (PTSD). *Cochrane Database of Systematic Reviews*(2007).

182. Wright,J.M. & Musini,V.M. First-line drugs for hypertension. *Cochrane Database of Systematic Reviews* Art. No.: CD001841 (2009).

183. Riaz,Y. *et al.* Surgical interventions for age-related cataract. *Cochrane Database of Systematic Reviews* **Issue 4. Art. No.: CD001323. DOI: 10.1002/14651858.CD001323.pub2.**, (2006).

184. Demicheli,V., Jefferson,T., Rivetti,A., & Price,D. Vaccines for measles, mumps and rubella in children. *Cochrane Database of Systematic Reviews* **Issue 4. Art. No.: CD004407. DOI: 10.1002/14651858.CD004407.pub2.**, (2005).

185. Arroll,B. *et al.* Antidepressants versus placebo for depression in primary care. *Cochrane Database of Systematic Reviews* **Issue 3. Art. No.: CD007954. DOI: 10.1002/14651858.CD007954.**, (2009).

186. Sinclair,D., Zani,B., Donegan,S., Olliaro,P., & Garner,P. Artemisinin-based combination therapy for treating uncomplicated malaria. *Cochrane Database of Systematic Reviews*(2009).

187. Shults,R.A., Nichols,J.L., Dinh-Zarr,T.B., Sleet,D.A., & Elder,R.W. Effectiveness of primary enforcement safety belt laws and enhanced enforcement of safety belt laws: a summary of the Guide to Community Preventive Services systematic reviews. *J safety Res* **35**, 189-196 (2004).

188. Ejemot, R.I., Ehiri,J.E., Meremikwu,M.M., & Critchley,J.A. Hand washing for preventing diarrhoea. *Cochrane Database of Systematic Reviews* **Issue 1. Art No.: CD004265. DOI: 10.1002/14651858.CD004265.pub2**, (2008).

189. Lavis,J., Oxman,A., Moynihan,R., & Paulsen,E. Evidence-informed health policy 1 - Synthesis of findings from a multi-method study of organizations that support the use of research evidence. *Implementation Science* **3**, 53 (2008).

190. Schmid,C., Stark,P., Berlin,J., & Lau,J. Meta-regression detected associations between heterogeneous treatment effects and study-level, but not patient-level, factors. *Journal of Clinical Epidemiology* **57**, 683-697 (2004).

191. Sampson,M., McGowan,J., Lefebvre,C., Moher,D., & Grimshaw,J. PRESS: peer review of electronic search strategies. 2008. Ottawa, Canadian Agency of Drugs and

Technology for Health (CADTH).
Ref Type: Report

192. Kendrick,D. *et al.* Preventing childhood falls at home: meta-analysis and meta-regression. [Review] [52 refs]. *Am. J. Prev. Med.* **35**, 370-379 (2008).

193. Simera,I., Moher,D., Hoey,J., Schulz,K.F., & Altman,D.G. The EQUATOR Network and reporting guidelines: Helping to achieve high standards in reporting health research studies. *Maturitas* **63**, 4-6 (2009).

194. Moher,D., Liberati,A., Tetzlaff,J., Altman,D.G., & PRISMA Group Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* **339**, (2009).

195. Liberati,A. *et al.* The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ* **339**, b2700 (2009).

196. Dans,A.L., Dans,L.F., Guyatt,G.H., & Richardson,S. Users' guides to the medical literature: XIV. How to decide on the applicability of clinical trial results to your patient. Evidence-Based Medicine Working Group.[see comment]. *JAMA* **279**, 545-549 (1998).

197. Glasziou,P. *et al.* Applying the results of trials and systematic reviews to individual patients.[see comment]. *ACP Journal Club* **129**, A15-A16 (1998).

198. Briss,P.A. *et al.* Developing an evidence-based Guide to Community Preventive Services--methods. The Task Force on Community Preventive Services. *American Journal of Preventive Medicine* **18**, Suppl-43 (2000).

199. Wang,S., Moss,J.R., & Hiller,J.E. Applicability and transferability of interventions in evidence-based public health. *Health Promotion International* **21**, 76-83 (2006).

