BETTER GUIDANCE FOR BETTER HEALTH SYSTEMS

DESIGNING A KNOWLEDGE TRANSLATION TOOL FOR THE DEVELOPMENT, APPRAISAL AND REPORTING OF HEALTH SYSTEMS GUIDANCE

By

DENIS E. AKO-ARREY, MHA, MPH.

A Thesis Submitted to the School of Graduate Studies in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy

McMaster University DOCTOR OF PHILOSOPHY (2015) Hamilton, Ontario

(Health Policy)

TITLE: Designing a knowledge translation tool for the development, appraisal and reporting of health systems guidance.

AUTHOR: Denis E. Ako-Arrey, MHA, MPH.

SUPERVISOR: Dr. Melissa C. Brouwers.

NUMBER OF PAGES: vii, 125

Abstract

Health systems guidance (HSG) can be defined as a set of options presented to policy-makers on how to address a particular health systems issue/challenge (Bosch-Capblanch et al, 2012). However, best strategies for developing HSG and translating it into policy are poorly understood at this time (Lavis et al, 2012). Additionally, there is currently no instrument that has the capacity to discriminate between higher quality HSG from those of lower quality. This thesis begins to address these gaps through three original scientific contributions that use a range of methodological approaches to design a tool that will be used to direct the development, appraisal and reporting of HSG. Taken together, the chapters present three stages conducted in a sequence:

Stage 1: A critical interpretive synthesis of the literature to generate a draft list of candidate concepts (items, criteria or domains) for the HSG tool, with their descriptions and a specific set of operational considerations to optimize their use.

Stage 2: Results from a survey conducted across the six World Health Organization (WHO) health regions to evaluate the importance of the candidate concepts, assess the appropriateness of their descriptions, and identify any missing components, in order to generate a beta version of the HSG tool.

Stage 3: Results from a survey conducted across the six WHO health regions to test the usability of the beta version of the HSG tool to determine its feasibility of application and ease of understanding, in order to generate version 1 of the HSG tool.

As a whole, the chapters presented in this thesis provide substantive, methodological and disciplinary contributions to the field of health systems research and in particular, about how to support the production, evaluation and reporting of high quality HSG for the purposes of strengthening health systems in low, middle and high income countries. The core deliverable of this program of research is version one of the HSG tool, the AGREE-HS (Appraisal of Guidelines Research and Evaluation – Health Systems).

Acknowledgements

First and foremost, I am extremely grateful to my Supervisor, Dr. Melissa Brouwers. Melissa, I want to thank you for believing in me and giving me the opportunity to complete this Health Policy PhD program. Your magnificent supervision, your formidable support and relentless mentorship, have had profound and long lasting impacts on my life academically, professionally and personally. For the past 4 years, I have been able to hone my skills as a researcher and expose myself to many invaluable opportunities thanks to your resolute commitment to my success. I earnestly pray that God will continue to bless you with bountiful wisdom.

I am highly indebted to my thesis committee members, Dr. John Lavis and Dr. Mita Giacomini for continuously permitting me to tap into their extensive expertise that provided the fuel to keep me on track, and steer this thesis into fruition. I want to also thank all the members of the AGREE-HS team for their un-abating guidance.

I will also like to thank Lydia Garland, Sharon Garden, and Laurie Cocking for their never-ending and timely administrative interventions that facilitated the achievements of my goals. I also express gratitude to McMaster University, KT-Canada and the Canadian Institute of health research (CIHR) for their financial assistance.

To my Family, Pa Joe, Mama Theresia, Pamela, Rita, Ayuk & Manyi, I know you are very proud of me and I want you to know that this is only the beginning of greater things to come. Your unfading love, your emotional, physical, and financial presence kept me focused on my objectives and grounded in life. Thank you all for creating the foundation upon which this challenging yet very fulfilling experience was built. This PhD thesis belongs to you, and to all the sons and daughter of the great lands of Afap & Ogomoko.

To Leni & DeMi, my pride and joy, the wind beneath my wings, the *raison d'être* of everything I do. A few years from now when you read this, may my hard work be an inspiration for your success in attaining your maximum potential, and may the sky always be your limit.

To Eva Manu, for always being by my side all these years, your unwavering support is the backbone upon which this was made possible. To all my friends, especially Daisy Tambe, Stephen Asong, Edward Asong, Chief Asonganyi, Keah Wubila, Vincent Nche, Godfrey Osakue, Dr. Magnus Ajong, Dr. Derick Akoku and Dr. Emmanuel Ngwakongwi, I would not have done this without you constantly cheering me on.

To all the students, staff and researchers at the center for health economics and policy analysis (CHEPA), and at Cancer Care Ontario's program for evidence-based care (PEBC), whose contributions to my studies made my experience a memorable one.

Finally, I want to specially dedicate this doctoral thesis to my late brother, Mathias Arrey Tambe alias "Bakayoko". God has you in his keeping; I have you forever in my heart. Rest assured that I can feel your invisible angelic watch over me from heaven. I Love You Lil Bro……

Table of content

List of figures, tables and appendices

Chapter 1 - Introduction

Chapter 2 - Health Systems Guidance Appraisal Concepts - A Critical Interpretive Synthesis.

Chapter 3 - Health Systems Guidance Appraisal Concepts - A Survey of Stakeholders.

Chapter 4 - Health Systems Guidance Appraisal Concepts - Usability Testing

Declaration of academic achievement

I, Denis Ako-Arrey, declare that I am the lead author of each of the chapters presented in this thesis. As the lead author, I was responsible for conceptualizing the theoretical and empirical formulations of each research project, literature review, study protocol and survey design, collecting and analyzing data as well as manuscript preparation. This doctoral dissertation and the manuscripts there-in are co-authored with my supervisor and members of my thesis committee who offered input and expertise during each phase of the research formulation and manuscript preparation and provided feedback on earlier drafts.

Chapter 1. Introduction

This doctoral dissertation is composed of an introductory chapter (Chapter 1), three chapters describing an original program of scientific research that uses different methodological approaches (Chapters 2, 3 & 4), and a concluding chapter (Chapter 5). In this chapter, I begin by presenting the rationale and the importance of this dissertation topic - designing a knowledge translation tool for the development, appraisal and reporting of health systems guidance (HSG) - by providing an understanding of health systems, addressing the need for stronger health systems, and focusing on the role that HSG can play in the process (and providing exemplars of such roles). I discuss some theoretical and empirical advancement that have been made in developing and implementing HSG while identifying the current gaps in the literature and then discuss the interface between HSG and clinical practice guidelines. Then I present the overall objectives of this dissertation, and provide a summary of the aims, methodologies and contributions detailed in Chapters 2, 3 and 4.

1.1. Understanding health systems

In the World Health Report produced by the World Health Organization (WHO) in the year 2000, a health system is defined as all organizations, people, actions and activities whose primary purpose is to promote, restore, and maintain health (WHO, 2000). This definition aligns with general systems theorist Ludwig von Bertalanffy's understanding of a "system" (1968): a collection or arrangement of parts interconnected with one another for a common purpose or goal. Therefore, just like in any other system, for a health system to be effective, it's numerous set of inter-connected components need to function together efficiently (Brinkerhoff, 2003). In 2012, the WHO's Alliance for Health Policy and Systems Research redefined a health system to refer to governance arrangements (e.g., policy, organizational or professional authority), financial arrangements (e.g., financing, funding, remuneration, or incentives) and delivery arrangements (e.g., to whom care is provided, by whom care is provided, or where care is provided) for health care and population health services and the broader context in which they are negotiated, implemented and reformed (Lavis et al, 2002; Hoffman et al, 2013; Lavis et al, 2015). According to the World Health Report (2000), overall health system goals include improving health as well as reducing health inequity in a way that is responsive; ensuring the most cost effective and efficient use of scarce resources; increasing access to and coverage for effective health interventions without compromising efforts to ensure provider quality and safety; protecting against health-related financial risk; and satisfying consumers' expectations (WHO, 2000). Failing to achieve these goals generates poor health outcomes and exacerbates health inequities. The *raison d'être* of a health system can be regarded ultimately to be better health (Arah, 2003).

Health systems in the world differ from one context to another or one region to another, with respect to design, capacity, priorities, resources etc. However, common to all is the need to deal with a wide variety of (often evolving) challenges ranging from simple to complex, such as, patterns of disease, care, and available treatment options (WHO 2007). There is the expectation that the health system will respond efficiently to these changes. Whether or how this response occurs depends on a number of factors. For example, increasing demands for health services, an aging population and rising health care costs associated with rapid technological advances are some factors putting a strain

on health systems of high-income, industrialized countries around the world (Snowdon & Cohen, 2011). In contrast, in low and middle-income countries, infectious diseases, fragile and fragmented health systems, shortfalls in the health workforce, limited health financing, lack of donor co-ordination, weak drug regulatory and supply systems, and weak information systems are among some reasons the achievement of health goals are impeded (Hanson et al, 2003; Chen, 2004; Travis et al, 2004). The capacity of the health system to provide essential health services in low and middle-income countries has been seriously compromised by decades of economic crises, structural adjustment programs and declining public expenditure (Schneider et al, 2006). Frenk & Moon (2013) noted that, as the world faces what can be referred to as global health systems challenges, there is a need to strengthen health systems.

1.2. The challenge: strengthening health systems

Health system strengthening refers to a broad range of initiatives and strategies aimed to enhance one or more of the functions of a health system. The goal of health systems strengthening is to produce better health through improvements in access to, coverage in, quality of, or efficiency within the system (Islam, 2007). In attempting to capture areas of the health system that require strengthening, it is essential to be clear about where challenges or problems arise, which parts of the health system are affected, where priorities need to be shifted, what will happen as a result, and by what means change can be monitored. For example, in a recent content analysis of HSG, Bosch-Capblanch and colleagues (2012) identified 283 recommendation statements related to strengthening health systems related to each of the 'building block' domains set out by the WHO (WHO, 2007). They found 83 (29%) of the recommendations focused on health system strengthening via improvements to governance and leadership (e.g., roles of governing bodies in producing or implementing recommendations); another 67 (24%) on access to effective medical products and technologies (e.g., vaccines); 53 (19%) on improved health workforce (e.g., training); 31 (11%) on enhanced information systems (e.g., data for measuring performance); 29 (10%) on redesigned service delivery (e.g., basic health care package) and 20 (7%) on innovation in financing (e.g., drug funding options).

There has been a growing recognition of the urgent need to strengthen health systems. For example, the Global Fund, GAVI Alliance and the World Bank, created the joint health systems strengthening financing platform. As a consequence, increased investment in health systems strengthening is driving demand for high quality recommendations based upon good quality evidence to inform this process. Demand for good quality evidence has, in turn, exposed the need for better investments in health systems research to create this evidence (Haines, 2011). Thus, research evidence and processes and tools, such as recommendations reflected in HSG that translate evidence into action by informing policies are required (Lavis et al, 2012; Haines, 2011).

1.3 Health systems guidance: a solution to strengthen and optimize health systems?

Guidance is a broad term that can be used to refer to a suite of advice tools such as guidelines, standards, benchmarks, targets, advisory reports, recommendations, options etc. Many authors have suggested definitions of HSG. Lomas et al (2005) defined health systems guidance (HSG) as sets of options presented to decision makers by neutral parties on how to respond to a particular health systems issue. HSG comprises (i) a summary, synthesis and interpretation of the research available on a health system

problem; and (ii) recommendations for concrete action to solve the problem that is informed by this evidence and contextualized factors to optimize implementability (Law 2008). Bosch-Capblanch et al (2012) further defined HSG as systematically developed statements produced at local, national or global levels in order to assist decisions made by clinical leaders, health care executives, and government policy-makers about appropriate options for addressing a health systems challenge, to assist with the implementation of these options, and to direct the monitoring and evaluation of the implementation efforts.

The main bodies tasked with developing HSG are international organizations (e.g., WHO, the Pan American Health Organization, GAVI etc.), national ministries of health (e.g., Public Health Agency of Canada supporting the H1N1 Pandemic Plan [PHAC, 2011], and special committees appointed by government ministries of health (e.g., Commission on the Future of Health Care in Canada [Health Canada, 2009]).

1.4. Health system guidance exemplars

Below are examples of HSG. The examples reflect the range of health system challenges addressed and the variability in who might create HSG.

1.4.1. WHO guidance on vaccine introduction

WHO developed HSG to provide recommendations for decision makers and program managers to assist them to make informed decisions related to planning the introduction of a vaccine into a national immunization program (WHO, 2014). The HSG suggests potential elements or benchmarks that can be used to assess the strength of the national immunization program to accommodate a new vaccine. Some of these elements are: a strong decision-making and accountability process that is transparent, coordinated and integrated with the overall health sector; well-performing or improving immunization programs to obtain the full benefit from existing vaccines; sufficient, well-trained and motivated health workforce; functional vaccine management, cold chain and logistics systems; safe immunization practices and monitoring and management of adverse events; high-quality disease surveillance and immunization coverage monitoring; and a financially sustainable program. It also provides recommendations on ways to mitigate the potential negative impacts this may have on the health system, as well as ways to maximize the opportunities this can provide to strengthen these systems.

1.4.2. WHO guidance on increasing access to health workers in rural and remote areas

WHO developed HSG to address the health systems challenge related to the acute shortages of trained, skilled and motivated health workers in rural and remote jurisdictions, especially in low-income countries (WHO, 2010). The HSG provides recommendations on a comprehensive set of strategies to help countries encourage health workers to live and work in remote and rural areas through improved attraction, recruitment and retention. There are 4 sets of recommendations; educational, regulatory, financial incentives and personal/professional support. Some educational recommendations include options like targeted admission policies to enroll students with a rural background and locating health professional schools, campuses and family medicine residency programs outside of capitals and other major cities. Some regulatory recommendations include options like introducing and regulating enhanced scopes of practice in rural and remote areas and providing scholarships, bursaries or other

education subsidies with enforceable agreements of return of service in rural or remote areas. Some financial incentives-related recommendations include options like using a combination of fiscally sustainable financial incentives like hardship allowance, grants for housing, free transportation, and paid vacations. Some personal/professional support recommendations include options like improving the living conditions of the health workers (sanitation, electricity, telecommunication, schools, etc.) and supporting career development programs, professional networks and providing senior posts in rural areas.

1.4.3. WHO guidance on task shifting

WHO issued guidance about task shifting to address the health system challenge related to human resources shortages in health services in general, and particularly related to the critical workforce shortages in the area of maternal and newborn health in low-income countries (WHO, 2012). The guidance promotes a more rational distribution of tasks and responsibilities among cadres of health workers (clinicians, nurses, midwives, lay health workers etc.) as a strategy for improving access and cost-effectiveness within health systems. For this HSG to be relevant and applicable, the research question had to address essential components of the intervention (e.g., training of lay health workers [who, how, where]), related actions (e.g., adaptations needed in the distributions of tasks to different providers), implementation issues (e.g., the preferences of potential clients), and the implications across other health system building blocks (e.g., adaptations to the health information sub-system that may be needed to capture the tasks undertaken by such workers).

1.4.4. Cancer Care Ontario (CCO) guidance on models of systemic therapy

While HSG have led efforts to support low-and middle-income countries (LMIC), system-level guidance has increasingly been seen in higher-income countries too. CCO is an advisory body to the province of Ontario, Canada on matters related to quality in cancer. It has a robust clinical practice guidelines initiative (i.e., the Program in Evidence-Based Care [PEBC]) that recently extended its reach to include a more systems-level perspective. As a result, through its PEBC, CCO has focused on system improvements and creating system efficiencies and has developed systems-level guidance for supporting the organization and delivery of cancer services in the province of Ontario, Canada. For example, CCO developed a Models of Systemic Therapy guidance document, which recommended a four-level provincial model to ensure safe and effective treatment is provided (Vandenberg et al, 2009). To that end, the document delineates clinician team phenotype (e.g., specialist versus primary care), institutional phenotype (community hospital versus academic hospital), equipment needs and safety requirements as a function of the complexity of the clinical scenario (e.g., complexity of systemic therapy treatment, complexity of patients) and the system context in which care will be delivered (e.g., rural versus urban).

1.5. Developing and implementing HSG: The challenges

There are challenges associated with developing and implementing HSG such as quality and types of evidence available, generalizability and adaptability of the evidence from one context to the next and the implications of this, and timeliness of being able to use and incorporate evidence into the discourse.

HSG has to use systematic and transparent approaches and needs to be informed

by the best available evidence (Pantoja, 2012). Guidance on health systems should include an articulation of the evidence underpinning the problems (needs of the population and health system bottlenecks), the interventions recommended to solve the problem (the formulation of policies) and implementation issues associated with the advised strategies; and each of these require different types of evidence and different approaches of synthesizing the evidence (Lavis, 2004). Lomas et al (2005) reported two distinct views on the role of science in HSG. On the one hand, science reveals universal truths and informs context-free guidance indicating what works in general under ideal conditions. On the other hand, there is context-sensitive guidance where evidence used for decision-making is of little importance if it does not adapt to the circumstances of its application. In developing, appraising and reporting HSG, it is crucial therefore to consider both - that is, guidance that respects both scientific integrity on the one hand and its implementability in a specific health system context on the other.

Optimizing health systems is a challenging task to which appropriate guidance can positively contribute. However, there are challenges to generating guidance for health systems. The wide range of study designs required to evaluate health system interventions, including those in which stakeholders have little experience; managing the complexity of the interconnections between the various health system elements and the health system context; the lack of capacity and knowledge about HSG and HSG development; the complicatedness related to involving decision-makers with diverse disciplinary and professional backgrounds; and the strong interests-based and ideological drivers of some health policies, represent some of the evidence, methodological and interpretation challenges HSG developers have to face (Haines, 2011).

HSG is highly context-specific, and may vary from one region to another. As a consequence, the available evidence that can be used to inform recommendations may have limited transferability to contexts outside those represented in the primary studies that compose the evidence base. For example, while evidence tends to be most often generated in high-income settings, it is often presented as a synthesis of "best available global evidence"; this generates concerns of applicability to local contexts in low-income contexts. Moreover, while coordination for guidance development is needed at the global level, decisions on the options recommended are usually taken at national level or, in federal jurisdictions, at the sub-national level (Lavis et al, 2012). Therefore, global guidance needs to be adaptable, ideally through national deliberative processes in which stakeholders develop context-specific guidance products that, in turn, result in policies for health system arrangements at local levels (Bosch-Capblanch et al, 2012). As a result, there is a need for HSG that is sensitive to the heterogeneity of the various health systems across the globe, and can be applied in a wide variety of contexts while also being applicable in specific circumstances.

An important consequence of systematically conducting context-specific assessments is that justifiably divergent HSG recommendations can emerge. For example, the Evidence-Informed Policy Network (EVIPNet) Africa teams assessed the pros and cons of different approaches to scaling up the agreed upon use of artemisinin-based combination therapy (ACT) for the treatment of malaria (Lavis & Panisset, 2010). While there was agreement across countries about the clinical option (ACT as the preferred first-line therapy), countries considered a series of contextual factors to determine the best way to achieve the implementation of this option in their setting. These factors included delivery arrangement options (e.g., who was permitted to prescribe and dispense ACT, who was to monitor cases), financial arrangement options

(e.g., subsidies for patients and/or providers), and governance options (e.g., licensing of drugs, changes to scope of practice). Some countries favored changes to governance arrangements (e.g., expanding the list of health care providers who can prescribe the drug and establishing regulatory protections against counterfeit medicines), while others favored changes to financial arrangements [e.g., drug subsidies] (Lavis & Panisset, 2010). Both option sets are legitimate recommendations given that they are aligned to the health systems context in which the action would be taken.

Another challenge that arises in the development of guidance for health systems is the timeliness in the production of the guidance in relation to the policy decision window (Bosch-Capblanch et al, 2012). While on one hand, the processes of research synthesis, of guidance development and of policy making usually take a long time to complete, on the other hand, policy decision making can occur in a relatively shorter time-spans, and in some cases, health system decisions are made in the absence of appropriate guidance. This tension requires that good quality HSG methodology needs to be timely, nimble, and usable by the broad range of health system stakeholders in order to create recommendations that are attractive to policy-makers who may be lured into other forms of evidence by mere convenience or by stakeholders' vested interests. Development of HSG has to resonate with needs, with available capacity to implement it, and with countries' priorities.

1.6. Clinical practice guidelines: The interface with HSG

Considerable effort in the knowledge translation (KT) research field has been dedicated to studying how decisions are made by clinicians and patients (and the public) and in designing and critically appraising clinical strategies to optimize decision-making so that the best health care options - based on evidence - are used in practice. The results of these efforts include clinical practice guidelines (CPGs), patient decision aids, and other knowledge tools that assist in the decision-making process (United Nations, 2011). CPGs have been shown to be effective tools to support decision-making in clinical settings. However, patients and providers do not operate in isolation – they operate within a health system; a system that has significant control over what options are available, to whom, when and how.

In contrast to HSG, there currently exist well-established methods to develop and appraise guidelines relevant to clinical practice. One such tool that has gained prominence with CPGs is the AGREE II tool (Brouwers et al, 2010a; Brouwers et al, 2010b; Brouwers et al, 2010c). This is a reliable, valid and user-accepted tool designed with the intention to evaluate the process of CPG development, and assess the quality of the guidelines in order to assist in patient, practitioner and policy decisions. The tool is also useful in directing the development and reporting of CPGs. It was designed for clinical and public health guidelines, including health promotion, public health, screening, diagnosis, and interventions. The AGREE II instrument (www.agreetrust.org) is listed among the online resources for clinical guidelines in the Cochrane Collaboration and has shown the capacity to discriminate between higher quality guidelines that follow technical documentation from those of lower quality. The uptake of AGREE II has been extensive and it has been used, in addition to CPG appraisal, to inform the CPG development and reporting standards, protocols and/or accreditation processes of organizations such as the Canadian Task Force on Preventive Health Care (Birtwhistle et al, 2012), National Institute for Health and Care Excellence (NICE, 2011), US Institutes

Of Medicine (Graham et al, 2011) and the Guidelines International Network [GIN – International] (Oaseem et al, 2012)…

Fundamentally, there exist key differences between CPGs and HSG (see Table 1 below); firstly, HSG provides recommendations for how the health system can best support a specific clinical action, public health action or health-related goal and not the clinical action itself. Secondly, there are concepts relevant to high quality and implementable HSG that a CPG does not address or does not address sufficiently (e.g., political soundness) and there are CPG concepts that are not relevant to HSG (e.g., clinical options). Thirdly, even where there are conceptually similar areas, how they are operationalized in existing tools, such as AGREE II, requires significant new efforts to make them applicable to the HSG context (e.g., use of different study designs; criteria and considerations to direct evaluation; and even fundamental issues such as agreement on what constitutes high quality process, methods, and reporting). Thus, CPGs, and their scientific and methodological evolution, may provide a conceptual map that could be used and tailored to improve the HSG enterprise with respect to issues of quality assessment and principles for development and reporting.

Table 1: Comparison of CPGs and HSG

Features	**CPGs**	**HSG**
Development processes	Methods relatively well defined. Processes protocolized, institutionalized and routinized.	Refinement of methods on-going. Methods not routinized or institutionalized.
Rationale for guidance	Clinical problems (i.e., diseases) needing guidance well defined and measureable.	Lack of an 'analytical' health systems framework to define and categorize problems.
Scope of guidance	Many examples by which to determine scope. Link between research evidence ("answering questions") and guidelines ("solving problems") relatively straightforward.	Scarcity of experience and examples of HSG on which to draw in determining scope. Links between research evidence and guidance more complex. There is a need to explicitly consider the potential for unintended system-wide consequences
Nature of evidence	Scope of research evidence quite well delimited. Widely accepted hierarchy of evidence.	More types of research evidence considered. More need for evidence on non-effectiveness issues (e.g., costs, process evaluations, stakeholders' views, equity), some of which have no established evidence hierarchy
Evidence assessment and	Existing tools are specifically designed for	More tools are needed for the different types of

presentation	research evidence on effectiveness and for clinical questions. Users' (i.e., clinicians') expertise to understand and process research evidence tends to be well established (i.e., clinicians).	evidence; few are well developed and/or validated. Users' (i.e., policy-makers' and stakeholders') expertise to understand and process evidence is extremely varied.
Link evidence into guidance and recommendations	Almost all evidence comes from research. Deliberative processes focused on research evidence, with clear standards by which to critical appraise and synthesize it.	The role of other types of evidence, including non-research information, is much more relevant. Deliberative processes focused on the full range of inputs that will influence decision-making
Monitor and evaluate recommendations	Well-defined approaches to evaluate the effects of clinical interventions in practice.	Complex measurement and evaluation approaches for health systems, services and programmes, with large attribution challenges.
Acceptance of guidance materials	Clinicians have grown receptive to guidelines over time (more normative).	Resistance from policy-makers and stakeholders to uptake of often complex and questionable guidance.

1.7. Statement of the problem and thesis objectives

Stakeholders are recognizing the value of HSG as a knowledge tool that can be used to assist in developing evidence-informed and implementable policy that is relevant to the health system and with the goal of optimizing health outcomes of a population (Bosch-Capblanch et al, 2012; Lavis et al, 2012; Lewin et al, 2012; Birtwhistle et al, 2012; NICE, 2011). Indeed, an international consensus by the WHO's health systems guidance task force in 2011 unanimously supported the creation of a tool to appraise the quality of HSG and further support its development and reporting (Bosch-Capblanch, 2011). Such an appraisal tool is crucial in order to determine acceptable quality thresholds and develop standards of practice. However, there is currently no universally acceptable gold standard approach for developing, appraising and reporting HSG. While there is general agreement about what HSG should cover (WHO, 2000), there is no consensus on how key elements should be optimized, the thresholds that should be used to differentiate HSG as a function of quality, and the optimal manner to develop and report these elements. Given the significant resource implications inherent in the application of recommendations emerging from HSG, this is an important gap to fill to ensure the most credible and highest quality recommendations are developed and implemented. Appraising the quality of HSG is important in establishing to what extent the guidance used state-of-the-art and validated methods during its development and whether or not it is balanced and reliable as it relates to the evidence that informs it (Bosch-Capblanch et al, 2012).

The objectives of the program of research represented in this dissertation begin to

address these gaps. **Specifically, the objectives of this program is to develop an instrument designed to assess the quality of HSGs, and that can also provide a methodological strategy for their development and inform what information should be reported in the HSG and how it should be reported.** The intended deliverable of this program of research is version one of the **A**ppraisal of **G**uidelines for **R**esearch and **E**valuation (AGREE) Health Systems (HS). The dissertation also aims to address the appropriateness of HSG, which is an important component in the uptake of guidance recommendations.

To accomplish these objectives, three original studies, each with a unique scientific contribution and methodological approach, were undertaken. A key principle for the design of an acceptable HSG appraisal tool (that can inform reporting and development) is to adhere to standard methodological quality criteria (e.g., usable, reliable, and valid) that confer on guidance the credibility to be adopted with confidence. Therefore, the strategy undertaken in this thesis was to tailor the methodological, conceptual and theoretical principles of measurement construction used to design the AGREE II tool (Cluzeau et al, 1999; AGREE Collaboration Writing Group, 2003; Brouwers et al, 2010a; Brouwers et al, 2010b) to develop a complementary, reliable, valid and useful tool for health systems, designed for national and international developers and users of HSG, that will be used to direct the development, appraisal and reporting of HSG.

1.8. Dissertation Structure

Three individual studies were undertaken to achieve the program of study objectives. The Study 1 (Chapter 2, entitled 'concept development and item generation') objective is to conduct a knowledge synthesis of the published and grey literatures to identify and synthesize the key insights from any papers that report on existing concepts (items, criteria or domains) currently used to describe, differentiate, or test the quality of HSG in order to generate those elements related to HSG development, appraisal or reporting. I apply a critical interpretive synthesis (CIS) approach to systematic review (Dixon-Wood et al, 2006) that enables an iterative, flexible and dynamic analysis of diverse bodies of literature (qualitative, quantitative and theoretical papers) to generate a candidate list of concepts. These are concepts considered to be a good fit for and important in HSG, constitute the foundational components of a tool for appraising HSG documents, and are useful for guidance development and reporting. I inductively develop a framework for HSG that shows relationships between these candidate concepts as well as between clusters of the concepts. For each concept, a refined description and a specific set of operational considerations to optimize its use are developed.

Building on the findings from the first study, the objective of Study 2 (Chapter 3, entitled 'concept evaluation and development of the draft instrument') is to evaluate the importance, value and relevant priority of the candidate concepts and definitions generated in the CIS and identify any missing components. I present findings from a survey of international stakeholders (policy-makers, HSG developers, researchers, healthcare professionals, health policy and health systems experts, health administrators/managers etc.) drawn from all six WHO health regions. Participants provide feedback on the candidate concepts in terms of coverage, overlap, content validity and clarity of their descriptions. Participants also nominate additional items, and with inputs from deliberations with health policy/health system experts. I perform a

reformulation of the selected concepts and eliminate overlapping concepts. This evaluation serves as an initial measure of the face validity of the concepts, and leads to appropriate refinements to the candidate list of the concepts and clusters of the concepts. Moreover, it reflects an integrated KT approach whereby the future users of the tool are able to inform its content and presentation so it is acceptable to them. The major contribution of this chapter is the beta-version of the tool complete with refined concepts, and operational descriptions.

Building on the first two studies, the objective of Study 3 (Chapter 4, entitled 'pretest and refinement of the draft instrument') is to test the evaluation capacity and usability of the draft (beta-version) tool, and to further test the face validity of the new instrument to determine its feasibility of application, ease of understanding and the anticipated value of the information it generates for users. I present findings from a survey of international stakeholders (policy-maker, HSG developers, researchers, healthcare professionals, health policy and health systems experts, health administrators/managers etc.) representing all six WHO health regions. Participants apply the beta version of the HSG tool to evaluate the quality of some WHO HSG documents and answer a series of survey questions about the appraisal process. The object of study and analysis for Chapter 4 is the HSG tool and not the HSG documents themselves. The major contributions of this chapter are: general feedback about the tool, perceptions of usefulness, appropriateness, and ease of application. Using these data, revisions to address the new gaps are made and a refined HSG tool is generated, version one of the AGREE-HS

Together, the chapters presented in this thesis provide substantive, methodological and disciplinary contributions to the field of health policy and health systems research in general, and particularly to the study of efforts that aim to support the development, appraisal and reporting of HSG. Substantively, this thesis generates new knowledge by producing a HSG tool that can be used by stakeholders from low, middle and high-income countries. It contributes a new theoretical framework that offers a more comprehensive approach to considering the concepts (items, criteria, domains) that are important to take into account during the process of development of HSG, that are essential for evaluating the quality of HSG, and that are useful for guidance reporting. The critical interpretive synthesis (CIS) in Chapter 2 provides a first attempt to generate concepts that are relevant to the design of a HSG tool. The survey in Chapter 3 provides a preliminary endeavor to understanding the value of these concepts and leads to the production of the beta version of the tool. The survey in Chapter 4 provides an initial effort to assess the ease of application of the new tool and leads to a refined version of the HSG tool, AGREE-HS (Version 1).

Methodologically, the three studies that I present in this thesis clearly build on each other in a logical and sequential manner. While the methodology is adapted from the approaches used to design the AGREE II tool, adaptation is performed in a number of key ways; first of all, the CIS feeding into this process is a novel approach to concept generation. In a budding field like HSG, the available papers are highly heterogeneous and methodologically diverse, and the CIS approach offers promise in allowing the analysis of these complex bodies of literature. Through a typically interpretive mode of inquiry, the CIS makes it possible to point out patterns and draw linkages between variables hence promoting concept and theory generation. Secondly, the purposeful recruitment of participants in the two surveys also provides a unique approach to sampling. I worked with international partners to make sure these studies resonated with

different people from different regions. For this thesis, a broad group of co-investigators, collaborators and participants are selected based on expertise and knowledge of particular WHO regions. This ensures that stakeholders in low, middle and high-income countries will find the new HSG tool useful and appropriate.

These studies also provide disciplinary contributions particularly to the nascent field of HSG and in general to health systems research and efforts to strengthen health systems by tackling some of its challenges. In contrast to clinical and patient contexts, the health systems/policy field has not optimized advancements in knowledge translation science or practice. By incorporating insights from CPGs, these studies take some important first steps towards adopting the rigor of the clinical world and appropriately adapting it for use in the policy world of HSG. Informing processes for development, appraisal and reporting of guidelines already exists in clinical practice, and this study is tackling an even more complex area, that of health systems, using similarly rigorous, but properly adapted approaches. This project, therefore, advances scholarship in the areas of health policy and knowledge translation by yielding a tool that will be used in developing guidelines for health systems that are different from CPGs. It constitutes a significant advancement towards the provision of high quality recommendations and modes of action needed to implement health systems delivery, financial and governance arrangements.

