An Investigation into the Quality of Clinical Practice Guidelines, the Developers who Produced Them, and an Inside Look Into the Guideline Development Process of Regulatory Bodies

Lyane Trépanier

Department of Educational and Counselling Psychology

McGill University, Montréal

June 2019

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy in Counselling Psychology

Table of Contents

List of Tables

List of Abbreviations

AMA	American Medical Association
APA	American Psychological Association
ASD	Autism Spectrum Disorders
CMA	Canadian Medical Association
CMQ	Collège des médecins du Québec (The College of Physicians of Quebec)
COI	Conflict of interest
CPA	Canadian Psychological Association
CPG	Clinical practice guidelines
EBP	Evidence-based practice
INESSS	Institut nationale d'excellence en santé et en services sociaux
IOM	Institute of Medicine
NICE	The National Institute of Clinical Excellence
OCCOQ	Ordre des conseillers et conseillères du Québec (The Ordre of Counsellors of Quebec)
OPPQ	Ordre des psychoéducateurs et psychoéducatrices du Québec (The Ordre of Psychoeducators of Quebec)
OPQ	Ordre des psychologues du Québec (The Order of Psychologists of Quebec)
OTSTCFQ	Ordre des travailleurs sociaux et des thérapeutes conjugaux et familiaux du Québec (The Ordre of Social Workers and Family Therapists of Quebec)
SIGN	The Scottish Intercollegiate Guidelines Network
WHO	World Health Organization
WoK	Web of Knowledge (Web of Science)

Abstract

Every day, clinical decisions are made in diverse practice and healthcare settings. Unlike procedural manuals that provide concise instructions on diagnostic screenings, appropriate interventions, treatment duration, and approved processes, clinical practice guidelines (CPGs) provide guidance intended to complement the user's own clinical judgment. Although they can be good resources for mental health and social service professionals, their quality determines their trustworthiness and usefulness in everyday practice. In medicine, we have assessed the quality of guidelines, but the assessment of psychology guidelines is still lagging. However, one systematic review demonstrated that there was a lack of clarity and consistency among guidelines intended for the treatment of depression (MacQueen et al., 2016). Therefore, more assessments of the quality of psychology guidelines are needed. It also remains that too little attention has been paid to the developers, specifically how the guideline development groups are composed and the nature of the expertise of those involved in guideline development. No studies, to our knowledge, have surveyed those who have participated in the development of practice guidelines to gain insight into their views of practices and processes. This is surprising given the time and resources involved in their development, as well as the potential impact to practitioners and the public. The Ordre des psychologues du Québec (OPQ; College of psychologists of Quebec), the most prolific producer of practice guidelines in Canada, has published five CPGs that are currently available to psychologists: *The Guidelines for the Evaluation of Dyslexia in Children* (2014), the *Guidelines for Autism Spectrum Disorder - Clinical Evaluation* (2012), the *Guidelines for the Assessment of Mental Retardation* (2007), the *Guidelines for the Assessment of a Child in Connection with a Request for Derogation to the Age of School Admission* (2006), and the *Guidelines for Expert Assessment Concerning Child Custody and Access Rights* (2006). These are intended to provide empirically supported guidance for psychologists in the areas of

assessment, diagnosis, general functioning, treatment, and other decision-making support. In Study 1, we evaluated the quality of these guidelines using a widely accepted assessment tool, the Appraisal of Guidelines for Research and Evaluation II (AGREE II) guideline evaluation instrument. Our results showed that there is a need for more methodological rigour in guideline development. Study 2 examined group composition as well as the expertise of guideline development committee members at the Order of Psychologists of Quebec (OPQ), as defined by academic research productivity. We analyzed the peer-reviewed publication productivity of committee members using PsycINFO and MEDLINE, and retrieved their *h*-index from Scopus and from Web of Knowledge. The findings revealed that there was a clear imbalance between clinical and research expertise, with only a small percentage of researchers represented on these committees. This highlights the need for improved group composition for future guideline development. For Study 3 we surveyed 40 CPG development committee members who worked on one of the 17 guidelines published by six Quebec regulatory bodies in the social sciences. We inquired about their knowledge about guideline development methodology and solicited their views on the quality of the guidelines they developed. The results show that the developers' familiarity with established development methods must be improved; group composition must be broadened; procedures for dealing with divergent views during the development process were vague or lacking; conflicts of interest were inadequately reported; and that the guidelines, although still currently available to practitioners, should be updated. These three studies show that careful planning and more stringent methodologies must be applied to the guideline development process.

Résumé

Des décisions cliniques sont prises dans divers milieux de soins de la santé tous les jours. Contrairement aux manuels de procédures qui offrent des instructions concises sur les dépistages diagnostiques, les interventions appropriées, la durée du traitement et les processus approuvés, les guides de pratique clinique (GPC) fournissent des recommandations qui servent à informer les décisions du clinicien. Bien qu'il s'agisse d'une ressource utile pour les professionnels de la santé mentale et des services sociaux, la valeur et l'utilité des GPCs sont dépendantes de leur qualité. En médecine, l'évaluation de la qualité des GPC est courante alors que ces efforts tardent dans le domaine de la psychologieMacQueen et al., 2016). Par conséquent, il est primordial d'évaluer davantage la qualité des recommandations psychologiques. De surcroît, peu d'attention a été accordée aux développeurs de guides, en particulier à la composition des groupes d'élaboration des guides cliniques et à la nature de l'expertise des personnes impliquées dans le développement des GPCs. À notre connaissance, aucune étude n'a été réalisée auprès de ceux qui ont participé à l'élaboration de guides de pratique clinique afin de mieux comprendre leur point de vue sur les pratiques et les processus de développement des guides. Ceci est surprenant compte tenu du temps et des ressources utilisés pour le développement des GPCs ainsi que l'impact potentiel qu'ils peuvent avoir sur les professionnels et le public. L'Ordre des psychologues du Québec (OPQ), le producteur de GPC le plus important au Canada, a publié cinq GPCs actuellement disponibles pour les psychologues: *L'évaluation de la dyslexie chez les enfants* (2014), *Les troubles du spectre de l'autisme l'évaluation clinique: Lignes directrices* (2012), *Lignes directrices pour l'évaluation du retard mental* (2007), *Lignes directrices pour l'évaluation d'un enfant en vue d'une demande de dérogation à l'âge d'admission à l'école* (2006), et les *Lignes directrices pour l'expertise en matière de garde d'enfants et des droits d'accès* (2006). Ceux-ci

ont pour but de fournir aux psychologues des recommandations pratiques empiriques dans les domaines de l'évaluation, du diagnostic, du fonctionnement général, et du traitement pouvant aider à la prise de décisions cliniques. La première étude décrite ici visait à évaluer la qualité de ces guides cliniques à l'aide d'un outil d'évaluation reconnu, le *Appraisal of Guidelines for Research and Evaluation II* (AGREE II). Nos résultats ont montré la nécessité d'une plus grande rigueur méthodologique dans l'élaboration de guides cliniques. La deuxième étude a examiné la composition des groupes de développement ainsi que l'expertise des membres des comités d'élaboration de guide publiés par l'OPQ, tel que défini par la productivité académique en recherche des membres des comités de rédaction de guides. Nous avons analysé la productivité des membres des comités en matière de publications examinées par des pairs par l'entremise de PsycINFO et MEDLINE, et nous avons récupéré leur indice-*h* à l'aide de Scopus et Web of Knowledge. Les résultats ont révélé un important déséquilibre entre l'expertise clinique et l'expertise en recherche des membres où seul un faible pourcentage de chercheurs ont participé sur ces comités. Cela souligne la nécessité d'améliorer la composition des groupes pour l'élaboration de guides cliniques futurs. Pour l'étude 3, nous avons interrogé 40 membres siégeant sur divers comités de développement et ayant travaillé sur l'un des 17 GPC publiées par six organismes de réglementation en sciences sociales au Québec. Nous avons examiné leurs connaissances en matière de méthodologie, leurs connaissances relatives à l'élaboration de guides cliniques, et avons sollicité leur avis sur la qualité des guides qu'ils ont élaboré. Les résultats montrent que la connaissance des développeurs en lien aux méthodologies de développement devrait être améliorée; que la composition des comités devrait être plus diversifiée; que les procédures permettant la gestion des points de vue divergents au cours du processus de développement étaient vagues ou inexistantes; que les conflits d'intérêts étaient signalés de manière inadéquate; et que es guides devraient être mises à jour. Ces trois études montrent qu'une

planification plus minutieuse et des méthodologies plus rigoureuses devraient être mises en place lors du processus d'élaboration des lignes directrices.

Acknowledgements

I would like to thank my thesis supervisor, Dr. Martin Drapeau and thesis committee member, Dr. Sylvie Beauchamp, for their invaluable direction, feedback, and support throughout my doctoral degree. I thank you both for your instrumental roles in the conception, development, and realization of my doctoral thesis. I would also like to thank all of the co-authors, and especially my fellow doctoral students Constantina (Tina) Stamoulos and Andrea Reyes, whose positive outlook and outstanding work ethic made this process fun and interesting, for their collaboration and support on the studies included in this thesis. Furthermore, I would like to thank my fellow student-researchers at the McGill Science and Practice in Psychology (SAAP) Research Lab for their collaboration and support, including Catherine Hébert who translated the Résumé of this thesis and the survey questions for Manuscript 3, from English to French. I would also like to acknowledge the personal funding support I received from the Social Sciences and Humanities Research Council (SSHRC) for the award of the Joseph-Armand Bombardier Doctoral Scholarship.

I would especially like to thank my loving spouse Slawek and our two amazing children Adrien and Anna, for their incredible patience, love, and support. Even in the most difficult moments, your sweet faces reminded me that I dare not quit, for in life our character is often tested as our little ones watch closely and learn from our example. Last but certainly not least, my deepest thanks go to my wonderful parents, Cécile and Louis Trépanier, who embody the qualities of love, persistence, and courage, and whose unwavering encouragement, support, and loving childcare at a moment's notice, made this work possible.

Contribution of Authors

The introduction (Chapter 1) and the three studies (Chapters, 2, 3, and 4) included in the present thesis, are all original research. I am the first author of each manuscript, as I contributed to the conception and design; conducted the data acquisition, analysis, and interpretation; completed the original draft of each paper; incorporated critical feedback from co-authors and reviewers; and prepared and submitted the final draft of each paper.

The first study was co-authored by Dr. Martin Drapeau who contributed to the conception, design, and analysis. He also provided editorial suggestions and reviewed the final version of this paper. Constantina Stamoulos and Andrea Reyes contributed to the conception, study design, data gathering, and reviewed the final version of the paper. Dr. Sylvie Beauchamp provided design feedback that lead to the expansion of the data gathered, adding invaluable information to the analysis, as well as reviewing the final version of the draft. Dr. Heather MacIntosh and Dr. Serge Larivée reviewed and provided editorial suggestions for the final draft of the study, and provided their approval of the final version for publication.

The second study was co-authored by Dr. Drapeau who contributed to the conception, design, analysis, editorial suggestions, and the final review of this paper. Constantina Stamoulos and Andrea Reyes contributed to the conception, study design, data gathering, and reviewed the final version of the paper. Dr. Beauchamp reviewed and provided editorial suggestions to the final version of the paper.

The third study was co-authored by Dr. Drapeau who contributed to the conception, design, analysis, provided editorial suggestions, and reviewed the final version of this paper. Constantina Stamoulos contributed to the conception, study design, and reviewed the final version of the paper. Dr. Beauchamp and Dr. Larivée reviewed and edited the final study and provided their approval for publication.

Chapter 1

Introduction: Literature Review

Clinical practice guidelines (CPGs) have proliferated over the past 25 years in medicine, and more recently in the social sciences (Alonso-Coello et al., 2010; Beauchamp, Duplantie, & Mercier, 2011; Burgers, Grol, Klazinga, Makela, & Zaat, 2003; Weisz et al., 2007). CPGs were first introduced in response to a growing number of medical specialties and clinical approaches that sometimes lacked adequate procedural and outcome information (Field & Lohr, 1992; Weisz et al., 2007).

Clinical Guidelines: A Brief History

In the early 1900's, health officials were mandated to raise public health standards in the wake of serious disease outbreaks such as yellow fever, typhoid, and tuberculosis (Weisz et al., 2007). Rising concern for public health and safety put pressure on the medical and scientific communities to monitor their treatment outcomes by collecting and organizing data, and standardizing practices (Flexner, 1910; Weisz et al., 2007). This new level of monitoring and record-keeping made it possible for others to replicate and critique reported outcomes, and address flawed methodologies that might otherwise lead to erroneous conclusions (Flexner, 1910).

Over the decades that followed, public demand for larger medical programs in schools, hospitals, and research centers grew. Traditional private medical practices became the standard channel for care, until the American Medical Association (AMA) began to view the practice of issuing medical licenses after completing medical school as inadequate (Blum, 1996; Field & Lohr, 1992; Weisz et al., 2007). As practices became increasingly specialized, the AMA recognized the need to update their credential and practice requirements, and thus established a

standardized system for medical specialties (Blum, 1996; Field & Lohr, 1992; Weisz et al., 2007).

However, as biomedical procedures proliferated within specialties, it became more evident that the act of creating a 'specialist' system did not guarantee quality and competency among physicians, as medical interventions for a given diagnosis still varied between professionals (Weisz et al., 2007). In response, medical professionals began to standardize their testing and treatment protocols in daily medical practice; thus, they became increasingly reliant on the efficient translation of knowledge (Weisz et al., 2007).

Treatment protocols in the form of clinical guidelines were the next logical step for broad and effective ways to share knowledge. One of the earliest guidelines produced was the American Pediatric Society's guideline for The Immunization of Children (1938); others soon followed (Margo, 2004). Change continued post-WWII with the expansion of government, as it assumed its role as provider, guarantor, and purchaser of public health and public health services (Weisz et al., 2007). By the end of the 1960's, physicians were required to manage and utilize increasingly complex medical and therapeutic options, leading to the need for more clinical guidance. In the 1970's, medical credentials were assumed to be evidence of professional competence (Blum, 1996). However, education was frequently supplemented by the opinion of experts to provide clinical guidance (Parry, Cape, & Pilling, 2003). It became traditional for physicians, especially young physicians, to rely on the opinion of experts (Oxman, Fretheim, & Schunemann, 2006; Weisz et al., 2007). For example, before the introduction of psychopharmacology guidelines, psychiatrists relied on planning meetings with other physicians which entailed lengthy discussions about a patient's psychological etiology, care, and best course of treatment such as if or how to medicate the patient (Shaner, 2001). As time passed, expert

medical opinions were increasingly seen as wide, diverse, and sometimes contradictory (Wiesz, et al. 2007). It became clear that soliciting informal expert opinion for guidance was insufficient, and a demand for more rigorous processes of examining the evidence increased. Thus, more attention was put toward the development of clinical guidelines, which gained increasing acceptance by medical professionals (Oxman et al., 2006; Woolf, Grol, Hutchinson, Eccles, & Grimshaw, 1999).

In the 1980's, the medical community published an increasing number of collectively produced guidelines in an effort to bring coherence to medical practice. However, some researchers reported that guideline protocols were frequently based on the opinion of topic experts, in combination with other forms of research evidence (e.g. systematic reviews) (Gould & Kendall, 2007; Pilling, 2008, 2009). By the late 1990's, significant variations in guideline recommendations were observed by healthcare providers (Alonso-Coello et al., 2010; Pilling, 2008; Sackett, Rosenberg, Gray, Haynes, & Richardson, 1996). This fostered a renewed interest in guideline development that focused on improving their quality (Pilling, 2008, 2009). Thus, physicians in diverse fields of specialization called for the standardization of treatment approaches that were both easily accessible and based on rigorous research evidence (Field & Lohr, 1992; Weisz et al., 2007; Woolf et al., 1999).