200. Green,L.W. & Glasgow,R.E. Evaluating the relevance, generalization, and applicability of research: issues in external validation and translation methodology. [Review] [91 refs]. *Evaluation & the Health Professions* **29**, 126-153 (2006).

201. Altman,D.G. *et al.* The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med* **134**, 663-694 (2001).

202. Lavis,J.N., Posada,F.B., Haines,A., & Osei,E. Use of research to inform public policymaking.[see comment]. *Lancet* **364**, 1615-1621 (2004).

203. Dans,A.M. *et al.* Assessing equity in clinical practice guidelines. [Review] [53 refs]. *Journal of Clinical Epidemiology* **60**, 540-546 (2007).

204. Tackling health inequalities in Europe: an integrated approach: EUROTHINE: Final report. 2007. Rotterdam, Erasmus University Medical Center.
Ref Type: Report

205. Curran,M. & MacLehose,H.G. Community animal health services for improving household wealth and health status of low income farmers. *Cochrane Database of Systematic Reviews* **Issue 2**, (2006).

206. Gastwirth,J.L. The estimation of the Lorenz curve and Gini index. *Review of Economics and Statistics* **54**, 306-316 (1972).

207. Ogilvie,D. & Petticrew,M. Reducing social inequalities in smoking: can evidence inform policy? A pilot study. *Tob Control* **13**, 129-131 (2004).

208. Johnson,S.M., Karvonen,C.A., Phelps,C.L., Nader,S., & Sanborn,B.M. Assessment of analysis by gender in the Cochrane reviews as related to treatment of cardiovascular disease. *J Womens Health (Larchmt)* **12**, 449-457 (2003).

209. Viswanathan,M. & Viswanathan,M. Tailoring systematic reviews to meet critical priorities in maternal health in the intrapartum period. [Review] [110 refs]. *Paediatric and Perinatal Epidemiology* **22 Suppl 1**, 10-17 (2008).

210. Browne,G., Gafni,A., Roberts,J., Byrne,C., & Majumdar,B. Effective/Efficient Mental Health Programs for School-Age Children: A Synthesis of Reviews. *Social Science and Medicine* **58**, 1367-1384, Apr (2004).

211. McCord,J.F. & Michelinakis,G. Systematic review of the evidence supporting intra-oral maxillofacial prosthodontic care. [Review] [61 refs]. *European Journal of Prosthodontics & Restorative Dentistry* **12**, 129-135 (2004).

212. Green,J.M., Hewison,J., Bekker,H.L., Bryant,L.D., & Cuckle,H.S. Psychosocial aspects of genetic screening of pregnant women and newborns: a systematic review. [Review] [171 refs]. *Health Technology Assessment (Winchester, England)* **8**, iii-iix (2001).

213. Pachler,J. & Wille-Jorgensen,P. Quality of life after rectal resection for cancer, with or without permanent colostomy. [Review] [58 refs][Update of Cochrane Database Syst Rev. 2004;(3):CD004323; PMID: 15266529]. *Cochrane Database of Systematic Reviews*CD004323 (2005).

214. Gourlay,S.G., Stead,L.F., & Benowitz,N.L. Clonidine for smoking cessation. [Review] [35 refs][Update of Cochrane Database Syst Rev. 2000;(2):CD000058; PMID: 10796479]. *Cochrane Database of Systematic Reviews*CD000058 (2004).

215. Mowatt,G. *et al.* Systematic review of the effectiveness and cost-effectiveness, and economic evaluation, of myocardial perfusion scintigraphy for the diagnosis and

management of angina and myocardial infarction. [Review] [138 refs]. *Health Technology Assessment (Winchester, England)* **8**, iii-iiv (2001).

216. Swanson,H.L. & Sachse-Lee,C. A meta-analysis of single-subject-design intervention research for students with LD. *Journal of learning disabilities* **33**, 114-136 (2000).

217. Wendel-Vos,G.C. *et al.* Physical activity and stroke. A meta-analysis of observational data. [Review] [55 refs]. *Int. J. Epidemiol.* **33**, 787-798 (2004).