References

1. AGREE Collaboration Writing Group: Cluzeau F, Burgers J, Brouwers M, Grol R, Mäkelä M, Littlejohns P, Grimshaw J, Hunt C (2003). Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: the AGREE project. Qual Safe Health Care. 2003;12:18-23.
2. Arah, O. A., Klazinga, N. S., Delnoij, D. M. J., Ten Asbroek, A. H. A., & Custers, T. (2003). Conceptual frameworks for health systems performance: a quest for effectiveness, quality, and improvement. International Journal for Quality in Health Care, 15(5), 377-398.
3. Birtwhistle R, Pottie K, Shaw E, Dickinson JA, Brauer P, Fortin M, Bell N, Singh H, Tonelli M, Connor Gorber S, Lewin G, Joffres M, Parkin P (2012): Canadian Task Force on Preventive Health Care: We're back! Can Fam Physician 2012;58: 13-15
4. Bosch-Capblanch X (2011) Handbook for Supporting the Development of Health System Guidance. Basel: Swiss Centre for International Health.
5. Bosch-Capblanch X, Lavis JN, Lewin S, Atun R, Røttingen JA, et al. (2012) Guidance for evidence-informed decisions about health systems: rationale for and challenges of guidance development. PLoS Med 9: e1001185.
6. Bosch-Capblanch X.(2011). Analytical framework for health systems guidance. Handbook for developing Health Systems Guidance; Swiss TPH. 2010.
7. Brinkerhoff, D (2003). Accountability and Health Systems: Overview, Framework and Strategies. U.S. Agency for International Development. Order No. TE 018. (www.who.int/management/partnerships/accountability/AccountabilityHealthSystemsOverview.pdf)
8. Brouwers M, Kho ME, Browman GP, et al for the AGREE Next Steps Consortium (2010c). AGREE II: Advancing guideline development, reporting and evaluation in healthcare. Can Med Assoc J. 2010 Dec;182:18. E839-E842.
9. Brouwers, M. C., Kho, M. E., Browman, G. P., Burgers, J. S., Cluzeau, F., Feder, G., ... & Makarski, J. (2010a). Development of the AGREE II, part 1: performance, usefulness and areas for improvement. Canadian Medical Association Journal, 182(10), 1045-1052.
10. Brouwers, M. C., Kho, M. E., Browman, G. P., Burgers, J. S., Cluzeau, F., Feder, G., ... & Makarski, J. (2010b). Development of the AGREE II, part 2: assessment of validity of items and tools to support application. Canadian Medical Association Journal, 182(10), E472-E478.
11. Chen LC. (2004). Harnessing the power of human resources for achieving the MDGs. Geneva: World Health Organization.
12. Cluzeau F, Littlejohns P, Grimshaw J, Feder G, Moran S (1999) Development and application of a generic methodology to assess the quality of clinical guidelines. Int J Qual Health Care. 1999;11:21-8.
13. Frenk, J., & Moon, S. (2013). Governance challenges in global health. New England Journal of Medicine, 368(10), 936-942.
14. Graham R, Mancher M, Wolman DM, Greenfield S, Steinberg E, eds (2011). Clinical practice guidelines we can trust. Washington, DC: National Academies Press; 2011
15. Haines, A., 2011: What types of evidence are needed to inform decisions about health systems? Available at

http://community.cochrane.org/multimedia/multimedia-cochrane-colloquia-and-meetings/colloquium-madrid-2011/)
16. Hanson K, Ranson KM, Oliveira-Cruz V, et al (2003). Expanding access to priority health interventions: a framework for understanding constraints to scaling up. Journal of International Development;15:1–14.
17. Health Canada (2009). Commission on the Future of Health Care in Canada: The Romanow Commission [homepage on the internet] c2002 [updated 2009 Apr 30; accessed 2013 Aug 13] Available from: http://www.hc-sc.gc.ca/hcs-sss/com/fed/romanow/index-eng.php.
18. Hoffman S, Rottingen J-A, Bennett S, Lavis J, Edge J, Frenk J (2013) Background Paper on Conceptual Issues Related to Health Systems Research to Inform a WHO Global Strategy on Health Systems Research - a working paper in progress. Geneva: World Health Organisation; 2013. http://www.phac-aspc.gc.ca/about_apropos/evaluation/reports-rapports/2010-2011/h1n1/mrap-mrpa-eng.php
19. Islam, M., ed. (2007). Health Systems Assessment Approach: A How-To Manual. Submitted to the U.S. Agency for International Development in collaboration with Health Systems 20/20, Partners for Health Reform plus, Quality Assurance Project, and Rational Pharmaceutical Management Plus. Arlington, VA: Management Sciences for Health.
20. Lavis JN, Lomas J, Hamid M, Sewankambo NK (2004) Assessing country-level efforts to link research to action. Bull World Health Organ 84(8): 620–628.
21. Lavis JN, Panisset U. EVIPNet Africa's first series of policy briefs to support evidence-informed policymaking. Int J Technol Assess. 2010; 26(2): 229-32.
22. Lavis JN, Ross SE, Hurley JE, Hohenadel JM, Stoddart GL, Woodward CA, Abelson J (2002). Examining the role of health services research in public policymaking. Milbank Quarterly, 80:125-54.
23. Lavis JN, Røttingen JA, Bosch-Capblanch X, Atun R, El-Jardali F, et al (2012). Guidance for evidence-informed policies about health systems: linking guidance development to policy development. PLoS Med ;9: e1001186
24. Lavis JN., Wilson, M. G., Moat, K. A., Hammill, A. C., Boyko, J. A., Grimshaw, J. M., & Flottorp, S. (2015). Developing and refining the methods for a 'one-stop shop'for research evidence about health systems. Health Research Policy and Systems,13,10.
25. Law, Susan (2008). Evidence Reconsidered for Health System Guidance. Retrieved on February 4th 2015 from http://www.ncchpp.ca/docs/Law_JASP_evidenceAN_nov08_2.pdf
26. Lewin S, Bosch-Capblanch X, Oliver S, et al. Guidance for evidence-informed policies about health systems: assessing how much confidence to place in the research evidence. PLoS Med. 2012; 9(3):e1001187.
27. Lomas, J., Culyer, T., Mccutcheon, C., Law, S., & Tetroe, J. (2005). Conceptualizing and Combining Evidence for Health System Guidance. Retrieved on October 4th 2013 from http://www.cfhi-fcass.ca/migrated/pdf/insightAction/evidence_e.pdf
28. National Institute for Health and Care Excellence - NICE - (2011). Process manual for accrediting producers of guidance, advice and recommendations for practice: a guide for producers and stakeholders. Last update: May 2013; Retrieved on Aug 30th 2013 from http://www.nice.org.uk/media/886/95/NICEAccreditationProcessManual.pdf
29. Pantoja, T (2012) Health Systems Guidance: The challenges of synthesizing health systems research. Conference proceedings from the Second Global

Symposium on Health Systems Research: Panel session on Health Systems Guidance: The challenge of translating health systems research into policies. (http://hsr2012.abstractsubmit.org/sessions/620/)
30. Public Health Agency of Canada - PHAC - (2011). Management Response and Action Plan – H1N1. [home page on the internet] [updated 2012 Nov 19; accessed 2013 Aug 13]. Available from:
31. Qaseem A, Forland F, Macbeth F, Ollenschlager G, Phillips S, van der Wees P (2012). Guidelines International Network: toward international standards for clinical practice guidelines. Ann Int Med. 2012 Apr;156(7):525-31.
32. Schneider, H., Blaauw, D., Gilson, L., Chabikuli, N., & Goudge, J. (2006). Health systems and access to antiretroviral drugs for HIV in Southern Africa: service delivery and human resources challenges. Reproductive health matters, 14(27), 12-23.
33. Snowdon, A., & Cohen, J. A. (2011). Strengthening health systems through innovation: lessons learned. Ivey Centre for Health Innovation.
34. Travis, P., Bennett, S., Haines, A., Pang, T., Bhutta, Z., Hyder, A. A., ... & Evans, T. (2004). Overcoming health-systems constraints to achieve the Millennium Development Goals. The Lancet, 364(9437), 900-906.
35. United Nations. The Millennium Development Goals Report (2011). Retrieved on March 02nd 2015 from: http://www.un.org/millenniumgoals/pdf/%282011_E%29%20MDG%20Report%202011_Book%20LR.pdf.
36. Vandenberg T, Coakley N, Nayler J, et al. A framework for the organization and delivery of systemic treatment. Curr Oncol. 2009 Jan;16(1):4-15.
37. von Bertalanffy, L. (1968). General system theory: Essays on its foundation and development, rev. ed. New York: George Braziller
38. WHO. 2007. Everybody's Business: Strengthening Health Systems to Improve Health Outcomes: WHO's Framework for Action. Geneva: World Health Organization.
39. World Health Organization - WHO - (2000). The World Health Report. Health systems: improving performance. Geneva, World Health Organization, 2000 (http://www.who.int/whr/2000/en/, accessed 9 February 2013)
40. World Health Organization - WHO- (2010). WHO recommendations: Increasing access to health workers in remote and rural areas through improved retention. Global policy recommendation. Geneva, 2010. Available from: http://whqlibdoc.who.int/publications/2010/9789241564014_eng.pdf?ua=1
41. World Health Organization (2012). WHO recommendations: optimizing health worker roles to improve access to key maternal and newborn health interventions through task shifting. Geneva, 2012. Available from: http://apps.who.int/iris/bitstream/10665/77764/1/9789241504843_eng.pdf.
42. World Health Organization - WHO - (2014). WHO recommendations: Principles and considerations for adding a vaccine to a national immunization programme. From decision to implementation and monitoring. Geneva, 2014. Available from: http://apps.who.int/iris/bitstream/10665/111548/1/9789241506892_eng.pdf?ua=1

Chapter 2: Health Systems Guidance Appraisal Concepts – A Critical Interpretive Synthesis

Denis Ako-Arrey[1]
Melissa Brouwers[1]
John N. Lavis[1,2]
Mita Giacomini[1]
The AGREE-HS team[*]

Word count:
597 (Abstract)
9110 (Full text)

1. McMaster University
2. Harvard T.H Chan School of Public Health

[*]The members of the AGREE-HS Scientific/advisory team include: Pierre Ongolo-Zogo (Cameroon); Jillian Ross, Mike Wilson, Padraig Warde and Sheila McNair (Canada); Yoalong Chen (China); Iván Darío Flórez Gómez (Colombia); Govin Permanand (Denmark); Fadi El-Jardali (Lebanon); Carmen Mihaela Dolea, Xavier Bosch-Capblanch and and Ulysses Panisset (Switzerland); Andy Haines, Francoise Cluzeau and Pablo Perel (United Kingdom); Luis Gabriel Cuervo (United States of America).

Preface

The paper presented in this chapter addresses an important theoretical gap in the literature focused on health systems guidance (HSG) and, specifically, on the concepts (items, criteria, domains) that are useful in evaluating the quality of HSG and further inform their development and reporting requirements. The work presented in this chapter generated a framework for HSG that is made up of 3 core domains and 30 concepts underpinning them, and complete with descriptions and operational definitions. This HSG framework represents those elements that define expectations of HSG to facilitate optimal informed decisions among policy-makers on health systems delivery, financial, and governance arrangements. Using a relatively new approach to knowledge synthesis of the available literature, a critical interpretative synthesis, the study provides a solid base for developers and users of HSG to reflect on the scope of crucial components of a good quality HSG. This chapter provides important background and data for the subsequent chapters 3 & 4.

I was responsible for conceptualizing and designing the study protocol and research instruments with my supervisor (Dr. Melissa Brouwers) and also received inputs from members of my supervisory committee (Dr. John Lavis and Dr. Mita Giacomini). I completed all data collection, analysis, and interpretation and I also drafted the paper. Through an iterative process and adopting an integrated knowledge translation approach of collaborating with researchers and knowledge users, members of our core and expanded scientific team contributed to refining the HSG framework as well as in providing feedback and suggestions on draft versions of the paper.

Abstract

Background: Health systems guidance (HSG) are systematically developed statements produced at global, national, and regional levels to assist with decisions about appropriate options for addressing a health systems challenge, including related changes in health systems arrangements (delivery, financial or governance arrangement); the implementation of these options; and the monitoring and evaluation of the implementation efforts (Bosch-Capblanch et al, 2012; Lewin et al, 2012). However, the development of HSG poses unique conceptual and methodological challenges related to the varied types of evidence that are relevant, valued and considered; the complexity of health systems internal and external relations that must be addressed; and the pre-eminence of contextual factors that directly influence the design and adoptability of recommendations (Bosch-Capblanch et al, 2012). With the rising trends that encourage bridging the gap between research and policy and practice, this is a significant gap in both the science and practice of knowledge translation. To address this gap, we are conducting a program of research with the international HSG community that aims to create a tool to support the appraisal of HSG and further enhance HSG development and reporting. The first step to this program of research, and the focus of this paper, was to conduct a knowledge synthesis of the published and grey literatures to determine concepts (items, criteria or domains) related to HSG development, reporting and quality.
Methods: We applied a critical interpretive synthesis approach to knowledge synthesis (Dixon-Wood et al, 2006) that enabled an iterative, flexible and dynamic analysis of diverse bodies of literature (qualitative, quantitative and theoretical papers) in order to generate a candidate list of concepts that will constitute the foundational components of the HSG tool. Using our review questions as compasses, we were able to guide the search strategy to look for papers based on their potential relevance to HSG appraisal, development and reporting. The search strategy included various electronic databases and sources (CINAHL, Cochrane Library, EMBASE, Google Scholar, Health Systems Evidence, Latin American and Caribbean Health Sciences Literature (LILACS), PUBMED, Virtual Health Library, Web of Science). We also searched subject specific journals, grey literature, including conference abstracts, a convenience sample of research reports, book chapters, unpublished data, dissertations and policy documents, and also contacted first authors of randomly selected papers. Screening the papers (titles, abstracts and full texts) as well as data extraction was completed independently and in duplicate with disagreements being resolved via consensus, and a narrative approach to data synthesis was executed.
Findings: We identified 43 papers that met eligibility criteria. No existing review was found on this topic and no HSG appraisal tool (draft or final version) was identified. Over one third of the authors implicitly or explicitly identified the need for a high quality tool aimed to systematically evaluate HSG and contribute to its development/reporting. We identified 30 concepts (items, criteria) that may be relevant to the appraisal of HSG and were able to cluster them into three meaningful categories (domains); process principles, content and context principles.
Conclusion: Development of high quality HSG will impact the type of recommendations being formulated, the degree to which they get implemented, the methods of dissemination, and the extent to which they impact on the usual operations of the health systems (WHO, 2003). Our study showed the role that the process principles, content and context principles play in the development, appraisal and reporting of HSG and demonstrated the link and resonance within and between their various concepts. We pointed out patterns and drew linkages between the concepts in order to show that the concepts and categories do not typically occur in a linear or independent manner.

1. Introduction / Background

Roemer (1991) defined a health system as the combination of resources, organizations, financing and management bodies that culminate in the delivery of health services to a population. A redefinition of the term health system was proposed by the World Health Organization (WHO, 2000) to consist of all organizations, people, actions and activities whose primary purpose is to promote, restore, and maintain health. This was further modified – to be more comprehensive and explicit – by a group convened by the WHO's Alliance for Health Policy and Systems Research, to refer to delivery, financial and governance arrangements for health care and population health services and the broader context in which they are negotiated, implemented and reformed (Hoffman et al, 2012).

Health systems guidance (HSG) are systematically developed statements produced at global, national, and regional levels to assist with decisions about appropriate options for addressing health systems challenges (including related changes in health systems arrangements), the implementation of these options, and the monitoring and evaluation of the implementation efforts (Bosch-Capblanch et al, 2012). The need to develop and use evidence-informed approaches to address national and global health system issues through the use of HSG has been articulated by many (Lavis et al, 2012). For example, the Millennium Development Goals (MDGs) reflect eight international goals (e.g., reduce child mortality, improve maternal health, and combat HIV/AIDS, malaria and other diseases, etc.) and were developed by the United Nations in 2000. However, their attainment has been hindered by weak health systems and lack of system-specific recommendations (Remme et al, 2010; Fryatt et al, 2010). Formulating recommendations to address root causes has the potential to clear the path towards better health outcomes in line with health goals (Murray & Frenk 2000).

The terms *guidelines* and *guidance* both refer to any document containing a recommendation on a course of action. However, in contrast to the word "*guidelines"*, the term "*guidance*" is used in the health system context in order to make explicit the difference between the process of supporting evidence informed judgments for health system issues from that of clinical judgments as would be expected with clinical practice guidelines (CPGs). As a policy-oriented product, HSG represents the whole body of knowledge that informs policy decisions on how health system issues should be classified or prioritized, appropriate health system governance, optimal financial arrangements, system organization, and the design and delivery of effective health programs and services (WHO, 2005; Bosch-Capblanch et al, 2010; WHO, 2011; Lavis el al, 2013).

The main bodies tasked with developing HSG are international and intergovernmental organizations (e.g., WHO, PAHO), local ministries of health, and special national committees or agencies providing support to ministries of health. Their production is usually linked with these high-level health sector entities, and less frequently in decentralized structures at the sub-national level. These organizations focus, at least in part, on strengthening health systems and getting the right mix of cost-effective programs, services and drugs to those who need them. While these bodies have demonstrated an interest in the development, appraisal and reporting of this type of guidance (Bosch-Capblanch et al, 2011), these intentions have so far not been matched by appropriate action or results, in part, due to the lack of experience in developing HSG.

Indeed, in comparison to CPGs, the development of HSG pose different conceptual and methodological challenges related to the varied types of evidence that are relevant, valued and considered; the complexity of health systems' internal and external relations that must be addressed; and the pre-eminence of contextual factors that directly influence the design and adoptability of recommendations (Bosch-Capblanch et al, 2012). As with CPGs, HSG statements or recommendations should be justified by assessments of the quality of evidence supporting them, the potential for unintended consequences, and by discussions of implementation and contextual issues. For example, the full implementation of HSG recommendations can be further hampered by "bottlenecks" like health system fragmentation and capacity limitations; these limitations should also be addressed in the guidance documents (Travis et al, 2004; Chopra, 2009).

The potential for positive impacts by HSG as a decision tool to improve health systems is great. However, as with CPGs (AGREE Next Steps, 2009), this potential is only as good as the quality of the HSG. Indeed, the ability to impact and optimize health system performance and efficiency through the development and adoption of HSG is hampered by the dearth of tools to guide their development, appraisal and reporting. As a consequence, this leaves knowledge users at a loss when choosing the highest quality and most appropriate guidance or in creating new guidance in circumstances where there is none, or navigating circumstances where the existing guidance is not credible or of poor quality. In contrast to the development, appraisal and reporting methods for CPGs, the development of HSG is still at a rudimentary stage (Bosch-Capblanch et al, 2012). There has been some work related to the use of ‘evidence briefs’ to assist policy-makers and stakeholders with working through a health-system problem, options for addressing it, and key implementation considerations, informed by the best available data and research evidence (e.g., SUPPORT, 2012). However, there is a need for systematically and transparently developed guidance that can feed into such context-specific documents. With the rising trends that encourage bridging the gap between research and policy and practice, this is a significant research gap in both the science and practice of knowledge translation. HSG can provide this bridge between research synthesis and policy needs for evidence.

The creation of high quality HSG requires tools to support the development and reporting of high quality guidance and tools able to differentiate between high and low quality reports. There is also value in establishing acceptable quality thresholds, and creating common methodologies and nomenclature among the HSG community (developers, users, and researchers). At present, however, there is no universally acceptable gold standard approach for appraising HSG, although there are some tools (for example the Handbook for Supporting the Development of Health System Guidance) to support their development (Bosch-Capblanch et al, 2011) and reporting. To address these gaps, we are conducting a program of research with the international health systems guidance community that aims to create a tool to support the appraisal of HSG and further enhance HSG development and reporting.

The first step to this program of research, and the focus of this paper, was to conduct a knowledge synthesis of the published and grey literatures to determine concepts (items, criteria or domains) related to HSG development, reporting and quality. The results of the synthesis are to provide the foundational components of our HSG tool and to serve as a conceptual status report for the research community.

2. Methodology

With the goal of generating a list of concepts that can be used to develop, appraise and report HSG, we conducted an initial search to identify existing concepts, tools, templates, or checklists that have been used or could be used to describe, differentiate or appraise the quality of HSG. The aim of this initial search was to determine if reviews on this topic exist, estimate an initial number of relevant papers available, highlight additional and useful search terms, and clarify inclusion and exclusion criteria and the most appropriate knowledge synthesis approach (Brettle, 2003; Heaton et al, 2012).

After contemplating the variety of knowledge synthesis approaches available, a critical interpretive synthesis (CIS) approach was considered the most appropriate for three reasons. First, CIS is a systematic approach that facilitates the analysis of complex and diverse bodies of literature including qualitative, quantitative and theoretical papers (Dixon-Woods et al, 2005a; Dixon-Woods et al, 2006a; Taylor et al, 2009; Barnett-Page & Thomas, 2009; Gysels et al, 2012; Entwistle et al, 2012; Kazimierczak et al, 2013). The available literature on HSG is highly heterogeneous and methodologically diverse, comprising of a mix of empirical qualitative and quantitative papers and non-empirical papers. A wider range of evidence and study designs are typical occurrences in health systems/health policy research and HSG is a nascent domain where good thinking is likely to be captured in expert opinion/views, editorial comment, policy documents, political statements, experiences of stakeholders, theoretical/discussion papers, and other colloquial forms of evidence. In contrast to CIS, conventional systematic reviews, for example, have been criticized for excluding forms of evidence traditionally considered as non-experimental (Dixon-Woods et al, 2005a; Dixon-Woods et al, 2006a; Dixon-Wood et al, 2006b; Talseth & Gilje, 2010; Morrison et al, 2012).

Second, the objective of a CIS is to develop new concepts and theories through a typically interpretive mode of inquiry. This is in contrast to more conventional systematic review approaches where the mode of inquiry is more aggregative and aimed at testing theories by collating, compiling, pooling, and summarizing common outcomes across a range of studies (Noblit & Hare, 1988; Dixon-Woods et al, 2005a; Dixon-Woods et al, 2005b; Dixon-Woods et al, 2006a; Heaton et al, 2012; Markoulakis & Kirsh, 2013; Entwistle et al, 2012; Kazimierczak et al, 2013). Since our review aims at generating a candidate list of items, criteria, or domains for HSG development, appraisal and reporting, this theory-generating approach that CIS promotes was deemed a good fit.

Third, CIS offers a more flexible, iterative, dynamic and reflective approach requiring investigators to assess the extent to which new information or data are provided with each additional paper considered. It applies a relatively loosely defined set of processes for critically analyzing and synthesizing literature (Dixon-Woods et al, 2005a; Dixon-Woods et al, 2006a; Dixon-Wood et al, 2006b; Barnett-Page & Thomas, 2009; Heaton et al, 2012; Morrison et al, 2012). This differs from the more conventional systematic review methods which have clearly stated study protocols, an exhaustive search of all available literature, standardized data extraction templates, explicit quality appraisal checklists, as well as pre-determined focused questions, strict inclusion and exclusion criteria and specified data boundaries (Dixon-Woods et al, 2005a; Dixon-Woods et al, 2006a; Dixon-Wood et al, 2006b; Barnett-Page & Thomas, 2009; Talseth & Gilje, 2010; Morrison et al, 2012).

With the unique conceptual and methodological challenges related to the development of HSG (Bosch- Capblanch et al, 2012), it was essential to have a review method that was iterative, flexible and dynamic. Our review was not simply aimed at summarizing this complex literature on HSG, so we sought to produce a logical and insightful interpretation of a purposefully sampled body of evidence in a comprehensive but not necessarily exhaustive fashion (Dixon-Woods et al, 2005a; Talseth & Gilje, 2010; Morrison et al, 2012).

2.1. Review question(s)

As per CIS methodological standards, our review questions served as compasses rather than anchors (Eakin & Mykhalovskiy, 2003; Dixon-Woods et al, 2005a; Talseth & Gilje, 2010) allowing for the concepts of HSG to be derived from synthesis of the literature and constantly modifying them in an iterative manner throughout the review.

Our guiding review questions were;

1. How is HSG quality defined in the literature?

2. What criteria (or tools or instruments, checklists, systems, etc.) have been used: to describe or define HSG quality or reporting requirements; to appraise HSG quality; or to differentiate between HSG on the basis of quality?

3. What methods have been used to develop these criteria and is there evidence that the criteria are reliable, valid and useable?

4. What methods have been used to address health system issues/challenges?

2.2 Literature search

In searching the literature, our goal was to select papers based on their potential relevance to HSG appraisal and quality, while including other papers that, though not directly relevant to health systems, were deemed important for the purpose of our review. Instead of including an exhaustive number of papers, our plan was to provide a comprehensive sampling frame of potentially relevant papers using emergent eligibility criteria (Markoulakis & Kirsh, 2013; Entwistle et al, 2012). As per the standards of a CIS, the boundaries of our inclusion and exclusion criteria were modifiable, dynamic and continuously shifting (Dixon-Woods et al, 2005a; Flemming 2009; Markoulakis & Kirsh, 2013). The goal was to populate the concept of HSG with new concepts rather than finding papers that reiterated ideas already captured in previously reviewed papers.

2.3 Eligibility criteria

1. Study content: (a) Papers that evaluated HSG or papers that report on criteria/tools that have been documented as important indicators of HSG quality. (b) Papers that reported on methods for addressing health system issues/challenges.
2. Time Frame: Papers are eligible if published in or after the year 2000 (when the first World Health Report on health systems [WHO, 2000], was published).

3. Context: Unrestricted. We sought papers that considered HSG evaluation and health system issues/challenges in various contexts (low, middle and high-income countries).
4. Study design: Any study design.
5. Language: English, French and Spanish

A combination of key and free text terms were used to search through various databases. Terms used were; health systems, health policy, guidelines, guidance, health services arrangement, health services organization, health system issues, health system challenges, tools, instruments, criteria, items, domains, evaluation, appraisal, quality, and standards.

The search strategy included various electronic databases and sources: CINAHL, Cochrane Library, EMBASE, Google Scholar, Health Systems Evidence, Latin American and Caribbean Health Sciences Literature (LILACS), PubMed, Virtual Health Library, and Web of Science. We also searched subject specific as well as regional electronic sources: Australia's National Health and Medical Research Council, Evidence Best Practices for Public Health, University of Massachusetts National Guidelines Clearinghouses, National Institute of Health and Care Excellence (NICE), WHO's Evidence Informed Policy Network (EVIPNet), WHO EURO's Health Evidence Network, and Guidelines International Network (GIN) directories.

We also searched for other grey literature, including conference abstracts (Global Symposium on Health Systems Research, Canadian Association for Health Services and Policy Research, Health Systems and Process Improvement Conference, Canada's Health Leadership Conference, Health Systems Reform in Asia Conference, International Society on Priorities in Health Care). Further, research reports, book chapters, unpublished data, dissertations and policy documents that were nominated by members of the team or found in unique holdings of health sciences libraries in Canada were also included. Additional papers were identified by manually searching bibliographies, while more were obtained by hand searching some key journals (i.e., Health Policy and Planning, Health Services Research and Policy, Health Research Policy and Systems, Global Healthcare Systems, Health Systems and Reform, and Health Policy). This was a complementary search strategy to account for papers not included in electronic databases or with search terms that do not allow them to be easily identified (Dickersin & Scherer, 1994; Flemming 2009).

Finally, we also contacted (through emails, phone, Skype or in person) experts, colleagues and members of our research team with a known interest on this topic, to identify additional papers (published, unpublished, or ongoing). To increase the scope of these key informants, we also asked initial contacts to refer us to others who can provide more information. Relevant papers were then imported to Endnote bibliographic software.

A two-step sampling process was used. First, a purposeful sampling approach (Seale, 1999; Annandale et al, 2007) was used to investigate the literature to determine the range of unique candidate concepts – or theoretical domains - associated with HSG. We stopped sampling at the saturation point where looking at new literature no longer contributed additional concepts (Annandale et al, 2007; Entwistle et al, 2012). Second, a theoretical sampling strategy was used to interrogate the literature relevant to each of the identified concepts. Theoretical sampling does not occur at a single point in the research

process but is a recurrent feature, which aligns it well with the dynamic CIS methodology. Similar to the first approach, we stopped sampling at the saturation point where looking at new literature no longer contributed additional descriptions of the identified concepts. The intent here was to thoroughly capture the depth of the concept across the literature in order to generate and develop the theoretical underpinnings of the ideas, rather than collecting numerous citations of identical concepts and/or descriptions. Screening titles, abstracts and full text was completed independently and in duplicates (DAA & SA[**]). Disagreements were resolved by consensus.

2.4 Quality appraisal

Given the diversity and complexity of the literature on HSG, considering that the available literature could not be hierarchically ordered in terms of study design importance or relevance, and in view of the fact that very little consensus exists as to whether to perform quality appraisal in a CIS methodology or what approach might be most appropriate (Sandelowski et al, 1997; Dixon-Wood et al, 2004; Dixon-Woods et al, 2006a; Markoulakis & Kirsh, 2013) this task was not undertaken. Instead, we chose to evaluate papers for inclusion based on a judgment of their relevance and likely contribution to concept development and theory (Gough, 2007; Barnett-Page & Thomas, 2009; Markoulakis & Kirsh, 2013; Kazimierczak et al, 2013). Our assumption was that some methodologically weak papers are theoretically and conceptually pertinent (Flemming 2009; Talseth & Gilje, 2010; Entwistle et al, 2012). The goal of our review was to generate key concepts relevant to the appraisal of HSG, so we applied a flexible relevance boundary in order to be as inclusive and comprehensive as possible and include papers that could contribute to this theory generation.

2.5 Data extraction and analysis

As a result of the nature of the data, we did not use a standardized data extraction template to retrieve the relevant elements of this review from all the retained papers. Instead, and as per CIS methodology, we gave latitude for a more narrative data retrieval approach. It was however possible to extract some key categories from all the papers including: author(s), and the year of publication, geographical location of the study or affiliation of the author, purpose of the study/paper, its relevance and a summary of its main conceptual contributions, and types of items, criteria or domains considered. Data were extracted independently and in duplicate (DAA & SA[**]) with disagreements resolved via consensus. The extracted data were then compiled in summary form.

Again due to the anticipated diversity of methodological designs and implementation of the different items, criteria, and domains that were extracted from the papers, a narrative approach to data synthesis was executed. The data extracted were sorted and categorized into groups with themes. Each group represented similar concepts. This systematic method of recording themes, and making connections between themes and the data collected within a comprehensive category system (Burnard, 1991; Graneheim & Lundman, 2004), is an advocated concept-development coding approach (Strauss & Corbin, 1998). The information was imported into the analysis software NVivo version 9 in order to analyze groups of themes that depicted similar patterns in information. We did not organize the concepts in any hierarchical order but reported the frequency that the concept was identified in this review.

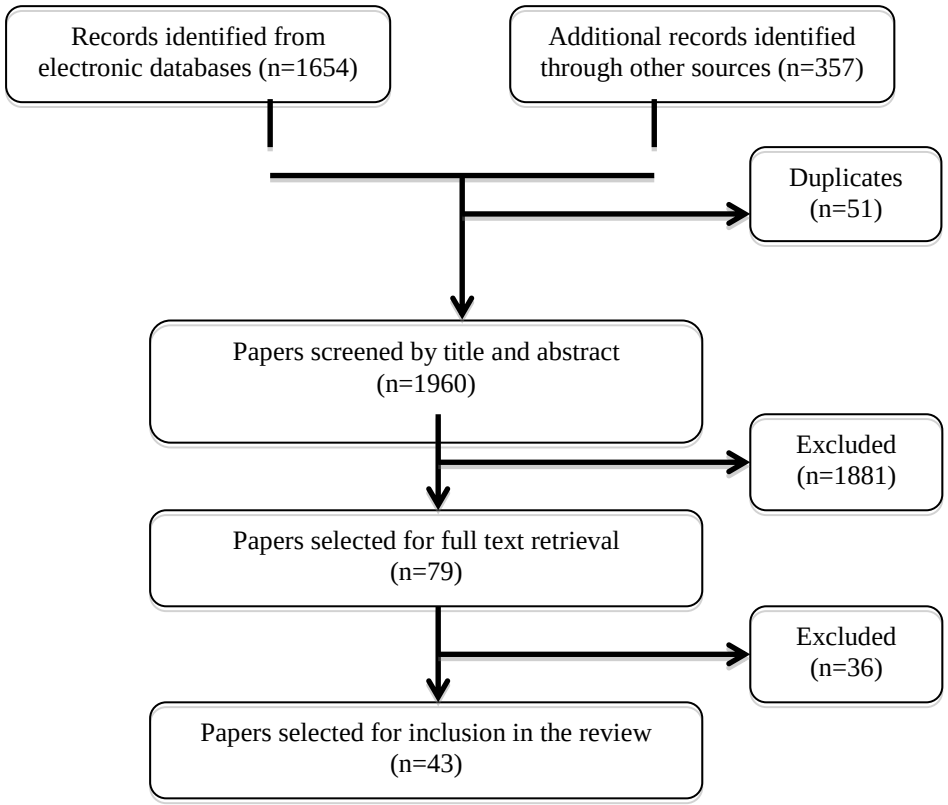
[**] DAA = Denis Ako-Arrey; SA = Saira Akram

Modeling from the health policy analysis triangle framework that was developed by Walt & Gilson (1994), we were able to connect the concepts (constructs) together into meaningful categories. The framework considers all the essential elements that interact to shape policy-making by demonstrating that health policy should focus on the processes contingent on developing and implementing change, on the content of health policy reform, on the context within which the policy is promulgated, as well as on the actors involved in the policy reform (Walt & Gilson 1994; Walt et al, 2008). The health policy analysis triangle framework is a highly simplified model of an extremely complex set of interrelationships between the different elements of the model (process, content, context and actors) with each element influencing or being influenced by the other (Buse et al, 2005; Walt et al, 2008). Development, appraisal and reporting of HSG play an important role in health policy making by providing options and recommendations to address a health systems issue. Therefore, similar to health policymaking, HSG development, appraisal and reporting can be seen to occur in a series of discrete yet interconnected components of process, content and context.

3. Results

No existing knowledge synthesis was found on the topic, and no existing HSG appraisal tool (draft or final version) was identified. We identified a total of 43 papers that met our eligibility criteria and reported on concepts (items, criteria, domains) considered directly or conceptually relevant to HSG and/or their quality (see Figure 1 for a flowchart of study selection).

Figure 1: Flowchart of study selection



Forty-seven percent of the retained studies were technical reports, 32% were concept papers, 13% were quantitative studies, and 8% used a mixed methods approach.

The authors held affiliations at the following organizations; 33% at universities, 25% at the WHO, 23% at research institutes, 12% at government agencies (United Kingdom's National Health Services [NHS], United States' Centre for Disease Control [CDC], United States Agency for International Development [USAID], United Kingdom's National Institute for Health and Care Excellence [NICE]) and 7% at national ministries of health. Fifty-eight percent (58%) of the papers were lead authored by an individual based in Europe, 30% in North America, 8% in Australia, 2% in Asia, and 2% in Africa. Over one third of the authors implicitly or explicitly identified the need for a high quality tool aimed to systematically evaluate HSG and contribute to development and reporting of HSG. Thirty concepts were identified that are considered to be a good fit for, and may be relevant to, the development, appraisal and reporting of HSG. Appendix I provides a description of the papers, their objectives, relevance to the HSG process, and the concepts extracted.

3.1. Process principles, content and context principles

Ostrom (2007) stated that identifying key elements (constructs) and relationships among them culminates in the generation of a theory, which connects these key constructs. We were able to organize the concepts and identify relationships within and across them. Through an iterative process, we clustered the concepts together into three meaningful categories (domains); process principles, content, and context principles.

3.1.1. Process principles

Process principles represent the methodological elements and the defining principles that demonstrate the development integrity of the HSG. It refers to the ways in which the HSG is initiated, developed and formulated and can be looked upon as the 'who' and 'how' of the guidance (Walt & Gilson, 1994; Buse et al, 2005). Here we find the procedures and principles that were employed by the HSG developers in coming up with the guidance recommendations. Process principles are critical in the development of HSG because the methodological strategy is crucial for guidance appraisal and for differentiating across HSG of varying quality. They are also important because they can articulate the tactics that could be utilized in the development and reporting of HSG to optimize quality. Process principles also depict the subjective and/or objective belief systems in place that represent the preferences (what ought to be) of individuals, groups, or populations on the course(s) of action for addressing a health systems challenge. Process principles also refer to how the guidance is made public in a way that is consistent and comparative as this can facilitate their comprehension, which will further enhance their uptake and aid in easing their application. Therefore, when developing, appraising or reporting HSG, a conceptual understanding of the process is fundamental. Table 2 below shows concepts for HSG process principles.