The development of CPGs was further invigorated in 1992 when the *Institute of Medicine* (IOM) (now the *National Academy of Medicine*) became the first large medical organization to emphasize their importance in daily medical practice. The IOM defined CPGs as "systematically developed statements to assist practitioner and patients' decisions about appropriate healthcare for specific clinical circumstances" (Field & Lohr, 1992). This definition guided guideline

policy and development for the two decades that followed (Alonso-Coello et al., 2010; Pilling, 2008; WHO, 2003, 2010, 2014).

Types of Clinical Practice Guidelines

CPGs come in different shapes and sizes, and may have different objectives. According to the *WHO Handbook for Guideline Development* (2014), there are four types of CPGs, and each one is designed to meet unique needs. There are *rapid advice guidelines*, which are typically produced in response to a public health emergency. These guidelines are typically published within one to three months from the time development begins. They are informed by *some* evidence but not by a full review of the literature, and are given a "review-by" date which represents a deadline for when they need to be reviewed and transformed into Standard Guidelines. *Standard guidelines* are produced in response to a request for guidance in a single clinical area. They take nine to 12 months to complete, and are prepared in consultation with experts in diverse but related field. A thorough systematic review of the evidence is required to support the recommendations in the guideline. The third type of guideline is a *full guideline* that provides full coverage of a health topic and includes recommendations to manage and treat all aspects of a condition. These can take two to three years to complete, and require a frequent and highly involved development group. The fourth and final type of guideline is the *compilation guideline* which is an aggregate of current recommendations from the WHO and other credible sources, but do not include any new recommendations. Guidelines can also be adapted to meet the needs of a particular setting or local culture (ADAPTE-Collaboration, 2009; Harrison, Légaré, Graham, & Fervers, 2010; WHO, 2012). Each of the guideline types listed above meet an important need. However, the greatest amount of time and resources have been invested in

the production of standard and full guidelines, and their appeal has expanded beyond the medical community in recent years.

Who Produces Practice Guidelines?

Relative to other healthcare professionals, physicians have been the most productive group for creating CPGs. They have done so, in part, as a means of preserving their professional autonomy in the face of growing government regulations and oversight (Timmermans & Kolker, 2004). As with the medical field, allied disciplines have advanced and increased in complexity. Physicians have been joined by other healthcare professionals, such as psychologists and social workers, amongst others, who are interested in developing CPGs as a means to provide access to empirically supported clinical guidance to inform their practice (Beauchamp, Drapeau, & Dionne, 2015; Beauchamp et al., 2011; Hollon et al., 2014). As such, CPGs facilitate practice by providing recommendations based on unbiased sources derived from scientific data that were compiled, analyzed, and organized into well-supported and properly referenced pronouncements (Ansari & Rashidian, 2012; Dozois, 2013; Dozois et al., 2014; Shaneyfelt, Mayo-Smith, & Rothwangl, 1999; WHO, 2003, 2014). Their recommendations are designed to assist medical professionals, counselors, educators, and social workers, and guide them toward decisions that best serve the needs of the client who receives their services.

Interest in the production and dissemination of high quality CPGs has expanded globally to various types of organization. They are currently published by the World Health Organization (WHO), governmental agencies in several countries, regulatory bodies, and professional associations (Ansari & Rashidian, 2012; García et al., 2014; Parry et al., 2003; SIGN50, 2015; WHO, 2003, 2010). In Canada, the Canadian Medical Association has been working to improve the quality of guidelines available to physicians across the country (Palda, Davis, & Goldman,

2007). Within Canada, Quebec stands out as one of the most prolific producers of practice guidelines for healthcare professionals among the provinces. Regulatory bodies that produce guidelines in Quebec include the *Collège des médecins du Québec* (The College of Physicians of Quebec; CMQ), the *Ordre des psychologues du Québec* (The Order of Psychologists of Quebec; OPQ), the *Ordre des travailleurs sociaux et des thérapeutes conjugaux et familiaux du Québec* (The Ordre of Social Workers and Family Therapists of Quebec), the *Ordre des psychoéducateurs et psychoéducatrices du Québec* (The Ordre of Psychoeducators of Quebec), and the *Ordre des conseillers et conseillères du Québec* (The Ordre of Counsellors of Quebec).

Guidelines Offer Several Benefits

Every day, clinical decisions are made in diverse practice and healthcare settings. Unlike procedural manuals that provide concise instructions on diagnostic screenings, procedural processes, appropriate interventions to be followed, and for how long, CPGs are intended to provide guidance in tandem with the user's own clinical judgment. They also have the potential to discourage outdated or ineffective approaches that would otherwise fail to reduce morbidity and mortality (Shaner, 2001).

Useful guidelines however, are not simple to produce. They require a significant investment of time, expertise, and financial resources for their planning, production, and implementation (Ansari & Rashidian, 2012; Beauchamp et al., 2011; Grimshaw et al., 1995; Hollon et al., 2014; WHO, 2010). Such an investment is justified when their production leads to clear and measurable benefits that health professionals value. CPGs are appreciated within a vast number of healthcare services including group therapy, social work, psychological testing, mental health, palliative care, cancer treatment, to name but a few (Bernard et al., 2008; Gordon & Cooper, 2010; Hudson, Quinn, O'Hanlon, & Aranda, 2008; Kendall, Taylor, Perez, & Taylor,

2008; OPQ, 2007; van Dijk, Oosterbaan, Verbraak, & van Balkom, 2013). The widening interest in CPGs can be explained as a function of their impact in the following intersecting areas.

Practice Guidelines Can Efficiently Enhance Practitioner Knowledge

Practice guidelines have practical implications for practitioners who are pressed for time and unable to search for, read, and analyze the latest research relevant to their area of expertise, in a timely manner (Graham, Mancher, Wolman, Greenfield, & Steinberg, 2011; Wolf, Hubbard, Faraday, & Forrest, 2011). The research shows that this is a common problem among health practitioners. Even the most innovative, scientifically supported, health interventions are at risk of not being considered and utilized by busy practitioners (Graham et al., 2011; Grimshaw, Eccles, Lavis, Hill, & Squires, 2012; Lesho, Myers, Ott, Winslow, & Brown, 2005; Martínez, Reyes, Lorenzo, & Menéndez, 2009; Wolf et al., 2011). For example, according to Davidoff and colleagues (1995), physicians would need to read at least 19 articles per day, in their field of expertise, in order to keep up with relevant research. This was in sharp contrast to the one-hour of reading per day that the physicians reported (Davidoff, Haynes, Sackett, & Smith, 1995). There are other reports that point to decision-makers from across all groups such as healthcare providers, patients, informal caregivers, managers, and policy-makers, who fail to use the latest research evidence on which to base their decisions (Straus, Tetroe, & Graham, 2009). CPGs have the potential to facilitate decision-making by translating scientific knowledge into well researched clinical process and interventions, a process often referred to as knowledge translation (Wollersheim, Burgers, & Grol, 2005). Indeed, the Canadian Institutes of Health Research (CIHR) define knowledge translation as:

> "a dynamic and iterative process that includes synthesis, dissemination, exchange and ethically-sound application of knowledge to improve the

> health of Canadians, provide more effective health services and products and strengthen the health care system. This process takes place within a complex system of interactions between researchers and knowledge users that may vary in intensity, complexity, and level of engagement depending on the nature of the research and the findings as well as the needs of the particular knowledge user." (CIHR, http://www.cihr-irsc.gc.ca/e/29418.html#2, accessed Dec 4, 2017).

It is therefore crucial that we address how knowledge is transferred and 'packaged' for end-user consumption. Once in the hands of the end-users, CPGs can offer clear benefits to the end client or patient.

Guidelines Can Improve Treatment Processes, Costs, and Outcomes

Practice guidelines are not simply a convenient source of clinical guidance, they have also demonstrated efficacy in improving treatment outcomes and reducing costs (Balogh et al., 2005; I. D. Graham et al., 2000; Hayward, Guyatt, Moore, McGibbon, & Carter, 1997; Hollon et al., 2014; Rutten et al., 2016; van Dijk et al., 2013). In a recent report by Johnson and colleagues (2018) examining the pre and post implantation effects of a pediatric asthma CPG on outcomes within the emergency department (ED), inpatient care, and intensive care unit (ICU) of a hospital over a two-year period, it was found that: there was a 15% reduction in wait and treatment times in the ER; patient length of stay was reduced from 1.5 days to 1.3 days; ED encounters requiring admission was reduced by 5%; admission requiring ICU was reduced by 10%; and that the total dollar charges were reduced by 18%, with the implementation of the CPG. The positive effects of clinical guidelines have been shown for a number of other conditions, such as hemodynamic

instability, arterial disease, traumatic brain injury, to name just a few (Balogh et al., 2005; Barriocanal, Lopez, Monreal, & Montane, 2016; Patel et al., 2016)

The impact of using mental health guidelines has also been studied. Van Dijk and colleagues (2013) conducted a cohort study in the Netherlands to identify the effects of the Dutch National Multidisciplinary Clinical-practice Guidelines for Anxiety Disorders implemented in a mental healthcare setting. The study examined whether adherence to CPGs for anxiety disorders yielded better results than non-adherence. The sample consisted of 181 outpatients with anxiety or hypochondriasis that were treated in a routine mental health setting. Patients were asked to complete questionnaires at the start of the program, and again one year later. Results demonstrated that the guidelines yielded superior treatment results with greater symptom reduction and increased patient satisfaction with treatment, after one year, when compared with patients whose mental health professional did not adhere to the clinical guidelines.

In another study, Köhler and colleagues (2012) investigated the efficacy of guideline recommended treatments for unipolar depression in a randomized controlled trial of 224 hospital inpatients. The patients were devided into treatment group and a control group. The treatment group was treated according to guideline recommended treatments for depression (e.g., treatment duration, neurophamacological interventions, and psychotherapy), and the control group was treated using various other methods to treat depression not found in the guidelines (e.g., treatment as usual) (Köhler, Hoffmann, Unger, Steinacher, & Fydrich, 2012). Participants completed depression scales at the time of admission, and at discharge, to assess for depression and treatment outcomes. The results showed that patients treated according to the guidelines were in remission from depression in 73% of cases, but only 59.6% patients who were not

treated according the guidelines reached remission at the time of discharge. Thus, improved outcomes were observed in depressed patients treated accorting to guideline recommended treatments (Köhler et al., 2012).

The above results highlight the potential benefits of practice guidelines when they are correctly implemented. However, when the implementation is not optimal, it can yield mixed results. Studies have shown that poorly targeted and inadequately implemented CPGs had little or no positive effect on health outcomes such as inpatient stay, length of treatment, treatment outcomes, and costs (Ellen, Brown, & Cockerill, 2009; Lesho et al., 2005; Weinberger, 2009). For example, Lesho and colleagues (2005) studied the effects of implementing an asthma CPG and a diabetes CPG, in a large medical setting. The researchers found that all outcomes improved by following the asthma guidelines, but not the diabetes guidelines. The diabetes guidelines improved the diagnostic and educational processes, but these did not translate to better disease outcomes. The authors point to several contributing factors for the poor outcomes including insufficient clinical resources, the added complication of comorbid diseases, and a lack of training of health professionals on how to implement the guideline's recommendations. Lesho and colleagues (2005) denied that the implantation strategies were to blame, owing to the fact that they used the same strategies for both guidelines. However, their explanation for the implementation failure strongly suggests that the treatment of diabetes has more complicated service delivery requirements as compared to the treatment of asthma. Thus, these results demonstrate that guidelines must be thoughtfully constructed with the input of all stakeholders who are best positioned to reflect the implementation needs of the healthcare providers in order to achieve better disease-specific outcomes.

Guidelines Reduce Practice Variation

Implementing and standardizing the best processes for interventions may improve health and medical care outcomes (Margo, 2004). Practice guidelines are seen as having the underlying purpose of increasing consistency of care, which leads to reduced treatment-outcome variability and contain healthcare costs (Field & Lohr, 1992; Mental Health Commission of Canada, 2012; Pilling, 2008; Sackett et al., 1996; Shiroiwa, Fukuda, Ikeda, Takura, & Moriwaki, 2016; Wolf et al., 2011). Studies have shown that the frequency with which certain procedures are performed varies dramatically between individual practitioners, specialties, health settings (such as hospitals and clinics), and geographical regions (Martínez et al., 2009; Menéndez, Ferrando, Vallés, & Vallterra, 2002; Schuh et al., 2017; Wolf et al., 2011). Patients with identical clinical presentations receive variable care depending on the clinician, hospital, or location – a problem that may be remedied with the use of clinical guidelines. For example, Martinez and colleagues (2009) reviewed several large cohort studies that examined the before and after effects of using guidelines developed by the American Thoracic Society (ATS) (Martínez et al., 2009). A high degree of variability was found in the quality of patient care between professionals and hospitals, that was explained by the degree of adherence to clinical guidelines (Martínez et al., 2009). They demonstrated that the implementation of, and adherence to, CPGs were positively associated with reduced mortality, fewer days to clinical stability, leading to lower healthcare costs.

Research on variations between practices was also conducted by geographical region. Schuh and colleagues (2017) examined how interventions using evidence-based supportive therapies differed between physicians located within the same clinical site, and physicians working in separate geographical location, when treating infants diagnosed with bronchiolitis, in 38 emergency departments of pediatric emergency research networks in Canada, the US,

Australia, the UK, New Zealand, Ireland, Spain, and Portugal. Schuh and her team found that only 30% of infants hospitalized with bronchiolitis received evidence-based supportive therapies, in the sample studied. Thus, practice variation exists globally, which highlights the importance of global initiatives like the ones advanced by the World Health Organization (WHO, 2003, 2010, 2014) for guideline production. The WHO's intention was in part to address practice variation and knowledge gaps in healthcare, thus assisting local organizations and practitioners in providing patients with a higher standard of care regardless of where they receive treatment.

Guidelines Serve as Educational Resources

Scientifically supported recommendations that are formulated as clearly written statements in well-organized documents have indisputable benefits as an educational tool (Hershenberg, Drabick, & Vivian, 2012; Miville et al., 2009). They serve as convenient, concise, and accessible reference tools that raise the level of service offered by clinicians-in-training. CPGs also benefit more experienced professionals, as they provide information on new or updated techniques and processes related to evaluation and treatment.

In sum, good guidelines serve multiple purposes and provide solutions to diverse challenges. However, much of the benefits they provide rest squarely on the quality of a CPG.

What Makes a Good Guideline?

The quality of a guideline is determined by many factors (Brouwers et al., 2010; Burgers, Cluzeau, Hanna, Hunt, & Grol, 2003; Grimshaw et al., 1995). Of prime importance to guidelines users is whether or not the recommendations provided therein are trustworthy (Graham et al., 2011; Ransohoff, Pignone, & Sox, 2013; Woolf et al., 1999). Guideline users should trust CPGs only if the recommendations accurately reflect strong underlying evidence and state how and when to follow them, and when they are counter-indicated (Ransohoff et al., 2013). Therefore

CPGs are an invaluable resource for medical professionals, psychologists, and social workers who are trained, and professionally and ethically required, to inform their practice with empirical evidence (Babione, 2010; Bergman, 1999).