218. Ramakrishnan,U., Aburto,N., McCabe,G., & Martorell,R. Multimicronutrient interventions but not vitamin a or iron interventions alone improve child growth: results of 3 meta-analyses. [Review] [80 refs]. *Journal of Nutrition* **134**, 2592-2602 (2004).

219. Matthews,S.J. & McCoy,C. Peginterferon alfa-2a: a review of approved and investigational uses. [Review] [149 refs]. *Clinical Therapeutics* **26**, 991-1025 (2004).

220. Parker,M.J., Gillespie,L.D., & Gillespie,W.J. Hip protectors for preventing hip fractures in the elderly. [Review] [37 refs][Update in Cochrane Database Syst Rev. 2005;(3):CD001255; PMID: 16034859].[Update of Cochrane Database Syst Rev. 2003;(3):CD001255; PMID: 12917903]. *Cochrane Database of Systematic Reviews*CD001255 (2004).

221. Heyn,P., Abreu,B.C., & Ottenbacher,K.J. The effects of exercise training on elderly persons with cognitive impairment and dementia: a meta-analysis. *Archives of Physical Medicine & Rehabilitation* **85**, 1694-1704 (2004).

222. Kenworthy,T., Adams,C.E., Bilby,C., Brooks-Gordon,B., & Fenton,M. Psychological interventions for those who have sexually offended or are at risk of offending. [Review] [127 refs][Update in Cochrane Database Syst Rev. 2008;(4):CD004858; PMID: 18843670]. *Cochrane Database of Systematic Reviews*CD004858 (2004).

223. Ejere,H., Alhassan,M.B., & Rabiu,M. Face washing promotion for preventing active trachoma. [Review] [18 refs]. *Cochrane Database of Systematic Reviews*CD003659 (2004).

224. Mattick,R.P., Kimber,J., Breen,C., & Davoli,M. Buprenorphine maintenance versus placebo or methadone maintenance for opioid dependence. [Review] [29 refs][Update in Cochrane Database Syst Rev. 2008;(2):CD002207; PMID: 18425880].[Update of Cochrane Database Syst Rev. 2003;(2):CD002207; PMID: 12804429]. *Cochrane Database of Systematic Reviews*CD002207 (2004).

225. Dinh-Zarr,T., Goss,C., Heitman,E., Roberts,I., & DiGuiseppi,C. Interventions for preventing injuries in problem drinkers. [Review] [46 refs][Update of Cochrane

Database Syst Rev. 2000;(2):CD001857; PMID: 10796829]. *Cochrane Database of Systematic Reviews*CD001857 (2004).

226. Grossman,P., Niemann,L., Schmidt,S., & Walach,H. Mindfulness-based stress reduction and health benefits. A meta-analysis. *Journal of Psychosomatic Research* **57**, 35-43 (2004).

227. Carter-Pokras,O. *et al.* Providing linguistically appropriate services to persons with limited English proficiency: a needs and resources investigation. [Review] [42 refs]. *American Journal of Managed Care* **10**, Spec-36 (2004).

228. Gundry,S., Wright,J., & Conroy,R. A systematic review of the health outcomes related to household water quality in developing countries. *Journal of Water and Health* **2**, 1-13 (2004).

229. Lees,J., Manning,N., & Rawlings,B. A culture of enquiry: research evidence and the therapeutic community. *Psychiatric Quarterly* **75**, 279-294 (2004).

230. Henry,C. & Ghaemi,S.N. Insight in psychosis: a systematic review of treatment interventions. [Review] [22 refs]. *Psychopathology* **37**, 194-199 (2004).

231. Hesse,M. Achieving abstinence by treating depression in the presence of substance-use disorders. *Addictive Behaviors* **29**, 1137-1141 (2004).

232. Coon,J.T. & Ernst,E. Complementary and alternative therapies in the treatment of chronic hepatitis C: a systematic review. [Review] [55 refs]. *Journal of Hepatology* **40**, 491-500 (2004).