Table 1: Process principles	
1. Prioritization	The guidance fits in properly and is consistent with current health system priority areas within all applicable system levels and sectors by targeting a priority topic/jurisdiction/population. The guidance addresses these specific local priority areas with a clearly documented/demonstrated need, and also informs policy decisions on how to further prioritize across competing areas. The origin of the mandate to develop the guidance is also reported (for example, guidance that is mandated by a top official

Table 1: Process principles	
	like the Minister of Health is considered to be of high priority).
2. Relevance	The guidance recommendations should be relevant to, appropriate to and valid for the health system issue being addressed and relevant to the target population. The recommendations are relevant to the setting within which the guidance will operate, the institutional needs of that system/sub-system, as well as local, national and potentially global needs.
3. Timeliness	The recommendations are available in a timely manner in relation to when the policy decisions are made or timely in relation to the health system issue being addressed. The guidance is timely and usable by the broad range of health systems stakeholders since some policy decisions are sometimes made within crucial corresponding time frames or as windows of opportunity open and close.
4. Scope	The guidance is comprehensive and covers all relevant/appropriate (direct and indirect) health system levels, sub-systems and sectors. This also includes the various relevant sub-systems/components (hospitals, regional health authorities, and public health units etc.) within the health system. Identifying the scope is important because these various components are interlinked, interdependent and interact at various interfaces for overall health system performance.
5. Transparency	Systematic, replicable and transparent processes are applied in developing and reporting the guidance. These processes are systematic and transparent enough for the methods of development/reporting of the guidance to be reproducible. In order to paint a clear picture to knowledge users and target populations, sufficient details on these processes are provided.
6. Evidence-based	The best available research evidence informs the recommendations. The type(s) of evidence that was used to generate the guidance is/are stated, and this can range from well-established scientific methodologies or it can also be non-experimental (for example, colloquial evidence, anecdotal evidence or preliminary models). The evidence is context sensitive enough to resonate with local realities
7. Stakeholder involvement	Alternative views on the policy issue and the complementary expertise of a multidisciplinary group of relevant stakeholders are considered in the development of the guidance. Guidance developers, those involved in the implementation and evaluation of the guidance, and those who will be affected by the guidance recommendations are involved in the development process.
8. Ethical	The recommendations reflect considerations of an ethical lens, and align with applicable ethical principles and values (for example equity, equality, human rights, liberty, efficiency, autonomy, dignity, beneficence, etc). The guidance adequately promotes fairness and equality in terms of age, ability, culture, gender, socioeconomic status, religion, occupation, language, ethnicity, race or sexual orientation among the target population.
9. Outcomes	The guidance describes all the anticipated effects/outcomes as well as the appropriate indicators that can be used to measure the effects/outcomes. Adequate rationale regarding the choice of the outcomes and the indicators selected is provided. Considering potential uncertainties that may result, alternative outcomes and outcome indicators are also

Table 1: Process principles	
	identified. Performance thresholds, targets and standards that are considered acceptable are also identified.
10. Competing interests	A declaration of competing interests (for example, financial, academic, professional etc.) by the guidance developers, whether direct or indirect, is/are made in advance. The author's positions, roles, affiliations are clearly stated. Any reported or identified conflicts of interest are managed, with a description of the approaches used to curb any influence clearly documented. It is also clear that the views of any funding body involved have not influenced the development process of the guidance.
11. Presentation	The recommendations are clear, succinct, unambiguous and presented in a readable and consistent format, with key recommendations easily identifiable. The guidance is presented in a manner that is uniform, user-friendly and easy to navigate. It contains an executive summary, full text, a complete list of relevant references, a glossary of terms, full meaning of abbreviations and contact information of authors. Words or phrases denote an aspirational rather than a mandatory intent.

We were able to point out patterns and draw linkages between the process principles concepts. We offer five examples here. First, guidance should be relevant to and developed for health system areas with a clearly demonstrated and documented need, and the feedback of appropriate stakeholders will further highlight priority areas, which may potentially lead to timely interventions. Second, systematic and transparent approaches have to be applied to search for and identify relevant evidence, which should also be available in a timely manner. Third, ensuring that stakeholders from all applicable health system and sub-system levels are involved in the HSG process, that relevant evidence is sought and that appropriate outcomes are chosen can enhance the comprehensiveness of the guidance recommendations. Fourth, having clearly defined and consistent outcomes, declaring and managing interests as well as engaging in a participatory approach that incorporates the various perspectives of multiple stakeholders is also important for transparency of HSG. Finally, the ethical lens applied will be impacted by the quality of the evidence available, the composition of the involved stakeholders, the outcomes/indicators selected, the health system levels/sector involved, and whether it is a priority area that is relevant to the setting.

3.1.2 Content

Content (the ‘what’ of the HSG) represents the topics, subjects and substance that demonstrate the content integrity of the guidance recommendations (Walt & Gilson, 1994; Buse et al, 2005). Here we find statements about the impetus of the endeavor that provides direction on the HSG objectives and goals. Keeping in mind that in the absence of an appropriate economic lens, even the best-designed HSG may not achieve adequate success, we also find here the myriad of economic factors that come with the HSG process and need to be taken into account. In addition, within the content of the HSG, it is also important to consider and reflect on the fact that not all guidance recommendations go into practice as planned and events may occur that may derail intended actions. We also find here assessment considerations, usually ongoing throughout the HSG process, and which refer to those elements that assist in determining whether the guidance process was properly followed and/or records the impact/outcomes of the HSG. The content of the HSG also articulates the operationalization of the

proposed recommendations and carries information on how to best reach the target users of the guidance. Table 3 below shows concepts for HSG content.

Table 2: Content	
12. Problem definition	The health systems challenge (for example, financial, governance, or delivery arrangements) and its causes are clearly articulated (including any links/integration with other policy problems on the government's agenda). The nature, causes, magnitude, frequency and intensity of the problem, the populations and jurisdictions that are affected are clearly described. Appropriate rationale exists to justify that either new guidance is needed or existing guidance of acceptable quality can be adapted and used to address the problem.
13. Operationalization	The recommended "solutions" are operationalized sufficiently with the conceptualization, operational guidance and the mode of delivery of the options clearly stated. For example, the guidance provides instructional support for their successful operation and staff training that corresponds with the guidance expectations. Training recommendations could be in the form of a course, a workshop, accompanying manuals or consultancy services that staff can refer to during the implementation phase in order to standardize practice. If technical assistance (research institutes, consulting firms, NGO's) is required, this is identified and documented.
14. Costs	The guidance clearly documents a tentative budget required to implement the guidance recommendations. The potential financial costs (including downstream costs) of the operation are stated so that decision makers can assess the feasibility of the guidance implementation and evaluate whether the cost of implementing the guidance will be worth its potential impacts.
15. Resources	The inputs and resources required to implement the recommendations are clearly defined and they have to be proportionate to the health system problem that is being addressed. Some of these resources could be time, infrastructure, administrative capacity, information, equipment, supplies, healthcare professionals, training etc. The guidance provides a description of the amount, frequency and duration of the inputs and resources required.
16. Effectiveness	The guidance reports whether the anticipated goals and objectives have been achieved elsewhere or in a similar setting/condition, either through evidence from evaluation studies done at other sites (if available), or from expert opinion. In describing this effectiveness the guidance makes projections on how and why the objectives and goals will be achieved in the current setting.
17. Cost-effectiveness	The recommendations are attentive to value for money considerations. Sound local or applicable evidence (wherever available) on the cost-effectiveness of the guidance recommendations are provided. These traditionally report costs, direct and indirect program inputs/resources and outcomes to guide health policy decisions and provide benchmark(s) or threshold(s) that the health system is willing to accept or support in relation to other competing health system priorities.
18.	Description of the potential unintended consequences (positive &

Table 2: Content	
Benefits/harms	negative) of the guidance is provided or an assessment/judgment of the potential benefits/harms are made. Descriptions of the populations or institutions that may experience significant impacts are identified.
19. Dissemination plan	Strategies for communicating the guidance are included with a clear dissemination framework, the mode of delivery, and the integrity of the avenue used for dissemination been properly reported. The proposed strategies for disseminating the guidance are tailored to the relevant audiences (for example, a formal written report, user-friendly summary, oral presentation, poster, press release, booklet, workbook, films, pocket card etc.).
20. Process evaluation	This involves recommendations for evaluating the structure and process of implementation as well as corresponding challenges. This evaluation examines the extent to which the guidance recommendations were implemented as planned, and also provides a way to monitor the process and make adjustments and improvements to implementation strategies. It documents the inputs, services and activities that were implemented, and can identify potential strengths, weaknesses, opportunities and threats to the implementation process.
21. Outcomes/impact evaluation	An assessment of the outcome/impact of the guidance is recommended to determine whether the course of action was a success or failure. There are recommendations on measuring the results, or outcomes of the guidance in a way that determines whether the changes observed in relation to the health system challenge being addressed can be attributed to the guidance. There are also recommendations for an impact evaluation to look at the short and long term deeper primary and secondary changes that resulted from the guidance.
22. Updating	Recommendations for periodic updates are made and the procedure to update the guidance is provided with explicit timelines on anticipated review, appropriate expiration date of the guidance and an explanation of the rational for the proposed time frames. Setting time frames for periodic updates ensures that guidance producers revisit the recommendations and respond accordingly to potential health system changes and emerging challenges. Also, the recommendations should be current, and the evidence (for example, systematic reviews) on which they are based is considered recent and up-to-date.

As with the process principles, we were also able to point out patterns and draw linkages between the content concepts. For example, the way the problem is defined can provide direction on the costs and resources required for implementation of the solution, reveal some potential unintended consequences and guide the operational plan. The operational plan can also provide hints on some potential unintended consequences. A clear problem definition will also inform the process and outcomes/impact evaluation of the HSG. Designing the operational options can inform the costs and resources required. The costs, resources and operational plan will in turn influence whether the HSG will be effective and/or cost-effective. The updating plan is also contingent on the health system issue being addressed, the effectiveness of the guidance recommendation, and the resources required.

3.1.3. Context principles

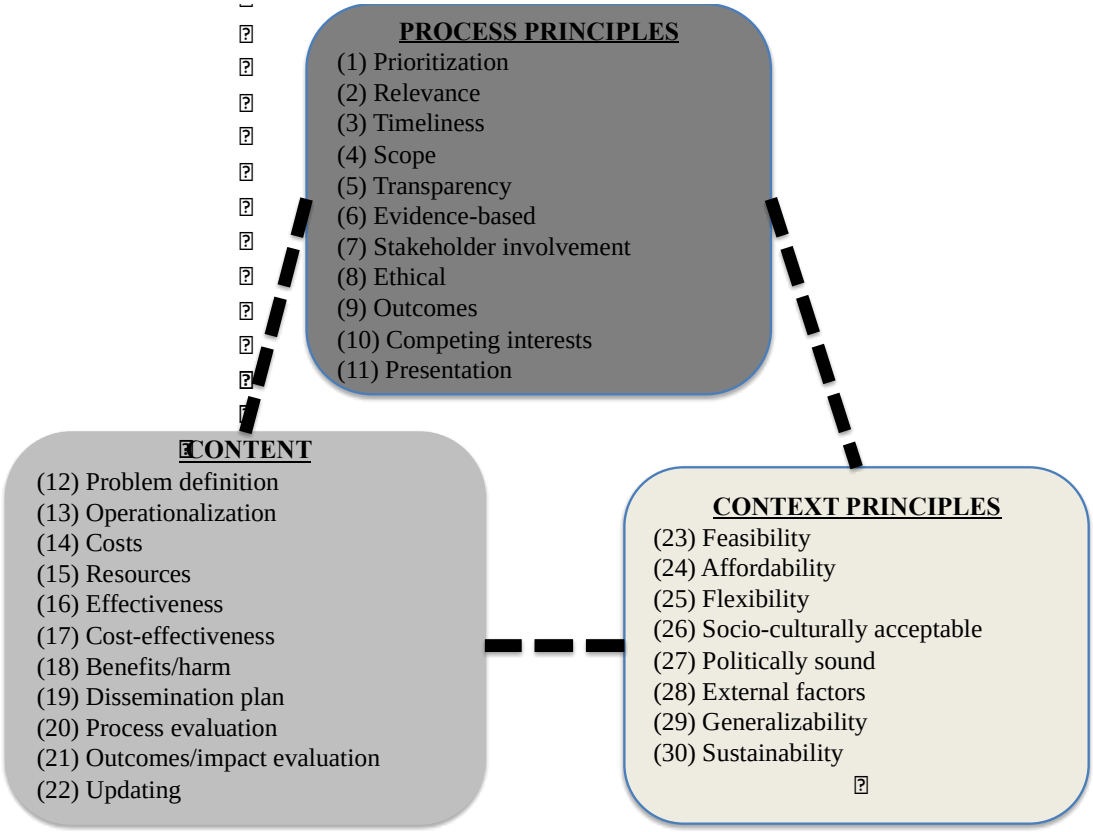
Context principles represent those systemic factors like the local technical, situational, structural, institutional political, and socio-cultural components of the health policy environment that can impact HSG recommendations (Walt & Gilson, 1994; Buse et al, 2005). Like health policymaking, HSG does not occur in a vacuum so it is important to pay attention to the variety of contextual factors that may have significant impact on how the guidance is developed, adapted and implemented by the end users (Collins, 1999; WHO, 2007). The process principles and the content of the HSG all have to be contextualized, therefore, context principles impact these other clusters of the concepts. Context principles refer to the usability in context and here we find concepts related to those system components that make up the setting within which the HSG is to be used. A clear overview of the context or setting that will be impacted by the guidance recommendations is essential. This will provide an understanding of why some guidance recommendations may work in some settings and not others. We find here factors that will enhance and facilitate the adherence to HSG recommendations as per protocol. We also find here factors that represent the values and moral fabric of the society and how this can either facilitate or impede the HSG recommendations. Table 4 below shows concepts for HSG context principles.

Table 3: Context principles	
23. Feasibility	The guidance recommendations are realistic and the actions are pragmatic. The guidance describes facilitators and barriers for implementation. It is clearly demonstrated that the implementation of the guidance is feasible within the proposed practice environment, and the recommendations match local capacities and expectations.
24. Affordability	The guidance recommendations are affordable within the financial structure and budgetary allocations of the health system. Potential sources of local government funding and donor organizations are identified. For policy issues in which there may be several sources of funding, the guidance also considers the level of coordination among the donors and between the donors and the local government.
25. Flexibility	The guidance is flexible and adaptable to the expertise of the user and the varying local conditions. It acknowledges the importance of professional judgment and discretion and provides recommendations that users can adapt in accordance with their own individual circumstances and needs. The recommendations steer away from the adoption of rigid approaches so as not to inappropriately or unnecessarily limit those in charge of applying them.
26. Socio-culturally acceptable	Considering the diversity of values in many regions, the recommendations are robust under societal and cultural scrutiny by adopting a socio-cultural perspective. It recognizes socio-cultural expectations and provides an understanding of the role that socio-cultural factors will play in the success of the guidance recommendations.
27. Politically sound	The political acceptability of the recommendations is considered in order to assess if they align with political interests/commitments. Implementation of guidance can stir swings in the national mood, lead to changes in the balance of organized forces, such as interest groups, or influence outcome of events within the government, for instance an election. Therefore options proposed that are in sync with the political

	climate may garner adequate support from top policy/government officials.
28. External factors	Determinants of health system performance that lie outside the formal architecture of the health system but will influence the performance of its functions are considered (for example, judicial system, social system, recession, corruption, state of the economy etc.). These are non-health system factors originating from other local institutional organizations that impact on the usual operations of the health system.
29. Generalizability	The recommendations are transferable to other settings with similar health system features (countries or regions); judgments are made about the applicability of the recommendations beyond its original context (setting or population) to ensure that contexts with similar institutional, socio-economic, and political demographics facing an identical health system challenges can adapt and use the guidance.
30. Sustainability	The guidance provides an indication of the sustainability of the effects of the recommendations to show that long-term outcomes can be continuously achieved and maintained at an acceptable level. Due to constantly evolving health system issues, looming budget cuts, fluctuating resources, rising costs of new technologies, an ageing population, shifting burdens of diseases etc., it is crucial to develop recommendations that will stand the test of time.

We were able to point out patterns and draw linkages between the context principles concepts as well. We offer some examples here. For HSG to be feasible and sustainable, it should be affordable, and resonate with local values (political and socio-cultural). Information on affordability, feasibility and sustainability may also determine whether the guidance recommendations can be transferable to other settings. Providing socio-culturally appropriate recommendations may also indicate which other health systems can adapt and use the HSG. Also, the HSG should be flexible enough to accommodate shifts in values (e.g., political, socio-cultural). Additionally, HSG that reflects the socio-cultural preferences of the target population, that is affordable, feasible, sustainable, and transferable to comparable settings, can amass support from politicians. The external factors that originate from other institutional systems may impact the feasibility and sustainability of the guidance, and may be influenced by the political and socio-cultural climate.

3.2. *Figure 2: Framework of health systems guidance concepts*



As is expected of a Critical Interpretive Synthesis (CIS), we were able to point out relationships between the concepts across the 3 clusters as shown in figure 2 above. We highlight some relationships below:

- Uptake of guidance can be enhanced if it is addressing a priority area for which evidence based reports of effectiveness and cost-effectiveness exist.
- Consultations with appropriate stakeholders are also crucial for feasibility of implementation and sustainability of the HSG because sometimes the individuals tasked with implementing guidance recommendations may not be committed to them and this may influence adherence. Their input into the HSG process and support for the recommendations may alleviate this concern.
- HSG outcome(s) chosen will influence the way the problem is defined and provides information that will be useful for evaluating the HSG.
- Including an ethical viewpoint will also impact how the problem is defined, and will influence cost effectiveness thresholds.
- Systematic and transparent processes may provide an impetus for donor involvement (affordability).
- Information on affordability can determine whether the health system issues can be addressed in a timely manner, and inform judgments on how to evaluate the process.
- The HSG should be flexible enough to accommodate constantly evolving evidence and changing health system priorities.
- The guidance should provide socio-culturally appropriate solutions that are relevant to the applicable levels/sectors of the health system.

- HSG that is in line with ethical principles, addresses priority issues, and is timely will be appealing to politicians.
- The external alignment of guidance may affect the effectiveness and impact the operational considerations.
- The outcomes selected, the system level/sectors involved or the ethical values in place may expose the external factors that may be pertinent.
- Some determinants of generalizability of the guidance are the stakeholders involved, the evidence used, transparency of the process, ethical lens considered and outcome/indicators selected

4. Discussion

This project is the first phase of a multistage approach to create an internationally useful HSG tool, AGREE for Health Systems (AGREE-HS) that will inform the development, reporting and appraisal of HSG. Modeling after the paradigm to create a tool for clinical practice guidelines (i.e., the AGREE II developed by Brouwers et al, 2010), our first step was to conduct a review of the published and grey literatures to identify concepts related to HSG quality. In this vein, it was our expectation that the receptiveness, adoption and diffusion of HSG recommendations depend on the perception of their quality, and with this study, we aimed to identify those core components of good quality HSG.

We found a total of 30 potential HSG appraisal concepts that have the capacity to discriminate between high and low quality guidance and direct their development and reporting. We found no existing tools to support HSG appraisal and found few studies describing concepts that were directly tested to appraise HSG. Indeed, the papers we examined reflected a variety of study designs and goals; none reflected methods used to develop appraisal methods. However, the data from our studies show a convergence of ideas in the HSG research community about what constitutes good and useable HSG. Together, these data can provide the foundation of a tool that guides HSG developers in the types of information important to report in a HSG document and tactics for optimal execution. These concepts also address issues of appropriateness and completeness as well as information, which are important components in the uptake of guidance recommendations.

One strength of this study is that we used a sound knowledge synthesis strategy, the critical interpretive synthesis (CIS) approach. We acknowledged the paucity of data related to HSG appraisal, as well as the diversity in the literature sources and types of available data. Also, given that HSG is not a highly bounded topic, we didn't rely on a narrow initial research question, instead refining it as the review progressed. The various methodological stages did not proceed as discrete entities, as there was a constant to-ing and fro-ing throughout the review and synthesis process; at times we were concurrently searching, sampling, critiquing and analyzing. The CIS approach is both systematic and iterative with an interpretive approach to analysis and synthesis of data that allowed us to capture and critically analyze an in-depth depiction of how to differentiate between HSG on the basis of quality, among other considerations. Through CIS methodology, we were able to further identify and include additional papers that were not directly related to HSG, but made valuable theoretical contributions to the process of appraising guidance (Dixon-wood et al 2005; Dixon-Wood et al, 2006).

A weakness of our study is in the nature of the CIS strategy, which exposes us to the risks that certain concepts may have been missed (calling saturation too soon, for example). If we have missed important elements, we anticipate they will be identified in the second stage of our research. Also, the health policy analysis triangle framework is made up of 4 elements; process, content, context and actors. But for our analysis, the original framework was modified as we included "actors" in the process principles. Actors (the "who" of policy) refer to local, national or international individuals or groups (governmental and non-governmental) involved in the policy process (Walt & Gilson, 1994; Walt et al, 2008). We found that "actors" could be incorporated under the process principles concept "participatory". Additionally and at first glance, it appears that there is the risk that some concepts may lead to potential contradictions. For example, in some jurisdictions, achieving an ethical HSG may not be aligned with achieving a feasible HSG (e.g., in contexts where there is a bias toward citizens because of race or sexual orientation). Currently, the data provides no guidance on how to reconcile a situation like this.

In our analysis, we showed the role that the process principles, content and context principles plays in the development, appraisal and reporting of HSG and demonstrated the link and resonance within and between their various concepts. We pointed out patterns and drew linkages between the concepts in order to show that the concepts and categories do not typically occur in a linear, independent or discrete manner. The interaction between these three clusters is an imperative consideration because they all influence the guidance process and can facilitate or impede the success of HSG recommendations.

Some of the concepts (criteria or items) and categories identified in this review may not be applicable to every jurisdiction or country so they are not intended to serve as a blueprint for all health systems to strictly apply. Guidance documents are administrative instruments that provide recommendations and implementation options typically in a step-by-step format, but do not have force of law and as such, allow for accommodations in approach (WHO, 2003; Health Canada, 2011). They are intended to provide recommendations on how to comply with governing statutes and regulations and assist staff and managers on how institutional mandates and objectives should be implemented in a manner that is fair, consistent and effective (Health Canada, 2011). Therefore, development of high quality HSG will impact the type of recommendations being formulated, the degree to which they get implemented, the methods of dissemination, and the extent to which they impact on the usual operations of the health system (WHO, 2003).

Different players in the HSG process play different roles at different times and under different circumstances. Developers of guidance may only be looking at process principles and content, while end users in the field may focus more on usability of the guidance in their individual contexts. Thus, for example, to optimize this dual perspective and facilitate the division of labor, there is interest at WHO to produce workbooks where the global guidance they develop (process principles and content), will be complemented by having a companion workbook for those at the receiving end (context principles). In this case, the role of the WHO will be the development of global guidance while the country's role (or health system's role) will be to take this global guidance and then apply it to their local context. This makes our framework of health systems guidance concepts quite practical.

Our study has elicited 30 unique concepts. It is unclear the extent to which HSG

stakeholders will view all concepts as equally important to our proposed tool, or, moreover, given the potential roles of the tool (development, appraisal, reporting), if there are varying levels of priority as a function of purpose. The next phase (stage 2) of our research program is aimed to create the beta version of the AGREE-HS by having international stakeholders prioritize these concepts and using those results to populate the tool. Our final phase (stage 3) will involve the usability testing of the beta version of the tool.

5. References

1. Abellán Perpiñán, J. M., Sánchez Martínez, F. I., & Martínez Pérez, J. E. (2009). Quality assessment of economic evaluations in health care: a checklist and user guide. *Revista Española de Salud Pública*, *83*(1), 71-84.
2. American Psychological Association. (2002) Criteria for practice guideline development and evaluation. American Psychologist, 57, 1048–1051.
3. Annandale, E., Harvey, J., Cavers, D., & Dixon-Woods, M. (2007). Gender and access to healthcare in the UK: a critical interpretive synthesis of the literature. *Evidence & Policy: A Journal of Research, Debate and Practice*, *3*(4), 463-486.
4. Arah, O. A., Klazinga, N. S., Delnoij, D. M. J., Ten Asbroek, A. H. A., & Custers, T. (2003). Conceptual frameworks for health systems performance: a quest for effectiveness, quality, and improvement. *International Journal for Quality in Health Care*, *15*(5), 377-398.
5. Arah, O. A., Westert, G. P., Hurst, J., & Klazinga, N. S. (2006). A conceptual framework for the OECD health care quality indicators project. *International Journal for Quality in Health Care*, *18*(suppl 1), 5-13.
6. Atun, R. (2012). Health systems, systems thinking and innovation. *Health policy and planning*, *27*(suppl 4), iv4-iv8.
7. Barnett-Page, E., & Thomas, J. (2009). Methods for the synthesis of qualitative research: a critical review. *BMC medical research methodology*, *9*(1), 59.
8. Bosch-Capblanch X (2010). Analytical framework for health systems guidance. Handbook for developing Health Systems Guidance; Swiss TPH.
9. Bosch-Capblanch X (2011) Handbook for Supporting the Development of Health System Guidance. Basel: Swiss Centre for International Health.
10. Bosch-Capblanch X, Lavis JN, Lewin S, Atun R, Røttingen JA, et al. (2012) Guidance for evidence-informed decisions about health systems: rationale for and challenges of guidance development. PLoS Med 9: e1001185.
11. Bosch-Capblanch X., Allen, C (2012). Health Systems Strengthening and Conflict: Transformation in Fragile States. Retrieved on August 17th 2013 from http://www.evidenceaid.org/wp-content/uploads/2013/05/MMI_for_Evidence_Aid_XBC_revised_by_MC.pdf
12. Brettle, A. (2003). "Information skills training: a systematic review of the literature." Health Info Libr J 20 Suppl 1: 3-9.
13. Brouwers et al (2009). Appraisal of Guidelines for Research and Evaluation. AGREE II Instrument. The AGREE Next Steps Consortium. [Accessed: 2013 Aug 23]. Available from: http://www.agreetrust.org/resource-centre/agree-ii/ - In the text, change AGREE Next steps, 2009 to Brouwers et al, 2009)
14. Brouwers MC, Kho ME, Browman GP, Burgers JS, Cluzeau F, Feder G, Fervers B, Graham ID, Grimshaw J, Hanna SE, Littlejohns P, Makarski J, Zitzelsberger L, for the AGREE Next Steps Consortium. AGREE II: Advancing guideline development, reporting and evaluation in health care. CMAJ 2010. DOI:10.1503/cmaj.090449.
15. Bryce, J., Requejo, J. H., Moulton, L. H., Ram, M., & Black, R. E. (2013). A common evaluation framework for the African Health Initiative. *BMC health services research*, *13* (Suppl 2), S10.
16. Burnard P(1991)A method of analysing transcripts in qualitative research. *Nurse Education Today*, Vol.11, 461-466.
17. Buse, K., Mays, N. and Walt, G. (2005) *Making Health Policy*, Berkshire: Open University Press.
18. Chopra M, Lawn J, Sanders D, Barron P (2009). Achieving the health millennium development goals for South Africa: challenges and priorities. Lancet 2009,

374:1023-1031.
19. Collins, Charles; Andrew Green; David Hunter. 1999. Health sector reform and the interpretation of policy context. Health Policy and Planning 47:69-83.
20. Davies, E., & Littlejohns, P. (2002). Views of Directors of Public Health about NICE Appraisal Guidance: results of a postal survey. *Journal of Public Health*, *24*(4), 319-325.
21. Dickersin, K., R. Scherer, et al. (1994). "Identifying relevant studies for systematic reviews." BMJ 309(6964): 1286-91.
22. Dixon-Woods, M., Bonas, S., Booth, A., Jones, D. R., Miller, T., Sutton, A. J., ... & Young, B. (2006b). How can systematic reviews incorporate qualitative research? A critical perspective. *Qualitative research*, *6*(1), 27-44.
23. Dixon-Woods, M., Cavers, D., Agarwal, S., Annandale, E., Arthur, A., Harvey, J., … Sutton, A. J. (2006a). Conducting a critical interpretive synthesis of the literature on access to healthcare by vulnerable groups. *BMC Medical Research Methodology*, *6*, 35. doi:10.1186/1471-2288-6-35
24. Dixon-Woods, M., Kirk, M. D., Agarwal, M. S., Annandale, E., Arthur, T., Harvey, J., ... & Dixon-Woods, M. (2005b). Vulnerable groups and access to health care: a critical interpretive review. *National Coordinating Centre NHS Service Delivery Organ RD (NCCSDO).*
25. Dixon-Woods, M., S. Agarwal, D. Jones, B. Young, and A. Sutton (2005a). Synthesising Qualitative and Quantitative Evidence: A Review of Possible Methods. Journal of Health Services Research Policy 10(1):45–53.
26. Dixon-Woods, M., Shaw, R., Agarwal, S. and Smith, J. 2004. The problem of appraising qualitative research. Quality and Safety in Health Care 13: 223-225
27. Eakin JM, Mykhalovskiy E: Reframing the evaluation of qualitative health research: reflections on a review of appraisal guidelines in the health sciences. J Eval Clin Pract 2003, 9:187-194.
28. Entwistle, V., Firnigl, D., Ryan, M., Francis, J., & Kinghorn, P. (2012). Which experiences of health care delivery matter to service users and why? A critical interpretive synthesis and conceptual map. *Journal of health services research & policy*, *17*(2), 70-78.
29. Etienne, C., Asamoa-Baah, A., & Evans, D. B. (2010). *Health systems financing: The path to universal coverage*. World Health Organization.
30. Flemming, K. (2010). Synthesis of quantitative and qualitative research: an example using Critical Interpretive Synthesis. *Journal of advanced nursing*, *66*(1), 201-217.
31. Fryatt R, Mills A, Nordstrom A (2010) Financing of health systems to achieve the health Millennium Development Goals in low-income countries. Lancet 375: 419–426.
32. Funk M., Drew N., Faydi E (n.d). Checklist for Evaluating a Mental Health Policy. Retrieved August 19th 2013 from http://www.who.int/mental_health/policy/WHOPolicyChecklist_forwebsite.pdf
33. German BACKUP Initiative - GIZ - (2011). Guidance on integrating gender-specific issues into Health Systems Strengthening activities. Retrieved on October 4th 2013 from http://www.genderingermandevelopment.net/files/images/pdf/Gender_Mainstreaming/giz2011-en-gender-dimensions-in-health-systems.pdf
34. German, R. R., Lee, L. M., Horan, J. M., Milstein, R. L., Pertowski, C. A., & Waller, M. N. (2001). Updated guidelines for evaluating public health surveillance systems: recommendations from the Guidelines Working Group. *MMWR. Recommendations and reports: Morbidity and mortality weekly report. Recommendations and reports/Centers for Disease Control*, (50), 1-35.