Licensing bodies and professional associations have placed greater emphasis on evidence-based practice for psychotherapy, in recent years. For instance, the 2004 APA Presidential Task Force on Evidence-Based Practice (EBP) was appointed to develop policy statements on this topic. They defined evidence-based practice in psychology as "the integration of the best available research with clinical expertise in the context of patient characteristics, culture, and preferences" (APA Pres. Task Force on Evidence-Based Practice. 2006, p.72). This document grants that clinical research evidence alone cannot capture the nuances and complexities of psychological service delivery. Therefore, recommendations must not only be evidence-based, but also reflect the diverse environmental, social, and cultural factors, best articulated by a diverse guideline development committee. In 2011, the Canadian Psychological Association (CPA) followed suit and released a statement with their own definition of EBPs as "the conscientious, explicit, and judicious use of the best available research evidence to inform each state of clinical decision-making and service delivery" (Dozois et al., 2014). Canadian psychologists are required to consider the empirical evidence to guide clinical decision-making of treatment plans that are informed by individual needs and client characteristics. Psychologists must also monitor, re-evaluate, and modify treatment according to the patient's response and new available evidence (Dozois, et al., 2014). Thus, CPGs are positioned to provide important clinical resources designed to provide psychologists with recommendations that are congruent with the CPA's and APA's definition of evidence-based practice.

However, the use of strong scientific studies alone does not guarantee a high quality practice guideline. There are other elements that producers must consider in the early planning stages of the development. The IOM has articulated eight important criteria that should be met in order for the complex development process to yield a successful product.

IOM's Standards of Trustworthiness: Eight Criteria for Quality Guidelines

In 2011, the IOM published a 290-page committee report titled *Practice Guidelines We Can Trust* (Graham et al., 2011). In it, the authors provided an update to the definition of clinical practice guidelines that they put forth in 1992, stating that "Clinical practice guidelines are statements that include recommendations intended to optimize patient care that are informed by a systematic review of evidence and an assessment of the benefits and harms of alternative care options" (Graham et al. 2011, p.4). The report provides standards that the committee defined as the processes and procedures for developing CPGs that should be considered essential to producing scientifically valid, transparent, and reproducible results (Graham et al. 2011). The IOM provided formal standards for trustworthy guidelines are comprised of eight key criteria.

1) Transparency of process

The first of these is *transparency of process*, referring to both the process of development and the funding bodies involved. All developers and reviewers should be listed in the guideline to include their names, titles, professional designations, and employers. Of course, all of the criteria within the standard should detail the processes used to meet them, with special attention to providing explicit details for identifying the research question, the search methods, the inclusion and exclusion criteria, and how conflicts were resolved in cases where there was disagreement between committee members about recommendations and their underlying evidence. Essentially, all the criteria should be detailed, somewhere in the guideline, about how

they were met.

2) Conflict of interest

The second IOM criterion is *conflict of interest* (COI). It specifies that only small minority development group members may have a COI. A COI includes any income received by a group member for his or her products or services that is related to a recommendation in the CPG. The authors also specify that the chair and co-chair should not have any COIs (Graham et al., 2011). It is worth noting that the WHO (2014) provides a more detailed explanation of a COI. They state that the involvement of developers and reviewers having a monetary advantage to support a particular recommendation, and anyone with non-financial advantages including academic, professional, and personal interests are in conflict of interest. Further, a COI is anything that may interfere with the objective assessment of a potential recommendation including, prior publication of a study being considered by the guideline, prior public declaration of a firm opinion in favour of or against a treatment, participation in past relevant editorial or judicial process, and having personal or professional affiliations with an organization which promotes a product or service that relates to the guideline (WHO, 2014). Meeting any of the above criteria indicates possible COI that threatens the credibility, and value, of a guideline. They are best avoided by requiring all members of guideline development groups to disclose any personal or professional COIs related to the guideline topic (Qaseem et al., 2012).

3) Guidelines development group composition

The third IOM criterion addresses *guideline development group composition*. It specifies that the development group should be composed of methods experts, clinicians, representatives of stakeholders, and affected populations. It should be determined in the planning phases what experience, expertise, and credentials a committee member should have. Additional

considerations for selection are the professional, geographic, social, and academic milieu from which they ought to be recruited. There should also be a standard process for deciding who would be responsible for inviting potential members to join the guideline development group, and by what established criteria. These requirements are not elaborated upon by the IOM**.**

The APA has also taken a special interest in the composition of guideline development groups and have thus produced a statement recognizing that guideline committees require careful consideration. The APA (2006) stated that clinical research evidence alone does not adequately describe all the nuances of psychological treatment, and recommendations must be coupled with a diverse guideline development panel (APA Presidential Task Force on Evidence-Based Practice. 2006, p.72).

Other authors have elaborated on the importance of stakeholder involvement in order to gain insight into their values, knowledge, and preferences, and promote patient engagement (Légaré et al., 2011). Légaré and colleagues (2011) propose various methods to facilitate consumer involvement such as organizing workshops, meetings and seminars, and providing opportunities for patient populations to be involvement in systematic reviews, focus groups, interviews, and public surveys. Further, the broad inclusion of experts and stakeholders is what Daniels and Sabin (2002) refer to as *procedural justice.* Procedural justice requires a transparent and fair process in which all relevant stakeholders are actively involved (Daniels & Sabin, 2002).

When it comes to the committee selection process, some may assume that the good judgment of those initiating the development process should be adequate. However, guidelines rarely contain information about how committee members were selected. This is potentially problematic because there are no established standards for selecting technical, scientific, and clinical experts, or experts in guideline production, as of yet. Thus less qualified individuals may

participate in the development of a guideline that would not have otherwise, if clear and objective criteria were put in place. We often speak of the scientific rigour needed for good evidence, but it is equally important to establish a rigorous process for identifying those qualified to judge the evidence and formulate the CPG's recommendations, to promote trustworthiness.

4) Systematic reviews

The fourth item requires that a guideline's evidence should be obtained from *systematic reviews*, which is considered essential to the process (Graham, et al., 2011). A systematic review is a process by which evidence from multiple studies are collected and analyzed using methods established prior to the formulation of the research question (Graham et al., 2011). A good guideline should provide a description of the systematic review process, who conducted it, the search engines and databases utilized including all other possible sources consulted, and when the search commenced and ended. The IOM also recommends seeking the services of guideline methodologists to conduct the systematic review, who are either part of the guideline development group, or an external group that is contracted to perform the systematic review who would in turn be guided by the groups scientist and clinicians (Graham et al., 2011).

5) Quality and strength of the evidence on the recommendations

The fifth item is the most stringent and detailed within the standard. It addresses the *quality and strength of the evidence on which the recommendations* are based. Specifically, trustworthy guidelines should explain the reasoning behind each recommendation and summarize the evidence in terms of the benefits and harms, characterizing the quantity and quality of the relevant evidence, and the role of the clinician's judgment (Graham et al. 2011). Further, a rating should be determined for the level of evidence (where expert opinion would be given the lowest rating, to meta-analysis of random controlled trials which would be assigned the

highest), and the strength of the recommendation. A description of the differences of opinions about the recommendations should also be included in the guideline. A good guideline will be based on high quality scientific evidence. Some authors emphasize the importance of determining the effect size obtained from randomized clinical trials (Schünemann et al., 2014) though others have stated that randomized clinical trials fail to capture the important nuances of studies in the social sciences (Gray, Plath, Webb, & Webb, 2009; Pawson, Boaz, Grayson, Long, & Barnes, 2003). Indeed, even when the best, most current, high quality research is systematically collected, there is still a measure of subjectivity when selecting which studies are most relevant and how they are to be translated into pronouncements (Bergman, 1999). That is why the overall quality of the final product is as vitally dependent on the selection and quality of the studies as it is on the quality and diversity of the 'experts' selected to be on the guideline development and review committees.

6) Articulating the recommendations

The sixth criterion is *articulating the recommendations*, which means how and when they should be used. In articulating a guideline's recommendations, the developers must find a balance between unambiguous and specific guidance, and not reach beyond the evidence or the scope of the guideline (Hussain, Michel, & Shiffman, 2009). When the behaviors are well specified using proper wording, guidelines are seen to be of greater value and more closely followed clinically (Michie & Johnston, 2004). The IOM (see also Hussain et al., 2009) provides the following formulation of recommendation: 1) Identify the critical recommendations in the guideline text using semantic indicators such as "The committee recommends…" or "If X should occur, then the clinicians should..."; 2) Use consistent semantic and formatting indicators such as wording, formatting, and typesetting; 3) Group recommendations together in a summary

section to facilitate their identification; 4) Recommendations must be clear, decidable, and executable; 5) Clearly assign evidence of quality and recommendations of strength with each recommendation (Graham et al., 2011, p.130). Articulation of recommendations are 1) provided in standardized form detailing precisely what the recommended action is, and under which circumstances it should be performed; 2) strong recommendations should be worded in such a way as to be left open for evaluation (Graham et al., 2011, p.131). As can be seen by the above requirements, finding the best manner in which to express a recommendation requires a degree of skill and expertise.

7) External reviews

The seventh item is *the external review,* also deemed essential to the process by the IOM (Graham et al., 2011). The external review should include all of the stakeholders and an explanation should be found within the guideline of all the edits that have been made in response to the reviewers' comments, as well as posted for public comment. According to the IOM, it is important that external reviewers not have a COI because their role is to provide unbiased professional suggestions and opinions. Other researchers have added that external reviewers should not be members of the guideline development group or involved in the systematic review (e.g., Hollon et al., 2014). External reviewers should include research and clinical experts on the topic of the guideline, representatives from provincial or federal agencies and professional organizations, and representatives from health plan, advocacy organizations and the general public (Hollon et al., 2014; Schünemann et al., 2014).

8) Updating guidelines

The eighth and final item addresses *updating* the guidelines. Guidelines should have a pre-determined date established to conduct a systematic review and a process for monitoring the

literature, in order to for a guidelines to maintain this reliability (Graham et al., 2011). This is an important point because the age of a guideline can influence its overall evaluation as it speaks to its validity or obsolesce. Other researchers suggest that guidelines developers should ask these three questions: 1) What constitutes sufficient evidence for an update to be initiated?; 2) Once sufficient evidence has been identified, what are the most efficient systematic methods to formulate recommendations?; 3) Once the update is complete, what is the most rapid way to disseminated it to target users? (see Shekelle, 2014). Factors shown to prompt an update include the release of new high quality empirical studies and the release of new drugs or interventions (Ahmadzai et al., 2013; Peterson, McDonagh, & Fu, 2011; Shekelle, 2014). One question not mentioned by these researchers is who should be involved in examining and selecting the new evidence and formulating the new recommendations, or how and by whom these experts will be selected for the update. These are important considerations, given that the frequency required for updates is normally every two to five years (Ahmadzai et al., 2013; Alderson, Alderson, & Tan, 2014; Shekelle, 2014; P. G. Shekelle, Ortiz, Rhodes, & et al., 2001).

To assist in meeting these standards, there are established guideline development organizations that have prepared comprehensive guidance on guideline development methods.

Guideline Development Efforts

Appropriate methodologies and rigorous strategies in the guideline development process are important for the successful implementation of the resulting recommendations and ease of use (Alonso-Coello et al., 2010; Ansari & Rashidian, 2012; Blozik et al., 2012; Cahill & Heyland, 2010; Grol, 2001; Norris, Holmer, Ogden, & Burda, 2011; Steinert, Richter, & Bergk, 2010). Numerous organizations have produced guidelines that provide a methodology for guideline development. These guidelines for guideline development or, ‘methods,’ were

designed in response to the uncertainty regarding the quality of guidelines recommended to professionals (Alonso-Coello et al., 2010; Ansari & Rashidian, 2012; Stamoulos, Reyes, Trepanier, & Drapeau, 2014). The methods provide a systematic development process that are congruent, in large part, with the standards proposed by the IOM (Eccles, Grimshaw, Shekelle, Schünemann, & Woolf, 2012; Graham et al., 2011). Guideline development organizations are comprised of scientists, clinicians, relevant stakeholders, and policy makers from diverse health disciplines and geographical regions who have formed working groups to develop methods intended for CPG developers to create user-friendly and scientifically rigorous guidelines (Beauchamp et al., 2015; NICE, 2007; SIGN 50, WHO, 2010, 2012). The methods are an important tool because many development committee members may have little prior experience and may not be familiar with the guideline development process and the methodological rigorous it requires (Kryworuchko, Stacey, Bai, & Graham, 2009; Norris et al., 2016). Further, members should be offered training and support to facilitate participation on development committees (Fretheim, Schunemann, & Oxman, 2006). To support this training, there are many guidelines development handbooks offered by several organizations. The following are a few of the more prolific guidelines development groups and organizations.

World Health Organization (WHO)

The World Health Organization Guideline (WHO) published the *Handbook for Guidelines Development* (2010) and a second edition in 2014. It contains information on guideline development, clinical practice, and public health policy. The handbook provides detailed procedural instructions on how to develop guidelines to the standard of the WHO. It specifies how to conduct the planning phase, the methodology, the processes and procedures, how to determine the composition of the development group, peer-review, and updating

information. The WHO guideline (2010, 2014) distinguishes itself from other "how to" guideline development manuals by placing a greater emphasis on assembling development teams and external reviewers from various countries and socioeconomic regions, as well as obtaining data from relevant research of various geographical and economic origins (WHO, 2010, 2014).

National Institute for Health and Clinical Excellence (NICE)

The National Institute of Clinical Excellence (NICE) is the main body responsible for the development of clinical guidelines in the United Kingdom. It produces guidance on the topics of quality standards, technology appraisals, clinical, public health and social care guidelines, and NICE implementation tools. They also produce guidance for mental health. Psychological interventions were included as key recommended interventions in 10 out of 13 NICE mental health guidelines (Pilling, 2008).

The NICE approach to development rests on its rigorous methodology, accountability, cost-effectiveness, clinical effectiveness, transparency, and stakeholder involvement (NICE, 2007). Developers must make use of the best possible evidence, from meta-analyses to expert consensus when evidence is lacking (Pilling, 2009). As a first step in the development process, the NICE guideline content must first be informed by a scoping review to identify the key questions. The questions identified are then brought to a multi-disciplinary group of clinicians, academics, guideline methodologists, and all potential service users, using a population, intervention, comparator, and outcomes (PICO) format (Pilling, 2009). This is followed by a systematic review of the literature. An important aspect of the NICE mental health guideline program is that they outline the methods used for forming recommendations, while ensuring any limitations of the evidence are clearly indicated (Pilling, 2012). NICE evaluates the benefits of the

guidelines recommendations by assessing the impact in terms of cost effectiveness, using different cost effectiveness measures.

Scottish Intercollegiate Guidelines Network (SIGN)

The Scottish Intercollegiate Guidelines Network (SIGN) develops and disseminates national clinical guidelines that provide evidence-based recommendations for clinical practice, with the objective of improving patient care in Scotland (SIGN; see http://www.sign.ac.uk). Its membership includes all medical specialties and professionals in allied medical fields, social services, and researchers. They have published the *SIGN Guideline Development Handbook*, with its most recent edition updated in 2015.

Their guidelines are developed by multidisciplinary groups representing diverse geographical locations and disciplines, throughout Scotland. To date they have published over 150 guidelines in the medical field, many of which have been superseded by more recent editions. What is notable about this guideline development organization is their attention to the validity of the guideline. In particular, they have clear indication of how their guidelines ranked in terms of how recently they were published. Guidelines published under three years bear a green checkmark, guidelines that are three to seven-years-old have a yellow question-mark with the caption "Some recommendations may be out of date," "Declarations of interest governance may not be in line with current policy," and guidelines over seven-years-old have an exclamation mark inside a red circle with the caption, "Use with caution" (SIGN; see http://www.sign.ac.uk/guidelines/published/numlist.html).