233. Ahmed,N., Ahmedzai,S., Vora,V., Hillam,S., & Paz,S. Supportive care for patients with gastrointestinal cancer. [Review] [30 refs]. *Cochrane Database of Systematic Reviews*CD003445 (2004).

234. Tan,A., Schulze,A., O'Donnell,C.P., & Davis,P.G. Air versus oxygen for resuscitation of infants at birth. [Review] [5 refs][Update in Cochrane Database Syst Rev. 2005;(2):CD002273; PMID: 15846632]. *Cochrane Database of Systematic Reviews*CD002273 (2004).

235. Hoenig,H. & Siebens,H. Research agenda for geriatric rehabilitation. [Review] [66 refs]. *American Journal of Physical Medicine & Rehabilitation* **83**, 858-866 (2004).

236. de,S.A., Jones,P.W., & Spencer,S.A. Does human milk reduce infection rates in preterm infants? A systematic review. [Review] [37 refs]. *Archives of Disease in Childhood Fetal & Neonatal Edition* **89**, F509-F513 (2004).

237. He,F.J. & MacGregor,G.A. Effect of longer-term modest salt reduction on blood pressure. [Review] [173 refs]. *Cochrane Database of Systematic Reviews*CD004937 (2004).

238. Turner,C., Spinks,A., McClure,R., & Nixon,J. Community-based interventions for the prevention of burns and scalds in children. [Review] [62 refs]. *Cochrane Database of Systematic Reviews*CD004335 (2004).

239. Moore,H. *et al.* Dietary advice for treatment of type 2 diabetes mellitus in adults. [Review] [314 refs][Update in Cochrane Database Syst Rev. 2007;(3):CD004097; PMID: 17636747].[Update of Cochrane Database Syst Rev. 2004;(2):CD004097; PMID: 15106237]. *Cochrane Database of Systematic Reviews*CD004097 (2004).

240. Nannini,L., Lasserson,T.J., & Poole,P. Combined corticosteroid and longacting beta-agonist in one inhaler for chronic obstructive pulmonary disease. [Review] [25 refs][Update in Cochrane Database Syst Rev. 2004;(3):CD003794; PMID: 15266502]. *Cochrane Database of Systematic Reviews*CD003794 (2003).

241. Stade,B., Shah,V., & Ohlsson,A. Vaginal chlorhexidine during labour to prevent early-onset neonatal group B streptococcal infection. [Review] [42 refs]. *Cochrane Database of Systematic Reviews*CD003520 (2004).

242. Rees,K., Taylor,R.S., Singh,S., Coats,A.J., & Ebrahim,S. Exercise based rehabilitation for heart failure. [Review] [78 refs]. *Cochrane Database of Systematic Reviews*CD003331 (2004).

243. Kulier,R., Boulvain,M., Walker,D., Candolle,G., & Campana,A. Minilaparotomy and endoscopic techniques for tubal sterilisation. [Review] [17 refs][Update of Cochrane Database Syst Rev. 2002;(3):CD001328; PMID: 12137624]. *Cochrane Database of Systematic Reviews*CD001328 (2004).

244. Zupan,J., Garner,P., & Omari,A.A. Topical umbilical cord care at birth. [Review] [67 refs][Update of Cochrane Database Syst Rev. 2000;(2):CD001057; PMID: 10796230]. *Cochrane Database of Systematic Reviews*CD001057 (2004).

245. Scott,I., Chan,J., Aroney,C., & Carroll,G. Local thrombolysis or rapid transfer for primary angioplasty for patients presenting with ST segment elevation myocardial infarction to hospitals without angioplasty facilities. *Internal Medicine Journal* **34**, 373-377 (2004).

246. Streiner,D.L. & Norman,G.R. *Health Measurement Scales: A Practical Guide to their Development and Use*(Oxford University PRess, London, 2003).

247. Buendia-Rodriguez,J.A., Sanchez-Villamil,J.P., Buendia-Rodriguez,J.A., & Sanchez-Villamil,J.P. Using systematic reviews for evidence-based health promotion: basic methodology issues. *Revista de Salud Publica* **8 Suppl 2**, 94-105 (2006).