35. Glaser, B. G., Strauss, A. L., & Strutzel, E. (1968). The discovery of grounded theory; strategies for qualitative research. *Nursing Research*, *17*(4), 364.
36. Gough D. (2007) Weight of evidence: a framework for the appraisal of the quality and relevance of evidence. In Applied and Practice-based Research. Edited by Furlong J, Oancea A. Special Edition of Research Papers in Education, 22 (2): 213-228.
37. Graneheim UH & Lundman B (2004). Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. Nurse Education Today 24, 105–112
38. Gysels, M. H., Evans, C., & Higginson, I. J. (2012). Patient, caregiver, health professional and researcher views and experiences of participating in research at the end of life: a critical interpretive synthesis of the literature. *BMC medical research methodology*, *12*(1), 123.
39. Health Canada (2009). Recognition and Use of Standards under the Medical Devices Regulations. Retrieved on October 1st, 2013 from http://www.hc-sc.gc.ca/dhp-mps/md-im/applic-demande/guide-ld/md_gd_standards_im_ld_normes-eng.php
40. Heaton J, Corden A, Parker G. 'Continuity of care': a critical interpretive synthesis of how the concept was elaborated by a national research programme. International Journal of Integrated Care 2012;12
41. Hoffman S, Rottingen J-A, Bennett S, Lavis J, Edge J, Frenk J (2012) *Background Paper on Conceptual Issues Related to Health Systems Research to Inform a WHO Global Strategy on Health Systems Research - a working paper in progress*. Geneva: World Health Organisation; 2013.
42. Hoffman S., Lavis J., Bennett S (2009). The Use of Research Evidence in Two International Organizations' Recommendations about Health Systems. *Healthcare Policy* Vol.5 No.1: 66–86.
43. International Centre for Allied Health Evidence - iCAHE - (n.d). The iCAHE guideline quality checklist. Retrieved on August 19th 2013 from http://www.biomedcentral.com/content/supplementary/1471-2288-14-63-s1.pdf
44. Islam, M. (2007). Health systems assessment approach: a how-to manual. Retrieved on August 17th 2013 from http://www.urc-chs.com/uploads/resourcefiles/Health_Systems_Assessment.pdf
45. Kazimierczak, K. A., Skea, Z. C., Dixon-Woods, M., Entwistle, V. A., Feldman-Stewart, D., N'dow, J. M., & MacLennan, S. J. (2013). Provision of cancer information as a "support for navigating the knowledge landscape": Findings from a critical interpretive literature synthesis. *European Journal of Oncology Nursing*, *17*(3), 360-369.
46. Lavis JN, Røttingen JA, Bosch-Capblanch X, Atun R, El-Jardali F, et al (2012). Guidance for evidence-informed policies about health systems: linking guidance development to policy development. PLoS Med ;9: e1001186
47. Lewin S, Bosch-Capblanch X, Oliver S, et al (2012). Guidance for evidence-informed policies about health systems: assessing how much confidence to place in the research evidence. PLoS Med. 2012; 9(3): e1001187
48. Lomas, J., Culyer, T., Mccutcheon, C., Law, S., & Tetroe, J. (2005). Conceptualizing and Combining Evidence for Health System Guidance. Retrieved on October 4th 2013 from http://www.cfhi-fcass.ca/migrated/pdf/insightAction/evidence_e.pdf
49. Markoulakis, R., & Kirsh, B. (2013). Difficulties for university students with mental health problems: A critical interpretive synthesis. *The Review of Higher Education*, *37*(1), 77-100.
50. Moga C, Guo B, Schopflocher D, Harstall C (2012). Development of a quality appraisal tool for case series studies using a modified delphi technique. Edmonton:

Institute of Health Economics (IHE)
51. Moher, D., Schulz, K. F., Simera, I., & Altman, D. G. (2010). Guidance for developers of health research reporting guidelines. *PLoS medicine*, *7*(2): e1000217.
52. Morrison, L. G., Yardley, L., Powell, J., & Michie, S. (2012). What design features are used in effective e-health interventions? A review using techniques from Critical Interpretive Synthesis. *Telemedicine and e-Health*, *18*(2), 137-144.
53. Murray C, Frenk J. A WHO framework for health system performance assessment. Bull World Health Organ2000; 78:717–31
54. Murray CJ, *Evans* DB (2003). Health System Performance Assessment: Debates, Methods, and Empiricism. Geneva: World Health Organization. Retrieved on August 17th 2013 from http://whqlibdoc.who.int/publications/2003/9241562455.pdf
55. Murray, C. J., & Frenk, J. (1999). *A WHO framework for health system performance assessment*. Evidence and Information for Policy, World Health Organization.
56. National Health and Medical Research Council - NHMRC - (2007). Standards and procedures for externally developed guidelines. Retrieved on August 16th 2013 from https://www.nhmrc.gov.au/_files_nhmrc/publications/attachments/nh56_nhmrc_standards_procedures_externally_developed_guidelines_131223.pdf
57. National Health Services - NHS - (2009a). How to use NICE guidance to commission high quality services. Retrieved October 4th 2013 from http://www.gserve.nice.org.uk/media/027/17/HowToCommissioningGuideFinal.pdf
58. National Health Services - NHS - (2009b). Methods for the development of NICE public health guidance (second edition). Retrieved October 4th 2013 from http://www.gserve.nice.org.uk/media/CE1/F7/CPHE_Methods_manual_LR.pdf
59. Noblit, George W. & Hare, R. Dwight (1988). Meta-ethnography:synthesising qualitative studies. London: Sage.
60. Orem, J. N., Wavamunno, J. B., Bakeera, S. K., & Criel, B. (2012). Do guidelines influence the implementation of health programs?–Uganda's experience. *Implementation Science*, *7*(1), 1-16.
61. Ostrom E (2007) Institutional rational choice: an assessment of the institutional analysis and development framework. In: Sabatier PA, editor. Theories of the policy process. 2nd edition. Boulder, CO: Westview Press; p. 21-64
62. Øvretveit J., Klazinga Niek (2008). Guidance on developing quality and safety strategies with a health system approach. Retrieved on August 17th 2013 from http://www.euro.who.int/__data/assets/pdf_file/0011/96473/E91317.pdf
63. Oxman, A. D., Fretheim, A., & Schünemann, H. J. (2006). Improving the use of research evidence in guideline development: 14. Reporting guidelines. *Health Res Policy Syst*, *4*(26.2006).
64. Oxman, A. D., Lavis, J. N., Lewin, S., & Fretheim, A. (2009). SUPPORT Tools for evidence-informed health Policymaking (STP) 1: What is evidence-informed policymaking?. *Health Research Policy and Systems*, *7*(Suppl 1), S1.
65. Peters, D. H., & Bennett, S. (2012). Better guidance is welcome, but without blinders. PLoS medicine, 9(3), e1001188.
66. Phillips, A. C., Lewis, L. K., McEvoy, M. P., Galipeau, J., Glasziou, P., Hammick, M., ... & Williams, M. T. (2013). Protocol for development of the guideline for reporting evidence based practice educational interventions and teaching (GREET) statement. *BMC medical education*, *13*(1), 9.
67. Remme JHF, Adam T, Becerra-Posada F, D'Arcangues C, Devlin M, et al. (2010) Defining Research to Improve Health Systems. PLoS Med 7(11): e1001000
68. Roemer M. I. (1991) National Health Systems of the World, Vol. 1: The Countries. New York: Oxford University Press; 1991. p 97.
69. Sandelowski, M., Docherty, S. and Emden, C. 1997. Focus on qualitative methods.

Qualitative meta-synthesis: issues and techniques. Research in Nursing and Health 20: 365-371

70. Savigny, D., & Adam, T. (Eds.). (2009). Systems thinking for health systems strengthening. *World Health Organization.*

71. Schünemann, H. J., Fretheim, A., Oxman, A. D., & WHO Advisory Committee on Health Research. (2006). Improving the use of research evidence in guideline development: 1. Guidelines for guidelines. *Health Res Policy Syst*, *4*(13), 1-6.
72. Seale, C. (1999) The quality of qualitative research, London: Sage Publications.
73. Shaw, C. D., & Kalo, I. (2002). A background for national quality policies in health systems. Retrieved on August 19th 2013 from http://apps.who.int/iris/handle/10665/107458
74. Sheikh, K., Gilson, L., Agyepong, I. A., Hanson, K., Ssengooba, F., & Bennett, S. (2011). Building the field of health policy and systems research: framing the questions. *PLoS Medicine*, *8*(8), e1001073.
75. Strauss, A., & Corbin, J. (1998). Basics of qualitative research: Grounded theory procedures and techniques (2nd ed.). Newbury Park, CA: Sage.
76. SUPPORT (2012) Supporting Policy relevant Reviews and Trials. Available: http://www. support-collaboration.org/index.htm. Accessed 30 January 2013
77. Swanson, R. C., Bongiovanni, A., Bradley, E., Murugan, V., Sundewall, J., Betigeri, A., ... & Labonté, R. (2010). Toward a consensus on guiding principles for health systems strengthening. *PLoS medicine*, *7*(12): e1000385
78. Talseth, A. G., & Gilje, F. L. (2011). Nurses' responses to suicide and suicidal patients: a critical interpretive synthesis*. *Journal of clinical nursing*, *20*(11- 12), 1651-1667.
79. Taylor, T. L., Killaspy, H., Wright, C., Turton, P., White, S., Kallert, T. W., ... & King, M. B. (2009). A systematic review of the international published literature relating to quality of institutional care for people with longer term mental health problems. *Bmc Psychiatry*, *9*(1), 55.
80. Travis P, Bennett S, Haines A, Pang T, Bhutta Z, Hyder A, Pielemeier N, Mills A, Evans T. Overcoming health-systems constraints to achieve the Millennium Development Goals. Lancet 2004; 364: 900-906"
81. Travis, P., Bennett, S., Haines, A., Pang, T., Bhutta, Z., Hyder, A. A., ... & Evans, T. (2004). Overcoming health-systems constraints to achieve the Millennium Development Goals. The Lancet, 364(9437), 900-906.
82. Van der Sluijs, J. P., Petersen, A. C., Janssen, P. H., Risbey, J. S., & Ravetz, J. R. (2008). Exploring the quality of evidence for complex and contested policy decisions. *Environmental Research Letters*, *3*(2), 024008.
83. Walt G, Gibson L (1994). Reforming the health sector in developing countries: The central role of policy analysis. *Health Policy and Planning*, 9: 353-70.
84. Walt, G., Shiffman, J., Schneider, H., Murray, S. F., Brugha, R., & Gilson, L. (2008). 'Doing'health policy analysis: methodological and conceptual reflections and challenges. *Health policy and planning*, *23*(5), 308-317.
85. WHO, World Health Organization. 2007. Sound Choices: Enhancing Capacity for Evidence-Informed Health Policy. Geneva.
86. Wilson, P. (2002). How to find the good and avoid the bad or ugly: a short guide to tools for rating quality of health information on the internet. *BMJ: British Medical Journal*, *324*(7337), 598.
87. World Health Organization - WHO - (2009). Practical guidance for scaling up health service innovations. Retrieved on October 5th 2013 from http://www.expandnet.net/PDFs/WHO_ExpandNet_Practical_Guide_published.pdf

88. World Health Organization - WHO - (2012). The Health Systems Assessment Approach: A How to Manual: Guidance on Assessing Health System Building Blocks. Retrieved on October 4th 2013 from http://www.healthsystemassessment.com/wp-content/uploads/2012/06/Section-3-Module-3-1-_HSAA-Manual-August-2012-11.pdf
89. World Health Organization (2000). The World Health Report 2000. Health systems: improving performance. Geneva, 2000. [accessed 9 July 2013] Available from: http://www.who.int/whr/2000/en/
90. World Health Organization (2005). Fifty-Eighth World Health Assembly, A58/22, 1-7. Geneva: Author; 2005. *WHA58.34: Ministerial Summit on Health Research*.
91. World Health Organization (2011) Health system strengthening: improving support to policy dialogue around national health policies, strategies and plans. Report by the Secretariat to the 64th World Health Assembly. A64/12. Geneva: World Health Organization. Retrieved on July 30th 2013 from: https://extranet.who.int/iris/restricted/bitstream/10665/3205/1/B128_8-en.pdf.
92. World Health Organization (2011). Health system strengthening: improving support to policy dialogue around national health policies, strategies and plans. Report by the Secretariat to the 64th World Health Assembly. A64/12. Geneva: World Health Organization. Retrieved on October 4th 2013 from http://apps.who.int/gb/ebwha/pdf_files/WHA64/A64_12-en.pdf.
93. World Health Organization. 2003. "Guidelines for WHO Guidelines." EIP/GPE/EqC/2003.1. Geneva: Author. Retrieved Jan 28, 2013. <http://whqlibdoc.who.int/hq/2003/EIP_GPE_ EqC_2003_1.pdf>.

Appendix I: Summary of selected papers				
Author(s)	**Title**	**Organization/ Location****	**Purpose of the study and relevance of its contribution**	**Concepts extracted**
Murray & Frenk, 2000	*A WHO framework for health system performance assessment*	WHO / Switzerland	This paper discusses how variations in health outcomes across different countries are related to differences in health system performance (design, content and management of health systems), and proposes a framework to assess and advance the understanding of health system performance	Timeliness; scope; evidence-based; stakeholder involvement; ethical; outcomes; operationalization; costs; feasibility; socio-culturally acceptable; politically sound; external factors
German et al, 2001	*Updated guidelines for evaluating public health surveillance systems*	CDC / USA	The paper aims to provide an operational framework and guidelines for evaluating the quality, efficiency and usefulness of a public health surveillance systems	Timeliness; scope; evidence-based; stakeholder involvement; problem definition; resources; outcomes/impact evaluation; feasibility; socio-culturally acceptable; generalizability; sustainability
Davies & Littlejohns, 2002	*Views of directors of public health about NICE appraisal guidance: results of a postal survey*	NICE / UK	The aim is to explore the view of Directors of Public Health with regards to the development, implementation and dissemination of appraisal guidance for health technologies within the UK health system	Prioritization; timeliness; scope; evidence-based; stakeholder involvement; presentation; feasibility; affordability; socio-culturally acceptable; politically sound
APA, 2002	*Criteria for practice guideline development and evaluation*	American Psycological Association / USA	The paper is designed to promote quality and consistency in practice guideline development and to describe the criteria by which practice guidelines are developed, evaluated and reviewed	Prioritization; relevance; transparency; evidence-based; stakeholder involvement; ethical; outcomes/impact evaluation; presentation; feasibility; flexibility; external factors
Shaw & Kalo, 2002	*A background for national quality policies in health systems*	WHO / Switzerland	The paper aims to outline some of the values, forms and concepts which affect national approaches for the improvement of quality as a central element for reform of health systems and health service delivery	Prioritization; transparency; evidence-based; stakeholder involvement; ethical; operationalization; resources; effectiveness; dissemination plan; affordability; socio-culturally

Appendix I: Summary of selected papers				
Author(s)	**Title**	**Organization/ Location****	**Purpose of the study and relevance of its contribution**	**Concepts extracted**
				acceptable; politically sound; external factors
Wilson, 2002	*How to find the good and avoid the bad or ugly: a short guide to tools for rating quality of health information on the internet*	European Commission / Belgium	The aim is to report methods, a set of criteria for good practice and tools for evaluating and rating the quality of health information	Ethical; feasibility
Arah et al, 2003	*Conceptual frameworks for health systems performance: a quest for effectiveness, quality, and improvement*	University of Amsterdam / Nederlands	The aims are to understand the underlying concepts of national and international performance frameworks for health systems (case studies: UK, Canada, Australia, the US, the WHO, and the OECD); to explore health system efficiency and performance indicators; and examine how and in what context the resultant performance data can be used to drive improvement	Prioritization; relevance; timeliness; scope; evidence-based; stakeholder involvement; ethical; outcomes; resources; effectiveness; outcomes/impact evaluation; feasibility; affordability; socio-culturally acceptable; politically sound; generalizability
Murray & Evans, 2003	*Health systems performance assessment: debates, methods and empiricism*	WHO / Switzerland	The aim is to explore the role that the WHO plays in providing advice to member states on how best to organize, manage and strengthen their health systems. It discusses how recommendations for clinical practice decisions differ from health system recommendations	Outcomes; evidence-based
Travis et al, 2004	*Overcoming health-systems constraints to achieve the millennium development goals*	WHO / Switzerland	The paper uses the Millennium Development Goals as a reference point to explore the advantages and disadvantages of approaches to health system strengthening through the lens of individual service or disease specific initiatives	Prioritization; evidence-based; stakeholder involvement; outcomes; operationalization; resources; cost-effectiveness; feasibility; affordability; external factors; generalizability; sustainability
Lomas et al, 2005	*Conceptualizing and combining*	Canadian Health	This review examines how guidance developers and policy-makers view evidence as well as how different forms of	Relevance; transparency; evidence-based; stakeholder

Appendix I: Summary of selected papers				
Author(s)	**Title**	**Organization/ Location****	**Purpose of the study and relevance of its contribution**	**Concepts extracted**
	evidence for health system guidance	Services Research Foundation / Canada	evidence can be combined to produce health system guidance.	involvement; outcomes; ethical; problem definition; effectiveness; feasibility; flexibility
Arah et al, 2006	*A conceptual framework for the OECD health care quality indicators project*	University of Amsterdam / Nederlands	The aim is to provide a sound conceptual framework that defines what is meant by quality of health care and to place it within a wider performance framework which acknowledges the key health policy goals adopted by the OECD and its member countries as they formally assess and 'incentivize' the performance of their health care systems	Relevance; timeliness; stakeholder involvement; ethical; problem definition; cost; effectiveness; feasibility; affordability; socio-culturally acceptable; politically sound; external factors; sustainability
Oxman et al, 2006	*Improving the use of research evidence in guideline development: reporting guidelines*	Norwegian Knowledge Centre for the Health Services / Norway	The aim is explore the standard formats for wide variety of WHO policies, recommendations or guidelines, and how these recommendations should be formulated and reported. It emphasizes that the information needed to judge the quality of guidance, determine its applicability and adaptability should be reported	Relevance; timeliness; transparency; evidence-based; ethical; presentation; outcomes; problem definition; operationalization; cost-effectiveness; outcomes/impact evaluation; dissemination plan; affordability; generalizability
Schünemann et al, 2006	*Improving the use of research evidence in guideline development: 1. guidelines for guidelines*	McMaster University / Canada	This report from the WHO Advisory Committee on Health Research is aimed at providing advice to the WHO on the use of more rigorous processes to ensure that the best available research evidence informs health care recommendations.	Prioritization; evidence-based; stakeholder involvement; ethical; outcomes; competing interests; presentation; problem definition; operationalization; costs; benefits/harm; process evaluation; outcomes/impact evaluation; updating; flexibility; generalizability
Islam, 2007	*Health systems assessment approach: a how-to manual*	USAID / USA	The aim of this report is to enable USAID Missions to assess a country's health system during early phases of program development or sector planning. Using a performance indicator-based and health indices approach, the report is designed to	Prioritization; relevance; timeliness; scope; evidence-based; stakeholder involvement; ethical; resources; effectiveness;

Appendix I: Summary of selected papers				
Author(s)	**Title**	**Organization/ Location****	**Purpose of the study and relevance of its contribution**	**Concepts extracted**
			provide a rapid and comprehensive assessment of key health systems functions (Governance, Health financing, Health service delivery, Human resources, Pharmaceutical management, Health information systems) and propose recommendations for improvement	affordability; sustainability
National Health and Medical Research Council (NHMRC), 2007	*Standards and procedures for externally developed guidelines*	Australian Government / Australia	The aims is to inform external persons and organizations of the procedures and key steps to be followed in developing, implementing and evaluating guidelines that are intended for submission to the NHMRC for approval	Prioritization; relevance; timeliness; evidence-based; stakeholder involvement; outcomes; presentation; problem definition; resources; effectiveness; cost-effectiveness; outcomes/impact evaluation; dissemination plan; feasibility; flexibility; socio-culturally acceptable; politically sound
Øvretveit & Klazinga, 2008	*Guidance on developing quality and safety strategies with a health system approach*	WHO / Denmark	The aim is to provide tools and approaches to help national policy advisers and policy-makers to create and implement a national quality strategy, drawing attention to the need for sustainable longer term public health measures in order to improve their health systems, and engage member states in a constructive dialogue	Prioritization; relevance; timeliness; scope; transparency; evidence-based; stakeholder involvement; ethical; presentation; operationalization; resources; process evaluation; outcomes/impact evaluation; feasibility; affordability; socio-culturally acceptable; politically sound; sustainability
Van der Sluijs et al,, 2008	*Exploring the quality of evidence for complex and contested policy decisions*	Utrecht University / The Nederlands	The aim is to provide a deeper understanding and increased awareness of the phenomenon of uncertainty and its policy implications, by discussing some key quality aspects of knowledge production and use especially in complex policy issues	Relevance; timeliness; evidence-based; stakeholder involvement; outcomes; problem definition; process evaluation; outcomes/impact evaluation; socio-culturally acceptable
Hoffman et al,	*The use of research*	McMaster	This study systematically compares health systems	Prioritization; relevance;

Appendix I: Summary of selected papers				
Author(s)	**Title**	**Organization/ Location****	**Purpose of the study and relevance of its contribution**	**Concepts extracted**
2009	*evidence in two international organizations' recommendations about health systems*	University / Canada	recommendations by international organizations (WHO and the World Bank) to the research evidence that was available at the time of their formulation. These recommendations about health systems (on technical guidance for example) have the potential to link research to action by acting as mediators between the best available research evidence and policy options.	transparency; evidence-based; stakeholder involvement; ethical; outcomes/impact evaluation; feasibility; affordability; politically sound
NHS, 2009a	*How to use NICE guidance to commission high-quality services*	NHS / UK	The guide is aimed for people involved in commissioning health and social care services and public health programs in the UK and provides guidance that can help to support the commissioning of new productive, efficient and high quality services, provides an implementation tool for planning and prioritizing services, and a framework for the evaluation or redesign of existing services and the decommissioning of ineffective interventions	Prioritization; relevance; scope; evidence-based; stakeholder involvement; ethical; presentation; problem definition; operationalization; cost-effectiveness; outcomes/impact evaluation; generalizability; sustainability
NHS, 2009b	*Methods for the development of NICE public health guidance*	NHS / UK	The paper describes the philosophical and methodological principles which govern the production of guidance for public health practice by NICE and the key components involved	Relevance; stakeholder involvement; ethical; outcomes; presentation; problem definition; generalizability
Oxman et al, 2009	*SUPPORT tools for evidence-informed health policymaking: what is evidence-informed policymaking?*	Norwegian Knowledge Centre for the Health Services / Norway	This article focuses on how to better use research evidence (what constitutes evidence?, what is its role in health policy?) to inform decisions about how best to organize health systems, including arrangements for delivering, financing and governing health services, and strategies for bringing about change	Prioritization; transparency; evidence-based; stakeholder involvement; ethical; outcomes; benefits/harms; politically sound
Perpiñán et al, 2009	*Quality assessment of economic evaluations in health care: a checklist and user guide*	Murcia University / Spain	The aim is to promote the efficiency in the process of incorporating new health technologies, as well as to guide their implementation in health systems by reporting an instrument composed of a user guide and a 12 criteria checklist in which a score is assigned to each items	Problem definition; cost; feasibility; socio-culturally acceptable; politically sound
Savigny &	*Systems thinking for*	Alliance for	This report offers a practical systems thinking approach to	Scope; relevance; transparency;

Appendix I: Summary of selected papers				
Author(s)	**Title**	**Organization/ Location****	**Purpose of the study and relevance of its contribution**	**Concepts extracted**
Adam, 2009	*health systems strengthening*	Health Policy and Systems Research, WHO	decipher the complexity of health systems, identify health systems challenges, and then applies that understanding to design better interventions to strengthen health systems and improve health	evidence-based; stakeholder involvement; ethical; outcomes; problem definition; cost; resources; effectiveness; cost-effectiveness; benefits/harms; process evaluation; outcomes/impact evaluation; feasibility; affordability; politically sound
WHO, 2009	*Practical guidance for scaling up health service innovations*	WHO / Switzerland	The aim is to identify general principles and make specific suggestions on the process of scaling up successfully tested health services innovations and discusses the strategic choices that facilitate and hinder the process	Timeliness; evidence-based; stakeholder involvement; ethical; operationalization; resources; effectiveness; outcomes/impact evaluation; feasibility; affordability; socio-culturally acceptable; politically sound; sustainability
Moher et al, 2010	*Guidance for developers of health research reporting guidelines*	Faculty of Medicine / University of Ottawa	The aim is to update and expand upon efforts to outline a strategy for developing reporting guidelines and shows that reporting guidelines is associated with improvements in the quality of reporting health research. An 18-step checklist on how to develop a reporting guideline is provided	Transparency; evidence-based; stakeholder involvement; presentation; problem definition; resources; dissemination plan
Swanson et al, 2010	*Toward a consensus on guiding principles for health systems strengthening*	Brigham Young University / USA	The paper proposes a list of ten guiding principles necessary for the development and communication of clear and consistent frameworks for policy, practice and evaluation with the overall goal of strengthening health system	Scope; transparency; evidence-based; stakeholder involvement; ethical; problem definition; operationalization; effectiveness; cost-effectiveness; outcomes/impact evaluation; affordability; socio-culturally acceptable; politically sound
Etienne et al, 2010	*Health systems financing: the path*	WHO / Switzerland	This report provides practical guidance on ways to finance health care by transforming available evidence based practices	Relevance; scope; evidence-based; ethical; outcomes;

Appendix I: Summary of selected papers				
Author(s)	**Title**	**Organization/ Location****	**Purpose of the study and relevance of its contribution**	**Concepts extracted**
	to universal coverage		into a menu of options for raising sufficient resources and removing financial barriers to access. Emphasis is placed on moving towards universal coverage to optimize health service provision.	feasibility; operationalization; flexibility; timeliness; cost; resources; effectiveness; cost-effectiveness; outcomes/impact evaluation; affordability; politically sound; external factors
German BACKUP Initiative (GIZ), 2011	*Guidance on integrating gender-specific issues into health systems strengthening activities*	Federal Ministry of Economic Co-operation and Development / Germany	The aim is to advise and assist organizations that are planning to apply to the German BACKUP Initiative for technical support on how to analyze and integrate gender-related issues into Health Systems Strengthening (HSS) activities. A checklist used to plan for technical support and develop applications that take into consideration specific gender dimensions in the different components of a health system is provided	Ethical; stakeholder involvement; socio-culturally acceptable; politically sound
Sheikh et al, 2011	*Building the field of health policy and systems research: framing the questions*	Public Health Foundation of India / India	This paper discusses the state-of-the-art in Health Policy and Systems Research (HPSR), addresses the current challenges and opportunities for the field and lays out what is needed to build capacity in HPSR and support local policy development and health systems strengthening	Prioritization; relevance; scope; stakeholder involvement; outcomes; problem definition; operationalization; feasibility; flexibility; socio-culturally acceptable; politically sound; generalizability
WHO, 2011	*Health system strengthening: improving support to policy dialogue around national health policies, strategies and plans*	WHO / Switzerland	This report reviews experiences with conducting and supporting policy dialogue for the development or renewal of comprehensive policies, strategies and plans to improve health service delivery, health outcomes and strengthen health systems	Prioritization; relevance; scope; evidence-based; stakeholder involvement; outcomes; operationalization; costs; resources; effectiveness; feasibility; affordability; politically sound; external factors; generalizability; sustainability
Atun, 2012	*Health systems, systems thinking and innovation*	Faculty of Medicine, London Imperial	The aim is to discusses factors that influence the achievement of health system performance and efficiency, and provide an understanding of why many well-intentioned policies and managerial decisions aimed at improving health systems do not	Scope; evidence-based; ethical; problem definition; operationalization; costs; resources; effectiveness;

Appendix I: Summary of selected papers				
Author(s)	**Title**	**Organization/ Location****	**Purpose of the study and relevance of its contribution**	**Concepts extracted**
		College / UK	achieve desired outcomes, but lead to unexpected or unintended consequence	feasibility; affordability; socio-culturally acceptable; politically sound; external factors
Bosch-Capblanch & Allen, 2012	*Health systems strengthening and conflict: transformation in fragile states*	Tropical and Public Health Institute / Switzerland	The aim is to inform programming, policy, advocacy, research and the civil society on the best ways to strengthen health systems. Main challenges encountered in producing methods to develop health systems guidance are discussed	Timeliness; evidence-based; costs; effectiveness; feasibility; politically sound
Bosch-Capblanch & al, 2012	*Guidance for evidence-informed policies about health systems: rationale for and challenges of guidance development*	Tropical and Public Health Institute / Switzerland	The aim is to assess extent to which the need for health systems guidance is part of national policies and plans and assess how guidance is currently formulated. Conceptual and the methodological challenges in the development and use of health systems guidance and ways to address them are discussed	Prioritization; relevance; timeliness; evidence-based; stakeholder involvement; ethical; outcomes; presentation; problem definition; cost-effectiveness; outcomes/impact evaluation; dissemination plan; feasibility; external factors
Lavis J et al, 2012	*Guidance for evidence-informed policies about health systems: linking guidance development to policy development*	McMaster University / Canada	The aim is to discuss the importance of contextual factors in shaping decisions about health systems and discusses the need to work through all the pros and cons of different options before adopting specific health systems guidance. It also shows the need for division of labor between national/global guidance & policy developers to support evidence- informed policymaking about health systems	Prioritization; relevance; scope; transparency; evidence-based; stakeholder involvement; problem definition; costs; effectiveness; outcomes/impact evaluation; feasibility; politically sound
Lewin et al, 2012	*Guidance for evidence-informed policies about health systems: assessing how much confidence to place in the research evidence*	Norwegian Knowledge Centre for the Health Services / Norway	The aim is to assess how much confidence to place in the types of evidence available on health systems interventions that inform judgments for health systems strengthening. The factors that are important when developing recommendations on policy options regarding health systems interventions are discussed	Relevance; scope; transparency; evidence-based; stakeholder involvement; ethical; operationalization; cost-effective; feasibility; affordability; socio-culturally acceptable; politically sound
Moga et al,	*Development of a*	Institute of	The aim is to outline the process of development of a checklist	Relevance; transparency;

Appendix I: Summary of selected papers				
Author(s)	**Title**	**Organization/ Location****	**Purpose of the study and relevance of its contribution**	**Concepts extracted**
2012	*quality appraisal tool for case series studies using a modified Delphi technique*	Health Economics / Canada	for quality appraisal of case series studies using a modified Delphi technique. Criteria and items that are useful for appraising the quality of case series are reported	stakeholder involvement; outcomes; problem definition
Orem et al, 2012	*Do guidelines influence the implementation of health programs? - Uganda's experience*	WHO / Uganda	The aim is to describe the processes of development, implementation, dissemination and evaluation of health planning, services management, and clinical guidelines within the health sector in Uganda, with the goal of understanding how these processes facilitate or abate the utility of guidelines	Prioritization; relevance; scope; evidence-based; stakeholder involvement; ethical; outcomes; problem definition; operationalization; costs; effectiveness; outcomes/impact evaluation; dissemination plan; generalizability
Peters & Bennett, 2012	*Better guidance is welcome, but without blinders*	John Hopkins School of Public Health / USA	This paper discusses the challenges related to engendering greater structure and systematization in the development of health system guidance and to the application of evidence to policy.	Evidence-based; stakeholder involvement; benefits/harms; feasibility; flexibility; politically sound; generalizability
WHO, 2012	*Guidance on assessing health system building blocks*	WHO / Switzerland	The aim is to provide an overview of the key opportunities and challenges facing a health system (health issues, systemic issues, political/policy issues), and how to address and assess them. It discusses indicators that can be used to assess each of the health systems building blocks and provides a how-to-manual	Prioritization; scope; evidence-based; stakeholder involvement; outcomes; presentation; operationalization; resources; outcomes/impact evaluation; feasibility; politically sound
Bryce et al, 2013	*A common evaluation framework for the African health initiative*	The Johns Hopkins Bloomberg School of Public Health / USA	The aim is to describe a common evaluation framework for the cross-site initiative to improve population health by strengthening health systems and evaluating the results. Some core elements, inputs and processes required for strengthening health systems in Africa are discussed	Stakeholder involvement; ethical; resources; cost-effectiveness; process evaluation; outcomes/impact evaluation; feasibility; politically sound; external factors; generalizability
Philips et al, 2013	*Protocol for development of the guideline for*	University of South Australia /	The aim is to develop reporting guidelines for evidence based practice educational interventions and teachings to enable their consistent and transparent reporting in health care. Criteria for	Evidence-based; stakeholder involvement; costs; effectiveness; outcomes/impact evaluation

Appendix I: Summary of selected papers				
Author(s)	**Title**	**Organization/ Location****	**Purpose of the study and relevance of its contribution**	**Concepts extracted**
	reporting evidence based practice educational interventions and teaching (GREET) statement	Australia	appraising practice education for health professional disciplines are provided	
International Centre for Applied Health Evidence (ICAHE), n.d	The iCAHE guideline checklist	University of South Australia / Australia	The aim is to provide a checklist (items, criteria and domains) to assist in the process of development and appraisal of guidelines	Relevance; timeliness; evidence-based; stakeholder involvement; presentation; problem definition; feasibility
Funk et al, n.d	*Checklist for evaluating a mental health policy*	WHO / Switzerland	The aim is to explore the processes that are likely to lead to the success of a mental health policy. A framework for the assessment of the quality of the processes and content of mental health policy recommendations is proposed	Prioritization; scope; transparency; evidence-based; stakeholder involvement; ethical; problem definition; resources; feasibility; affordability; external factors; transferability

** *Organization/location of the lead authors as reported in the paper*

Chapter 3: Health Systems Guidance Appraisal Concepts – A Survey of Stakeholders

Denis Ako-Arrey[1]
Melissa Brouwers[1]
John N. Lavis[1,2]
Mita Giacomini[1]
The AGREE-HS team[*]

Word count:
567 (Abstract)
4627 (Full text)

1. McMaster University
2. Harvard T.H Chan School of Public Health

[*]The members of the AGREE-HS Scientific/advisory team include: Pierre Ongolo-Zogo (Cameroon); Jillian Ross, Mike Wilson, Padraig Warde and Sheila McNair (Canada); Yoalong Chen (China); Iván Darío Flórez Gómez (Colombia); Govin Permanand (Denmark); Fadi El-Jardali (Lebanon); Carmen Mihaela Dolea, Xavier Bosch-Capblanch and and Ulysses Panisset (Switzerland); Andy Haines, Francoise Cluzeau and Pablo Perel (United Kingdom); Luis Gabriel Cuervo (United States of America).

Preface

The paper presented in this chapter builds on the work presented in chapter 2. Through a systematic and formal knowledge synthesis process, chapter 2 identified 30 concepts, reflected in 3 domains that together comprise the HSG framework. The framework created the foundation for a knowledge translation tool that aims to direct the development, reporting and evaluation of HSG. This paper addresses an important empirical gap in that it engaged a large international health systems/health policy community and requested participants to provide comprehensive and systematic feedback on the importance of the items and domains that comprise the framework. The sampling approach of respondents used in this study ensured participants were recruited from all the WHO health regions so that the perspectives of health systems/health policy experts was represented from low, middle and high-income countries. The data presented in this chapter directly led to the refinement of the HSG concepts and their descriptions and generated a beta version of the HSG tool comprised of 4 domains and 32 concepts. Moreover, this chapter also shows that, stakeholders perceive that the HSG tool could be useful in all three of its intended purposes - development of HSG, for their appraisal and for their reporting. This chapter provides important background and data for the subsequent chapter 4; the testing of the beta version of the HSG tool.

I was responsible for conceiving and designing this study along with my supervisor (Dr. Melissa Brouwers), and received guidance from members of my supervisory committee (Dr. John Lavis and Dr. Mita Giacomini), as well as from members of our core and expanded scientific team. I completed all recruitment, data collection, analysis, and interpretation and I also drafted the paper. Through an iterative process, my supervisor, members of my supervisory committee and members of our scientific team provided comments and suggestions on earlier drafts that were incorporated into revisions.

Abstract

Background: Health systems guidance (HSG) provides a transparent approach to developing and implementing recommendations on the possible courses of action to address health system challenges and thereby strengthen health systems. (Bosch-Capblanch, 2011; Lavis et al, 2012). However, there is a paucity of methods to direct the HSG development process; there's a lack of an appropriate conceptual model to appraise HSG quality; and there is a dearth of best practice strategies for reporting guidance recommendations (Bosch-Capblanch, 2011). We are conducting a multi-stage program of research with the input of international HSG experts in order to create a reliable, valid and useful tool for the appraisal of HSG that can also be used to support HSG development and reporting. To fully understand the HSG landscape, Stage 1 of this research project was a critical interpretive synthesis (CIS) aimed at generating a candidate list of concepts (items/criteria/domains) that will comprise a potential HSG tool. From this review, we identified 30 candidate HSG quality criteria and confirmed that no existing evaluation tool exists for HSG. Stage 2 of this program of research, and the focus of this paper, was to conduct a structured survey in order to evaluate the importance, value and priority of the 30 candidate concepts and their definitions generated from the CIS and identify any missing components.
Methods: A strategically selected group of global policy-makers, HSG developers, knowledge users and research communities (stratified by geography & expertise) assessed the 30 candidate concepts for appropriateness to, relevance to, and priority for health system decisions and HSG. Individuals from the six World Health Organization (WHO) regions with relevant expertise were purposefully sampled for this study. Each candidate concept was provided with an operational definition and a 7-point scale (1, strongly disagree to 7, strongly agree) was used to rate each concept. For each of the concepts we asked 4 key questions relating to whether the concept is (i) a core component of HSG, (ii) important in development of HSG, (iii) important in appraisal of HSG, and (iv) important in reporting of HSG. Descriptive analyses were computed.
Findings: Fifty participants were invited to complete the survey and we had a response rate of 82%. The mean responses and standard deviations for each concept within each key question were universally favorable. There was also an overall agreement about the need for a high quality tool to systematically direct the development, appraisal and reporting of HSG. Considerable qualitative feedback from the respondents and deliberations with some experts was also received regarding refinements and changes to the wordings of the concepts and their descriptions. From this feedback, some of the concepts were merged into one, while other concepts were split into two. Two new concepts found to be useful for the end users of the HSG also emerged and were added into the list. This led to the creation of a Beta (draft) version of the HSG tool complete with 32 items (and their descriptions) and 4 domains.
Conclusion: A key strategy for the production of an acceptable HSG tool is to adhere to standard methodological quality criteria (e.g., usable, reliable, and valid) that confer guidance the credibility to be used and adapted. This study adds to the existing literature by moving from the generated HSG quality criteria (concepts) to providing a foundation for a knowledge tool and a common analytic framework for health systems that can ultimately improve the HSG enterprise.