SIGN has a strong focus on implementation support, stating that they wish to improve the implementability of its recommendations. This means greater efforts toward patient and care-

giver involvement in the guidelines development process, of which there is a section is devoted on their website.

Canadian Medical Association (CMA)

The Canadian Medical Association (CMA) produced a handbook on CPGs titled *The Canadian Medical Association Handbook on Clinical Practice Guidelines* (Davis, Goldman, & Palda, 2007). The handbook is intended for Canadian physicians across specialties and includes an introduction to clinical practice guidelines, how to develop CPGs, how to perform their adaptations, how to implement them, and how to evaluate them (Palda et al., 2007). It states that guideline developers and implementers may be healthcare practitioners, administrators, health organizations, and policy-makers. Experienced guidelines developers and implementers may find within the handbook innovations from the international communities that apply to their work (Palda, et al., 2007). However, absent from this document is the method by which potential committee members are to be searched for and selected for invitation to participate on guideline development committees.

Institut national d'excellence en santé et en services sociaux (INESSS)

Quebec is leading other Canadian provinces in the area of promoting guidelines. *The Institut national d'excellence en santé et en services sociaux* (INESSS) of Quebec was created on January 19, 2011. INESSS created a taskforce whose purpose was to provide a methodology for guideline development and dissemination in the social and human sciences (Beauchamp, Drapeau, Dionne, et al., 2015). The organization's mission includes promoting guideline adoption among end-users, and to ensure that they are included in the public health plan. Chief among their recommendations for developing quality guidelines is the emphasis on the proper

analysis and synthesis of different types of data when formulating recommendations, and instructions on how to grade the recommendations for practice (Beauchamp, et al., 2015).

Guideline International Network (G-I-N)

Although not a guideline development group, the Guidelines International Network (G-I-N), founded in 2002, claims to be the largest international guidelines library in the world (G-I-N: see, http:// g-i-n.net). They are a network of individual organizations for guideline developers to collaborate. The organization's website states that their objectives are to promote the systematic development, and international collaboration in guidelines activities to avoid duplication efforts, and to facilitate information sharing, education and knowledge transfer among guidelines developers.

Comparative Assessment of Guideline Handbooks

In a recent study, Ansari and Rashidian (2012) reviewed 19 guidelines development handbooks (see Table 1), including those proposed by NICE, WHO, SIGN, and CMA. The authors included handbooks that were produced by national and international organizations as well as professional bodies working in guideline development. The researchers identified 27 tasks (see Table 2) considered to be important for an evidence-based guideline development process. They submitted these to an international panel of experts asking them to weight each task between zero and five to reflect how important the task was (Ansari & Rashidian, 2012). Their results showed that all 27 tasks were addressed in only three handbooks. Further, of the 27 tasks, 15 (see Table 2 in bold) were deemed "necessary" in 75% of the handbooks. However, the researchers observed significant variation between the guidelines for depth and quality of information on each task. They noted that the tasks that received the least amount of attention were *considering ethical issues* and *piloting.* They proposed that ethical considerations were

poorly addressed, such that only eight handbooks briefly addressed the topic of ethics (Ansari & Rashidian, 2012). The researchers also showed that five of the 19 handbooks did not expand on methodology at all.

The above study by Ansari and Rashidian (2012) illustrates the significant variability between trusted development resources that are ultimately created to give developers a reliable process to create trustworthy guidelines. It is worth noting that the more frequently consulted development tools, the NICE, SIGN50, WHO, and CMA, were among the top five of the 19 examined. Nevertheless, guideline development handbooks have variable emphasis on methodology, which inevitably impacts the quality of guidelines currently available.

Problems with Current Clinical Practice Guidelines

The potential benefits of guidelines are only as good as the quality of the guidelines themselves (Burgers, Cluzeau, et al., 2003; Gordon & Cooper, 2010). The availability of guideline development handbooks and CPGs do not automatically lead to improved health outcome (Bergman, 1999; Cabana, 1999; Ward & Grieco, 1996). Despite the large volume of published guidelines, the methodology is often inadequately defined and varies greatly within and between organizations (Rosenfeld, Shiffman, & Robertson, 2009).

This apparent lack of methodological rigour has led researchers to evaluate the quality of practice guidelines; thus they have been closely scrutinized over the past 15 years. Although some researchers have found that the quality of guidelines has improved over the last decade, they acknowledge that there is a continued need from improvement (Alonso-Coello et al., 2010). In fact, there are several reports that show there is a high variability in the quality of CPGs, many of which were cited as ‘poor’ in quality (Al-Ansary et al., 2013; Graham, Beardall, Carter, Tetroe, & Davies, 2003; Hasenfeld & Shekelle, 2003; Michie & Johnston, 2004; Ruszczyński,

Horvath, Dziechciarz, & Szajewska, 2016).

When provided with poor guidelines, clinicians learn that they cannot rely upon guidelines to make treatment decisions, and can develop a bias against the potential usefulness of any and all guidelines. Kalies and colleagues (2017) wrote a paper that was based on a German national survey on the critical attitudes and beliefs towards guidelines amongst palliative care professionals for adults with incurable cancer. The purpose was to evaluate critical attitudes and beliefs that could negatively impact the implementation of new guidelines and to evaluate the differences within professional groups and medical specializations. The responses of the 1031 respondents showed that skepticism regarding the quality of existing guidelines was high. Almost half expressed doubts about the usefulness of the guidelines, and about a third of professionals indicated that guidelines restrict their treatment options. The main barrier identified by the authors was the respondents' high skepticism about the quality of the guideline and its effective implementation (Kalies et al., 2017). Furthermore, the respondents viewed routine treatments as difficult to change to newer ones. The researchers also reported that the guidelines were viewed as not being up to date (Kalies et al., 2017).

In the social sciences, the development and use of CPGs arose out of a need for practitioners in the field to ground their practice in scientific evidence (Dozois, 2013; Drapeau & Hunsley, 2014; Gambrill, 2003; Ionita & Fitzpatrick, 2014). Social science professionals such as psychologists and social workers, like physicians, are required to provide scientifically supported therapies and interventions. In the past, psychologists and other mental health professionals have used the term 'guidelines' when referring to treatment recommendations based on personal or expert clinical experience and on non-systematic (e.g., narrative) reviews (Parry et al., 2003). Unfortunately, despite a move toward scientifically supported recommendations, there continues

to be a mixed reception for CPGs among psychologists and other mental health workers, as more guidelines become available (Nathan, 1998; Shaner, 2001).

Another factor that impacts the content and the quality of guideline recommendations are with COIs. Shnier and colleagues (2016) examined Canadian guidelines produced by the CMA and the financial relationships of guideline authors with pharmaceutical companies. The researchers considered 350 authors from 28 guidelines and found over 400 instances of financial COIs, as some authors appeared on multiple guidelines (Shnier, Lexchin, Romero, & Brown, 2016). Further, that same study found 75% of the guidelines analyzed had at least one author who disclosed financial COI, and 54% of guidelines has at least one author who disclosed financial COIs with drug manufacturers. Other studies have also found significant COI problems with guidelines authors (Pincus et al., 2017).

There is no question that handbook developers have acted with noble intentions by responding to the needs of CPG developers, who in turn have invested significant time and effort toward producing documents intended to facilitate clinical practice. However, the research has shown that there is a need to verify their quality with regards to the consistency of scientifically validated recommendations, trustworthiness, methodology, and the presence of potential COIs, for guidelines already available to practitioners.

Evaluating the Quality of Guidelines With AGREE II

Given the vast numbers of guidelines that are still actively promoted as reference tools in various disciplines, some of which are five to 10 years old or older, it is worth prioritizing these for evaluation in the interest of public health and professional confidence. Evaluation instruments designed to rate the quality of guidelines were also created to appraise ease of use, methodology, grade the evidence upon which the recommendations are made, and monitor

conflicts of interest (AGREE, 2003; Brouwers et al., 2010; GRADE, 2004). The literature on guideline evaluations continues to expand and simultaneously calls for more rigour in guideline development (Al-Ansary et al., 2013; Pilling, 2008, 2009). The Appraisal of Guidelines for Research and Evaluation (AGREE) is the most widely used instrument for evaluating the equality of guidelines, and is generally considered to be the gold standards of evaluation tools.

Appraisal of Guidelines for Research & Evaluations (AGREE) II

Although there are different appraisal tools developed by other organizations (SIGN, NICE), the AGREE II is widely used (Alonso-Coello et al., 2010; Brouwers et al., 2010). The first version of AGREE was developed out of a need to reduce variability in guideline quality. AGREE was published in 2003 by a group of international guideline developers and researchers called the AGREE Collaboration (AGREE, 2003). The AGREE items were further refined, and a new item was added, "The strength and limitations of the body of evidence are clearly described," and "guideline was piloted among end users" was deleted from the newer AGREE II in 2010 (see www.agreetrust.org). This version of the instrument contains 23 items grouped into six quality domains with a 7-point Likert scale to score each item in the following domains: 1) Scope and purpose of the guideline; 2) Stakeholder involvement; 3) Rigour of development (methodology); 4) Quality and integration of evidence for development; 5) Attention to the psychosocial, cultural, and community characteristics; 6) Applicability and ease of use; 7) Clarity of the presentation; and 8) Editorial independence of the authors. The AGREE II is widely used today and considered to be the gold standard for quality assessment of CPGs in the medical and human services professions (Alonso-Coello et al., 2010; Brouwers et al., 2010). AGREE II instrument is intended to assists in the development of a quality guideline, but it does not assess its clinical recommendations (Burgers et al., 2012). The AGREE Trust recommends

two to four independent reviewers per guideline with a coordinator to manage the reviewers and maintain the anonymity between reviewers. AGREE II does not provide a threshold for the acceptance or rejection of a guideline based on quality.

The AGREE II Clarifies How Guidelines Must Improve

As more and more guidelines become available, healthcare practitioners must assess guidelines recommended to them by their professional associations or other. Some may glance or read through them to get an overall subjective sense of their applicability and usefulness, while others use them in good faith. The AGREE II instrument provides a means to objectively assess a guideline's quality through empirically validated methods (Brouwers et al., 2010).

Unsurprisingly, the AGREE II scale is widely used by researchers. For example, Burgers and colleagues (2003) evaluated 86 clinical guidelines developed by 10 European countries and Canada using the AGREE instrument. The researchers found significant differences between guideline developers and agencies. Government agencies had the highest scores. Scoring differences between government agencies and professional societies were significant on "editorial independence," "rigor of development" and "clarity of presentation." There were no differences between government organizations and professional societies on "stakeholder involvement," " applicability," and "scope of purpose" (Burgers, Cluzeau, et al., 2003). In line with this, Fervers and others (2005) found that guidelines developed by government-supported organizations had higher mean scores than by professional societies. Overall domain scores were better for "rigor of development' and 'clarity of presentation."

In another study, Patel and colleagues (2016) examined the quality of 24 CPGs on traumatic brain injury published between April 2013 and December 2015. Five independent reviewers conducted assessments using the AGREE II instrument. Overall, the CPGs were found to score

high on ‘clarity and presentation,’ ‘scope and purpose,’ and ‘rigor of development.’ This finding agreed with many other guideline assessments reported in the literature (Berrigan, Marshall, McCullagh, Velikonja, & Bayley, 2011; Irani, Rashidian, Yousefi-Nooraie, & Soltani, 2011). However, other researchers did not find the domain ‘rigour of development’ to be as robust (Berrigan et al., 2011; Irani et al., 2011; MacQueen et al., 2016; Watine & Bunting, 2008). The weaker domains were ‘stakeholder involvement’ and ‘editorial independence,’ which were found to be insufficiently described (Patel et al., 2016). The finding that ‘stakeholder involvement’ and ‘editorial independence’ was of poor quality, is congruent with findings on other guidelines assessed in the literature using the AGREE II instrument (Berrigan et al., 2011; MacQueen et al., 2016; Ye, Liu, Cui, & Liu, 2016). Berrigan and colleagues (2011) evaluated the quality of clinical practice guidelines for mild traumatic brain injuries and its symptoms using the AGREE II. As is the case with the other studies, AGREE scores were highest for the ‘scope and purpose,’ and ‘clarity and presentation’ domains. Lower scores were obtained for the domains ‘rigour of development,’ ‘stakeholder involvement,’ ‘editorial independence,’ and ‘applicability’. There was high variation between the quality of the guidelines, in that half of the guidelines evaluated scored well (70%) for ‘rigour of development’ while the other half scored 25% on average.

More troubling perhaps, some guidelines in use today do not pass the minimum requirement for usability. Ye and colleagues (2016) appraised the quality of seven CPGs for stress ulcers using AGREE II, and found that the overall quality was relatively low and that the use of these guidelines was not recommended. Of the seven CPGs examined, the highest scores were for ‘clarity of presentation,’ and the lowest were for ‘editorial independence,’ ‘rigour of development,’ ‘stakeholder involvement,’ and no points were given for editorial independence

(Ye et al., 2016). This is a pattern that is repeated throughout the literature on CPGs evaluated using AGREE.

There is a vast number of studies that evaluate the quality of medical guidelines, in the literature. Only recently have researchers turned their attention toward mental health guidelines. Of the mental health guidelines that were studied, many were found to need improvement. For example, researchers evaluated CPGs for mild to severe depression, published between 2004 and 2014, using the AGREE II instrument (MacQueen et al., 2016). They investigated the presence and quality of 2nd-line treatment recommendations for patients who do not respond to the 1st-line recommendations of SSRIs. None of the guidelines provided recommendations for non-responders to first-line SSRI treatment. Treatment recommendations varied between CPGs. The AGREE II ratings for 'stakeholder involvement' in CPG development, 'editorial independence,' and 'rigor of development' were rated low (MacQueen et al., 2016).

Guidelines for cognitive assessments have also been evaluated. Trepanier and colleagues (2017) assessed the quality of the OPQ's *Guidelines for the Evaluation of Dyslexia in Children* using the AGREE appraisal instrument (Trepanier, Stamoulos, Reyes, 2017). The authors reported that the guideline obtained the highest scores for the domain 'scope and purpose' and 'clarity of presentation' adding that there was still room for improvement in these area. The lowest domain scores were for 'rigour of development' and 'applicability.' The researchers noted that the guidelines did not provide adequate details of the evidence-gathering process and why recommendations were proposed, nor did the guideline offer details about how the recommendations could be put into practice addressing facilitators and barriers.

Discussion

The quality and usefulness of practice guidelines depend upon multiple considerations throughout the guideline development process, from establishing the key research questions to guideline dissemination and utilization. A useful guideline is constructed flexibly to account for the complex relationships between the practitioners, the interventions, delivery setting, and the needs of each patient (Hollon et al., 2014). It may be argued that the two most critical stages of the process are gathering and synthesizing the evidence, and selecting and coordinating the development group. Although we have seen an improvement in the quality of practice guidelines in recent years among some guideline development groups, there is still much room for improvement.

CPGs have become an important resource for health professionals in medicine and social services, with initiatives to develop and update guidelines taking place globally (WHO, 2012). Their primary purpose is to provide evidence-based guidance at a time when reading primary research has become a time-consuming endeavor. Well-constructed, easy to use guidelines, that are based on the latest research synthesized by topic experts, give health practitioners the opportunity to focus more on caregiving and less on reading high volumes of research papers daily (Davidoff et al., 1995).