1. Introduction

The ways in which health systems are designed, how they operate, how they are governed, how financial arrangements within them are set, and how services are delivered, can have significant impacts on the health of individuals and communities (WHO, 2004). Strengthening health systems is increasingly seen as a foundation for optimizing and maintaining improvements in population health outcomes, as well as in improving the patient experience and keeping per capita costs manageable (Coker et al. 2004; McKee et al. 2009; Samb et al. 2010; Atun, 2012). However, the achievement of health goals in several countries and regions has been hindered by a variety of challenges ranging from weak and dysfunctional health system features like existing delivery, financial and governance arrangements (Travis et al, 2004; Hoffman et al, 2012), through influences on the policy process that compromise efforts like institutions, interests and ideas (Pierson, 1993), to context-specific features (political, social, cultural and economic) that run counter to goals (Szreter & Woolcock, 2004; WHO, 2008). Improving the suitability of health systems to deliver health care and public health interventions is, therefore, an essential and complex task that can be supported by appropriate health systems guidance (HSG).

HSG are systematically developed statements produced at global, national and regional levels (e.g., by the World Health Organization, ministries and departments of health, and special committees supporting ministries and departments of health) that provide possible courses of action to address these challenges and thereby strengthen health systems (Bosch-Capblanch et al, 2012). For example, HSG recommendations can help to determine appropriate ways to frame the problem of a population not having access to a primary care physician (e.g., supply, distribution or payment problem), to outline viable options for health system arrangements that will strengthen primary care (e.g., financial and governance arrangements), to identify alternative implementation strategies that will get cost-effective programs, services and drugs to those who need them, to monitor implementation efforts and to evaluate their impacts (WHO, 2011). The quality of HSG may therefore impact the type of recommendations being formulated, the degree to which they get implemented, the methods of dissemination, and the extent to which they impact the usual operations of the health system (WHO, 2003). Higher quality guidance has the capacity to contribute to higher quality policy decisions (Bosch-Capblanch et al, 2011; Lavis et al, 2012), which in turn will better optimize health impacts through well-functioning health systems (WHO, 2010).

Clinical practice guidelines (CPGS) – guidance documents that target clinical questions and provide recommendations relevant to (primarily) clinician and patient decisions – could be considered conceptually equivalent knowledge tools to HSG. However, while considerable advancements have been made regarding the science and practice of CPG development, appraisal and reporting, the same cannot be said for HSG. Optimizing health systems is a challenging task to which appropriate guidance can positively contribute, but there is a paucity of methods to direct the HSG development process, there's a lack of an appropriate conceptual model to appraise HSG quality, and there is a dearth of best practice strategies for reporting guidance recommendations.

Indeed, no specific tool exists that is designed to evaluate HSG or to differentiate high quality from low quality HSG, or to contribute to guidance development and reporting (Bosch-Capblanch et al, 2011). As more groups want to rely on the innovation of HSG, coupled with increasing pressures to demonstrate value for money, there has been an international call to action to create a tool and accompanying resources to support their use and ensure that the most valid, credible and implementable guidance is identified and applied in health systems (Bosch-Capblanch, 2011; Birtwhistle et al, 2012; Bosch-Capblanch, et al 2012).

We are conducting a multi-stage program of research with the input of international HSG experts in order to create a reliable, valid and useful tool for the appraisal of HSG that can also be used to support HSG development and reporting. To fully understand the HSG landscape, stage 1 of this research project was a review aimed at generating a candidate list of concepts (items/domains/criteria) that will comprise a potential HSG appraisal tool. We completed the review (using a critical interpretive synthesis - CIS - approach) of the existing literature in order to identify any published studies that report on existing criteria currently used to describe, differentiate or test the quality of HSG (Ako-Arrey et al, 2015). It was our expectation that the receptiveness, adoption and diffusion of HSG recommendations depend on the perception of their quality, and with this review, we aimed to identify those core elements of a good quality HSG. From this review, we identified 30 candidate HSG quality criteria (concepts), clustered into 3 domains, and confirmed that no existing evaluation tool (draft or final version) exists for HSG.

Applying standardized measurement design techniques for item generation, validation and reduction (Streiner & Norman, 2003), the overall goal of this study (stage 2) was to have the intended users of a HSG appraisal tool evaluate the importance, value and priority of the 30 candidate concepts and definitions generated from the CIS and identify any missing components. The purpose of this paper is to report on this stage of the program of research.

2. Methodology

The approach used for this study was a structured survey targeted at international stakeholders in guidance development and implementation processes. Collaborators and co-investigators in our research team selected individuals to invite for participation in the study. The selected individuals made up a master list that served as the population from which participants were purposively sampled for the survey. Individuals from the six World Health Organization (WHO) regions with expertise in health system/health policy either as knowledge users (clinical leaders, healthcare executives, and policy-makers), as well as, HSG developers, health policy and health systems experts, and researchers were eligible to be included in this master list. Our goal was to be inclusive and as geographically dispersed as possible. We also pooled together the authors listed from various HSG-related documents to further identify additional candidate participants. These target users and stakeholders were engaged in the systematic evaluation of the candidate concepts for appropriateness to, relevance to, and priority for health system

decisions and HSG. Participants from our master list, stratified by location (WHO health region) and expertise (production and use of HSG) were invited to participate in the survey. Letters of invitation, describing the study, were e-mailed to candidate participants to solicit their participation (see appendix I for sample letter of invitation). Individuals who agreed to participate were e-mailed a password-protected unique identifier to log into a Web-based study platform (LimeSurvey®) to complete the structured survey. We also accommodated the requests of participants who preferred print packages of research materials. Our letters of invitation to participants outlined the purpose of the study, definition of key terms, likely time commitment and the survey process. Up to four reminders were sent out to the invited participants over the 4 months study period (June to September 2014). See Appendix II for a sample of the survey questionnaire.

Each candidate concept was accompanied by an operational definition and considerations for scoring. For each of the 30 concepts, participants were asked to rate their agreement with the following four key questions (measures);

1. This concept is a defining feature (core component) of HSG.
2. This is an important concept to address in the DEVELOPMENT process of HSG.
3. This is an important concept in the APPRAISAL of HSG to differentiate between higher and lower quality guidance documents.
4. This is an important concept to be REPORTED in HSG.

A 7-point scale (1, strongly disagree to 7, strongly agree) was used to rate each of the concepts. Participants were provided with the opportunity to suggest refinements and modifications to each of the candidate concepts (i.e., labels, definitions, etc.) and to suggest additional concepts not addressed in the list.

Participants were then asked to rate their overall agreement about the need for a high quality tool aimed to systematically appraise HSG and contribute to HSG development and reporting. Demographic questions that captured the participants' gender, affiliation/organization, role/expertise and years of experience were also included in the survey. Some members of the scientific research team as well as some selected health services/systems researchers, pilot-tested the survey to enable refinements before it was distributed to consenting participants.

Survey responses were downloaded into Microsoft Excel spreadsheets and analyzed using Excel and SPSS. Overall descriptive analyses were calculated for each of the rated concepts (mean, standard deviation (SD), mode, median, and range). Items that 80% or more of the respondents rated favorably on each of the four measures (between 5 and 7 on the response scale) were maintained. Those that did not meet this threshold were prioritized for refinement or deletion. Final decisions regarding the concepts were made through consensus of the core and extended members of the scientific team (AGREE-HS team). Additional concepts nominated by the participants were reviewed by the scientific team and reworked to align with the style and format of the other candidate concepts. Written feedback was reviewed and a thematic analysis was done.

3. Results

The total number of participants invited to complete the survey regarding the importance, value and priority of the thirty candidate concepts was 50, and the total number of respondents who completed the survey was 41, for a response rate of 82%. As seen on table 1 below, three quarters of the respondents were males (76%). Respondents represented all six World Health Organization regions with the Americas and Europe most represented (66%) and the Eastern Mediterranean and South East Asia least represented (8%). In terms of expertise, our respondents represented a variety of health systems/health policy roles either at national health ministries or international health agencies, and others were health services/systems researchers either within academia or with applied research institutes. Our respondents also had health systems/policy experience ranging from over one year to over 20 years.

Table 1: Demographic details
Gender Males = 31 (76%) Females = 10 (24%)
WHO Health Region Americas = 17 (41%) Europe = 10 (25%) Africa = 7 (17%) Western Pacific = 4 (10%) Eastern Mediterranean = 2 (5%) South-East Asia = 1 (3%)
Role/Expertise Director = 8 (20%) Manager = 4 (10%) Technical adviser = 9 (21%) Researcher (academia) = 8 (20%) Researcher (applied) = 12 (29%)
Years of Experience 1-4 years = 5 (12%) 5-9 years = 12 (29%) 10-19 years = 11 (27%) 20+ years = 13 (32%)

Table 2 below reports the participants' ratings (mean and standard deviation) for each of the concepts to the four key questions that were asked in the survey;

1. Concept is a core component (C) of HSG
2. Concept is important in Development (D) of HSG
3. Concept is important in Appraisal (A) of HSG
4. Concept is important in Reporting (R) of HSG.

Table 2: Means and standard deviations for each concept based on the four outcome measures

Concepts	Core (C)		Development (D)		Appraisal (A)		Reporting (R)	
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
1. Priority	6.3	0.9	6.3	0.9	5.9	1.0	6.0	0.9
2. Relevant	6.6	0.8	6.5	0.7	6.2	1.1	6.2	0.9
3. Timely	5.8	1.0	6.0	1.1	5.6	1.2	5.8	0.9
4. Comprehensive	6.1	0.9	6.3	0.7	6.1	0.9	6.3	0.6
5. Systematic & Transparent	6.5	0.7	6.7	0.4	6.6	0.5	6.6	0.6
6. Evidence-based	6.6	0.6	6.5	0.8	6.6	0.6	6.5	0.9
7. Participatory	6.4	0.9	6.7	0.5	6.6	0.6	6.5	1.0
8. Ethical	6.5	0.6	6.3	0.7	6.2	0.8	6.3	0.8
9. Outcomes oriented	6.3	0.9	6.5	0.8	6.5	0.8	6.4	0.8
10. Interests managed	6.6	0.7	6.7	0.5	6.7	0.5	6.7	0.5
11. Clearly presented	6.3	0.9	6.4	0.8	6.2	0.8	6.3	0.9
12. Defined problem	6.5	0.5	6.5	0.6	6.3	0.9	6.6	0.6
13. Operational options	5.6	1.2	5.9	1.0	5.9	1.0	5.9	1.0
14. Costs	6.0	1.0	6.2	0.8	5.8	1.1	5.9	0.9
15. Resources	6.1	1.0	6.2	1.0	5.9	1.2	6.0	1.0
16. Effectiveness	6.2	1.2	6.3	0.8	6.0	1.1	6.1	1.1
17. Cost-effectiveness	6.2	1.2	6.1	1.0	5.9	1.3	6.1	1.1
18. Benefits/harms weighting	6.2	1.1	6.3	1.0	6.2	1.0	6.1	1.5
19. Dissemination plan	6.3	0.9	6.3	0.9	6.0	1.1	6.2	1.1
20. Process evaluation	6.2	1.1	6.4	0.9	6.1	1.1	6.2	1.1
21. Outcomes/impact evaluation	6.2	0.8	6.4	0.8	6.2	0.9	6.1	1.1
22. Updating plan	6.1	0.9	6.1	0.9	5.9	1.0	5.9	1.0
23. Feasible	6.4	0.6	6.4	0.7	6.5	0.8	6.3	1.1
24. Affordable	5.5	1.4	5.5	1.4	5.3	1.5	5.5	1.4
25. Flexible	5.6	0.9	5.8	0.9	5.4	1.1	5.4	1.1
26. Socio-cultural alignment	5.9	1.0	6.0	1.0	5.8	1.9	5.9	1.0
27. Political alignment	4.9	1.5	5.0	1.4	4.7	1.6	5.0	1.4
28. External alignment	5.6	1.2	5.8	0.9	5.3	1.4	5.6	1.3
29. Transferable	5.6	1.1	5.7	1.1	5.6	1.2	5.7	1.1
30. Sustainable	5.8	1.0	5.8	1.0	5.6	1.3	5.7	1.3

As can be seen in Table 2 above, ratings were universally favorable and, for each concept, there was consistency in the mean ratings across the four metrics (i.e., core, development, appraisal and reporting). For the *core* metric (C), mean ratings fell between 4.9 (political alignment) and 6.6 (interests managed, evidence-based, and relevant). For the *development* metric (D), mean ratings fell between 5.0 (political alignment) and 6.7 (systematic & transparent, participatory, and interests managed). For the *appraisal* metric (A), mean ratings fell between 4.7 (political alignment) and 6.7 (interests managed). And

for the *reporting* metric (R), mean ratings fell between 5.0 (political alignment) and 6.7 (interests managed). Standard deviations for all the 30 concepts across all the four outcome measures were small, suggesting consistency in responses across participants.

"Political alignment" was the least favorable concept; for two of the measures, *core* and *appraisal*, it did not reach the mean threshold of 5.0, scoring 4.9 and 4.7, respectively. Nonetheless, the members of our scientific team considered that this concept was important, and in view of the fact that it had only missed the threshold slightly, upon deliberation, the final consensus decision was to include it in the tool.

We also present the results below (table 3) to show the respondents' overall agreement about the need for a high quality tool to systematically direct the DEVELOPMENT of HSG, to systematically direct the APPRAISAL of HSG and to systematically direct the REPORTING of HSG.

Table 3: Overall agreement on the need of a HSG tool

Overall Agreement	Mean	SD	Mode	Median	Range
Overall agreement about the need for a high quality tool to systematically direct the DEVELOPMENT of HSG	6.6	0.7	7	7	2
Overall agreement about the need for a high quality tool to systematically direct the APPRAISAL of HSG	6.6	0.6	7	7	2
Overall agreement about the need for a high quality tool to systematically direct the REPORTING of HSG	6.3	1.0	7	7	5

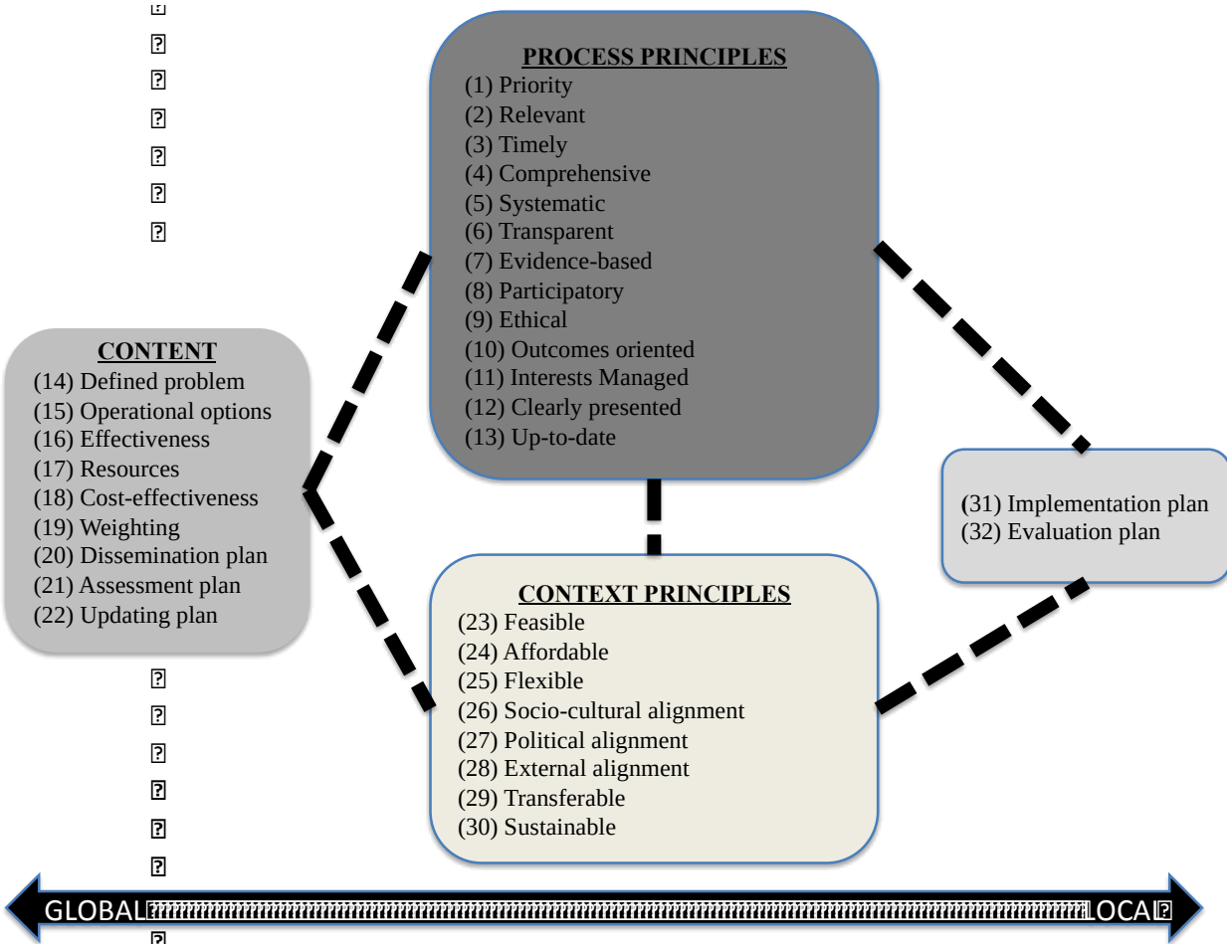
As shown in table 3 above, overall agreement about the need for a high quality tool to systematically direct the development, appraisal and reporting of health systems guidance was overwhelmingly high and consistent among the survey participants. The means were above 6, the standard deviations were low, and the modes and medians were 7 for all three categories. The range of 2 recorded for the development and appraisal categories was also low. For the reporting category, a range of 5 was recorded as a result of one outlier who rated their overall agreement for this category with a value of 2.

Considerable feedback was also provided by the participants regarding refinements and changes to the wordings of the concepts and their descriptions. Using the results and feedback from the survey, reconsiderations of the raw data from the review, and a series of meetings with the core and expanded members of the team (n=11), the concept labels and descriptors were refined. For example, two of the concepts that emerged from the CIS (costs and resources) were merged to represent one concept (resources). Also, “process evaluation" and "outcomes/impact evaluation" were merged

into “assessment plan”. Additionally, two of the concepts were split into two: “updating plan” became “updating plan” and “up-to-date”, while “systematic and transparent" became “systematic” and “transparent”. Appendix III shows a table comparing the original labels and the new labels after the refinement process.

The feedback from the survey and deliberations with members of the scientific team also led to the modification of the AGREE-HS framework that shows relationships between the concepts as well as relationships between clusters of the concepts (figure 1 below). Building from our previous study (Ako-Arrey et al, 2015), we clustered the concepts together into four meaningful categories (domains); Process principles, content principles, context principles and implementation/evaluation plan.

Figure 1: Framework of health systems guidance concepts



In contrast to the original version of the framework, for this version, a double-headed arrow was added to depict the division of labor between roles at the global level and roles at the local level. At the local level, an additional category was added to represent the need for end users to design a detailed implementation and evaluation plan for their individual contexts. The implementation plan represents the development of a

strategic plan by the end users to put the guidance recommendations into action. The evaluation plan entails the development of a monitoring and evaluation strategy for the process of implementation as well as the outcomes/impacts of the guidance recommendations. This brought the number of items to a total of 32 clustered into 4 domains.

The final beta version of the Health System Guidance appraisal tool concepts (labels and definitions) is presented in table 4 below. With the approval of the Appraisal of Guidelines for Research and Evaluation (AGREE) consortium, we have named the tool AGREE-for Health Systems (AGREE-HS).

Table 4: Beta version of the AGREE for Health Systems (AGREE-HS) tool

Process principles	
1. Priority	The guidance is properly aligned with current health system priorities from the perspective of topic, jurisdictional focus (e.g., all low-and middle-income countries, sub-Saharan Africa), health system level and population. The expression of the need and origin of the mandate for the guidance is clear.
2. Relevant	The guidance recommendations are relevant to, appropriate to and valid for the health system challenge, system or sub-system needs, the target population(s), and the setting in which they will operate.
3. Timely	The recommendations are available in a timely manner in relation to when the policy decisions are made or timely in relation to the health system issue being addressed.
4. Comprehensive	The guidance is comprehensive and covers all relevant/appropriate (direct and indirect) health system levels (e.g., district), sub-systems (e.g., mental health) and sectors (e.g., acute care)
5. Systematic	Systematic processes are applied in developing the guidance according to a specific plan and/or explicit methodologies.
6. Transparent	A transparent and reproducible approach in the development and reporting of the guidance is demonstrated.
7. Evidence-based	The best available and ideally most contextually relevant evidence informs the recommendations.
8. Participatory	The health system guidance team is comprised of multidisciplinary/multi-sectoral membership and includes those with an interest, stake or responsibility in the development, implementation and evaluation of the recommendations.
9. Ethical	The recommendations are considered within the lens of an ethical framework and align with applicable ethical principles and values (e.g. equity, equality, human rights, liberty, efficiency, autonomy, dignity, beneficence, etc). The guidance adequately promotes fairness and equality in terms of age, ability, culture, gender, socioeconomic status, religion, occupation, language, ethnicity, race or sexual orientation among the target population.
10. Outcomes	The guidance describes all the anticipated effects/outcomes as well as

oriented	the appropriate indicators, performance thresholds, targets and standards that can be used to measure the effects/outcomes.
11. Interests managed	A declaration of competing interests from the guidance developers (e.g. financial, academic, professional, etc.) is identified and the strategies to manage them are described. It is also clear that the views of any funding body involved have not influenced the development process of the guidance.
12. Clearly presented	The recommendations are clear, user-friendly, succinct, unambiguous and presented in a readable and consistent format, with key recommendations easily identifiable.
13. Up-to-date	The recommendations are current and the evidence (e.g. systematic reviews) on which they are based is considered up-to-date.
Content	
14. Defined problem	The health system challenge and its causes are clearly articulated; specifically, the nature, causes, and magnitude, frequency or intensity of the problem, the populations and jurisdictions that are affected are clearly described.
15. Operational options	The recommended "solutions" are operationalized sufficiently with the conceptualization, operational guidance and the mode of delivery of the options clearly stated.
16. Effectiveness	Evidence of recommendation's effectiveness are described including methods used, context where tested, and results.
17. Resources	The inputs to and/or the costs of the implementation processes (amounts, frequency, duration) are described and are commensurate to the health systems issue; specifically, money, time, infrastructure, administrative capacity, information, equipment, supplies, healthcare professionals, training, etc. are considered.
18. Cost-effectiveness	The recommendations are attentive to value for money considerations with relevant cost-effectiveness evidence of recommendations described.
19. Benefits/harms weighting	Descriptions and/or judgments of the potential intended and unintended consequences (positive & negative) of the guidance on the population and/or the system are provided.
20. Dissemination plan	Methods for communicating guidance are clearly described and framed within an overall dissemination strategy.
21. Assessment plan	This involves high-level recommendations for assessing the structure and process of the implementation process as well as an assessment of the outcome/impact of the guidance to determine whether the course of action was a success or failure.
22. Updating plan	Recommendations for periodic updates are made and the procedure to update the guidance is provided with explicit timelines on anticipated review, appropriate expiration date of the guidance and an explanation of the rationale for the proposed time frames.

Context principles	
23. Feasible	The guidance recommendations are realistic and the actions are pragmatic. The guidance describes facilitators and barriers for implementation.
24. Affordable	The guidance recommendations are affordable within the financial structure and budgetary allocations of the health system.
25. Flexible	The guidance is flexible and adaptable to the expertise of the user and the varying local conditions in the context where implementation will take place.
26. Socio-cultural alignment	The recommendations adopt a socio-cultural perspective and are robust under societal and cultural scrutiny.
27. Political alignment	The political acceptability of the recommendations is considered and the degree of alignment with political interests and commitments are described.
28. External alignment	Determinants of health system performance that lie outside the formal architecture of the health system but will influence the performance of its functions are considered and described (for example, judicial system, social system, recession, corruption, state of the economy etc.).
29. Transferable	A description of the degree to which recommendations are transferable to other similar or different regions and contexts is provided.
30. Sustainable	The anticipated sustainability and maintenance of long-term outcomes is described.
Implementation and evaluation plan	
31. Implementation plan (end-users)	This involves the development of a strategic plan by end-users at the local level to describe the process of moving the recommendations into action. The plan may include a description of inputs, services and activities that are required for implementation; identification of the strengths, weaknesses, opportunities and threats to the implementation process; and allocation of responsibilities and duties. Designing an implementation strategy will facilitate adherence and compliance to planned activities, and enhance efficiency.
32. Evaluation plan (end-users)	A strategy for the monitoring and evaluation of the implementation strategy/process and/or outcomes of the guidance in a way that determines whether the changes observed in relation to the health system challenge being addressed can be attributed to the guidance is provided. There are also recommendations for an impact evaluation to look at the short and long term deeper primary and secondary changes that resulted from the guidance as well as corresponding challenges.

4. Discussion

Through a structured survey of relevant stakeholders from all six World Health Organization regions as well as feedback from members of our scientific team, we found that all the concepts met our criteria of being rated favorably as it relates to each of our outcomes: being a core-component of HSG, being important to the HSG development process, being an important appraisal criteria to differentiate HSG as a function of quality, and being important to be reported upon in HSG. Moreover, the idea of the need for an instrument that will be used to direct the development, appraisal and reporting of good quality HSG was well supported. Feedback from the survey and interactions with the scientific team led to reformulation of the labels of the concepts and their operational definitions in order to generate the beta version of the AGREE-HS tool as shown on table 4. These deliberations also led to the reformulation of the AGREE-HS framework (figure 1)

A key strategy for the production of an acceptable HSG tool is to adhere to standard methodological quality criteria (e.g., usable, reliable, and valid) that confer on guidance the credibility to be used and adapted. This study adds to the existing literature by moving from the generated HSG quality criteria (concepts) to providing a foundation for a knowledge tool and a common analytic framework for health systems that can ultimately improve the HSG enterprise. It enables the creation of a vehicle for facilitating informed decision-making about HSG at various levels and promoting a culture of informed HSG developers and consumers.

Some strengths of this study: Firstly, it involved a multidisciplinary blend of international participants recruited based on geography and expertise in order to cover various perspectives and jurisdictions. Secondly, it involved an iterative collaborative process with members of our core and expanded team comprised of investigators and collaborators with an extensive knowledge in health systems and policy research. Thirdly, it involved a high quality approach adapted from the methodological, conceptual and theoretical principles of measurement construction used to design a complementary tool, AGREE II, which aims to facilitate the development, appraisal and reporting of clinical practice guidelines. Fourthly, consensus among members of the scientific team with regards to an acceptable threshold (between 5 and 7 on the response scale) was reached prior to carrying out the survey.

Some weaknesses of this study: First, the small sample size meant that we did not have sufficient power to conduct a factor analysis. While not part of the scope of this study, a factor analysis is an important step in the development of a measurement tool as it provides data to determine the number of domains in the tool, how the assessment scores should be calculated, and whether the items in the instrument cluster empirically (Streiner, 2008). This is being considered for a future study. Secondly, it would have been interesting to see whether there was any variation in the ratings of the concepts that match directly into specific roles/expertise or jurisdictions. Again, while this was not the focus of this study, the small response ranges and small standard deviations suggest little variability in perceptions across participants. However, we could not test this directly.

Thirdly, while an 82% response rate is excellent, we have little information about the demographic characteristics of non-responders and/or the reasons for not responding. However, we do know that of the non-responders, 44% were invitees from the Americas, 34% were from Europe, 11% from Africa and 11% from the western pacific region.

The next steps of this program of research involve: testing and refining the new tool to assess its measurement performance and usability by using the beta version of AGREE-HS to evaluate the quality of actual HSG documents; developing a user manual to serve as a guide to support the use of the tool; promoting the use of the tool internationally to groups who develop HSG and collate HSG in on-line system directories; and developing an on-line training program to orient new users to AGREE-HS to ensure its optimal use.

5. References

1. Atun, R. (2012). Health systems, systems thinking and innovation. *Health policy and planning*, *27*(suppl 4), iv4-iv8.
2. Birtwhistle R, Pottie K, Shaw E, Dickinson JA, Brauer P, Fortin M, Bell N, Singh H, Tonelli M, Connor Gorber S, Lewin G, Joffres M, Parkin P (2012) Canadian Task Force on Preventive Health Care: We're back! Can Fam Physician: 58: 13-15
3. Bosch-Capblanch X (2011) Handbook for Supporting the Development of Health System Guidance. Basel: Swiss Centre for International Health.
4. Bosch-Capblanch X, Lavis JN, Lewin S, Atun R, Røttingen JA, et al. (2012) Guidance for evidence-informed decisions about health systems: rationale for and challenges of guidance development. PLoS Med 9: e1001185.
5. Coker RJ, Atun R, McKee M. (2004) Health care system frailties and public health control of communicable diseases on the European Union's new eastern border. The Lancet 363: 1389–92.
6. Hoffman S, Rottingen J-A, Bennett S, Lavis J, Edge J, Frenk J (2012) Background Paper on Conceptual Issues Related to Health Systems Research to Inform a WHO Global Strategy on Health Systems Research - a working paper in progress. Geneva: World Health Organisation; 2012.
7. Lavis JN, Røttingen JA, Bosch-Capblanch X, Atun R, El-Jardali F, et al. Guidance for evidence-informed policies about health systems: linking guidance development to policy development. PLoS Med. 2012;9: e1001186
8. McKee M, Suhrcke M, Nolte E et al. 2009. Health systems, health, and wealth: a European perspective. The Lancet 373: 349–51.
9. Pierson, P. (1993) When Effect Becomes Cause: Policy Feedback and Political Change. World Politics 45(July):595–628.
10. Samb B, Desai N, Nishtar S et al. (2010). Prevention and management of chronic disease: a litmus test for health-systems strengthening in low-income and middle-income countries. The Lancet 376:1785–97.
11. Streiner DL, Norman GR (2008). Health measurement scales: a practical guide to their development and use. (4th ed). Oxford University Press. Oxford, UK
12. Streiner, D. L. (2012). 10 Figuring Out Factors: The Use and Misuse of Factor Analysis. *A Guide for the Statistically Perplexed: Selected Readings for Clinical Researchers*, 110.
13. Szreter, S., & Woolcock, M. (2004). Health by association? Social capital, social theory, and the political economy of public health. *International Journal of Epidemiology*, *33*(4), 650-667.
14. Travis, P., Bennett, S., Haines, A., Pang, T., Bhutta, Z., Hyder, A. A., ... & Evans, T. (2004). Overcoming health-systems constraints to achieve the Millennium Development Goals. *The Lancet*, *364*(9437), 900-906.
15. Walt G, Gibson L (1994). Reforming the health sector in developing countries: The central role of policy analysis. *Health Policy and Planning*, 9: 353-70.
16. World Health Organization (2003). "Guidelines for WHO Guidelines." EIP/GPE/EqC/2003.1. Geneva: Author. Retrieved July 28, 2013.

http://whqlibdoc.who.int/hq/2003/EIP_GPE_ EqC_2003_1.pdf

17. World Health Organization - WHO - (2008) The WHO Commission on Social Determinants of Health: *Closing the Gap in a Generation: Health Equity Through Action on the Social Determinants of Health: Commission on Social Determinants of Health Final Report*. World Health Organization (Ed.). World Health Organization.
18. World Health Organization - WHO - (2011). Health system strengthening: improving support to policy dialogue around national health policies, strategies and plans. Report by the Secretariat to the 64th World Health Assembly. A64/12. Geneva: World Health Organization. Retrieved on July 30th 2013 from: https://extranet.who.int/iris/restricted/bitstream/10665/3205/1/B128_8-en.pdf.
19. World Health Organization - WHO - (2014). Social Determinants of Health: Health Systems. Retrieved on November 13th 2014 fromhttp://www.who.int/social_determinants/themes/healthsystems/en/

Appendix I: Sample email invitation for stage 2

Sample email
Subject: Research study invitation from McMaster University

Email text:

Dear {FIRSTNAME},

You are being invited to participate in our Health Systems Guidance Appraisal Tool (HSG-AT) project. We sincerely welcome your participation.

The HSG-AT project is funded by the Canadian Institutes of Health Research (CIHR), and Drs Melissa Brouwers and John Lavis both of McMaster University, Hamilton, Canada are Principal Investigators.

The goal of the project is to design a reliable, valid and useful tool for the appraisal of Health Systems Guidance (Guidelines for health system's financial, organizational, delivery and governance arrangements) that can also be used to support HSG development and reporting. With the aim of generating a candidate list of concepts (items, criteria, domains) for the tool, we have conducted a review of the existing literature in order to identify any studies that report on existing concepts currently used to describe, differentiate, or test the quality of Health Systems Guidance (HSG).

Participation in this survey should take no longer than 50 minutes of your time. You may decline to answer any question and may withdraw from the study for any reason, at any time, without penalty**.** To ensure the confidentiality of your responses we have provided you with a unique identifier code to login.

If you have any questions regarding the study, please feel free to contact Denis Ako-Arrey at: akoarrde@mcmaster.ca or call (905) 379-0110 at any time.

Any questions regarding your rights as a participant may be addressed to the McMaster Research Ethics Board by calling collect the Office of Research Ethics Officer at (905) 525-9140 Ext 23142 or by emailing ethicsoffice@mcmaster.ca

Thank you, for your time and consideration.