As CPGs proliferated, researchers discovered that a significant proportion of guidelines were of poor quality. Efforts to provide guideline developers with tools to facilitate the development of high-quality guidelines were put forth by organizations dedicated to healthcare management. The SIGN network was one of the first major organizations to provide health professionals with clinical guidance in the early 1990's. They were later followed by the NICE institute, which was originally set up in the UK in 1999 as a national organization to manage clinical care, until it added guideline development to its mandate in 2005. Although other

countries have followed suit, SIGN and NICE remained the two of the most recognized names in guideline development until the WHO published their own development manuals in 2012 and 2014.

By this time, hundreds of guidelines were already available through the Canadian Medical Association (Palda et al., 2007), the National Guideline Clearing House in the US (AHRQ; see, http://www.ahrq.gov), and other professional associations. With the availability of guideline development manuals, researchers saw a slight trend toward improved quality over time. However, even more recent research has shown that poor quality guidelines are still endorsed by professional organizations. A large number of comparison studies confirmed that, not only were CPGs highly variable in quality, there were important shortcomings shown to be consistent among them. When guidelines were evaluated using the AGREE II instrument, it was found that most of them scored highest in the domains of 'scope of purpose' and 'clarity of presentation,' and lowest for 'rigor of development' and 'editorial independence.' Hence, on the whole, guideline end-users were able to identify the purpose of the guidelines, the population they were intended to benefit, and that the recommendations were clearly presented. It also means that guideline methodology was either inadequately described, poorly conducted, and experts who participated in the development did not disclose their COIs adequately, or failed to recuse themselves when a COI was at risk. Instead they must rely on the insight of the committee experts. This highlights the importance of selecting the most qualified experts to be on guideline development panels, with a balance of scientists, who are able to synthesize the literature, and clinicians who deal with the target population on a daily basis. The inclusion of guideline methodologists was strongly suggested, in order to meet the high standards of an AGREE II evaluation by independent reviewers.

The most surprising finding in the literature was the glaring absence of methodology for selecting balanced teams of developers. Specifically, how to set up a fair and democratic process for finding the top experts on a given topic, and how to assess their expertise through an objective vetting process, prior to formally inviting them to be on a development panel. Development groups would also want to consider using experts in guideline methodology.

When we consider creating guidelines for mental health, we must keep in mind that patient buy-in is critically important. Stakeholder involvement with the patient populations was also a common critique of guidelines, indicating that there was a lack of opportunity for patients to have a voice during CPG development process. Reasons given for not including lay-people were their lack of experience in understanding the scientific literature and having only subjective experience to contribute. The defense for including patients on development teams is precisely the insight that may be garnered from first-hand experiential difficulties with a particular disorder, and the broader consequences they face, which can impact a patient's ability to follow recommendations.

However, practitioner 'buy in' is also very important. The most frequently cited barriers for guideline adoption among practitioners were negative attitudes toward evidence-based practice, and lack of training (Pagoto et al., 2007). Other factors found to influence guideline adherence include clinician disagreement with the content of the guideline, perceived lack of relevance to the patient population or their practice, and resource limitations (Mosavianpour, Collet, Sarmast, & Kissoon, 2016). Researchers suggest there may be three major problematic areas: 1) Guideline knowledge (lack of awareness and familiarity); 2) Attitudes (lack of agreement, outcome expectancy; self-efficacy, and lack of motivation); 3) Behaviour (contextual barriers such as environmental and guideline related factors) (Mosavianpour et al., 2016).

Although these should all be addressed at some point, the development process lays the foundation of a guideline worthy of trust. Today, we have at our disposal good development tools that provide clear guidance for the development process. Future research should endeavor to explain how development groups have been selected thus far, and to find the best methods for selecting guideline development committee members, going forward.

References

AGREE Collaboration. (2003). Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: The AGREE project. *Quality & Safety in Health Care, 12,* 18-23. doi: 10.1136/qhc.12.1.18

AGREE Next Steps Consortium. (2017). *The AGREE II instrument.* [Electronic version]. Retrieved from https://www.agreetrust.org/wp-content/uploads/2017/12/AGREE-II-Users-Manual-and-23-item-Instrument-2009-Update-2017.pdf

Ahmadzai, N., Newberry, S. J., Maglione, M. A., Tsertsvadze, A., Ansari, M. T., Hempel, S., . . . Shanman, R. (2013). A surveillance system to assess the need for updating systematic reviews. *Systematic Revie*ws, *2,* 1-4. doi: 10.1186/2046-4053-2-104

Al-Ansary, L. A., Tricco, A. C., Adi, Y., Bawazeer, G., Perrier, L., Al-Ghonaim, M., . . . Straus, S. E. (2013). A systematic review of recent clinical practice guidelines on the diagnosis, assessment and management of hypertension. *PLoS ONE, 8,* e53744. doi:10.1371/journal.pone.0053744

Alderson, L. J., Alderson, P., & Tan, T. (2014). Median life span of a cohort of National Institute for Health and Care Excellence clinical guidelines was about 60 months. *Journal of Clinical Epidemiology, 67,* 52-55. doi: 10.1016/j.jclinepi.2013.07.012

Alonso-Coello, P., Irfan, A., Solà, I., Gich, I., Delgado-Noguera, M., Rigau, D., . . . Schunemann, H. (2010). The quality of clinical practice guidelines over the last two decades: A systematic review of guideline appraisal studies. *Quality and Safety in Health Care, 19,* e58. doi:10.1136/qshc.2010.042077

Ansari, S., & Rashidian, A. (2012). Guidelines for guidelines: Are they up to the task? A comparative assessment of clinical practice guideline development handbooks. *PLoS ONE, 7,* e49864. doi:10.1371/journal.pone.0049864

Babione, J. M. (2010). Evidence-based practice in psychology: An ethical framework for graduate education, clinical training, and maintaining professional competence. *Ethics & Behavior, 20,* 443-453. doi: 10.1080/10508422.2010.521446

Balogh, Z., Caldwell, E., Heetveld, M., D'Amours, S., Schlaphoff, G., Harris, I., & Sugrue, M. (2005). Institutional practice guidelines on management of pelvic fracture-related hemodynamic instability: Do they make a difference? *Journal of Trauma and Acute Care Surgery, 58,* 778-782. doi: 10.1097/01.ta.0000158251.40760.b2

Barriocanal, A. M., Lopez, A., Monreal, M., & Montané, E. (2016). Quality assessment of peripheral artery disease clinical guidelines. *Journal of Vascular Surgery, 63,* 1091-1098. doi:10.1016/j.jvs.2015.12.040

Beauchamp, S., Drapeau, M., & Dionne, C. (2015). The development of practice guidelines in the social and human sciences. *Canadian Psychology/Psychologie canadienne, 56,* 357-367. doi:10.1037/a0038811

Beauchamp, S., Drapeau, M., Dionne, C., & Duplantie, J. P. (2015). *Institut national d'excellence en santé et en services sociaux (INESSS): Cadre d'élaboration des guides de pratique dans le secteur des services sociaux* [Policy statement]. Retrieved from INESSS website: https://www.inesss.qc.ca/fileadmin/doc/INESSS/Rapports/ServicesSociaux/INESSS_Cadre_elaboration_guides_pratique_servicessociaux.pdf

Beauchamp, S., Duplantie, J. P., & Mercier, C. (2011). *Development of practice guidelines in social care: A position of the INESSS social care practice guideline committee* [Policy statement]. Retrieved from Regroupement provinciale des comités usagers website: http://www.rpcu.qc.ca/pdf/english/INESSS_Practice_Guidelines _Social_Care.pdf

Bergman, D. A. (1999). Evidence-based guidelines and critical pathways for quality improvement. *Pediatrics*, *103,* 225-232.

Bernard, H., Burlingame, G., Flores, P., Greene, L., Joyce, A., Kobos, J. C., . . . McEneaney, A. M. S. (2008). Clinical practice guidelines for group psychotherapy. *International Journal of Group Psychotherapy, 58,* 455-542. doi: 10.1521/ijgp.2008.58.4.455

Berrigan, L., Marshall, S., McCullagh, S., Velikonja, D., & Bayley, M. (2011). Quality of clinical practice guidelines for persons who have sustained mild traumatic brain injury. *Brain Injury, 25,* 742-751. doi:10.3109/02699052.2011.580317

Blozik, E., Nothacker, M., Bunk, T., Szecsenyi, J., Ollenschläger, G., & Scherer, M. (2012). Simultaneous development of guidelines and quality indicators – how do guideline groups act? *International Journal of Health Care Quality Assurance, 25,* 712-729. doi:10.1108 /09526861211270659

Blum, D. J. (1996). The evolution of physician credentialing into managed care selective contracting. *American Journal of Law and Medicine, 22,* 173-179.

Brouwers, M. C., Kho, M. E., Browman, G. P., Burgers, J. S., Cluzeau, F., Feder, G., . . . The AGREE Next Step Consoritum (2010). AGREE II: Advancing guideline development, reporting and evaluation in health care. *Canadian Medical Association Journal, 182,* E839-E842. doi:10.1503/cmaj.090449

Burgers, J. S., Anzueto, A., Black, P. N., Cruz, A. A., Fervers, B., Graham, I. D., . . . Yawn, B. P. (2012). Adaptation, evaluation, and updating of guidelines. *Proceedings of the American Thoracic Society, 9*, 304-310. doi:10.1513/pats.201208-067ST

Burgers, J. S., Cluzeau, F. A., Hanna, S. E., Hunt, C., & Grol, R. (2003). Characteristics of high-quality guidelines. *International Journal of Technology Assessment in Health Care, 19*, 148-157. doi: 10.1017/S026646230300014X

Burgers, J. S., Grol, R., Klazinga, N. S., Makela, M., & Zaat, J. (2003). Towards evidence-based clinical practice: An international survey of 18 clinical guideline programs. *International Journal for Quality in Health Care, 15*, 31-45. doi: 10.1093/intqhc/15.1.31

Cabana, M. D. (1999). Why don't physicians follow clinical practice guidelines?: A framework for improvement. *Journal of the American Medical Association, 282*, 1458-1465. doi: 10.1177/2333794X19847026

Cahill, N. E., & Heyland, D. K. (2010). Bridging the guideline–practice gap in critical care nutrition: A review of guideline implementation studies. *Journal of Parenteral and Enteral Nutrition, 34*, 653-659. doi:10.1177/0148607110361907

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

Davidoff, F., Haynes, B., Sackett, D., & Smith, R. (1995). Evidence based medicine. *British Medical Journal, 310*, 1085-1093.

Davis, Goldman, J., & Palda, V. (2007). A guide to Canadian Medical Association handbook on clinical practice guidelines. *Canadian Medical Association Journal, 177*, 1221-1226. doi: 10.1503/cmaj.070880

Dozois, D. J. A. (2013). Psychological treatments: Putting evidence into practice and practice into evidence. *Canadian Psychology/Psychologie canadienne, 54,* 1-11. doi:10.1037/a0031125

Dozois, D. J. A., Mikail, S. F., Alden, L. E., Bieling, P. J., Bourgon, G., Clark, D., . . . Johnston, C. (2014). The CPA presidential task force on evidence-based practice of psychological treatments. *Canadian Psychology/Psychologie canadienne, 55,* 153-160. doi:10.1037/a0035767

Drapeau, M., & Hunsley, J. (2014). Where's the science? Introduction to a special issue of Canadian Psychology on science in psychology. *Canadian Psychology/Psychologie canadienne, 55,* 145-152. doi: 10.1037/a0037321

Eccles, M. P., Grimshaw, J. M., Shekelle, P., Schünemann, H. J., & Woolf, S. (2012). Developing clinical practice guidelines: Target audiences, identifying topics for guidelines, guideline group composition and functioning and conflicts of interest. *Implementation Science, 7,* 1-8. doi: 10.1037/a0035767

Ellen, M., Brown, A., & Cockerill, R. (2009). Do clinical practice guidelines influence length of stay in ontario acute care hospitals? *Healthcare Quarterly, 12*, 64-71. doi:10.1080/10437797.2003.10779115

Flexner, A. (1910). *Medical education in the United States and Canada: A report to the Carnegie Foundation for the advancement of teaching* (Bulletin No. 4). New York, NY: The Merrymount Press. Retrieved from http://archive.carnegiefoundation.org/pdfs/elibrary/Carnegie_Flexner_Report.pdf

Fretheim, A., Schunemann, H. J., & Oxman, A. D. (2006). Improving the use of research evidence in guideline development: 5. group processes. *Health Resources and Policy Systems, 4*, 17-35. doi:10.1186/1478-4505-4-17

Gambrill, E. (2003). Evidence-based practice: Sea of change or the emperor's new clothes? *Journal of Social Work Education, 39*, 3-23. doi: 10.1080/10437797.2003.10779115

García, L. M., McFarlane, E., Barnes, S., Sanabria, A. J., Alonso-Coello, P., & Alderson, P. (2014). Updated recommendations: An assessment of NICE clinical guidelines. *Implementation Science, 9*, 72-91. doi: 10.1186/1748-5908-9-72

Gordon, J., & Cooper, B. (2010). Talking knowledge — practising knowledge: A critical best practice approach to how social workers understand and use knowledge in practice. *Practice, 22*, 245-257. doi: 10.1080/09503153.2010.495768

Gould, N., & Kendall, T. (2007). Developing the NICE/SCIE guidelines for dementia care: The challenges of enhancing the evidence base for social and health care. *British Journal of Social Work, 37*, 475-490. doi: 1093/bjsw/bcm020

Graham, I. D., Beardall, S., Carter, A., Tetroe, J., & Davies, B. (2003). The state of the science and art of practice guidelines development, dissemination and evaluation in Canada. *Journal of Evaluation in Clinical Practice, 9*, 195-202.

Graham, I. D., Mancher, M., Wolman, D. M., Greenfield, S., & Steinberg, E. (2011). *Clinical practice guidelines we can trust*. Washington, DC: Institue of Medicine National Academies Press.

Graham, I. D., Evans, W. K., Logan, D., O'Connor, A., Palda, V., McAuley, L., . . . Harrison, M. B. (2000). Canadian oncologists and clinical practice guidelines: A national survey of attitudes and reported use. *Oncology, 59*, 283-290. doi: 10.1046/j.1365-2753.2003.003

85.x

Gray, M., Plath, D., Webb, S. A., & Webb, S. (2009). *Evidence-based social work: A critical stance.* London, England: Routledge.

Grimshaw, J., Eccles, M. P., Lavis, J. N., Hill, S. J., & Squires, J. E. (2012). Knowledge translation of research findings. *Implementation Science, 7,* 50-58. doi: 10.1186/1748-5908-7-50

Grimshaw, J., Freemantle, N., Wallace, S., Russell, I., Hurwitz, B., Watt, I., . . . Sheldon, T. (1995). Developing and implementing clinical practice guidelines. *Quality in Healthcare, 4,* 55-63.

Grol, R. (2001). Successes and failures in the implementation of evidence-based guidelines for clinical practice. *Medical care, 39,* II-46-II-54.

Harrison, M. B., Légaré, F., Graham, I. D., & Fervers, B. (2010). Adapting clinical practice guidelines to local context and assessing barriers to their use. *Canadian Medical Association Journal, 182,* E78-E84. doi: 10.1503/cmaj.081232

Hasenfeld, R., & Shekelle, P. G. (2003). Is the methodological quality of guidelines declining in the US? Comparison of the quality of US Agency for Health Care Policy and Research (AHCPR) guidelines with those published subsequently. *Quality and Safety in Health Care, 12,* 428-434. doi: 10.1136/qhc.12.6.428

Hayward, R. S., Guyatt, G. H., Moore, K.-A., McGibbon, K. A., & Carter, A. O. (1997). Canadian physicians' attitudes about and preferences regarding clinical practice guidelines. *Canadian Medical Association Journal, 156,* 1715-1723.

Hershenberg, R., Drabick, D. A., & Vivian, D. (2012). An opportunity to bridge the gap between clinical research and clinical practice: Implications for clinical training. *Psychotherapy, 49*, 123-131.

Hollon, S. D., Areán, P. A., Craske, M. G., Crawford, K. A., Kivlahan, D. R., Magnavita, J. J., . . . Bufka, L. F. (2014). Development of clinical practice guidelines. *Annual Review of Clinical Psychology, 10*, 213-241. doi: 10.1146/annurev-clinpsy-050212-185529

Hudson, P., Quinn, K., O'Hanlon, B., & Aranda, S. (2008). Family meetings in palliative care: Multidisciplinary clinical practice guidelines. *BMC Palliative Care, 7*, 1-6. doi: 10.1186/1472-684X-7-12

Hussain, T., Michel, G., & Shiffman, R. N. (2009). The Yale guideline recommendation corpus: A representative sample of the knowledge content of guidelines. *International Journal of Medical Informatics, 78*, 354-363. doi: 10.1016/j.ijmedinf.2008.11.001

Ionita, G., & Fitzpatrick, M. (2014). Bringing science to clinical practice: A Canadian survey of psychological practice and usage of progress monitoring measures. *Canadian Psychology/Psychologie canadienne, 55*, 187-196. doi: 10.1037/a0037355

Irani, S., Rashidian, A., Yousefi-Nooraie, R., & Soltani, A. (2011). Evaluating clinical practice guidelines developed for the management of thyroid nodules and thyroid cancers and assessing the reliability and validity of the AGREE instrument. *Journal of Evaluation in Clinical Practice, 17*, 729-736. doi: 10.1111/j.1365-2753.2010.01561.x

Kalies, H., Schöttmer, R., Simon, S. T., Voltz, R., Crispin, A., & Bausewein, C. (2017). Critical attitudes and beliefs towards guidelines amongst palliative care professionals – Results from a national survey. *BMC Palliative Care, 16*, 20-24. doi:10.1186/s12904-017-0187-y

Kendall, T., Taylor, E., Perez, A., & Taylor, C. (2008). Guidelines: Diagnosis and management of attention-deficit/hyperactivity disorder in children, young people, and adults: Summary of NICE guidance. *British Medical Journal, 337*, 751-753. doi: 10.1136/bmj.a1239

Köhler, S., Hoffmann, S., Unger, T., Steinacher, B., & Fydrich, T. (2012). Adherence to guidelines and effectiveness of inpatient treatment for unipolar depression. *International Journal of Psychiatry in Clinical Practice, 16,* 103-112. doi: 10.3109/13651501.2011.638142

Kryworuchko, J., Stacey, D., Bai, N., & Graham, I. D. (2009). Twelve years of clinical practice guideline development, dissemination and evaluation in Canada (1994 to 2005). *Implementation Science, 4*, 49-56. doi:10.1186/1748-5908-4-49

Légaré, F., Boivin, A., van der Weijden, T., Pakenham, C., Burgers, J., Légaré, J., … Gagnon, S. (2011). Patient and public involvement in clinical practice guidelines: A knowledge synthesis of existing programs. *Medical Decision Making*, *31,* 45–74. doi:10.1177/0272989X11424401

Lesho, E. P., Myers, C. P., Ott, M., Winslow, C., & Brown, J. E. (2005). Do clinical practice guidelines improve processes or outcomes in primary care? *Military Medicine*, *170*, 243-246. doi: 10.7205/MILMED.170.3.243

MacQueen, G., Santaguida, P., Keshavarz, H., Jaworska, N., Levine, M., Beyene, J., & Raina, P. (2016). Systematic review of clinical practice guidelines for failed antidepressant treatment response in major depressive disorder, dysthymia, and sub-threshold depression in adults. *The Canadian Journal of Psychiatry, 62*, 11–23. doi:10.1177/0706743716664885

Margo, C. E. (2004). Quality care and practice variation: The roles of practice guidelines and public profiles. *Survey of ophthalmology, 49,* 359-371. doi:10.1016/j.survophthal.2004.02.010

Martínez, R., Reyes, S., Lorenzo, M. J., & Menéndez, R. (2009, April). *Impact of guidelines on outcome: The evidence.* Paper presented at the Seminars in Respiratory and Critical Care Medicine, Valencia, Spain. Retrieved from https://www.thiemeconnect.com/products/ejournals/html/10.1055/s-0029-1202936

Menéndez, R., Ferrando, D., Vallés, J. M., & Vallterra, J. (2002). Influence of deviation from guidelines on the outcome of community-acquired pneumonia. *Chest, 122,* 612-617.

Mental Health Commission of Canada. (2012). *Changing directions, changing lives: The mental health strategy for canada. Calgary.* Calgary, AB: Author. Retrieved from https://www.mentalhealthcommission.ca/sites/default/files/MHStrategy_Strategy_ENG.pdf

Michie, S., & Johnston, M. (2004). Changing clinical behaviour by making guidelines specific. *British Medical Journal, 328,* 343-345.

Miville, M. L., Duan, C., Nutt, R. L., Waehler, C. A., Suzuki, L., Pistole, M. C., . . . Corpus, M. (2009). Integrating practice guidelines into professional training: Implications for diversity competence. *The Counseling Psychologist, 37,* 519-563. doi: 10.1177/0011000008323651

Mosavianpour, M., Collet, J.-P., Sarmast, H., & Kissoon, N. (2016). Barriers to the implementation of sepsis guideline in a Canadian pediatric tertiary care centre. *Journal of Nursing Education and Practice, 6,* 34. doi:10.5430/jnep.v6n12p34

Nathan, P. E. (1998). Practice guidelines: Not yet ideal. *American Psychologist, 53*(3), 290. doi: 10.1037/0003-066X.53.3.290

Norris, S. L., Holmer, H. K., Ogden, L. A., & Burda, B. U. (2011). Conflict of interest in clinical practice guideline development: A systematic review. *PLoS ONE, 6*. doi:10.1371/journal .pone.0025153

Norris, S. L., Meerpohl, J. J., Akl, E. A., Schünemann, H. J., Gartlehner, G., Chen, Y., & Whittington, C. (2016). The skills and experience of GRADE methodologists can be assessed with a simple tool. *Journal of Clinical Epidemiology, 79,* 150-158. doi:10.1016/j.jclinepi.2016.07.001

Oxman, A. D., Fretheim, A., & Schunemann, H. J. (2006). Improving the use of research evidence in guideline development: Introduction. *Health Resources and Policy Systems, 4*, 12-18. doi: 10.1186/1478-4505-4-12

Palda, V. A., Davis, D., & Goldman, J. (2007). A guide to the Canadian Medical Association handbook on clinical practice guidelines. *Canadian Medical Association Journal, 177*, 1221-1226. doi:10.1503/cmaj.070880

Parry, G., Cape, J., & Pilling, S. (2003). Clinical practice guidelines in clinical psychology and psychotherapy. *Clinical Psychology & Psychotherapy, 10,* 337-351. doi:10.1002/cpp. 381

Patel, A., Vieira, M. M. C., Abraham, J., Reid, N., Tran, T., Tomecsek, K., . . . Staton, C. A. (2016). Quality of the development of traumatic brain injury clinical practice guidelines: A systematic review. *PLoS ONE, 11*, e0161554. doi:10.1371/journal.pone.0161554

Pawson, R., Boaz, A., Grayson, L., Long, A., & Barnes, C. (2003). *Types and quality of knowledge in social care* (Report No. 3). London, England: Queen Mary University Press. Retrieved from www.kcl.ac.uk/sspp/departments/politicaleco nomy/research/cep/pubs/papers/assets/wp18.pdf

Peterson, K., McDonagh, M. S., & Fu, R. (2011). Decisions to update comparative drug effectiveness reviews vary based on type of new evidence. *Journal of Clininical Epidemiology, 64*, 977-984. doi: 10.1016/j.jclinepi.2010.11.019

Pilling, S. (2008). History, context, process, and rationale for the development of clinical guidelines. *Psychology and Psychotherapy: Theory, Research and Practice, 81*, 331-350. doi:10.1348/147608308X324923

Pilling, S. (2009). Developing evidence-based guidance – implications for systemic interventions. *Journal of Family Therapy, 31*, 194-205. doi:10.1111/j.1467-6427.2009.00462.x

Pilling, S. (2012). *Developing clinical guidelines for adults: Handbook of Evidence-Based Practice in Clinical Psychology*. Hobogen, New Jersey: John Wiley & Sons, Inc.

Pincus, D., Kuhn, J. E., Sheth, U., Rizzone, K., Colbenson, K., Dwyer, T., … Wasserstein, D. (2017). A systematic review and appraisal of clinical practice guidelines for musculoskeletal soft tissue injuries and conditions. *The American Journal of Sports Medicine, 45*, 1458–1464. doi: 10.1177/0363546516667903

Qaseem, A., Forland, F., Macbeth, F., Ollenschläger, G., Phillips, S., & van der Wees, P. (2012). Guidelines international network: Toward international standards for clinical practice guidelines. *Annals of Internal Medicine, 156*, 525-531. doi:10.7326/0003-4819-156-7-201204030-00009

Ransohoff, D. F., Pignone, M., & Sox, H. C. (2013). How to decide whether a clinical practice guideline is trustworthy. *Journal of the American Medical Association, 309*, 139-140. doi: 10.1001/jama.2012.156703

Rosenfeld, R. M., & Shiffman, R. N. (2009). Clinical practice guideline development manual: a quality-driven approach for translating evidence into action. *Otolaryngology Head and Neck Surgery, 148,* 1-55. doi: 10.1016/j.otohns.2009.04.015

Ruszczyński, M., Horvath, A., Dziechciarz, P., & Szajewska, H. (2016). Cow's milk allergy guidelines: A quality appraisal with the AGREE II instrument. *Clinical & Experimental Allergy, 46,* 1236-1241. doi: 10.1111/cea.12784

Rutten, G. M., Degen, S., Hendriks, E. J., Braspenning, J. C., Harting, J., & Oostendorp, R. A. (2016). Adherence to clinical practice guidelines for low back pain in physical therapy: Do patients benefit? *Physical Therapy, 90,* 1111-1122. doi: 10.2522/ptj.20090173

Sackett, D. L., Rosenberg, W. M. C., Gray, J. A. M., Haynes, R. B., & Richardson, W. S. (1996). Evidence based medicine: What it is and what it isn't. *British Medical Journal, 312,* 70-72. doi:10.1136/bmj.312.7023.71

Schuh, S., Babl, F. E., Dalziel, S. R., Freedman, S. B., Macias, C. G., Stephens, D., . . . Plint, A. C. (2017). Practice variation in acute bronchiolitis: A pediatric emergency research networks study. *Pediatrics,* e20170842. doi: 10.1542/peds.2017-0842

Schünemann, H. J., Wiercioch, W., Etxeandia, I., Falavigna, M., Santesso, N., Mustafa, R., . . . al., e. (2014). Guidelines 2.0: Systematic development of a comprehensive checklist for a successful guideline enterprise. *Canadian Medical Association Journal, 186,* E123-E142. doi:10.1503/cmaj.131237

Shaner, R. (2001). Practical psychotherapy: How practice guidelines can rescue psychotherapy in public systems. *Psychiatric Services, 52,* 1021-1025. doi:10.1176/appi.ps.52.8.1021

Shaneyfelt, T. M., Mayo-Smith, M. F., & Rothwangl, J. (1999). Are guidelines following guidelines?: The methodological quality of clinical practice guidelines in the peer-

reviewed medical literature. *Journal of the American Medical Association, 281*, 1900-1905. doi:10.1001/jama.281.20.1900

Shekelle. (2014). Updating practice guidelines. *Journal of the American Medical Association, 311*, 2072-2074. doi:10.1001/jama.2014.4950

Shekelle, P. G., Ortiz, E., Rhodes, S., & et al. (2001). Validity of the agency for healthcare research and quality clinical practice guidelines: How quickly do guidelines become outdated? *Journal of the American Medical Association, 286*, 1461-1467. doi:10.1001/jama.286.12.1461

Shiroiwa, T., Fukuda, T., Ikeda, S., Takura, T., & Moriwaki, K. (2016). Development of an official guideline for the economic evaluation of drugs/medical devices in Japan. *Value in Health. 20*, 372-378. doi 10.1016/j.jval.2016.08.726

Shnier, A., Lexchin, J., Romero, M., & Brown, K. (2016). Reporting of financial conflicts of interest in clinical practice guidelines: A case study analysis of guidelines from the Canadian Medical Association Infobase. *BMC Health Services Research, 16*, 1-8. doi:10.1186/s12913-016-1646-5

Scottish Intercollegiate Guidelines Network (SIGN). (2001). *SIGN 50: A guideline developer's handbook*. Edinburgh, Scottland: Author

Scottish Intercollegiate Guidelines Network (SIGN). (2015). *SIGN50: A guideline developer's handbook* [Electronic version]. Retrieved from https://sign.ac.uk/assets/sign50_2015.pdf

Stamoulos, C., Reyes, A., Trepanier, L., & Drapeau, M. (2014). The methodological quality of clinical practice guidelines in psychology: A pilot study of the guidelines developed by the Ordre des Psychologues du Québec. *Canadian Psychology/Psychologie canadienne, 55*, 177-186. doi:10.1037/a0037287

Steinert, T., Richter, D., & Bergk, J. (2010). Sources of bias in guideline development. *Psychiatric Services, 61,* 946. doi: 10.1176/appi.ps.61.9.946

Straus, S. E., Tetroe, J., & Graham, I. D. (2009). Defining knowledge translation. *Canadian Medical Association Journal, 181,* 165-168. doi:10.1503/cmaj.081229

The ADAPTE Collaboration (2009). *The ADAPTE process: Resource toolkit for guideline adaptation* (Version 2.0). Retrieved from http://wwwg-i-n.net.

Timmermans, S., & Kolker, E. S. (2004). Evidence-based medicine and the reconfiguration of medical knowledge. *Journal of Health and Social Behaviour, 45,* 177-193.