Yours sincerely,

Dr. Melissa Brouwers (Associate Professor, McMaster University and Provincial Director, Program in Evidence-Based Care, Cancer Care Ontario)

Dr. John Lavis (Professor and Director of the Program in Policy Decision-Making, McMaster University)

Denis Ako-Arrey (PhD Candidate at McMaster University)

Click here to enter the survey:
{SURVEYURL}

If you do not want to participate in this survey and don't want to receive any more invitations please click the following link:
{OPTOUTURL}

Appendix II: AGREE-HS survey of stakeholders' questionnaire

Thank you for taking the time to participate in this survey for the creation of the Health Systems Guidance Appraisal Tool – AGREE-HS. There are 34 questions in this survey

1. Prioritization: The HSG should fit in properly and be consistent with current health system priority areas. Variations exist (sometimes considerable) across health systems in the areas they consider as priority. Generating guidance that addresses specific local priority areas will lead to improvements in crucial population health outcomes. The guidance should also be able to inform policy decisions on how to further prioritize across these competing areas where improvements are most warranted. Therefore, demonstrating how the guidance recommendations fit into the country's existing national health strategy will enhance development and uptake of HSG. Guidance should therefore be developed only for health system areas with a clearly demonstrated and documented need. It is also important to report the origin of the mandate to develop the guidance. This is because guidance that is mandated by a top official (e.g the Minister of Health) will also be considered to be of high priority.
Using a 7 points scale (1 =strongly disagree, 2=disagree, 3=disagree slightly, 4=neither disagree or agree, 5=agree slightly, 6=agree and 7=strongly agree), please rate this candidate item/criteria (prioritization) on the four traits below

Please choose the appropriate response for this concept	**1**	**2**	**3**	**4**	**5**	**6**	**7**
This concept is a defining feature (core component) of HSG							
This is an important concept to address in the DEVELOPMENT process of HSG							
This is an important concept in the APPRAISAL of health system guidance to differentiate between higher and lower quality guidance documents							
This is an important concept to be REPORTED in health system guidance							

2. Relevance: The guidance should be relevant to the health system policy issue being addressed, relevant to the target population, and relevant to the context within which the guidance will operate. There should be a statement on the impetus for the endeavor, and this should be relevant to the demonstrated needs of the population, the institutional needs of the system, as well as local, national and potentially global needs. The HSG recommendations also need to be responsive to the expectations of the local population and the local institutions Overall, the guidance should be relevant to the public health developmental plan and priorities of the region or country. It should also be clear whether new guidance is needed or whether existing guidance of acceptable quality can be adapted and used. It is important to document contextual factors that may impact development, implementation, or outcomes. This will provide an understanding of why some guidance recommendations may work in some settings and not others.
Using a 7 points scale (1 =strongly disagree, 2=disagree, 3=disagree slightly, 4=neither

disagree or agree, 5=agree slightly, 6=agree and 7=strongly agree), please rate this candidate item/criteria (relevance) on the four traits below

Please choose the appropriate response for this concept	**1**	**2**	**3**	**4**	**5**	**6**	**7**
This concept is a defining feature (core component) of HSG							
This is an important concept to address in the DEVELOPMENT process of HSG							
This is an important concept in the APPRAISAL of health system guidance to differentiate between higher and lower quality guidance documents							
This is an important concept to be REPORTED in health system guidance							

3. Timeliness: Good quality HSG needs to be timely and usable by the broad range of health system stakeholders in order to be attractive to policy-makers who may be lured into other, and potentially biased, forms of evidence by mere convenience or by stakeholders' vested interests. The HSG needs to be available in a timely manner in relation to when the policy decisions are made. Prompt attention to population and health system needs in a timely manner is a good health system quality indicator as it may lead to better health outcomes. Producing HSG in a timely way as windows of policy opportunity open and close is critical especially where there is limited time available for decision-making like in the case of a crisis. At the same time the potential impacts of wrong decisions can be huge given that sometimes, health policy decisions need to be taken even before conclusive supporting evidence is available. Producing HSG under these time pressures can be a potential indicator of the quality of the recommendations. Using a 7 points scale (1 =strongly disagree, 2=disagree, 3=disagree slightly, 4=neither disagree or agree, 5=agree slightly, 6=agree and 7=strongly agree), please rate this candidate item/criteria (timeliness) on the four traits below

Please choose the appropriate response for this concept	**1**	**2**	**3**	**4**	**5**	**6**	**7**
This concept is a defining feature (core component) of HSG							
This is an important concept to address in the DEVELOPMENT process of HSG							
This is an important concept in the APPRAISAL of health system guidance to differentiate between higher and lower quality guidance documents							
This is an important concept to be REPORTED in health system guidance							

4. Scope: Since health policy recommendations produce different results in a variety of settings, and given variations in health systems performance across jurisdictions, it is important that HSG cover all relevant (direct and indirect) health system levels and sectors. Specifically, this should include the various sub-systems/components (hospitals, regional health authorities, and public health units) within the health system. These various components are interlinked, interdependent and interact at various interfaces;

guidance that recognizes this is more apt to be accepted and implementable than guidance that does not consider coverage adequately. Overall health system performance is dependent on the performance of the individual sub-components, and HSG provide recommendations along this continuum going from broad policy directions to specific operational guidance at the sub-system level. So health system strengthening recommendations should cover the different levels and sectors. There is value that international level health system guidance considers one or more of the WHO building blocks as this a common frame for problems at this level.
Using a 7 points scale (1 =strongly disagree, 2=disagree, 3=disagree slightly, 4=neither disagree or agree, 5=agree slightly, 6=agree and 7=strongly agree), please rate this candidate item/criteria (scope) on the four traits below

Please choose the appropriate response for this concept	**1**	**2**	**3**	**4**	**5**	**6**	**7**
This concept is a defining feature (core component) of HSG							
This is an important concept to address in the DEVELOPMENT process of HSG							
This is an important concept in the APPRAISAL of health system guidance to differentiate between higher and lower quality guidance documents							
This is an important concept to be REPORTED in health system guidance							

5. Transparency: The processes for development and implementation of the recommendations should be transparent enough for the methods to be reproducible. Who was involved, which evidence was used, what was considered, how it was considered, how the recommendations were formulated, how they will be implemented, how it will be evaluated are all aspects of guidance development that need to be readily available where appropriate. Therefore, sufficient details regarding the process of development and implementation of HSG provided. This will enable paint a clear picture to knowledge users and target populations on the processes involved in guidance development, which will enhance its reliability and further promote uptake. Using a 7 points scale (1 =strongly disagree, 2=disagree, 3=disagree slightly, 4=neither disagree or agree, 5=agree slightly, 6=agree and 7=strongly agree), please rate this candidate item/criteria (transparency) on the four traits below

Please choose the appropriate response for this concept	**1**	**2**	**3**	**4**	**5**	**6**	**7**
This concept is a defining feature (core component) of HSG							
This is an important concept to address in the DEVELOPMENT process of HSG							
This is an important concept in the APPRAISAL of health system guidance to differentiate between higher and lower quality guidance documents							
This is an important concept to be REPORTED in health system guidance							

6. Evidence-based: Guidance recommendations on best practices for addressing a health systems challenge have to be informed by the best quality and available evidence. Ideally, the evidence used needs to be objective reliable, rigorous, valid and up to date. The evidence that is used needs to be context sensitive enough to resonate with local realities. Producing HSG that is backed by locally generated evidence of feasibility and cost-effectiveness enhances the possibility of successful acceptability and implementation. In practice however, local and national guidance development is typically derived from global guidance that relies on global evidence and data and some derivative products (workbooks). Local actors are then prompted to identify local data and evidence and to convene local processes for contextualizing the guidance. HSG developers also need to pay attention to technical issues reported in the studies related to data collection methods, sampling, and project designs if applicable. Ensuring that the evidence used was methodologically rigorous will enhance the validity of the HSG recommendations. While evidence used in health policy decision making usually comes from well–established scientific methodologies, it is common to find that recommendations on a course of action is based on evidence that may be uncertain, conflicting or controversial, and HSG developers need to take this into consideration. Given the diversity in the quality of evidence that is used in health policy decision-making, sometimes recommendations can be based on facts, expert opinion, tentative assumptions, educated guesses, preliminary models, or good anecdotal evidence; the quantity and quality of evidence of recommendations should be well explicit and transparent. Systematic approaches have to be applied to search for and identify relevant research that supports recommendations. A clear description of the systematic process applied to identify and use this body of knowledge should be reported.
Using a 7 points scale (1 =strongly disagree, 2=disagree, 3=disagree slightly, 4=neither disagree or agree, 5=agree slightly, 6=agree and 7=strongly agree), please rate this candidate item/criteria (evidence-based) on the four traits below

Please choose the appropriate response for this concept	**1**	**2**	**3**	**4**	**5**	**6**	**7**
This concept is a defining feature (core component) of HSG							
This is an important concept to address in the DEVELOPMENT process of HSG							
This is an important concept in the APPRAISAL of health system guidance to differentiate between higher and lower quality guidance documents							
This is an important concept to be REPORTED in health system guidance							

7. Stakeholder involvement: In the development of the guidance recommendations, a rigorous set ideas and contributions from a multidisciplinary group of key stakeholders should be sought. Relevant stakeholders in HSG development may include members from; international organizations, national and local government, research institutes, universities, civil society organizations, non-governmental organizations, citizen groups, professional organizations, media etc. HSG developers, those to be involved in the implementation and evaluation of the guidance, and those who will be affected by the guidance recommendations should be involved in the process. A participatory approach

that involves multiple stakeholders and incorporates their various perspectives into the guidance development process is important. This minimizes the potential error of overlooking some crucial dimensions of addressing the health system issue, provides insights into potential obstacles related to implementation, and creates a culture of stakeholders working towards the same goal and using as part of the solution. It also serves as a forum for considering the value systems of the actors, and resolving disagreements of opinion among various groups. Broad consultations with various stakeholders not only enhance the production of robust and useful HSG, but also improve the dissemination and implementation of the guidance recommendations. Engagement of stakeholders should begin early on from conceptualization and continue to till the end of the process. The guidance documents should make reference to the stakeholders involved, their expertise and roles, as well as their contributions to the goals and objectives of the guidance.
Using a 7 points scale (1 =strongly disagree, 2=disagree, 3=disagree slightly, 4=neither disagree or agree, 5=agree slightly, 6=agree and 7=strongly agree), please rate this candidate item/criteria (stakeholder involvement) on the four traits below

Please choose the appropriate response for this concept	**1**	**2**	**3**	**4**	**5**	**6**	**7**
This concept is a defining feature (core component) of HSG							
This is an important concept to address in the DEVELOPMENT process of HSG							
This is an important concept in the APPRAISAL of health system guidance to differentiate between higher and lower quality guidance documents							
This is an important concept to be REPORTED in health system guidance							

8. Ethical: An ethical lens should be applied throughout the development and reporting of guidance recommendations. The guidance should align with moral values and focus on setting ethical standards and promoting the "good" by bringing to focus principles like equity, equality, human rights, liberty, efficiency, autonomy, beneficence etc. The guidance should adequately accommodate fairness and equality in terms of age, ability, culture, gender, socioeconomic status, religion, occupation, language, ethnicity, race and sexual orientation. It should incorporate respect for human rights, autonomy and dignity as well as promote fairness in the distribution of health among the target population.
Using a 7 points scale (1 =strongly disagree, 2=disagree, 3=disagree slightly, 4=neither disagree or agree, 5=agree slightly, 6=agree and 7=strongly agree), please rate this candidate item/criteria (ethical) on the four traits below

Please choose the appropriate response for this concept	**1**	**2**	**3**	**4**	**5**	**6**	**7**
This concept is a defining feature (core component) of HSG							
This is an important concept to address in the DEVELOPMENT process of HSG							
This is an important concept in the APPRAISAL of health system							

guidance to differentiate between higher and lower quality guidance documents							
This is an important concept to be REPORTED in health system guidance							

9. Outcomes: The guidance should describe all the expected effects/outcomes anticipated as well as the appropriate indicators that can be used to measure the effects/outcomes. Adequate backing should be provided as to the choice of the outcomes and the indicators selected. This should be supported by evidence, politics or societal factors. Considering potential uncertainties that may result, alternative outcomes and outcome indicators should also be identified. Also it is important to describe performance thresholds, targets and standards that are considered acceptable. This conveys information on what the guidance recommendations must achieve to be considered successful.

Using a 7 points scale (1 =strongly disagree, 2=disagree, 3=disagree slightly, 4=neither disagree or agree, 5=agree slightly, 6=agree and 7=strongly agree), please rate this candidate item/criteria (outcomes) on the four traits below

Please choose the appropriate response for this concept	**1**	**2**	**3**	**4**	**5**	**6**	**7**
This concept is a defining feature (core component) of HSG							
This is an important concept to address in the DEVELOPMENT process of HSG							
This is an important concept in the APPRAISAL of health system guidance to differentiate between higher and lower quality guidance documents							
This is an important concept to be REPORTED in health system guidance							

10. Competing interest: A declaration of competing interests by the guidance developers, whether direct or indirect, should be made. It should also be clear that the views of any funding body involved have not influenced the content of the guidance. Full disclosures of any potential benefits to the guidance developers should be indicated. Any perceived conflict of interests to the policy-makers or the staff involved in implementing the guidance should also be identified. These interests should be reported and addressed, with a description of the approaches used to curb any influenced clearly documented. The author's positions, roles, affiliations should be clearly stated.
Using a 7 points scale (1 =strongly disagree, 2=disagree, 3=disagree slightly, 4=neither disagree or agree, 5=agree slightly, 6=agree and 7=strongly agree), please rate this candidate item/criteria (competing interests) on the four traits below

Please choose the appropriate response for this concept	**1**	**2**	**3**	**4**	**5**	**6**	**7**
This concept is a defining feature (core component) of HSG							
This is an important concept to address in the DEVELOPMENT process of HSG							

This is an important concept in the APPRAISAL of health system guidance to differentiate between higher and lower quality guidance documents							
This is an important concept to be REPORTED in health system guidance							

11. Presentation: The recommendations should be specific, direct, explicit and distinct. Also the key recommendations should be easily identifiable. The guidance should be presented in a manner that is clear, consistent, user-friendly and easy to navigate. An executive summary, as well as full texts, should be readily available to the target audience. A complete list of relevant references, a glossary of terms, full meaning of abbreviations and contact information of authors should be available. Because guidance is intended to be flexible, words or phrases that denote an aspirational rather than a mandatory intent should be employed. Using a 7 points scale (1 =strongly disagree, 2=disagree, 3=disagree slightly, 4=neither disagree or agree, 5=agree slightly, 6=agree and 7=strongly agree), please rate this candidate item/criteria (presentation) on the four traits below

Please choose the appropriate response for this concept	**1**	**2**	**3**	**4**	**5**	**6**	**7**
This concept is a defining feature (core component) of HSG							
This is an important concept to address in the DEVELOPMENT process of HSG							
This is an important concept in the APPRAISAL of health system guidance to differentiate between higher and lower quality guidance documents							
This is an important concept to be REPORTED in health system guidance							

12. Problem definition: This refers to the framing of the policy challenge that the guidance aims to address (financial, governance, delivery, or organizational challenges). Whether it is focusing on an event, changes in health indicators, or feedback from another problem, the problem definition provides a statement of the policy issue or what triggered the need for guidance. In defining the problem guidance developers describe the nature, causes, magnitude, frequency and intensity of the problem, the populations and jurisdictions that are affected. A clearly defined problem more readily leads to guidance with a clearly stated purpose, objectives, and goals that readily align and provide a solution to the problem. HSG developers also need to provide appropriate rationale that justifies the development of guidance. In framing the policy problem that the HSG is intended for, it is important to consider alternative viewpoints of the policy issue and include linkages with other policy problems on the agenda. A clear articulation of the problem and its causes is critical as it is a key determinant for the remaining stages in the development of HSG.
Using a 7 points scale (1 =strongly disagree, 2=disagree, 3=disagree slightly, 4=neither disagree or agree, 5=agree slightly, 6=agree and 7=strongly agree), please rate this candidate item/criteria (problem definition) on the four traits below

Please choose the appropriate response for this concept	**1**	**2**	**3**	**4**	**5**	**6**	**7**
This concept is a defining feature (core component) of HSG							
This is an important concept to address in the DEVELOPMENT process of HSG							
This is an important concept in the APPRAISAL of health system guidance to differentiate between higher and lower quality guidance documents							
This is an important concept to be REPORTED in health system guidance							

13. Operationalization: The conceptualization and operationalization of the proposed recommendations for addressing the health system challenge should be well articulated. For example, useful HSG provides instructional support for their successful operation, and staff training that is commensurate with guidance expectations. Training recommendations could be in the form of a course, a workshop, accompanying manuals or consultancy services that staff can refer to during the implementation phase in order to standardize practice; there should be a link between the recommendations and the availability of these options. The operational strategies or mode of delivery of the options and how they will be put into practice should be well described. If technical assistance (research institutes, consulting firms, NGO's) is required, this should be identified and documented. Using a 7 points scale (1 =strongly disagree, 2=disagree, 3=disagree slightly, 4=neither disagree or agree, 5=agree slightly, 6=agree and 7=strongly agree), please rate this candidate item/criteria (operationalization) on the four traits below

Please choose the appropriate response for this concept	**1**	**2**	**3**	**4**	**5**	**6**	**7**
This concept is a defining feature (core component) of HSG							
This is an important concept to address in the DEVELOPMENT process of HSG							
This is an important concept in the APPRAISAL of health system guidance to differentiate between higher and lower quality guidance documents							
This is an important concept to be REPORTED in health system guidance							

14. Costs: The tentative budget required to implement the guidance recommendations should be clearly documented. It is crucial to state the potential costs (including downstream costs) of the operation so that decision makers can assess feasibility of the HSG implementation, as well as for assess the potentials of expanding or generalizing to other settings. This information will also help decision makers assess whether the cost of implementing the guidance will be worth its potential impacts.
Using a 7 points scale (1 =strongly disagree, 2=disagree, 3=disagree slightly, 4=neither disagree or agree, 5=agree slightly, 6=agree and 7=strongly agree), please rate this candidate item/criteria (costs) on the four traits below

Please choose the appropriate response for this concept	**1**	**2**	**3**	**4**	**5**	**6**	**7**
This concept is a defining feature (core component) of HSG							
This is an important concept to address in the DEVELOPMENT process of HSG							
This is an important concept in the APPRAISAL of health system guidance to differentiate between higher and lower quality guidance documents							
This is an important concept to be REPORTED in health system guidance							

15. Resources: The inputs and resources required to implement the recommendations need to be clearly defined. Some of these could be time, infrastructure, administrative capacity, information, equipment, supplies, healthcare professionals, training, the population etc. The guidance should provide a description of the amount, frequency and duration of the inputs and resources required for implementation.
Using a 7 points scale (1 =strongly disagree, 2=disagree, 3=disagree slightly, 4=neither disagree or agree, 5=agree slightly, 6=agree and 7=strongly agree), please rate this candidate item/criteria (resources) on the four traits below

Please choose the appropriate response for this concept	**1**	**2**	**3**	**4**	**5**	**6**	**7**
This concept is a defining feature (core component) of HSG							
This is an important concept to address in the DEVELOPMENT process of HSG							
This is an important concept in the APPRAISAL of health system guidance to differentiate between higher and lower quality guidance documents							
This is an important concept to be REPORTED in health system guidance							

16. Effectiveness: The guidance should report whether these recommendations have been effective in similar settings. They should report if the objectives were achieved, the range of the effectiveness and the sustainability of their effectiveness. In describing this effectiveness the guidance should make projections on how and why the objectives and goals will be achieved in the current setting. Effectiveness of HSG recommendations is a good health system performance indicator.
Using a 7 points scale (1 =strongly disagree, 2=disagree, 3=disagree slightly, 4=neither disagree or agree, 5=agree slightly, 6=agree and 7=strongly agree), please rate this candidate item/criteria (effectiveness) on the four traits below

Please choose the appropriate response for this concept	**1**	**2**	**3**	**4**	**5**	**6**	**7**
This concept is a defining feature (core component) of HSG							
This is an important concept to address in the DEVELOPMENT process of HSG							
This is an important concept in the APPRAISAL of health system guidance to differentiate between higher and lower quality guidance							

documents							
This is an important concept to be REPORTED in health system guidance							

17. Cost-effectiveness: Implementation of effective guidance recommendations has not been occurring mainly because of scarce resources, but also due to the fact that there is a dearth of convincing evidence from economic evaluations. Economically effective HSG options assist policy-makers in deciding which interventions represent the best value for money, thereby helping them properly allocate scarce resources and defining healthcare priorities. Economic evaluations traditionally report costs, direct and indirect program inputs and outcomes to guide health policy decisions and provide benchmark(s) or threshold(s) that the health system is willing to accept or support. Projecting the cost effectiveness of HSG will demonstrate whether the addressing a particular health system challenge compares favorably with other health system issues. Good quality guidance provides cost-effective recommendations in order to support and sustain a healthcare system. Using a 7 points scale (1 =strongly disagree, 2=disagree, 3=disagree slightly, 4=neither disagree or agree, 5=agree slightly, 6=agree and 7=strongly agree), please rate this candidate item/criteria (cost-effectiveness) on the four traits below

Please choose the appropriate response for this concept	**1**	**2**	**3**	**4**	**5**	**6**	**7**
This concept is a defining feature (core component) of HSG							
This is an important concept to address in the DEVELOPMENT process of HSG							
This is an important concept in the APPRAISAL of health system guidance to differentiate between higher and lower quality guidance documents							
This is an important concept to be REPORTED in health system guidance							

18. Benefits/harms: Judgments about the potential benefits, harms or a risk assessment of the HSG should be made. There should be considerations of the anticipated potential unintended consequences of the guidance and the populations or institutions that may experience significant impact should be identified in formulating the recommendations. Using a 7 points scale (1 =strongly disagree, 2=disagree, 3=disagree slightly, 4=neither disagree or agree, 5=agree slightly, 6=agree and 7=strongly agree), please rate this candidate item/criteria (benefits/harms) on the four traits below

Please choose the appropriate response for this concept	**1**	**2**	**3**	**4**	**5**	**6**	**7**
This concept is a defining feature (core component) of HSG							
This is an important concept to address in the DEVELOPMENT process of HSG							
This is an important concept in the APPRAISAL of health system guidance to differentiate between higher and lower quality guidance documents							

This is an important concept to be REPORTED in health system guidance							

19. Dissemination plan: Having an a priori dissemination framework of the HSG recommendations is crucial, with the mode of delivery, and the integrity of the avenue used for dissemination been properly covered. The proposed strategies for disseminating the HSG should be tailored to relevant audiences (a formal written report, policy brief, oral presentation, poster, press release, booklet, workbook, films, pocket card etc.). The dissemination plan should therefore contain how to best reach the target users of the guidance, in what form the guidance will be published. Having a solid dissemination plan will increase the likelihood that guidance recommendations will be incorporated into policy, programs, or practice. Using a 7 points scale (1 =strongly disagree, 2=disagree, 3=disagree slightly, 4=neither disagree or agree, 5=agree slightly, 6=agree and 7=strongly agree), please rate this candidate item/criteria (dissemination plan) on the four traits below

Please choose the appropriate response for this concept	**1**	**2**	**3**	**4**	**5**	**6**	**7**
This concept is a defining feature (core component) of HSG							
This is an important concept to address in the DEVELOPMENT process of HSG							
This is an important concept in the APPRAISAL of health system guidance to differentiate between higher and lower quality guidance documents							
This is an important concept to be REPORTED in health system guidance							

20. Process evaluation: Recommendations for evaluating the structure and process of HSG development and implementation as well as corresponding challenges should be made. This evaluation will examine to the extent to which the HSG recommendations were implemented as planned. This provides a way to monitor the quality and make adjustments and improvements to implementation strategies. This evaluation will look at what was done, how it was done, who was involved, who was reached, what inputs and resources where used and how they were used. It documents the services and activities that were implemented, as well as the policies and factors in place that either facilitated or hindered the process. Therefore, it can report on the strengths, weaknesses, opportunities and threats of the HSG and provide opportunities for understanding why the recommendations were successful or not.
Using a 7 points scale (1 =strongly disagree, 2=disagree, 3=disagree slightly, 4=neither disagree or agree, 5=agree slightly, 6=agree and 7=strongly agree), please rate this candidate item/criteria (process evaluation) on the four traits below

Please choose the appropriate response for this concept	**1**	**2**	**3**	**4**	**5**	**6**	**7**
This concept is a defining feature (core component) of HSG							
This is an important concept to address in the DEVELOPMENT							

process of HSG							
This is an important concept in the APPRAISAL of health system guidance to differentiate between higher and lower quality guidance documents							
This is an important concept to be REPORTED in health system guidance							

21. Outcomes/impact evaluation: An important quality criteria is being able to measure the results, or outcomes of the guidance recommendations in a way that determines whether the changes observed in relation to the health system challenge been addressed can be attributed to the HSG recommendations. The HSG should therefore provide recommendations for an outcome evaluation in order to assess the relationship between implementation of the guidance and observed effects. In this same vein, while the outcome evaluation will measures the change that has occurred as a result of the guidance, an impact evaluation will look at the long term deeper primary and secondary changes that resulted from the guidance. The impact evaluation will paint a picture of how the guidance might have affected the health system and the target population on a broader
scale.

Using a 7 points scale (1 =strongly disagree, 2=disagree, 3=disagree slightly, 4=neither disagree or agree, 5=agree slightly, 6=agree and 7=strongly agree), please rate this candidate item/criteria (outcome/impact evaluation) on the four traits below

Please choose the appropriate response for this concept	**1**	**2**	**3**	**4**	**5**	**6**	**7**
This concept is a defining feature (core component) of HSG							
This is an important concept to address in the DEVELOPMENT process of HSG							
This is an important concept in the APPRAISAL of health system guidance to differentiate between higher and lower quality guidance documents							
This is an important concept to be REPORTED in health system guidance							

22. Updating: The HSG documents should provide recommendations in the process of periodical updates as priority focus changes and new evidence is identified. The procedure to update the HSG should be provided with explicit timelines on anticipated review. An appropriate expiration date of the guidance should therefore be stated with an explanation of the rational for the proposed time frames. Setting time frames for periodic updates ensures that guidance producers revisit the recommendations and respond accordingly to potential health system changes.

Using a 7 points scale (1 =strongly disagree, 2=disagree, 3=disagree slightly, 4=neither disagree or agree, 5=agree slightly, 6=agree and 7=strongly agree), please rate this candidate item/criteria (updating) on the four traits below

Please choose the appropriate response for this concept	**1**	**2**	**3**	**4**	**5**	**6**	**7**
This concept is a defining feature (core component) of HSG							
This is an important concept to address in the DEVELOPMENT process of HSG							
This is an important concept in the APPRAISAL of health system guidance to differentiate between higher and lower quality guidance documents							
This is an important concept to be REPORTED in health system guidance							

23. Feasibility: The guidance recommendations should be pragmatic, realistic and should resonate with local social, political and economic conditions. It should be clearly demonstrated that the implementation of the guidance is feasible within the proposed practice environment. Recommendations should match local capacities and expectations. The goals (short, medium or long term) that are proposed should be feasible enough to lead to reasonable health system outcomes. The guidance should also describe facilitators and barriers to its implementation. Using a 7 points scale (1 =strongly disagree, 2=disagree, 3=disagree slightly, 4=neither disagree or agree, 5=agree slightly, 6=agree and 7=strongly agree), please rate this candidate item/criteria (feasibility) on the four traits below

Please choose the appropriate response for this concept	**1**	**2**	**3**	**4**	**5**	**6**	**7**
This concept is a defining feature (core component) of HSG							
This is an important concept to address in the DEVELOPMENT process of HSG							
This is an important concept in the APPRAISAL of health system guidance to differentiate between higher and lower quality guidance documents							
This is an important concept to be REPORTED in health system guidance							

24. Affordability: Upon considerations of costs and resources, the guidance should be affordable within the financial structure and budgetary allocations of the health system. Potential sources of local government funding and donor organizations should be identified. Donors can play a role in financing, advocacy, technical support and delivery of guidance recommendations, so it is important to identify those groups that can be potentially involved. For policy issues in which there may be several sources of funding, guidance should also assess the level of coordination among the donors and between the donors and the local government. Proper coordination will enhance the efficient use of available funding. Using a 7 points scale (1 =strongly disagree, 2=disagree, 3=disagree slightly, 4=neither disagree or agree, 5=agree slightly, 6=agree and 7=strongly agree), please rate this candidate item/criteria (affordability) on the four traits below

Please choose the appropriate response for this concept	**1**	**2**	**3**	**4**	**5**	**6**	**7**

This concept is a defining feature (core component) of HSG							
This is an important concept to address in the DEVELOPMENT process of HSG							
This is an important concept in the APPRAISAL of health system guidance to differentiate between higher and lower quality guidance documents							
This is an important concept to be REPORTED in health system guidance							

25. Flexibility: HSG should acknowledge the importance of professional judgment and discretion and provide recommendations that users can adapt in accordance with their own individual circumstances and needs. They should therefore be flexible enough so as not to limit inappropriately or unnecessarily limit those that apply them.
Using a 7 points scale (1 =strongly disagree, 2=disagree, 3=disagree slightly, 4=neither disagree or agree, 5=agree slightly, 6=agree and 7=strongly agree), please rate this candidate item/criteria (flexibility) on the four traits below

Please choose the appropriate response for this concept	**1**	**2**	**3**	**4**	**5**	**6**	**7**
This concept is a defining feature (core component) of HSG							
This is an important concept to address in the DEVELOPMENT process of HSG							
This is an important concept in the APPRAISAL of health system guidance to differentiate between higher and lower quality guidance documents							
This is an important concept to be REPORTED in health system guidance							

26. Socio-culturally acceptable: The guidance should adopt a socio-cultural perspective. Given the multicultural and social diversity in many regions, the guidance needs to systematically accommodate the diversity of values. It needs to recognize cultural expectations and provide an understanding of the role that culture plays in the success of the guidance recommendations. The socio-cultural factors that may potentially impact the development and implementation of the guidance need to be reported. The guidance should therefore provide socio-culturally appropriate solutions to health system issues. HSG that is insensitive to socio-cultural differences will compromise improvements in population health outcomes. Using a 7 points scale (1 =strongly disagree, 2=disagree, 3=disagree slightly, 4=neither disagree or agree, 5=agree slightly, 6=agree and 7=strongly agree), please rate this candidate item/criteria (socio-culturally acceptable) on the four traits below

Please choose the appropriate response for this concept	**1**	**2**	**3**	**4**	**5**	**6**	**7**
This concept is a defining feature (core component) of HSG							
This is an important concept to address in the DEVELOPMENT process of HSG							

This is an important concept in the APPRAISAL of health system guidance to differentiate between higher and lower quality guidance documents							
This is an important concept to be REPORTED in health system guidance							

27. Politically sound: For some knowledge users of HSG like managers and policy-makers, it is crucial that the guidance recommendations be politically pragmatic. Developers of HSG have to be aware whether there is political will to achieve the strategies and desired goals. Development and implementation of HSG can stir swings in the national mood, can lead to changes in the balance of organized forces, such as interest groups, and can influence outcome of events within the government, for instance an election. Therefore options proposed should be in sync with political commitments and interests in order for them to garner adequate support from top policy officials.
Using a 7 points scale (1 =strongly disagree, 2=disagree, 3=disagree slightly, 4=neither disagree or agree, 5=agree slightly, 6=agree and 7=strongly agree), please rate this candidate item/criteria (politically sound) on the four traits below

Please choose the appropriate response for this concept	**1**	**2**	**3**	**4**	**5**	**6**	**7**
This concept is a defining feature (core component) of HSG							
This is an important concept to address in the DEVELOPMENT process of HSG							
This is an important concept in the APPRAISAL of health system guidance to differentiate between higher and lower quality guidance documents							
This is an important concept to be REPORTED in health system guidance							

28. External factors: In developing and implementing guidance, those factors that lie beyond the formal fabric of the health system yet will influence the performance of its functions should be considered. These are aspects of the organization of other local institutional systems (effective judicial system, recession, tolerance of corruption, ethical codes of conduct, social system etc.). These external factors that originate from other systems but may affect the effectiveness of guidance recommendations should be noted. This will further enhance intra- and inter-sectorial collaboration within the entire system.
Using a 7 points scale (1 =strongly disagree, 2=disagree, 3=disagree slightly, 4=neither disagree or agree, 5=agree slightly, 6=agree and 7=strongly agree), please rate this candidate item/criteria (external factors) on the four traits below

Please choose the appropriate response for this concept	**1**	**2**	**3**	**4**	**5**	**6**	**7**
This concept is a defining feature (core component) of HSG							
This is an important concept to address in the DEVELOPMENT process of HSG							
This is an important concept in the APPRAISAL of health system guidance to differentiate between higher and lower quality guidance documents							

This is an important concept to be REPORTED in health system guidance							

29. Generalizability: When producing HSG, judgments should be made about the applicability of the recommendations beyond its original context or setting or population. This will ensure that settings with similar institutional, socio-economic, political demographics facing an identical health system challenges can adapt and use these HSG. Uptake of guidance recommendations in a wide variety of settings is a good quality indicator. Using a 7 points scale (1 =strongly disagree, 2=disagree, 3=disagree slightly, 4=neither disagree or agree, 5=agree slightly, 6=agree and 7=strongly agree), please rate this candidate item/criteria (generalizability) on the four traits below

Please choose the appropriate response for this concept	**1**	**2**	**3**	**4**	**5**	**6**	**7**
This concept is a defining feature (core component) of HSG							
This is an important concept to address in the DEVELOPMENT process of HSG							
This is an important concept in the APPRAISAL of health system guidance to differentiate between higher and lower quality guidance documents							
This is an important concept to be REPORTED in health system guidance							

30. Sustainability: One goal of a quality assessment of HSG is to achieve a continuous and sustainable improvement in the health system. Therefore, providing an indication of the sustainability of the effects of the guidance recommendations is important. Due to constantly evolving health system issues, looming budget cuts, fluctuating resources, rising costs of new technologies, an ageing population, shifting burdens of disease etc, it is crucial to develop recommendations that will stand the test of time and be maintained at an acceptable level. Using a 7 points scale (1 =strongly disagree, 2=disagree, 3=disagree slightly, 4=neither disagree or agree, 5=agree slightly, 6=agree and 7=strongly agree), please rate this candidate item/criteria (sustainability) on the four traits below

Please choose the appropriate response for this concept	**1**	**2**	**3**	**4**	**5**	**6**	**7**
This concept is a defining feature (core component) of HSG							
This is an important concept to address in the DEVELOPMENT process of HSG							
This is an important concept in the APPRAISAL of health system guidance to differentiate between higher and lower quality guidance documents							
This is an important concept to be REPORTED in health system guidance							

31. Overall Agreement: Using a 7 points scale (1 =strongly disagree, 2=disagree, 3=disagree slightly, 4=neither disagree or agree, 5=agree slightly, 6=agree and 7=strongly agree), please rate your overall agreement about the need for a high quality tool aimed to systematically appraise health system guidance (HSG) and contribute to the development and reporting of HSG.