Trepanier, L., Stamoulos, C., & Reyes, A. (2017). Assessing the quality of the OPQ's guidelines for the evaluation of dyslexia in children using the Appraisal of Guidelines for Research and Evaluation (AGREE) II: A brief report. *Canadian Psychology/Psychologie canadienne, 58,* 211-217. doi:10.1037/cap0000121

van Dijk, M. K., Oosterbaan, D. B., Verbraak, M. J., & van Balkom, A. J. (2013). The effectiveness of adhering to clinical-practice guidelines for anxiety disorders in secondary mental health care: the results of a cohort study in the Netherlands. *Journal of Evaluation in Clinical Practice, 19*, 791-797. doi: 10.1111/j.1365-2753.2012.01851.x

Watine, J. C., & Bunting, P. S. (2008). Mass colorectal cancer screening: Methodological quality of practice guidelines is not related to their content validity. *Clinical Biochemistry, 41*, 459-466. doi:10.1016/j.clinbiochem.2007.12.020

Weinberger, M. (2009). Seventeen years of asthma guidelines: Why hasn't the outcome improved for children? *The Journal of Pediatrics, 154*, 786-788. doi: 10.1016/j.peds,2009.01.003

Weisz, G., Cambrosio, A., Keating, P., Knaapen, L., Schlich, T., & Tournay, V. J. (2007). The emergence of clinical practice guidelines. *Milbank Quarterly, 85*, 691-727. doi:10.1111/j.1468-0009.2007.00505.x

World Health Organization (WHO). (2003). *Global programme on evidence for health policy: Guidelines for WHO guidelines.* Geneva, Switzerland: Author. Retrieved from https://apps.who.int/iris/bitstream/handle/10665/68925/EIP_GPE_EQC_2003_1.pdf?sequence=1&isAllowed=y

World Health Organization (WHO). (2010). *WHO handbook for guideline development.* Geneva, Switzerland: Author. Retrieved from http://www.who.int/hiv/topics/mtct/grc_handbook_mar2010_1.pdf

Wold Health Organization (WHO). (2012). *WHO handbook for guideline development*. Geneva, Switzerland: Author. Retrieved from https://apps.who.int/iris/bitstream/handle/10665/75146/9789241548441_eng.pdf?sequence=1

Wolf, J. S., Hubbard, H., Faraday, M. M., & Forrest, J. B. (2011). Clinical practice guidelines to inform evidence-based clinical practice. *World Journal of Urology, 29*, 303-309. doi: 10.1007/s00345-011-0656-5

Wollersheim, H., Burgers, J., & Grol, R. (2005). Clinical guidelines to improve patient care. *Netherlands Journal of Medicine, 63*, 188-192.

Woolf, S. H., Grol, R., Hutchinson, A., Eccles, M., & Grimshaw, J. (1999). Clinical guidelines: Potential benefits, limitations, and harms of clinical guidelines. *British Medical Journal, 318*, 527-536. doi: 10.1136/bmj.318.7182.527

Ye, Z. K., Liu, Y., Cui, X., & Liu, L. (2016). Critical appraisal of the quality of clinical practice guidelines for stress ulcer prophylaxis. *PLoS ONE, 11*, e0155020. doi:10.1371/journal.pone.0155020

Table 1.

Guideline Development Handbooks

Handbook	Publication Year	Country of Origin	Website
Advisory Committee on Health Research (ACHR)	2006	International	www.who.int/rpc/advisory_committee
American Society of Clinical Oncology (ASCO)	2011	USA	www.asco.org
American Society of Clinical Oncology (ASCO)	2011	USA	www.asco.org
Australian Health Policy Institute (AHPI)	2008	Australia	healthpolicystudies.org.au
Canadian Medical Association (CMA)	2007	Canada	www.accesscopyright.ca
Cancer Care Ontario (CCO)	2011	Canada	www.cancercare.on.ca
Council of Europe (CE)	2001	International	www.social.coe.int
International Diabetes Federation (IDF)	2003	International	www.idf.org
National Institute for Health and Clinical Excellence (NICE)	2009	UK	www.nice.org.uk
National Health and Medical Research Council (NHMRC)	1998	Australia	www.health.gov.au/nhmrc
New Zealand Guideline Group	2001	New Zealand	www.nzgg.org.nz

(NZGG)			
Regional Centre for Quality of Health Care (RCQHC)	2003	Regional	www.rcqhc.org
Royal College of Psychiatrists (RCP)	1994	UK	www.rcpsych.ac.uk
Scottish Intercollegiate Guidelines Network (SIGN)	2008	Scotland	www.sign.ac.uk
Swiss Centre for International Guidelines Network (SCIH)	2011	Swiss	http://www.swisstph.ch/
The Chartered Society of Physiotherapy (CSP)	2006	UK	www.csp.orp.org.uk
U.S. Preventative Services Task Force (UPSTF)	2008	USA	www.preventiveservices.ahrq.gov
World Confederation of Physiotherapy (WCPT)	2006	International	www.wcpt.org
World Health Organization (WHO)	2012	International	www.who.int
World Stroke Organization (WSO)	2009	International	www.world-stroke.org

Table 2.

Guideline development tasks identified by Shebnam Ansari and Arash Rashidian 2012

Task	Definition
1.Selecting the guideline topic	The process and criteria for selecting and prioritizing topics.
2.Determining the guideline scope	A framework that describes the epidemiology of the disease or condition and the aspects of care and the settings is covered by the guideline.
3. Preparing the work plan	An objective search of important and relevant databases and search engines for existing guidelines.
4. Identifying relevant existing guidelines	Objective appraisal of existing guidelines e.g. by using AGREE.
5. Appraising relevant existing guidelines	Objective appraisal of existing guidelines e.g. by using AGREE.
6. Adapting existing guidelines	Describing guideline adaption methods.
7. Involving consumers (patients…)	Contribution of the target population (patients, public, etc.) in relevant tasks.
8. Forming guideline development group	Describing the composition of guideline development group, including all relevant stakeholders.
9. Managing conflict of interests	Declaration of guideline development group members competing interests.
10. Running guideline development group	Describing how to run a GDG (meetings, agenda items, chairing, responsibilities and roles).
11. Developing clinical questions	Developing clinical question according to an objective approach, e.g. PICO framework.
12. Systematic search for evidence	Systematic searches of important bibliographic databases using.

	sensitive key words
13. Selecting relevant evidence for search results	Inclusion and exclusion criteria for selecting the evidence.
14. Appraising identified research evidence	Appraising identified evidences using objective instruments (for example CASP tools).
15. Evidence synthesis and analysis	Describing synthesis approaches of primary studies, including meta-analysis etc.
16. Conducting economic evaluation	Describing the process of identifying, selecting and synthesizing economic evaluation data.
17. Making group decisions	Using clear and objective consensus development techniques (e.g. voting, Delphi…).
18. Grading available evidence	Appraising and summarizing the quality and strength of recommendations.
19. Consider ethical issues	Discussing the approaches used for considering ethical issues in the guideline development process.
20. Creating recommendations	Interpreting the evidence to make recommendations and the wording and format of recommendations.
21. Final stakeholder consultation	Final consultation with stakeholders before publishing the guideline.
22. Publishing formats	Describing different publication formats (full guideline, quick reference guides, information for patient, wed-based publication).
23. Guideline implementation strategies	Describing how the recommendations can be put into Practice.
24. Piloting the developed guidelines	Describing a process of pre-testing a guideline in the field before its final release.
25. Assessment the potential impacts	The cost and resource implications of implementing the guideline in

of guideline implementation	practice.
26. Developing clinical audit and evaluation criteria	Describing monitoring and auditing criteria and indicators to assess guideline implementation.
27. Updating recommendations and correcting potential errors	Describes the process, timeline, frequency and criteria for updating recommendations or correcting errors.

(*Note:* Tasks and definitions published by S. Ansari and A. Rashidian, (2012). *Guidelines for Guidelines: Are they up to the task? A comparative assessment of clinical practice guideline development handbooks*.)

Brief Overview

Chapter 1 showed that practice guidelines offer several service delivery and cost-saving benefits; they also require vast amounts of time and resources to produce. However, the literature also showed that there is some variability in the quality of practice guidelines that are available to practitioners in the medical field, and only a small number of studies were conducted to evaluate psychology guidelines. With the exception of our study (Trepanier et al., 2017), which was part of this larger project, no studies to date have examined the quality of the clinical practice guidelines produced by the Ordre des psychologues du Québec (OPQ). In the following first manuscript, we examined the quality of five OPQ guidelines. Although the OPQ produces many types of guidelines, such as Guidelines for Keeping Records, and an Explanatory Guide to the Code of Ethics of Quebec Psychologists for example, the five guidelines in this study were selected because they comprised all of the available clinical practice guidelines (for assessment and treatment), produced by the OPQ at the time of this study. Manuscript 1 (Chapter 2) provides a thorough assessment of each guideline, explaining specific strengths and weaknesses in quality.

Next, manuscript 2 (Chapter 3) is presented, which examined the group composition of the OPQ guideline development committees, and the relevant research productivity of committee members. Lastly, manuscript 3 (Chapter 4) reports the findings of a survey used to investigate the views and experiences of guideline development committee members who worked on one of the practice guidelines produced by the OPQ along with five other Quebec regulatory bodies.

Chapter 2

Manuscript 1

(Paper to be submitted)

Assessing the Quality of Five Clinical Practice Guidelines Using the Appraisal of Guidelines for Research and Evaluation (AGREE) II.

Trepanier, Lyane[1], Stamoulos, Constantina[1], Reyes, Andrea[1], MacIntosh, Heather[2], Beauchamp, Sylvie [1,3], Larivée, Serge[4], & Drapeau, Martin[1,5]

McGill University, Montréal, Canada

Author Note

Department of Counselling Psychology, McGill University[1], School of Social Work, McGill University[2], Montreal West Island Integrated University Health and Social Services Center[3], School of Psychoeducation, University of Montreal[4], Department of Psychiatry, McGill University[5]

Lyane Trepanier's was supported by the SSHRC Joseph-Bombardier Canada Graduate Scholarship Doctoral Award. The work of Martin Drapeau was supported by the Social Sciences and Humanities Research Council (SSHRC) Investigator Award. The work of Constantina Stamoulos was supported by SSHRC Doctoral Award and the Fonds Québécois de la Recherche sur la Société et la Culture (FQRSC) Doctoral Research Training Scholarship, and the work of

Andrea Reyes was supported by the FQRSC Doctoral Research Training Scholarship and the McGill GPS Pilot Project Recruitment Award. No funding body had any input into any aspect of this research.

Correspondence address: Dr. Martin Drapeau, MPPRG – McGill University, 3700 McTavish, Montreal, Quebec, H3A 1Y2, Canada. E-mail: martin.drapeau@mcgill.ca

Abstract

Clinical Practice Guidelines (CPGs) have been shown to improve healthcare services and clinical outcomes. Although CPGs are useful resources for mental health and social service professionals, they are helpful only to the degree that they are developed according to the most rigorous standards. However, multiple studies have demonstrated significant variability between guidelines with regards to specific indicators of quality. The Ordre des psychologues du Québec (OPQ; College of psychologists of Quebec) has published a number of CPGs that are currently available to psychologists, including the *Guidelines for the Evaluation of Dyslexia in Children* (2014), the *Guidelines for Autism Spectrum Disorder - Clinical Evaluation* (2012), the *Guidelines for the Assessment of Mental Retardation* (2007), the *Guidelines for the Assessment of a Child in Connection with a Request for Derogation to the Age of School Admission* (2006), and the *Guidelines for Expert Assessment Concerning Child Custody and Access Rights* (2006). These CPGs are intended to provide empirically supported guidance for psychologists in the areas of assessment, diagnosis, general functioning, treatment, and other decision-making support. The aim of this study was to evaluate the quality of these guidelines using the Appraisal of Guidelines for Research and Evaluation II (AGREE II) guideline evaluation instrument. Our results show that although there have been some modest improvements in quality of the CPGs over time, there are important methodological inadequacies in all five CPGs that must be addressed in future guideline updates. The findings of this study demonstrate the need for more methodological rigour in guideline development.

Keywords: AGREE II, clinical practice guidelines, assessment, autism, cognitive assessments, dyslexia, mental retardation, child custody, psychological evaluation.

Assessing the Quality of Five Clinical Practice Guidelines Using the Appraisal of Guidelines for Research and Evaluation (AGREE) II.

Clinical practice guidelines (CPGs) are a useful and an often-critical resource for health practitioners in diverse clinical settings. They are used by a number of healthcare service providers including, but not limited to, physicians, social workers, and psychologists (Bernard et al., 2008; Gordon & Cooper, 2010; Hudson, Quinn, O'Hanlon, & Aranda, 2008; Kendall, Taylor, Perez, & Taylor, 2008; OPQ, 2007; van Dijk, Oosterbaan, Verbraak, & van Balkom, 2013). Contrary to what is often believed, CPGs are not procedural manuals that specify precise instructions and strict procedural processes; rather, they provide general guidance intended to enhance the users' knowledge and on which they can base their clinical judgment. A high quality guideline is grounded in scientific evidence, which has been synthesized from multiple peer-reviewed sources, with well-formulated recommendations (Graham, Mancher, Wolman, Greenfield, & Steinberg, 2011). As such, it is an important knowledge synthesis and dissemination tool that brings science to those who are expected to use it: practitioners. Guidelines also serve to discourage outdated and ineffective approaches that have been shown to be ineffective or put patients at risk (Shaner, 2001).

The production of high quality guidelines requires a significant investment of time, expertise, and financial resources to support the multiple stages involved in the planning, production, and implementation of the guideline (Ansari & Rashidian, 2012; Beauchamp, Duplantie, & Mercier, 2011; Hollon et al., 2014; WHO, 2010). Such investments are justified when the use of the guidelines lead to clear and measurable benefits to end-users and their patients. Thus, the widening interest and motivation to produce CPGs are due to several factors. They enhance practitioner knowledge in a convenient format that is well suited to a fast pace environment (Graham et al., 2011; Wolf, Hubbard, Faraday, & Forrest, 2011), they improve

treatment processes, outcomes, and costs (Balogh et al., 2005; Graham et al., 2000; Hollon et al., 2014; Rutten et al., 2016; van Dijk et al., 2013), they have been shown to reduce practice variation between practitioners and treatment locations (Mental Health Commission of Canada, 2012; Pilling, 2008; Shiroiwa, Fukuda, Ikeda, Takura, & Moriwaki, 2016; Wolf et al., 2011), and can also serve as practical and expedient educational resources for practitioners and clinicians-in-training (Hershenberg, Drabick, & Vivian, 2012; Miville et al., 2009). For example, in psychology, some universities use CPGs developed by regulatory bodies such as Quebec's College of psychologists (Ordre des psychologues du Québec; OPQ) and other institutions in psychology as part of their clinical training programs (Stamoulos, Reyes, Trepanier, & Drapeau, 2014). Therefore, good guidelines serve multiple purposes and provide solutions to diverse challenges.

Practice guidelines are typically published by regulatory bodies, governmental agencies, and professional associations. Licensing boards and regulatory bodies are charged with the responsibility of overseeing the ethical and responsible conduct of their members, and to ensure that the practitioners under their jurisdiction observe the highest clinical assessment and treatment standards. The regulatory bodies that produce guidelines in Quebec include the *Collège des médecins du Québec* (the Quebec College of Physicians; CMQ), the *Ordre des psychologues du Québec* (Order of Psychologists of Quebec; OPQ), the *Ordre des travailleurs sociaux et des thérapeutes conjugaux et familiaux du Québec* (the Order of Social Workers and Family Therapists of Quebec), the *Ordre des psychoéducateurs et psychoéducatrices du Québec* (the Ordre of Psychoeducators of Quebec), and the *Ordre des conseillers et conseillières du Québec* (the Order of Counsellors of Quebec). All of these Orders have demonstrated interest in producing guidelines for their professionals to have accessible guidance at their disposal, with some producing guidelines on a regular basis.

However, the usefulness of guidelines depends upon their quality (Burgers, Cluzeau, Hanna, Hunt, & Grol, 2003; Gordon & Cooper, 2010). Hence, with the growing interest in the development of CPGs and their increasing dissemination, there is a need to assess the quality of the work we, as health and psychosocial professions, have produced. Healthcare providers may wish to undertake their own assessment of a guideline before adopting its recommendations in their practice. Policy makers may be interested in conducting structured guideline assessments when determining which guidelines should be recommended for practice or to inform policy. Educators may want to assess guidelines to help enhance critical appraisal skills amongst professionals and to teach core competencies in guidelines development and reporting. Finally, guidelines developers who follow a structured and rigorous development methodology, may want to construct an internal assessment to ensure that their guidelines are sound, or evaluate guidelines from other groups for potential adaptation to their own context (ADAPTE-Collaboration, 2009; Graham, et al., 2011).