Please choose the appropriate response for this concept	**1**	**2**	**3**	**4**	**5**	**6**	**7**
There is a need for a tool to direct the DEVELOPMENT of HSG							
There is a need for a tool to direct how to APPRAISE the quality of HSG							
There is a need for a tool to direct what to REPORT in HSG							

32. Additional concepts (items, criteria, or domain): Are there any additional concepts (items, criteria, domains) not covered in this survey, that you have identified to be relevant to directing the appraisal, development and reporting of HSG? Please write your answer here: [open ended response]

33. Any Comments? Having completed this survey, do you have any comments on the process and on the content of the candidate concepts presented? Please provide any comments or concerns you may have had when completing the survey or on the description/definition of the various concepts.

Please write your answer here: [open ended response]

34. Demographic questions

Please take a minute to answer the following demographic questions. Please write your answer(s) here: [open ended response]

- Gender
- Affiliation or organization
- Role or position
- Years of Experience
- Please indicate (Yes or No) if you will be interested to participate in, and be contacted for the next phases of this project

Thank you for completing our survey. Your valuable feedback and time is highly appreciated.

Appendix III: Comparison of the original concept labels and the new concept labels		
Original concept labels (CIS)	**Action**	**New concept label (post survey)**
Prioritization	Refined	Priority
Relevance	Refined	Relevant
Timeliness	Refined	Timely
Scope	Changed	Comprehensive
Transparency	Split into 2	Systematic
		Transparent
Evidence-based	None	Evidence-based
Stakeholder involvement	Changed	Participatory
Ethical	None	Ethical
Outcomes	Refined	Outcomes oriented
Competing interests	Refined	Interests managed
Presentation	Refined	Clearly presented
Problem definition	Refined	Defined problem
Operationalization	Changed	Operational options
Costs	Merged	Resources
Resources		
Effectiveness	None	Effectiveness
Cost-effectiveness	None	Cost-effectiveness
Benefits/harms	Refined	Benefits/harms weighting
Dissemination plan	None	Dissemination plan
Process evaluation	Merged	Assessment plan
Outcomes/impact evaluation		
Updating	Split into 2	Updating plan
		Up-to date
Feasibility	Refined	Feasible
Affordability	Refined	Affordable
Flexibility	Refined	Flexible
Socio-culturally acceptable	Refined	Socio-cultural alignment
Politically sound	Refined	Political alignment
External factors	Refined	External alignment
Generalizability	Changed	Transferable
Sustainability	Refined	Sustainable

Chapter 4: Health Systems Guidance Appraisal Concepts – Usability Testing

Denis Ako-Arrey[1]
Melissa Brouwers[1]
John N. Lavis[1,2]
Mita Giacomini[1]
The AGREE-HS team[*]

Word count:
480 (Abstract)
3432 (Full text)

1. McMaster University
2. Harvard T.H Chan School of Public Health

[*]The members of the AGREE-HS Scientific/advisory team include: Pierre Ongolo-Zogo (Cameroon); Jillian Ross, Mike Wilson, Padraig Warde and Sheila McNair (Canada); Yoalong Chen (China); Iván Darío Flórez Gómez (Colombia); Govin Permanand (Denmark); Fadi El-Jardali (Lebanon); Carmen Mihaela Dolea, Xavier Bosch-Capblanch and and Ulysses Panisset (Switzerland); Andy Haines, Francoise Cluzeau and Pablo Perel (United Kingdom); Luis Gabriel Cuervo (United States of America).

Preface

The paper presented in this chapter builds on the works presented in chapters 2 and 3, and addresses an important empirical gap by having intended users apply the HSG tool to actual HSG documents, and by seeking their comprehensive feedback on its performance and the experience. In this study, the selection of a HSG documents to evaluate was based on a mix of topics - delivery, financial and governance arrangements - in order to ensure that the HSG tool can be useful for a broad range of health systems challenges. The sampling approach ensured participants were recruited from all the WHO health regions and reflected the range of potential users. The work presented in this chapter led to the refinement of the beta version of the HSG tool. The final deliverable, AGREE-HS Version 1.0, is comprised of 4 domains and 32 concepts, with each item answered using a 5-point scale.

I was responsible for conceiving and designing this study along with my supervisor (Dr. Melissa Brouwers), and received guidance from members of my supervisory committee (Dr. John Lavis and Dr. Mita Giacomini), as well as from members of our core and expanded scientific team. I completed all recruitment, data collection, analysis, and interpretation and I also drafted the paper. Through an iterative process, my supervisor and members of my supervisory committee provided comments and suggestions on various drafts of the paper that were incorporated in revised versions of the paper.

Abstract

Background: Health systems guidance (HSG) provides recommendations, informed by a synthesis, interpretation and contextualization of available research evidence, aimed to solve health systems problems. They are designed for policy-makers. Currently, there is no gold standard aimed to evaluate HSG and guide its development and reporting. Earlier research identified thirty candidate HSG quality criteria (Ako-Arrey et al 2015a) that were judged by an international community of HSG researchers and users to be relevant to HSG (Ako-Arrey et al 2015b). Based on these data, a beta version of a tool, the AGREE-HS, was designed to enable evaluation of HSG and inform its development and reporting. The objective of this study was to test the beta version of the AGREE-HS, to assess its performance and to gather additional feedback for its final revision.
Methods: We selected an international group of stakeholders, representing each of the World Health Organization (WHO) regions, and with expertise in health systems and/or health policy, to participate. Participants were asked to (a) review one of 3 HSG documents (b) apply the beta version of AGREE-HS to appraise the document and (c) answer a series of questions about the appraisal process using the performance assessment scale. The AGREE-HS is comprised of 32 items; for each item, users indicate if the concept was covered (reported or documented) in the HSG document using a binary yes/no response scale. To gather feedback about the AGREE-HS experience, participants were asked to rate (a) for each AGREE-HS item, their agreement it was easy to understand and to apply; (b) for the whole AGREE-HS tool, their agreement it is a usable tool to direct development, reporting and evaluation of HSG; and (c) the appropriateness of the yes/no response scale. Survey responses were analyzed using Excel and SPSS and descriptive analyses were carried out. For exploratory purposes, mean AGREE-HS scores were calculated for each of the three evaluated HSGs.
Findings: Thirty-five participants were invited to complete the survey and we had a response rate of 74%. A favorable consensus was reached with participants agreeing that the AGREE items were easy to understand and easy to apply (all mean scores were above 5 on a scale of 7). In contrast, the appropriateness of the response scale (yes/no) was rated less favorably (mean score of 4.1). The overall agreement was high for the usability of AGREE-HS to systematically direct the development (85%), appraisal (92%) and reporting (81%) of HSG. Analysis of exploratory investigations showed that the quality of the HSGs reviewed varied as a function of the AGREE-HS item and specific report assessed. The feedback from the survey and deliberations with members of our scientific team led to the refinement of the AGREE-HS tool complete with 32 concepts (and their descriptions), 4 domains and a 5-points response scale.
Conclusion: The refined AGREE-HS tool (version 1.0) defines expectations of HSG and facilitates informed decisions among policy-makers on health systems delivery, financial, and governance arrangements.

1. Introduction

There is a growing need to strengthen health systems as countries face various health systems challenges which present governments, voters and consumers with increasingly difficult choices (Hurst, 2000; Frenk & Moon, 2013). Even between nations of comparable economic and social levels, wide variations in health outcomes exist partly as a result of how the health systems perform in response to these challenges (Preston, 1986; World bank, 1993). As a result, recommendations are increasingly being prepared to inform decision making related to how to tackle these challenges in terms of the health system delivery options (e.g. how the quality and safety of care is monitored), financial options (e.g. resources allocation), and governance arrangements (e.g. consumer and stakeholder involvement). The health systems recommendations are usually packaged in what is referred to as health systems guidance (HSG) developed by organizations, ministries and donors at local, national or global level (Lavis et al, 2002; Lomas et al, 2005; Bosch-Capblanch et al, 2012).

However, there has been limited understanding on how best to develop HSG and to translate it to policy while accounting for the complexity of health systems and the varied contexts in which health systems are embedded (Lavis et al, 2012). There are conceptual and methodological issues unique to HSG that have compromised scientific advancement in this area. This has included a historical lack of common nomenclature, no universally agreed upon definition of HSG quality or health system strengthening, and complexities related to the nature of available evidence about health systems etc. (Bosch-Capblanch et al, 2012). These challenges and viable strategies to overcome them are poorly understood at this time. As a consequence, there has been little progress at developing a tool for assessing the quality of HSG or to inform its development and reporting requirements. Currently, there is no instrument that has the capacity to discriminate between higher quality guidance that follows technical documentation from those of lower quality.

Overall, the health services research, health system and health policy communities lack a framework to ensure that optimal HSG are produced and implemented. There is, therefore, a need for a HSG tool that is of high methodological quality and that is developed through a transparent, evidence-informed decision-making processes. This tool will provide a conceptual model to direct the process of guidance development, provide gold standard quality criteria for appraising HSG and provide best practice strategies for reporting HSG.

More recently, there has been some progress to address these gaps in the health systems literature. Our program of research aims to improve the health systems enterprise by creating an instrument designed to direct the development, reporting and assessment of HSG. To this end, the program unfolded in three stages with each stage informing the next. In stage 1 (Ako-Arrey et al 2015a), the aim was to conduct a knowledge synthesis, using critical interpretive synthesis methods, to determine concepts (items, criteria or domains) related to HSG development, reporting and quality. From this review, we identified 30 candidate HSG quality criteria clustered into 3 domains. In stage

2 (Ako-Arrey et al 2015b), international stakeholders rated the importance, value and priority of the 30 candidate concepts and definitions generated from stage 1. In general, the mean ratings for all the candidate concepts were universally favorable, and a beta-version of the HSG tool (AGREE-HS) was generated. The results of these studies was a comprehensive framework that describes defining features of HSG and created a nomenclature upon which to describe, define and report these elements. Moreover, these studies provided the foundations that enabled the establishment of the beta version of the AGREE-HS.

The overall objectives for stage 3, this study, was to (1) test the usability and performance of the beta-version of the AGREE-HS tool, (2) further test the face validity of the HSG concepts and their definitions, and (3) test the anticipated value of the information it generates for users. The purpose of this paper is to report on stage 3 of the program of research.

2. Methodology

2.1. Overview

The object of analysis in this study was the beta version of AGREE-HS. Using structured survey methods, we sought data on stakeholders experience on applying the beta version of AGREE-HS on existing HSG. For exploratory purposes, we sought to assess the quality of the targeted HSG.

2.2. Participants

In this study, purposeful sampling targeting two key criteria, geography and knowledge user expertise, was used. We sought representation from each of the six World Health Organization (WHO) regions and from stakeholders involved in both the use and the production of HSG. We identified candidate participants from nominations by our international advisory panel team (scientific team) and from attendees at relevant international conferences (Global Symposium of Health Systems Research and the Guidelines International Network Symposium). Our goal was to recruit 5 participants from each of the WHO regions, and an additional 5 from the Americas health region.

2.3. Materials

2.3.1 *AGREE HS:* The beta version of AGREE-HS is designed to evaluate the quality of HSG and facilitate high quality development and reporting of these documents. It is comprised of 32 operationally defined items clustered into 4 domains. Each item is accompanied by an operational definition and a binary response scale (yes/no); users rate whether the concept reflected in each item is documented in the HSG being assessed. AGREE-HS scores can be compared within and across different HSG by adding up the total number of yes/no responses. For this survey, only 30 of the AGREE-HS items were used to appraise the HSG. Two of the items (implementation plan and evaluation plan) were excluded for this exercise as they only refer to the end users, and how they can

design a detailed implementation and evaluation plan at the local level for their individual contexts.

2.3.2 HSG Documents: For this study, we purposefully chose 3 WHO HSG documents from the McMaster University Health Forum's Health System Evidence database (www.healthsystemsevidence.org). We purposively sampled HSG documents to ensure that we had a mix of: 1) guidance addressing health system arrangements as the principal focus and addressing health system arrangements indirectly as a way to get the right mix of programs, services and drugs to those who need them, and 2) delivery, financial and governance arrangements. Multiple participants rated each HSG document; however the document was not a variable investigated in this study.

2.3.3 Feedback and demographics: The participants were asked to rate their overall agreement on the usability of the AGREE-HS tool as an instrument to systematically direct the development of HSG, to direct the appraisal of HSG and to direct what needs to be reported in HSG (Yes/No/Uncertain response scale). The participants were also asked to rate the performance of the AGREE-HS tool; were the concepts easy to understand, easy to apply, and was the Yes/No scale appropriate? (7-point scale, strongly disagree-strongly agree). The participants were asked to provide any additional comments on the survey process, on the content of the candidate concepts (operational descriptions/definitions) presented, and on the AGREE-HS tool (perceptions of its usefulness, appropriateness, ease of application). Demographic questions that captured the participants' gender, affiliation/organization, role/position, years of experience and previous participation in HSG development were also included.

2.4. Procedure

We sent letters of invitation to candidate participants soliciting their co-operation. In these letters, we provided a brief description and purpose of the study, likely time commitment, the survey process, the expected output as well as conditions for participating (see appendix II for sample letter of invitation). Ensuring a balance in terms geography and expertise, each individual who agreed to participate was purposefully assigned one HSG document. These participants were also e-mailed a password-protected unique identifier login to a Web-based study platform (LimeSurvey®) to complete a structured survey. Participants were asked to (a) review the HSG document to which they were assigned (b) review the beta-version of the AGREE-HS tool provided with complete descriptions of the items (c) apply the beta-version of the AGREE-HS tool to appraise the HSG document to which they were assigned (d) answer a series of questions about the appraisal process (i.e. feedback) and (e) provide demographic information. Over a period of 3 months (January to March 2015), we sent 2 reminders to the invited participants. The survey was initially pilot-tested by some members of the scientific research team as well as some selected health services/systems researchers, to enable refinements prior to its distribution to consenting participants. See Appendix III for a sample of the survey questionnaire.

2.5. Analysis

Survey responses were downloaded and analyzed using Microsoft Excel spreadsheets. Appropriate descriptive statistics were calculated for each of the question groups. The AGREE-HS scores were calculated (percentages of the yes/no responses) and compared within and across the HSGs. Usability of AGREE-HS was assessed by calculating the percentages of yes/no/uncertain responses for each of the development, reporting and evaluation metrics. Overall means were calculated on participants' ratings that the instrument was easy to understand, the instrument was easy to use, and the rating scale was appropriate. We reviewed the qualitative feedback received and performed a thematic analysis. Final decisions regarding the concepts and the generation of a refined AGREE-HS tool were made through consensus of the members of our scientific team.

3. Results

Thirty-five invitations to participate in the usability testing of the beta-version of the AGREE-HS tool were distributed, and 26 complete surveys were returned (response rate of 74%). Table 1 below shows the demographic details of the survey participants. While geographic representation was achieved, the majority of the respondents were men. In terms of expertise, most of our respondents were researchers in academia with expertise in health policy and health systems research, and we also had researchers from applied research institutes, as well as directors, managers and technical advisers either at local ministries of health or national or international health agencies. Participants' years of experience in their roles/position ranged from 3 to 40, and their years of health systems experience ranged from 2 to 33 years. Two-thirds of our respondents had not participated in the development of a HSG document.

Table 1:Demographic details
Gender Males = 19 (73%) Females = 7 (27%)
WHO Health Region Americas = 6 (23%) Europe = 5 (19%) Africa = 4 (15%) Western Pacific = 4 (15%) Eastern Mediterranean = 4 (15%) South-East Asia = 3 (12%)

Role/Expertise
Director = 5 (19%) Manager = 3 (12%) Technical Adviser = 6 (23%) Researcher (Academia) = 9 (34%) Researcher (Applied) = 3 (12%)
Years of Experience
1-4 years = 5 (19%) 5-9 years = 7 (27%) 10-19 years = 7 (27%) 20+ years = 7 (27%)
Years of Health Systems Experience
1-4 years = 10 (39%) 5-9 years = 10 (39%) 10-19 years = 2 (7%) 20+ years = 4 (15%)
Participation in HSG development
Yes = 9 (34%) No = 17 (66%)

Below in table 2 are the findings when respondents were asked to state their agreement (Yes/No) whether each of the 30 concepts on the beta-version of AGREE-HS was covered in the HSG document that they were provided with.

Table 2: Coverage of the AGREE-HS concepts in 3 HSG documents						
	HSG Document X**		**HSG document Y****		**HSG document Z****	
	Yes (%)	***No (%)***	***Yes (%)***	***No (%)***	***Yes (%)***	***No (%)***
1. Priority	88	12	100	0	100	0
2. Relevant	88	12	100	0	100	0
3. Timely	88	12	90	10	88	12
4. Comprehensive	37	63	70	30	75	25
5. Systematic	88	12	100	0	75	25
6. Transparent	88	12	90	10	12	88
7. Evidence-based	88	12	100	0	25	75
8. Participatory	88	12	90	10	25	75
9. Ethical	37	63	70	30	75	25
10. Outcomes oriented	37	63	60	40	75	25
11. Interests managed	100	0	80	20	0	100
12. Clearly presented	100	0	100	0	75	25
13. Up-to-date	100	0	90	10	75	25
14. Defined problem	100	0	100	0	100	0
15. Operational options	75	25	70	30	88	12
16. Effectiveness	88	12	70	30	12	88
17. Resources	25	75	30	70	88	12
18. Cost-effectiveness	12	88	30	70	63	37
19. Benefits/harm weighting	88	12	70	30	88	12
20. Dissemination plan	75	25	80	20	25	75
21. Assessment plan	25	75	10	90	63	37
22. Updating plan	63	37	70	30	0	100
23. Feasible	37	63	90	10	100	0
24. Affordable	75	25	70	30	63	37
25. Flexible	88	12	80	20	75	25
26. Socio-cultural alignment	75	25	80	20	75	25
27. Political alignment	100	0	70	30	88	12
28. External alignment	50	50	40	60	12	88
29. Transferable	75	25	80	20	25	75
30. Sustainable	75	25	30	70	88	12
OVERALL MEAN %	**72**	**28**	**74**	**26**	**62**	**38**

**The object of analysis was not the HSG, so we have denoted them here simply as HSG document X, Y & Z

While the primary objective of analysis of this study was the performance of the AGREE-HS itself and not the HSG document, table 4 above provides a summary of how the documents performed using this tool. As can be seen, quality varied as a function of the AGREE-HS item and varied as a function of the HSG being evaluated. For example, across the 3 HSG documents reviewed, higher quality was seen (as reflected with higher percentage of Yes responses) with the AGREE-HS concepts; priority (88%, 100%, and 100% for HSG document X, Y and Z, respectively), relevant (88%, 100%, 100%), timely (88%, 90%, 88%), and defined problem (100%, 100%, 100%). In contrast, lower quality was seen (as reflected with higher percentage of No responses) with the AGREE-HS concepts; cost-effectiveness (12%, 30%, 63% for HSG document X, Y and Z, respectively), assessment plan (25%, 10%, 63%), and external alignment (50%, 40, 12%).

Below in Table 3, we present the participants' responses relating to the performance of the beta version of the AGREE-HS tool. We recorded a favorable consensus (5-7 on the response scale) when respondents were asked whether the concepts were easy to understand (overall mean value of 5.9) as well as when they were asked whether the concepts were easy to apply (overall mean value of 5.6). The standard deviations and the ranges were low, and the modes and medians were both 6 for these 2 categories. For the question relating to the appropriateness of the scale, a favorable consensus was not obtained as we recorded a mean value of 4.1.

Table 3: Ease of applicability/understanding of AGREE-HS and appropriateness of the Yes/No scale

	Mean	SD	Mode	Median	Range
The AGREE-HS concepts are easy to understand	5.9	1.0	6	6	2
The AGREE-HS concepts are easy to apply	5.6	1.0	6	6	3
The scale (Yes/No) was appropriate	4.1	1.5	5	5	3

Table 4 below shows the overall agreement of the usability of AGREE-HS as a knowledge translation tool to systematically direct the DEVELOPMENT of HSG, to systematically direct the APPRAISAL of HSG and to systematically direct the REPORTING of HSG. The overall agreement among the respondents was high for all the 3 categories. The development, appraisal and reporting categories recorded an overall consensus of 85%, 92% and 81%, respectively.

Table 4: Overall agreement on the usability of AGREE-HS

	Yes	No	Uncertain
AGREE-HS can systematically direct the DEVELOPMENT of HSG	22 (85%)	0	4 (15%)
AGREE-HS can systematically direct the APPRAISAL of HSG	24 (92%)	0	2 (8%)
AGREE-HS can systematically direct what to REPORT in HSG	21 (81%)	0	5 (19%)

We received substantial qualitative feedback from the survey respondents mainly regarding the overall usability of the AGREE-HS tool, about the concepts and their descriptions/definitions, and about the Yes/No rating scale that was used. In general, the tool garnered support from the respondents who found it a useful instrument to differentiate HSG documents on the basis of quality and to support the development and reporting of HSG. The core members of the research team (n-4) reviewed the feedback received regarding the descriptions of the concepts and considered these for further refinement of the AGREE-HS tool. In terms of the scale, participants found that, in instances where there was more than one statement describing a concept, it was harder to apply the tool if the HSG document only partially covered that concept, making it challenging to squeeze the appraisal into the 2-items (Yes/No) response scale. To allow for these nuances to be considered, the respondents suggested either a 3-item response scale (Yes/No/Partially) or a Likert scale (strongly disagree to strongly agree). Feedback was incorporated into the tool to create the version 1.0 of the AGREE-HS tool (see appendix I).

4. Discussion

The findings from our survey of health policy/health systems stakeholders from the six WHO health regions showed a favorable consensus on the usability of the HSG candidate concepts generated from our previous study (Ako-Arrey, et al, 2015a) as it relates to their ease of understanding and ease of application. No additional concept was nominated, and no existing AGREE-HS concept was deleted. Also, we did not obtain a favorable consensus with regards to the Yes/No scale that was used to rate whether the concepts were covered (reported or documented) in the HSG documents. We also recorded that overall, there was a high agreement for the need for a HSG appraisal tool that can also further enhance HSG development and reporting requirements. This corroborates earlier positive findings on the face validity of the AGREE-HS tool obtained from our other study (Ako-Arrey et al, 2015b). Furthermore, we found that for all the 3 HSG documents that participants were asked to evaluate against the AGREE-HS tool, an absolute consensus (100%) was recorded in each case for some of the items, but in general, not all the 30 concepts were covered in any of the documents reviewed. The findings and the comments received on the appropriateness of the scale as well as the overall feedback on the tool and the appraisal exercise were reviewed by members of the scientific team and incorporated into the refinement of the beta-version of the AGREE-HS tool. The refined AGREE-HS version (appendix I) is made up of 32 items clustered

into 4 domains, and each answered with a 5-point response scale (1, strongly disagree to 5, strongly agree).

Some strengths of this study: first, the study applied a meticulous research approach, was conducted based on the findings/feedback of our two previous studies and incorporated the input of a broad-based advisory group. Secondly, we recruited qualified participants worldwide to ensure that the study resonates with low, middle and high-income countries. Thirdly, the blend of the 3 HSG documents selected for appraisal ensured that we were covering guidance related to delivery, financial, and governance arrangements in a health system. The 3 selected documents also ensured that there was a mix of HSG where the primary focus was programs, services and drugs or the health-system arrangements that determine whether programs, services and drugs get to those who need them. Lastly, we asked a wide variety of broad questions that permitted an understanding of the various dimensions of the usefulness of the tool as well as potential areas where issues may arise.

Some weaknesses of this study: first, participants who have more experience with HSG either as developers or users may have found the tool easier to apply as compared to respondents with little or no such experience. It would have been practical to provide guiding materials or a user manual, which further describes the concepts in detail with concrete examples, and provides explanations on how to use the tool. Secondly, because we had a small sample size, we could not conduct a factor analysis in order to see whether the AGREE-HS concepts cluster empirically. In the same line, our study was not powered enough to do a sub-group analysis to see whether there was a relationship between the survey responses and expertise or geography.

The next steps of our research program involve developing a user manual with explanations and detailed examples, as well as developing an on-line training program that will be useful for potential users of the tool. We will also proceed with further usability testing, reliability testing, validity testing and refinement of the AGREE-HS version 1.0 in order to generate the alpha-version ready for international unveiling and branding. Our goal is that, through this project we will contribute to bolster collaborations among global experts with a wide array of expertise, working towards a common health research goal; that of creating better quality and more implementable HSG that will improve critical decision-making and lead to stronger health systems for the benefit of patients and populations.

5. References

1. Ako-Arrey Denis., Brouwers Melissa., Lavis JN., & Giacomini Mita on behalf of the AGREE-HS Group (2015a). Health System Guidance Appraisal Concepts – A Critical Interpretive Synthesis.
2. Ako-Arrey Denis., Brouwers Melissa., Lavis JN., & Giacomini Mita on behalf of the AGREE-HS Group (2015a). Health System Guidance Appraisal Concepts – A Survey of Stakeholders
3. Bosch-Capblanch X, Lavis JN, Lewin S, Atun R, Røttingen JA, Dröschel D, Beck L, Abalos E, El-Jardali F, Gilson L, Oliver S, Wyss K, Tugwell P, Kulier R, Pang T, Haines A. Guidance for evidence-informed decisions about health systems (2012). Rationale for and challenges of guidance development. PLoS Medicine 2012; 9(3): e1001185.
4. Frenk, J., & Moon, S. (2013). Governance challenges in global health. *New England Journal of Medicine*, 368(10), 936-942.
5. Jeremy Hurst (2000). Challenges for health systems in Member Countries of the Organisation for Economic Co-operation and Development. *Bulletin of the World Health Organization*, 2000, 78 (6).
6. Lavis JN, Ross SE, Hurley JE, Hohenadel JM, Stoddart GL, Woodward CA, Abelson J (2002). Examining the role of health services research in public policymaking. *Milbank Quarterly*, 80:125-54.
7. Lavis JN, Røttingen JA, Bosch-Capblanch X, Atun R, El-Jardali F, Gilson L, Lewin S, Oliver S, Ongolo-Zogo P, Haines A (2012). Guidance for evidence-informed policies about health systems: 2) Linking guidance development to policy development. PLoS Medicine, 9(3): e1001186. doi:10.1371/journal.pmed.1001186. [PMID: 22427746]
8. Lomas Jonathan, Tony Culyer, Chris McCutcheon, Laura McAuley, & Susan Law. Conceptualizing and combining evidence for health system guidance: Final report. Canadian Health Services Research Foundation = Fondation canadienne de la recherche sur les services de santé, 2005.
9. Preston, SH (1986). Mortality and development revisited. *Population Bulletin of the United Nations* 1986, 18:34-40.
10. World Bank. World Development Report (1993): Investing in Health. New York: Oxford University Press for the World Bank, 1993.

Appendix I: Version1.0 of the AGREE for Health Systems (AGREE-HS) tool	
Process principles	
1. Priority	The guidance is properly aligned with current health system priorities from the perspective of topic, jurisdiction, health system level and population. The expression of the need and origin of the mandate for the guidance is clear.
2. Relevant	The guidance recommendations are relevant to, appropriate to and valid for the health system challenge, system or sub-system needs, the target population(s), and the setting in which they will operate.
3. Timely	The recommendations are available in a timely manner in relation to when the policy decisions are made or timely in relation to the health system issue being addressed.
4. Comprehensive	The guidance is comprehensive and covers all relevant/appropriate (direct and indirect) health system levels (e.g., district), sub-systems (e.g., mental health) and sectors (e.g., acute care)
5. Systematic	Systematic processes are applied in developing the guidance according to a specific plan and/or explicit methodologies.
6. Transparent	A transparent and reproducible approach in the development and reporting of the guidance is demonstrated.
7. Evidence-based	The best available and ideally most contextually relevant evidence informs the recommendations.
8. Participatory	The health system guidance team is comprised of multidisciplinary/multi-sectoral membership and includes those with an interest, stake or responsibility in the development, implementation and evaluation of the recommendations.
9. Ethical	The recommendations are considered within the lens of an ethical framework and align with applicable ethical principles and values (e.g. equity, equality, human rights, liberty, efficiency, autonomy, dignity, beneficence, etc.). The guidance adequately promotes fairness and equality in terms of age, ability, culture, gender, socioeconomic status, religion, occupation, language, ethnicity, race or sexual orientation among the target population.
10. Outcomes oriented	The guidance describes all the anticipated effects/outcomes as well as the appropriate indicators, performance thresholds, targets and standards that can be used to measure the effects/outcomes.
11. Interests managed	A declaration of competing interests from the guidance developers (e.g. financial, academic, professional, etc.) is identified and the strategies to manage them are described. It is also clear that the views of any funding body involved have not influenced the development process of the guidance.
12. Clearly presented	The recommendations are clear, user-friendly, succinct, unambiguous and presented in a readable and consistent format, with key recommendations easily identifiable.
13. Up-to-date	The recommendations are current and the evidence (e.g. systematic reviews) on which they are based is considered up-to-date.

Content	
14. Defined problem	The health system challenge and its causes are clearly articulated; specifically, the nature, causes, and magnitude, frequency or intensity of the problem, the populations and jurisdictions that are affected are clearly described.
15. Operational options	The recommended "solutions" are operationalized sufficiently with the conceptualization, operational guidance and the mode of delivery of the options clearly stated.
16. Effectiveness	Evidence of recommendation's effectiveness are described including methods used, context where tested, and results.
17. Resources	The inputs to and/or the costs of the implementation processes (amounts, frequency, duration) are described and are commensurate to the health systems issue; specifically, money, time, infrastructure, administrative capacity, information, equipment, supplies, healthcare professionals, training, etc. are considered.
18. Cost-effectiveness	The recommendations are attentive to value for money considerations with relevant cost-effectiveness evidence of recommendations described.
19. Benefits/harm weighting	Descriptions and/or judgments of the potential intended and unintended consequences (positive & negative) of the guidance on the population and/or the system are provided.
20. Dissemination plan	Methods for communicating guidance are clearly described and framed within an overall dissemination strategy.
21. Assessment plan	This involves high-level recommendations for assessing the structure and process of the implementation process as well as an assessment of the outcome/impact of the guidance to determine whether the course of action was a success or failure.
22. Updating plan	Recommendations for periodic updates are made and the procedure to update the guidance is provided with explicit timelines on anticipated review, appropriate expiration date of the guidance and an explanation of the rationale for the proposed time frames.
Context principles	
23. Feasible	The guidance recommendations are realistic and the actions are pragmatic. The guidance describes facilitators and barriers for implementation.
24. Affordable	The guidance recommendations are affordable within the financial structure and budgetary allocations of the health system.
25. Flexible	The guidance is flexible and adaptable to the expertise of the user and the varying local conditions in the context where implementation will take place.
26. Socio-culturally aligned	The recommendations adopt a socio-cultural perspective and are robust under societal and cultural scrutiny.
27. Political alignment	The political acceptability of the recommendations is considered and the degree of alignment with political interests and commitments are described.
28. External factors	Determinants of health system performance that lie outside the formal architecture of the health system but will influence the performance of its functions are considered and described (for example, judicial system, social system, recession, corruption, state of the economy etc.).

29. Transferable	A description of the degree to which recommendations are transferable to other similar or different regions and contexts is provided.
30. Sustainable	The anticipated sustainability and maintenance of long-term outcomes is described.
Implementation of HSG	
31. Implementation plan (end-users)	This involves the development of a strategic plan by end-users at the local level to describe the process of moving the recommendations into action. The plan may include a description of inputs, services and activities that are required for implementation; identification of the strengths, weaknesses, opportunities and threats to the implementation process; and allocation of responsibilities and duties. Designing an implementation strategy will facilitate adherence and compliance to planned activities, and enhance efficiency.
32. Evaluation plan (end-users)	A strategy for the monitoring and evaluation of the implementation strategy/process and/or outcomes of the guidance in a way that determines whether the changes observed in relation to the health system challenge being addressed can be attributed to the guidance is provided. There are also recommendations for an impact evaluation to look at the short and long term deeper primary and secondary changes that resulted from the guidance as well as corresponding challenges.

Appendix II: Sample email invitation for stage 3

Sample email
Subject: Research study invitation from McMaster University

Email text:

Dear {FIRSTNAME},

You are being invited to participate in our Health Systems Guidance Appraisal Tool (HSG-AT) project. We sincerely welcome your participation.

The HSG-AT (otherwise referred to as AGREE-HS) project is funded by the Canadian Institutes of Health Research (CIHR), and Drs Melissa Brouwers and John Lavis both of McMaster University, Hamilton, Canada are Principal Investigators. This project is co-ordinated by Denis Ako-Arrey who is a doctoral candidate in the Health Policy PhD program at Mcmaster University, Canada.

The goal of the project is to design a reliable, valid and useful tool for the appraisal of Health Systems Guidance that can also be used to support HSG development and reporting (Guidelines for health system's financial, delivery and governance arrangements).

This project unfolds in 3 Stages. We have completed a review of the literature (**Stage 1**), in which we generated a list of candidate concepts (items, criteria, domains) that are considered to be a good fit for, and may be relevant in the creation of the HSG Appraisal Tool (HSG-AT). We have also completed a structured survey (**Stage 2**), in which respondents systematically evaluated these candidate concepts for appropriateness to, relevance to, and priority for health system decisions and HSG. This led to the creation of a draft version of the Health Systems Guidance Appraisal Tool (HSG-AT), complete with thirty (30) quality appraisal concepts and their descriptions.

Given your background, we are seeking your participation for **Stage 3** of this project. During this stage, you will be required to (a) review one HSG document we will provide (b) apply the draft version of the HSG-AT to appraise the document and (c) answer a series of questions about the appraisal process. It is note-worthy to add that the object of analysis for this project is the HSG-AT and not the HSG document.

We are providing a stipend of $CAD200 as compensation upon your participation and receipt of your data. You may decline to answer any question(s) and may withdraw from the study for any reason, at any time**.** To ensure the confidentiality of your responses we are providing you with a unique identifier code to login into an online-based survey platform (Limesurvey). See the link at the end of this email.

We are allowing about six (6) weeks for respondents to complete the survey. If for any reason(s) this time requirement does not fit into your schedule, please let us know so we can accommodate.

If you have any questions regarding the study, please feel free to contact Denis Ako-Arrey at: akoarrde@mcmaster.ca or call (905) 379-0110 at any time.

Any questions regarding your rights as a participant may be addressed to the McMaster Research Ethics Board by calling collect the Office of Research Ethics Officer at (905) 525-9140 Ext 23142 or by emailing ethicsoffice@mcmaster.ca

Thank you, for your time and consideration.

Yours sincerely,

Dr. Melissa Brouwers (Associate Professor, McMaster University and Provincial Director, Program in Evidence-Based Care, Cancer Care Ontario)

Dr. John Lavis (Professor and Director of the Program in Policy Decision-Making, McMaster University)

Denis Ako-Arrey (PhD Candidate at McMaster University)

To participate, please click on the link below.

Click here to do the survey:
{SURVEYURL}

If you do not want to participate in this survey and don't want to receive any more invitations please click the following link:
{OPTOUTURL}

Appendix III: AGREE-HS usability testing survey questionnaire

Thank you for taking the time to participate in this survey for the creation of the Health Systems Guidance Appraisal Tool – AGREE-HS.

There are 37 questions in this survey

1. Priority
The guidance is properly aligned with current health system priorities from the perspective of topic, jurisdictional focus (e.g., all low-and middle-income countries, sub-Saharan Africa), health system level and population. The expression of the need and origin of the mandate for the guidance is clear.

Please state your agreement whether the criterion (Priority) was covered in this HSG document. Please choose **only one** of the following:

- ○Yes
- ○No

2. Relevant
The guidance recommendations are relevant to, appropriate to and valid for the health system challenge, system or sub-system needs, the target population(s), and the setting in which they will operate.

Please state your agreement whether the criterion (Relevant) was covered in this HSG document. Please choose **only one** of the following:

- ○Yes
- ○No

3. Timely
The recommendations are available in a timely manner in relation to when the policy decisions are made or timely in relation to the health system issue being addressed.

Please state your agreement whether the criterion (Timely) was covered in this HSG document. Please choose **only one** of the following:

- ○Yes
- ○No

4. Comprehensive
The guidance is comprehensive and covers all relevant/appropriate (direct and indirect) health system levels (e.g., district), sub-systems (e.g., mental health) and sectors (e.g., acute care)

Please state your agreement whether the criterion (Comprehensive) was covered in this HSG document. Please choose **only one** of the following:

- ○Yes
- ○No

5. Systematic
Systematic processes are applied in developing the guidance according to a specific plan and/or explicit methodologies.

Please state your agreement whether the criterion (Systematic) was covered in this HSG document. Please choose **only one** of the following:

- ○Yes
- ○No

6. Transparent
A transparent and reproducible approach in the development and reporting of the guidance is demonstrated.

Please state your agreement whether the criterion (Transparent) was covered in this HSG document. Please choose **only one** of the following:

- ○Yes
- ○No

7. Evidence-based
The best available and ideally most contextually relevant evidence informs the recommendations.

Please state your agreement whether the criterion (Evidence-based) was covered in this HSG document. Please choose **only one** of the following:

- ○Yes
- ○No

8. Participatory
The health system guidance team is comprised of multidisciplinary/multi-sectoral membership and includes those with an interest, stake or responsibility in the development, implementation and evaluation of the recommendations.

Please state your agreement whether the criterion (Participatory) was covered in this HSG document. Please choose **only one** of the following:

- ○Yes
- ○No

9. Ethical
The recommendations are considered within the lens of an ethical framework and align with applicable ethical principles and values (e.g. equity, equality, human rights, liberty, efficiency, autonomy, dignity, beneficence, etc). The guidance adequately promotes fairness and equality in terms of age, ability, culture, gender, socioeconomic status, religion, occupation, language, ethnicity, race or sexual orientation among the target population.

Please state your agreement whether the criterion (Ethical) was covered in this HSG document. Please choose **only one** of the following:

- ○Yes
- ○No

10. Outcomes oriented
The guidance describes all the anticipated effects/outcomes as well as the appropriate indicators, performance thresholds, targets and standards that can be used to measure the effects/outcomes.

Please state your agreement whether the criterion (Outcomes oriented) was covered in this HSG document. Please choose **only one** of the following:

- ○Yes
- ○No

11. Interests managed
A declaration of competing interests from the guidance developers (e.g. financial, academic, professional, etc.) is identified and the strategies to manage them are described. It is also clear that the views of any funding body involved have not influenced the development process of the guidance.

Please state your agreement whether the criterion (Interests managed) was covered in this HSG document. Please choose **only one** of the following:

- ○Yes
- ○No

12. Clearly presented
The recommendations are clear, user-friendly, succinct, unambiguous and presented in a readable and consistent format, with key recommendations easily identifiable.

Please state your agreement whether the criterion (Clearly presented) was covered in this HSG document. Please choose **only one** of the following:

- ○Yes
- ○No

13. Up-to-date
The recommendations are current and the evidence (e.g. systematic reviews) on which they are based is considered up-to-date.

Please state your agreement whether the criterion (Up-to-date) was covered in this HSG document. Please choose **only one** of the following:

- ○Yes
- ○No

14. Defined problem
The health system challenge and its causes are clearly articulated; specifically, the nature, causes, and magnitude, frequency or intensity of the problem, the populations and jurisdictions that are affected are clearly described.

Please state your agreement whether the criterion (Defined problem) was covered in this HSG document. Please choose **only one** of the following:

- ○Yes
- ○No

15. Operational options
The recommended "solutions" are operationalized sufficiently with the conceptualization, operational guidance and the mode of delivery of the options clearly stated.

Please state your agreement whether the criterion (Operational plan) was covered in this HSG document. Please choose **only one** of the following:

- ○Yes
- ○No

16. Effectiveness
Evidence of recommendation's effectiveness are described including methods used, context where tested, and results.

Please state your agreement whether the criterion (Effectiveness) was covered in this HSG document. Please choose **only one** of the following:

- ○Yes
- ○No

17. Resources
The inputs to and/or the costs of the implementation processes (amounts, frequency, duration) are described and are commensurate to the health systems issue; specifically, money, time, infrastructure, administrative capacity, information, equipment, supplies, healthcare professionals, training, etc. are considered.

Please state your agreement whether the criterion (Resources) was covered in this HSG document. Please choose **only one** of the following:

- ○Yes
- ○No

18. Cost-effectiveness
The recommendations are attentive to value for money considerations with relevant cost-effectiveness evidence of recommendations described.

Please state your agreement whether the criterion (Cost-effectiveness) was covered in this HSG document. Please choose **only one** of the following:

- ○Yes
- ○No

19. Benefits/harms weighting
Descriptions and/or judgements of the potential intended and unintended consequences (positive & negative) of the guidance on the population and/or the system are provided.

Please state your agreement whether the criterion (Benefits/harm weighting) was covered in this HSG document. Please choose **only one** of the following:

- ○Yes
- ○No

20. Dissemination plan
Methods for communicating guidance are clearly described and framed within an overall dissemination strategy

Please state your agreement whether the criterion (Dissemination plan) was covered in this HSG document. Please choose **only one** of the following:

- ○Yes
- ○No

21. Assessment plan
This involves high-level recommendations for assessing the structure and process of the implementation process as well as an assessment of the outcome/impact of the guidance to determine whether the course of action was a success or failure.

Please state your agreement whether the criterion (Assessment plan) was covered in this HSG document. Please choose **only one** of the following:

- ○Yes
- ○No

22. Updating plan
Recommendations for periodic updates are made and the procedure to update the guidance is provided with explicit timelines on anticipated review, appropriate expiration date of the guidance and an explanation of the rationale for the proposed time frames.

Please state your agreement whether the criterion (Updating plan) was covered in this HSG document. Please choose **only one** of the following:

- ○Yes
- ○No

23. Feasible
The guidance recommendations are realistic and the actions are pragmatic.The guidance describes facilitators and barriers for implementation.

Please state your agreement whether the criterion (Feasible) was covered in this HSG document. Please choose **only one** of the following:

- ○Yes
- ○No

24. Affordable
The guidance recommendations are affordable within the financial structure and budgetary allocations of the health system.

Please state your agreement whether the criterion (Affordable) was covered in this HSG document. Please choose **only one** of the following:

- ○Yes
- ○No

25. Flexible
The guidance is flexible and adaptable to the expertise of the user and the varying local conditions in the context where implementation will take place.

Please state your agreement whether the criterion (Flexible) was covered in this HSG document. Please choose **only one** of the following:

- ○Yes
- ○No

26. Socio-culturally aligned

The recommendations adopt a socio-cultural perspective and are robust under societal and cultural scrutiny.

Please state your agreement whether the criterion (Socio-culturally aligned) was covered in this HSG document.

Please choose **only one** of the following:

- ○Yes
- ○No

27. Political alignment
The political acceptability of the recommendations is considered and the degree of alignment with political interests and commitments are described.

Please state your agreement whether the criterion (Political alignment) was covered in this HSG document. Please choose **only one** of the following:

- ○Yes
- ○No

28. External factors
Determinants of health system performance that lie outside the formal architecture of the health system but will influence the performance of its functions are considered and described (for example, judicial system, social system, recession, corruption, state of the economy etc.).

Please state your agreement whether the criterion (External factors) was covered in this HSG document**.** Please choose **only one** of the following:

- ○Yes
- ○No

29. Transferable
A description of the degree to which recommendations are transferable to other similar or different regions and contexts is provided.

Please state your agreement whether the criterion (Transferable) was covered in this HSG document. Please choose **only one** of the following:

- ○Yes
- ○No

30. Sustainable
The anticipated sustainability and maintenance of long-term outcomes is described.

Please state your agreement whether the criterion (Sustainable) was covered in this HSG document. Please choose **only one** of the following:

- ○Yes
- ○No

31. Overall quality
Please rate your overall agreement on the USABILITY of the Health System Guidance Appraisal Tool (HSG-AT) as an instrument to systematically appraise health system guidance (HSG) and contribute to the development and reporting of HSG

Please choose the appropriate response for each item:

	Yes	Uncertain	No
The AGREE-HS can help direct the DEVELOPMENT of health system guidance	○	○	○
The AGREE-HS can help direct how to APPRAISE the quality of health system guidance	○	○	○
The AGREE-HS can help direct what to REPORT in health system guidance	○	○	○

32. Additional Feedback
Overall, the HSG-AT concepts were easy to understand. Please provide any comment(s) if necessary. Please choose **only one** of the following:

- ○Strongly Disagree
- ○Disagree
- ○Disagree Slightly
- ○Neither Disagree or Agree
- ○Agree Slightly
- ○Agree
- ○Strongly Agree

Make a comment on your choice here: [open ended response]

33. Additional Feedback
Overall, the HSG-AT concepts were easy to apply. Please provide any comment(s) if necessary. Please choose **only one** of the following:

- ○Strongly Disagree
- ○Disagree
- ○Disagree Slightly
- ○Neither Disagree or Agree
- ○Agree Slightly

- ○Agree
- ○Strongly Agree

Make a comment on your choice here: [open ended response]

34. Additional Feedback

Overall, the scale (YES/NO) used in this survey was appropriate. Please provide any comment(s) if necessary**.** Please choose **only one** of the following:

- ○Strongly Disagree
- ○Disagree
- ○Disagree Slightly
- ○Neither Disagree or Agree
- ○Agree Slightly
- ○Agree
- ○Strongly Agree

Make a comment on your choice here: [open ended response]

35. Additional comments
Having completed this survey, do you have any comments on the process and on the content of the candidate concepts presented? Please provide any comments or concerns you may have had when completing the survey or on the description/definition of the various concepts.
Please write your answer here: [open ended response]

36 Additional comments
Do you have any additional feedback about the tool? For example perceptions of its usefulness, appropriateness, ease of application etc.
Please write your answer here: [open ended response]

37. Demographic questions
Please take a minute to answer the following demographic questions. Please write your answer(s) here: [open ended response]

- Gender
- Affiliation or organization
- Role or position
- Years of Experience
- Years of health system experience
- Please indicate (Yes or No) if you have ever participated in the development of a HSG document

Thank you for completing our survey. Your valuable feedback and time is highly appreciated.

Chapter 5. Conclusion

In this thesis, I presented three original studies in Chapters 2-4 that collectively describe a program of research aimed to design and evaluate a knowledge translation tool for appraising guidance for health systems and further contribute to support their development and reporting requirements. In this concluding chapter, I first highlight the principal findings that can be drawn from each of the individual studies and from the thesis in general. I then discuss the major substantive, methodological and disciplinary contributions of the thesis to the field. Next, I discuss the strengths and limitations of the thesis. Finally, I provide implications that this thesis has on future research and provide recommendations for next steps.

5.1. Principal findings

This thesis presents an important first step towards the development of a set of concepts (items, criteria or domains) with which to assess the quality of HSG (guidelines for health system's financial, organizational, delivery and governance arrangements), and further enhance their development and reporting. It is believed that this is the first time a systematic examination of the conceptual underpinnings of HSG has been undertaken. In Chapter 2, I presented a review of the relevant literature in which I used a critical interpretative synthesis approach to develop a framework for HSG comprising 3 core domains and 30 concepts that are considered important in the development, appraisal and reporting of HSG. The 43 papers that met eligibility criteria reflected a variety of study designs and were the sources from which a candidate list of 30 concepts, complete with descriptions and operational definitions, was initially generated. Through an iterative process, the concepts were clustered into 3 meaningful categories (domains) in order to show relationships between the concepts within each domain, as well as across the domains. The 3 domains of the framework are process principle, content and context principles.

Under the process principles domain, the concepts that were clustered are those related to how the need for the guidance originated, who was involved in its development, and how they contributed. As a whole, this domain depicts the development integrity of the HSG and provides a conceptual understanding of the procedure. The 11 candidate concepts under this domain are: prioritization, relevance, timeliness, scope, transparency, evidence-based, stakeholder involvement, ethical, outcomes, competing interests and presentation.

Under the content domain, the concepts that were clustered are those that depict the content integrity of the recommendations and are components related to the constituents, make-up and structure of the guidance documents. The eleven candidate concepts under this domain are: problem definition, operationalization, costs, resources, effectiveness, cost-effectiveness, benefits/harms, dissemination plan, process evaluation, outcomes/impact evaluation, and updating.

Under context principles domain, the concepts that were clustered are those that

represent systemic and contextual components of the health system that can influence HSG recommendations and their adoptability. The eight candidate concepts under this domain are: feasibility, affordability, flexibility, socio-culturally acceptable, politically sound, external factors, generalizability, and sustainability.

Together, these 30 candidate concepts provided the foundation of the HSG tool and demonstrate interconnectedness between its three core domains. Specifically, it is proposed that HSG evolves through a series of interconnected stages or phases starting from when the challenges that the health systems face initially rise to the attention of policy-makers and the guidance developers identify and describe these issues. It is further proposed that guidance developers, together with policy-makers, then locate and describe options for addressing these challenges, prescribe selected recommendations and options, get the buy-ins from influential stakeholders, implement the recommendations and then monitor and evaluate the implementation efforts. We hypothesized that each of the concepts could be relevant to the goals of better HSG development, improved HSG reporting, and valid process by which HSG could be assessed.

In Chapter 3, I presented a survey of international stakeholders with expertise in health policy and health systems and asked them to use a likert scale to systematically test our hypothesis and evaluate the 30 candidate concepts generated from the study in Chapter 2 and identify any missing concepts. Data was collected from 41 participants (82% response rate) representing the six WHO health regions. Overall, the mean ratings and standard deviations showed that the participants viewed the candidate concepts positively. Approval ratings were recorded for each of the candidate concepts for all the four outcome measures that were investigated; (i) concept being a core component of HSG (ii) concept being important in HSG development (iii) concept being important in the appraisal of HS and (iv) concept being important in reporting of HSG. There was also a high consensus from the respondents on the overall agreement about the need for a HSG tool. These survey results along with substantial qualitative feedback from the participants and members of our scientific team led to the refinement of the candidate concepts' labels and descriptions and generated a beta version of the HSG tool (AGREE-HS) comprised of four domains and 32 items that was ready for initial testing.

In Chapter 4, I presented another survey of international health policy and health system experts from the six WHO health regions during which participants were asked to test the performance and usability of the AGREE-HS tool generated in Chapter 3. To this end, participants used the beta version of the AGREE-HS tool to evaluate the quality of a HSG document and then completed a survey questionnaire about the evaluation process and the usability of the AGREE-HS tool. Thirty-five stakeholders were recruited for this survey and completed responses were received from 26 participants (74% response rate). Three international HSG documents were selected for this exercise based on their coverage of either a delivery and/or financial and/or governance arrangement issue. Participants rated that the concepts were easy to understand and the tool was easy to apply. Moreover, the findings showed that participants were in support of the need for a tool for HSG development, appraisal and reporting, which further confirmed the findings reported in Chapter 3. However, participants were less in favor of the use a Yes/No scale

for some of the questions related to the usability of the tool, and suggested a modification to a broader more user-friendly scale. The findings from the survey, the qualitative feedback from the participants and consultations with members of our scientific team led to further refinement of the beta version of the AGREE-HS tool to generate the version 1.0 of AGREE-HS. Specifically, while no additional items were added or deleted, some of the wording and descriptions were refined and the response scale was switched to a 5-points likert scale (1, strongly disagree to 5, strongly agree).

5.2. Study Contributions

Collectively, the three original scientific contributions in this thesis represent a first attempt to fill significant research gaps by designing a tool for the evaluation of the quality of HSG developed at local, national or international levels, and by further supporting the process of development and reporting of guidance related to health systems issues. While considerable advancements have been made regarding the science and practice of clinical practice guidelines (i.e., guidance documents that target clinical questions and provide recommendations relevant to clinician and patient decisions) the same cannot be said for HSG. Good quality research on health systems is relatively scarce (El-Jardali et al, 2012), encompasses a wide variety of research methods that may be more and less familiar with stakeholders (Lavis et al, 2015) and often lacks enough detail to understand how the investigated systems' components influence the reported findings (Blanchet et al, 2014). Also, there tends to be a degree of uncertainty about their effects (Green & Bennett, 2007), with implementation issues (rather than the process of development) usually perceived as more relevant. For example, policy decisions about a health system often encompass multiple interventions packaged into a particular policy with little evidence to support the multiple interventions (Bosch-Capblanch et al, 2012). Consequently, understanding the various factors that may inform HSG recommendations, say in contrast to clinical scenarios, is relatively new. So far, there has been little (or no) theoretical or empirical work that seeks to incorporate all these health systems factors in order to optimize guidance recommendations. This paved the need for rigorous research on the HSG enterprise, and this thesis begins to provide foundations to this area of scholarship and particularly on how to enhance HSG recommendations. The work presented in this thesis consists of substantive contributions that provide a better theoretical and empirical understanding of HSG development, appraisal and reporting; methodological contributions providing a range of approaches that can be adopted by others for designing a guidance appraisal tool; as well as disciplinary contributions to the field of knowledge translation, health systems and health policy.

5.2.1. Substantive contributions

The research done in Chapter 2 contributes a new framework for HSG concepts that depicts the comprehensive array of essential elements of a HSG document. The framework consists of interconnected concepts that are considered important for, and can be used to, appraise HSG quality as well as support their development and reporting. This framework ensures that HSG is relevant to the health system's developmental plan and priorities of the region or country as knowledge users may be more interested in

responses to questions like "what can work in our environment" (Peters & Bennett, 2012). It is important to recognize that HSG recommendations vary with respect to relevance or importance across individuals in the health system. As such, actors in the field can prioritize those AGREE-HS concepts that are more or less applicable to their particular roles or context. For example, technical staff may tend to focus more on the feasibility of implementation, managers may focus on what is least expensive, while, policy-makers may pay more attention to cost-effectiveness and political soundness (Bosch-Capblanch & Allen, 2012). Other crucial considerations are contextual characteristics like path dependency, socio-cultural norms, economy, and country history, which may impact how, and whether health systems achieve good health efficiently (Balabanova et al, 2011). The common analytic framework for health systems that the model for HSG concepts provides, gives direction on how one should approach, reflect upon, adapt, adopt and interpret key factors in a systematic fashion to enable the development and assessment of high quality and justifiable health system options based on evidence.

Chapter 3 also provides substantial contribution with the generation of a beta version of the AGREE-HS tool, complete with 32 concepts and their descriptions, and the multiple roles the tool can play (development, reporting, and appraisal). Contributions from health systems and health policy experts from all over the world provided a face validation of the importance of these concepts and the value of the AGREE-HS tool. This tool can provide much needed support to enhance local and global capacities for staff charged with developing and implementing guidance on a wide variety of health system topics. It promotes a common knowledge of the key components of a health system, a common nomenclature to communicate by, and highlights which parts of a health system are critical in enhancing health system strengthening.

Chapter 4 further provides a substantive contribution related to the design of a refined version of AGREE-HS tool (version 1.0) and further confirms the face validation of the concept and usefulness and applicability of the tool in low, middle and high-income countries. The strengthening of research for health through the dissemination and translation of knowledge, the promotion of health research governance, and the monitoring of standards within research practices, are all key components for health system strengthening that need to thrive in order to responds to the needs of the population. An instrument like AGREE-HS can be a key driver to help ensure that good quality recommendations for appropriate options/action are available for adoption and through their application will establish consistency in standards of practice across various settings. There are many conceptual and contextual challenges associated with the production and use of HSG, so it is important to have gold standard quality criteria that will render HSG credible. HSG provides recommendations on how to approach health system challenges, so having a tool that can be used to appraise the quality of the guidance being produced is important in maximizing health system efficiency.

5.2.2. Methodological contributions

The three original research studies presented in this thesis contribute a clear methodological approach to the development of an appraisal tool in general and one focused on health systems issues in particular. The methodology was logically sequential with Chapter 2 informing Chapter 3, and Chapters 2 and 3 informing Chapter 4. The approach used in these 3 studies to design the AGREE-HS tool are adapted from the methodological, conceptual and theoretical principles of measurement construction used to design the AGREE II tool, but this adaptation is performed in a number of original ways. Chapter 2 presents a first attempt at using a critical interpretive synthesis (CIS) approach to concept generation for the purposes of designing an appraisal tool. With the design of the original AGREE tool (Cluzeau et al, 1999), a traditional systematic review was undertaken for concept generation, as that was the most common knowledge synthesis method available at the time. For the AGREE REX tool (Brouwers et al, 2015), a realist review was conducted given the study objectives demonstrating some advances to knowledge synthesis techniques and a better alignment to its overall study goals. Carried further, for the AGREE-HS, to match the study goals, a CIS was the most appropriate and contemporary method of knowledge synthesis. Given the quantity, quality and comprehensiveness of studies on HSG and the heterogeneity in the available literature, the CIS approach was well suited for the review and analysis of this type of data. CIS applies a relatively loosely defined set of processes for critically analyzing and synthesizing literature The objective of a CIS is to develop new concepts and theories through a typically interpretive mode of inquiry and was well suited for this study as our goal was to generate a candidate list of concepts that will serve as the foundation for the HSG tool. The use of the CIS approach in this study provides a good methodological foundation for current and future scholars engaged in the design of appraisal tools in other fields (old or emerging) where literature is scarce and/or diverse.

Chapters 3 and 4 both present a first endeavor to survey international health systems and health policy experts on the topic of HSG and on the design of a global HSG tool that will serve the purpose of this community. The recruitment of experts in health system/health policy either as, knowledge users (clinical leaders, healthcare executives, policy-makers), HSG developers, health policy and health systems experts, as well as researchers from all the WHO health regions was a distinctive sampling approach that can ensure that the AGREE-HS tool will garner sufficient support from around the world. Members of the advisory group and participants for the survey were chosen in order to bring a jurisdictional and setting perspectives into the discussion and to provide a diversity of views from international producers and users of HSG. Applying this integrated KT science approach provided a forum for the design of the tool, co-created between target HSG producers/user communities and the research communities. Considering the huge role that context plays in HSG, and with a goal to create a tool that was relevant and generalizable to varied circumstances, ensuring these geographical and multi-perspective representational goals were achieved was very important. Chapter 4 presents the first time HSG documents have been appraised using an evaluation tool designed specifically for this purpose. While AGREE II has shown capacity to discriminate between high and poor quality guidelines, the instrument has a clear focus

on clinical practice guidelines (CPGs) and it is, therefore, not well adapted for the purpose of HSG. Additionally, while the strategic selection of respondents ensures the need for the tool to resonate with countries at all income levels, the decision to use AGREE-HS to appraise three HSG documents that cover 3 distinct health systems areas (rural retention of health workers, task shifting of health services and introduction of a new vaccine into a country's immunization program), further ensures that the tool is designed to address the needs of a wide variety of health systems challenges.

5.2.3. Disciplinary contributions to the field of health systems research

The three original research studies presented in this thesis also provide disciplinary contributions particularly to the emerging field of HSG and generally to the field of health systems research and the global efforts to address health systems challenges by developing appropriate guidance recommendations. This thesis comprised of an integration of study designs and insights from the fields of health services research, health systems research, knowledge translation (KT), political science and particularly clinical practice. The HSG enterprise seems to be where the CPG enterprise was approximately 20 years ago; an exciting field with the potential to make an impact on health systems and outcomes but requiring research on methods and knowledge translation processes to optimize its potential. This project addresses this status by integrating insights into HSG from the rigorous practice that is common in the clinical field. Also, by working closely with a large group of international stakeholders, these studies sought to strengthen the relationships between researchers, HSG developers, HSG users and system leaders to promote and enhance the use of guidance in health systems and policy development. System leaders and policy-makers can use the AGREE-HS and the HSG process as a vehicle by which evidence and context can be better understood; its strengths and limitations can be more easily identified; and its application to health system challenges can be more readily undertaken. Development, appraisal and/or reporting of HSG are the activities by which these capacities can be advanced within the discipline. Additionally, as it relates to the scientific advancement, KT research in the area of health systems and health policy is relatively understudied compared to advancements made targeting clinicians and patient. This thesis was geared to develop a tool to measure the quality and implementability of HSG, a KT intervention. By generating a new knowledge translation (KT) tool, AGREE-HS, this thesis had advanced the science and practice of KT.

5.3. Strengths and limitations

Taken together, the three studies presented in this thesis have several strengths.

(1) The main focus of this thesis is HSG, which is increasingly being utilized as a health policy product that can provide appropriate recommendations for addressing a health systems challenge. As local, national and international organizations like the WHO and ministries of health are producing HSG, so have the interests in supporting their development, appraisal and reporting being rising. Consequently, this thesis can be considered an important contribution to the literature and an advancement that fills a high priority gap in the field. Supporting countries and organizations worldwide in developing

and using high quality guidance is a practical relevance of this thesis that I articulate in each of the studies.

(2) Considering the reality that the research on HSG is relatively scarce, and by using a multidisciplinary approach which incorporates insights from clinical practice, health services research, health systems research, knowledge translation, and political science, the three original studies in this thesis provide important steps in further enriching the research relevant to the HSG enterprise. Additionally, the three studies capitalize on advances already achieved by these other fields, relying on these existing theories and practices and leveraging these to create a useable tool that balances rigor with feasibility in a nascent field like HSG; this will further accelerate the process of uptake and implementation of guidance recommendations within the HSG community.

(3) The major deliverable for this thesis was AGREE-HS, a tool that will enhance the development, appraisal and reporting of HSG at all levels of government and organizations whether locally, nationally and internationally. The AGREE-HS tool will be the first tool of its kind to facilitate the timely use of relevant and high quality HSG by policy-makers internationally. This unique contribution to the field of health systems will promote health systems strengthening efforts globally and further curb the growing concerns around how to address health systems challenges. These studies contribute to providing the means to the ultimate goal of creating better quality and more implementable HSG that will lead to stronger and more sustainable health systems for the benefit of populations and patients.

(4) The mix of methodological approaches used in the 3 studies all constitutes a first for studying HSG. Our methodology was sequential (one stage led to the next), differentiated (each stage represented a distinctive study required to move to the next stage) and cumulative (each study produced data that fed into the overall process). The novel use of the critical interpretive synthesis approach to knowledge synthesis of concepts for the design of an appraisal tool, made it possible to review and analyze a methodologically diverse body of literature. The use of structured surveys of experts drawn from all WHO health regions not only represented a first attempt to systematically collect their viewpoints on this topic, but also presented the possibility to reach a wider audience and amass global endorsement for the new tool. The methodology and research approach used in these studies will be of interest to health policy and health systems scholars, while the findings of this thesis will be particularly relevant to producers and knowledge users of HSG.

(5) The three studies involved an iterative, highly collaborative process and several deliberations with members of our core and expanded scientific team comprised of investigators and collaborators with an extensive knowledge in health systems and policy research. This led to the comprehensive nature of the tool and the clarity of the concise definitions of each of the concepts. Using our integrated KT approach, these studies provide an important step in facilitating a culture of stakeholders within the health systems and policy communities who are receptive to the role of evidence, understand the benefits and limitations of evidence, and are more apt to use evidence in the types of decisions they make.

As a whole, this thesis also has some limitations.

(1) The nature of the HSG enterprise meant that the literature available is sparse,

diverse and complex. Consequently, the literature that led to the creation of AGREE-HS was chosen mostly based on a judgment of their likely contribution to the design of the tool, because, some methodologically poor papers appeared theoretically and conceptually pertinent. Given the rudimentary nature of the HSG field, decisions about which papers to include were made based on relevance rather than methodological rigor.

(2) While HSG is produced at local, national and international level, in this thesis, the focus was more on guidance that is being produced globally and adapted for local use. Global guidance is produced typically for universal use by low, middle and high-income countries and may therefore not consider the contextual characteristics (economic, social, political, cultural etc.) of the different income groups. It would have therefore been useful to draw on HSG from local, national and international sources produced for or by all the income groups so as to adequately tailor the AGREE-HS tool to suit the specific needs of the various local and national actors.

(3) Some study design decisions framed in the research protocol can be considered limitations in some of the individual studies For example, the small sample sizes for Chapters 3 and 4 meant that neither a factor analysis (to determine whether the concepts in the tool cluster empirically) or a sub-group analysis (to determine if variations were recorded across the respondents in terms of geography or expertise) could be performed. The studies were designed in this manner mainly to initially investigate generalizability of the findings and also to meet the expectations and timeline of a doctoral thesis; these limitations will be addressed in next steps of this program of research.

(4) HSG is not the only "vehicle" that drives policy making by helping decision makers address health systems challenges. So while HSG is considered important in policy making, it would be better to articulate it not as an independent instrument, but one that is complemented by others like evidence briefs, policy briefs, deliberative dialogues, etc. Moreover, as in the CPG realm, knowledge tools such as HSG or evidence briefs are likely to be more impactful and lead to greater change if coupled with effective implementation interventions. This will be worth investigating further. Nonetheless, the findings from these studies can be adapted and applied to support the process of development, appraisal and reporting of other policy instruments.

5.4. Future research and next steps

While this thesis has addressed numerous gaps in the research literature by mounting a foundation for HSG development, appraisal and reporting, some important areas for future research were also identified:

Additional assessments (testing the importance of the concepts, clarity and appropriateness of the revised response scale, criterion validity, reliability, usability and validity testing of the tool etc.) of the version 1.0 of AGREE-HS will be performed with a larger international sample in order to produce sufficient power for the performance of factor analysis. This is an important step in the development of a measurement tool, as it will provide data to determine the number of domains in the tool and how the assessment scores should be calculated. A larger sample size will also allow for sub-group analyses to be performed across the different experts from the different regions.

A current status report of existing HSG documents will also be performed. Here we will strategically select HSG covering a wide variety of topics and produced in low, middle and high-income countries as well as from international organizations, and evaluate their quality using the AGREE-HS tool in order to generate AGREE-HS scores for these internationally diverse documents. This will provide a useful baseline from which future HSGs can be compared and used as one means of building capacity, strength and expertise in this area.

We will also target the development and reporting goals of the AGREE-HS tool for further investigation by bringing it specifically to the HSG developers in the field to see how it performs (is it a good organizational framework, how does it complement tools/methods that already exist, can it enhance confidence, speed and reliability to HSG development, can it be used to refine current methodological strategies for HSG development etc.).

The findings from these three original studies are intended for publication in local and international peer-reviewed journals to facilitate access for global benefits. The findings are also intended for dissemination through reports and presentations at local and international meetings, seminars, workshops and conferences. This research project and the scientific team involved are well positioned within the HSG research, development, and user communities to facilitate the uptake of the findings by relevant stakeholders. The AGREE Enterprise Website (www.agreetrust.org), will be the "home" for users to have public access to the AGREE-HS and any related projects. The AGREE-HS tool will be promoted so it can be used to facilitate development, reporting and evaluation of HSG within national and international agencies and organizations that are making forays into the area of HSG. It will also be promoted internationally to groups who already develop HSG and who collate HSG within on-line system directories. In addition, an on-line training program and a user-manual will be developed to orient new users to the new AGREE-HS tool and to ensure its optimal application. The AGREE-HS tool will be formally integrated into the Knowledge Translation course of the Health Research Methodology graduate program at McMaster University, Canada. It will also be promoted to similar courses offered by Universities in Canada and abroad.

While more work still needs to be done to further refine and promote the tool, this thesis has paved the path to understanding how the use of HSG can be optimized by enhancing their development, reporting and especially their quality. This thesis and its major deliverable, the AGREE-HS tool will begin to provide an avenue that enables debate and critical appraisal about HSG through the evaluation of guidance documents and it will also provide a framework that can be used in the development and reporting of future HSG efforts and ultimately contribute to stronger health systems globally.

References

1. Balabanova D, McKee M, Mills A. 2011. Good Health at Low Cost 25 Years On: What Makes a Successful health System?. London: London School of Hygiene and Tropical Medicine.
2. Blanchet, K., Palmer, J., Boggs, D., Palanchoke, R., Jama, R., & Girois, S. (2014). Advancing the application of systems thinking in health: analysing the contextual and social network factors influencing the use of sustainability indicators in a health system-a comparative study in Nepal and Somaliland.Health Res Policy Syst, 12, 46.
3. Bosch-Capblanch X, Lavis JN, Lewin S, Atun R, Røttingen JA, et al. (2012) Guidance for evidence-informed decisions about health systems: rationale for and challenges of guidance development. PLoS Med 9: e1001185.
4. Bosch-Capblanch X., Allen, C (2012). Health Systems Strengthening and Conflict: Transformation in Fragile States. Retrieved on August 17th 2013 from http://www.evidenceaid.org/wp-content/uploads/2013/05/MMI_for_Evidence_Aid_XBC_revised_by_MC.pdf
5. Brouwers, M., Bhattacharyya O., Kastner M., Hanna, Steve., Gagliardi Anna., Graham, Ian., et al (2015). Innovations to enhance the capacity of practice guidelines to improve health and healthcare systems: AGREE- REX (Recommendation EXcellence). Under review.
6. Cluzeau F, Littlejohns P, Grimshaw J, Feder G, Moran S (1999) Development and application of a generic methodology to assess the quality of clinical guidelines. Int J Qual Health Care. 1999;11:21-8.
7. El-Jardali, F., Lavis, J. N., Ataya, N., & Jamal, D. (2012). Use of health systems and policy research evidence in the health policymaking in eastern Mediterranean countries: views and practices of researchers. Implement Sci,7(2).
8. Green, A., & Bennett, S. (2007). Sound choices: enhancing capacity for evidence-informed health policy. Retrieved on March 3rd 2015 from http://www.who.int/alliance-hpsr/resources/Alliance_BR.pdf.
9. Peters, D. H., & Bennett, S. (2012). Better guidance is welcome, but without blinders. PLoS medicine, 9(3), e1001188.