The Quality of Current Practice Guidelines

Because the potential benefits of guidelines depend heavily on the quality of the guidelines themselves (Burgers, Cluzeau, Hanna, Hunt, & Grol, 2003; Gordon & Cooper, 2010), appropriate methodologies and rigorous strategies in the guideline development process are important to provide high quality recommendations, ease of use, and successful implementation (Alonso-Coello et al., 2010). Several reports have highlighted the high degree of variability in the quality of CPGs, many of which were cited as ‘poor’ in quality (Al-Ansary et al., 2013; Graham, Beardall, Carter, Tetroe, & Davies, 2003; Hasenfeld & Shekelle, 2003; Ruszczyński, Horvath, Dziechciarz, & Szajewska, 2016). For these reasons clinicians risk not trusting guidelines (Kalies et al., 2017).

Good CPGs are developed with transparency with the majority of the developers having few or no conflicts of interest (COI), which when present, are well documented (Graham et al., 2011). They have a knowledgeable development group composed of relevant stakeholders (e.g., guideline users, caregivers, and affected populations) (Brouwers et al., 2010; Graham et al., 2011), and contain recommendations derived from evidence collated from systematic reviews and discerned by experts in the field for their quality and strength. Finally, quality CPGs are submitted to an external review process, and have established criteria for guideline updates (Brouwers et al., 2010; Graham et al., 2011).

Social science professionals such as psychologists, psychoeducators, and social workers, like other professionals, could benefit from quality guidelines, since they are professionally required to provide scientifically supported therapies and interventions. However, there has been a mixed reception for CPGs among psychologists and other mental health workers (Bennett, Courtney, Duda, Henderson, & Szatmari, 2018; van Dijk et al., 2013), and recent research has demonstrated a need to improve the quality of guidelines made available to psychologists (Stamoulos et al., 2014; Trepanier, Stamoulos, & Reyes, 2017). For example, Stamoulos and colleagues (2014) examined the methodological rigour of five CPGs developed by the OPQ on different broad tasks that are considered important to the development process (see Ansari & Rashidian, 2012), to determine if they were consistent with generally agreed-upon guideline development recommendations. They found that CPG developers failed to utilize almost all the tasks deemed 'essential' for the development process. Only a small number of tasks that were 'generally important,' but not essential, were used for the development of the CPGs studied. Trepanier and colleagues (2017) evaluated the quality of the OPQ's *Guidelines for the Evaluation of Dyslexia in Children* using the Appraisal of Guidelines for Research and Evaluation II (AGREE II) guideline evaluation instrument, which is a widely-used tool in guideline evaluation

research (AGREE, 2017; Brouwers, et al., 2010). Their results showed that this guideline was also developed with important methodological flaws, such as not describing the search methods used to support the underlying evidence to support the evidence (Alonso-Coello et al., 2010), and suggested improvements in the development methodology. These studies raise important questions about the quality of the guidelines currently available to psychologists or other practitioners. Many other CPGS are available to psychologists, but have yet to be assessed.

Hence, the present study assessed the quality of five OPQ guidelines that are currently available to Quebec psychologists using the AGREE II appraisal instrument (AGREE, 2017). In a previous study, we assessed the OPQ guidelines titled *Guidelines for the Evaluation of Dyslexia in Children* (2014) (see Trepanier et al., 2017). With the current study, we further examined those guidelines and examined four other CPGs developed and published by the OPQ that are currently available to Quebec psychologists: the *Guidelines for Autism Spectrum Disorder - Clinical Evaluation* (2012) (jointly developed with the Quebec College of Physicians), the *Guidelines for the Assessment of Mental Retardation* (2007), the *Guidelines for the Assessment of a Child in Connection with a Request for Derogation to the Age of School Admission* (2006), and the *Guidelines for Expert Assessment Concerning Child Custody and Access Rights* (2006) (jointly developed with the Order of Social Workers and Family Therapists of Quebec), using the AGREE II assessment instrument (see below).

In addition to evaluating the above guidelines, we aimed to examine the relationship between appraised guideline quality, indicated by AGREE II scores, and the guidelines' development costs. We also examined the date of publication of the listed references in each of the five OPQ guidelines to determine the recency of the studies on which the GPG's recommendations may have been based.

Method

AGREE II Ratings

The *Appraisal of Guidelines for Research and Evaluation* (AGREE) is the most widely used internationally and is considered to be the gold standard for quality assessment of practice guidelines in the medical and social service professions (Alonso-Coello et al., 2010; Brouwers et al., 2010). This instrument was developed by scientists and clinicians from diverse disciplines and geographical regions who formed an official independent body, named the AGREE Research Trust. The first version of the AGREE instrument was published in 2003 (AGREE, 2003). Thereafter, the AGREE items were refined, and a new item was added to the newer AGREE II in 2009 (see www.agreetrust.org). The AGREE II appraises the process of practice guideline development. It contains 23 items (see www.agreetrust.org), graded on a 7-point Likert scale from 1 (Strongly disagree) to 7 (Strongly agree), which are grouped into six Domains: 1) *Scope and purpose* of the guideline, includes items used to rate the overall aim of the guideline and specific health questions and references, relevant to the target populations; 2) *Stakeholder involvement* focuses on the extent to which the guideline was developed by appropriate stakeholders and represents the views of the intended users; 3) *Rigour of development* addresses the guideline's development process and methodology used; 4) *Clarity of presentation* assesses the language, structure, and format of the guideline, different management options, and the clarity of recommendations; 5) *Applicability* rates the ease of use of the guideline; and; 6) *Editorial independence of the authors* focuses on the degree to which the formulation of the recommendations are not influenced by competing interests.

In this study, three trained raters independently applied the AGREE II instrument to the five OPQ guidelines listed above. In preparation for this study, the raters received intensive training in using the detailed criteria listed in the manual (AGREE, 2017). They continued to practice using other sample guidelines until all pairs of evaluators reliably achieved an interclass

correlation coefficient (ICC) of at least .74 or higher using the scale. A minimum score of .74 is considered a "good" cut-off (see Cicchetti, 1994; Trepanier et al., 2017).

Once trained, the evaluators proceeded to assess each guideline independently, then digitally entered their score directly into their AgreeTrust web guideline-rater account (see www.agreetrust.org/AGREE II). Once entered, the scores were automatically calculated by the AgreeTrust platform (see *Data analysis* below). The evaluators then conducted a *consensus* session comparing their ratings, and arrived at a single score for each item, referred hereafter as the *consensus score*.

A supplementary inquiry was then performed to complement the AGREE assessments of the OPQ guidelines. Through Quebec's Access to Information Act, we requested the following information from the OPQ for all four guidelines in this study as well as for the guidelines on dyslexia previously assessed by Trepanier and colleagues (2017): 1) Development procedures that were provided to persons on the guideline development committees; 2) Supplemental materials that may not have been included in the guidelines and that describe the method(s) used to develop the guidelines; 3) Information about whether or not any systematic literature search was completed; if such a search had been done, we requested information on the method(s) used, and any available supplementary material describing the methods used, including the results of the systematic searches that would have supported the recommendations; and 4) The results of any impact studies or assessments completed after the guidelines became available to end-users, in order to determine the effects of the guideline. Although this additional step is not required for the AGREE II, we chose to consider any additional development methods used to enhance the quality of the CPG, not detailed in the guidelines, for a deeper analysis of the development process. This supplemental information was to be considered by raters when conducting the AGREE ratings.

For each guideline, we examined the item scores, and the Domain scores generated by the AgreeTrust website. The Domain scores are expressed as a percentage, and are calculated by summing up all of the item scores within each of the six Domains, then scaling the total as a percentage of the maximum possible for that Domain (see AGREE, 2017), as follows: ([Obtained scores – Minimum possible score] / [Maximum possible score – Minimum possible score] x 100).

The Domain scores are used to determine whether a guideline should be recommended, or for comparing guidelines. The AGREE Consortium has not set a minimum Domain score or patterns of scores across Domains to differentiate between high quality and poor quality guidelines (AGREE, 2017); however a score of 60% is often used as such a cut-off (Barriocanal, Lopez, Monreal, & Montane, 2016; Chang et al., 2016; Middleton, Kalogeropoulos, Middleton, & Drapeau, 2018). In addition to reporting the Domain scores computed by the AGREE platform, and to facilitate comparisons with other studies, the *mean* scores of each independently rated item and the *consensus* score of each item were also examined, replicating the method by Trepanier and colleagues (2017).

Development Budgets and Costs vs. AGREE II Score

Through Quebec's Access to Information Act, we contacted the OPQ to obtain information on the costs and budgets for each of the guidelines. The aim was to examine whether or not there is a relationship between the cost of producing a guideline and the quality of the guideline.

Reference Publication Dates

The year of publication of all the listed references from each of the five guidelines were entered into a spreadsheet to perform a comparative analysis of the references' year, first to the date of the guidelines' publication, and second, to the present year (2018). We examined the *mean* age of the publications listed in the reference section of each guideline and provided descriptive statistics of the following: 1) the *mean* age of the listed references relative to the year

of the guideline's publication, and 2) the *mean* age of the listed references from the current year (2018) – since all five guidelines are still recommended for use by the OPQ at the time of this writing, the latter might illuminate the user regarding the recency of the science on which the recommendations were based.

Results

The OPQ did not provide us with any additional information in response to our request. The OPQ informed us that they were unable to find any information on the development procedures that may have been provided to persons on the guideline development committees, including search methods and criteria used, as well as the results of the systematic searches that would have supported the recommendations. Further, the OPQ informed us that they did not perform any impact studies or assessments after the guidelines became available to the public. Therefore, none of the guidelines were evaluated with consideration to any additional supporting materials.

Our request, through Quebec's Access to Information Act, for information about the costs and budget assigned to the guidelines, only led to information about the Dyslexia guideline. The OPQ reported that for this most recent guideline, the total cost for its development was $67,321. No information was provided as to what these costs included. For the other four guidelines, the OPQ informed us that costs and spending details could not be found in their records. Therefore, we were not able to compare the development costs of each guideline to their respective AGREE quality scores.

In what follows, for the sake of clarity, the AGREE assessment of each guideline is presented separately, along with the analysis pertaining to the age of the references reported in each guideline.

Guidelines for the Evaluation of Dyslexia in Children **(2014)**

AGREE II assessment. As reported by Trepanier and colleagues (2017), the inter-rater reliability on the individual AGREE items was excellent with a *mean* intra-class coefficient of (2, 1) of .92 (*SD* = .025) (see Table 1). Only two of the six AGREE domains scored above our minimum cut-off of 60%; one domain received a marginal scaled domain score of 59% (see Table 2 & 3 for detailed results of the domain and *consensus* scores). The following is a brief description of the scores received for this guideline (for a full interpretation of the results, see Trepanier et al., 2017).

For Domain 1 (scope and purpose), all raters agreed that the scope and purpose of the guideline were well explained, and the target populations were clearly identified, as reflected in the Domain 1 scores (100%) and *consensus* score (7 on 7) (Trepanier, et al., 2017). The Domain 2 (stakeholder involvement) quality score was 59%, slightly below the 60% cut-off, and the group *consensus* score was 4.7 on 7. In Domain 3 (rigour of development), major weaknesses were identified that resulted in a scaled quality score of 51% (see Tables 2 & 3). Domain 4 (clarity of presentation) relates to the language, structure, and format of the guideline and it received one of the higher Domain scores (65%) with a *consensus* score of 5 on 7. Domain 5 (applicability) examines the guideline's treatment of possible barriers and facilitators to implementation, strategies to improve uptake, and resource implications of applying the guideline. This Domain scored the lowest with a scaled score of 28% and a *consensus* score of 2.3 on 7. Finally, Domain 6 (editorial independence) is concerned with the formulation of recommendations not being unduly biased with competing interests and it rated relatively low (56% and 4.5 on 7) (Trepanier et al., 2017).

Reference and publication dates. *The guidelines for the evaluation of Dyslexia in children* had the most listed (*n*=150) references of the five guidelines. It also had the greatest *mean* difference (20.1 years) between the CPG's year of publication and the year of the

references (see Table 4). The *mean* difference between the current year (2018) and the references years was 24.1.

***Guidelines for Autism Spectrum Disorder - Clinical Evaluation* (2012)**

The inter-rater reliability on the individual AGREE items was excellent with a *mean* intra-class coefficient of .90 (*SD* = .04) (Table 1).

AGREE II assessment. For Domain 1 (scope and purpose), raters agreed that the scope and purpose of the guideline were well explained, as reflected in the scores; the Domain score was 100%, and the *consensus* score was 7 on 7.

Domain 2 (stakeholder involvement) scored 43% (*consensus*: 3.3 on 7) and did not meet the minimum cut-off of 60%. It was determined that although many essential stakeholders were among the members of the development group (e.g. psychologists, pediatric psychiatrists, and pediatricians), other key stakeholders typically involved in the diagnostic process for ASD in children (e.g., audiologists, speech pathologists, occupational therapists, special educators) were noticeably absent. Further, the views of the patients and their primary caregivers were also necessary to earn full points in this AGREE Domain.

Domain 3 (rigour of development) received an AGREE quality score of 10% and the *consensus* discussion led to a 1.5 on 7 (see Table 2). Points were deducted from AGREE item seven, "systematic methods were used to search for evidence," which received a *mean* score of 1 on 7 (see Table 3). There was an overall lack of detailed information on the strategies used to search for the evidence that supported each recommendation in the guideline. The guideline highlighted the critical importance of experience and expertise in ASD assessments, but did not indicate what constitutes expertise (i.e., minimum number of supervised evaluations or specific competencies). Item 12 received a *mean* score of 1 on 7 because there were almost no explicit links (text citations) between the recommendations and the supporting evidence (see Table 3).

The guideline's external review process was vaguely described in the acknowledgments (Item 13). For item 14, no statement was provided that indicated a timeline for when the guideline should be updated, nor a methodology for updating procedures. The raters commented that although the guideline mentioned that "more recent evidence or novel research would render the guideline no longer applicable and inadmissible" (OPQ, 2012, p. 38), this places the responsibility of knowing when newer evidence has been published on the CPG end-users, who in-turn may have a false sense of assurance that the guideline producer would update it accordingly, thus affecting trustworthiness. It was noted that the *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.; *DSM–5*; American Psychiatric Association [APA], 2013), hds been out for several years, but the guideline was still available and promoted to psychologists at the time of this study, without an addendum.

Domain 4 (clarity of presentation) received a moderate scaled domain score of 67% and a *consensus* score of 4 on 7. A higher rating for this domain would have been obtained if the treatment and management options were also provided. Potential comorbid disorders were also clearly listed and presented as necessary considerations during the evaluation of ASD. However, it was found that although it presented different options for the diagnosis of ASD but not its treatment, item 16 received a *mean* score of 4 on 7, which was downgraded in the *consensus* discussion to a 2 on 7, owing to the strict adherence to the item's criteria which rates the management and the treatment of a condition – (*Note*: the raters recognized that this was not the objective of this CPG).

Domain 5 (applicability) received a low quality score of 18% (*consensus* = 2/7). The score for this Domain would have been higher if there were suggestions for coordinating an interdisciplinary team for diagnosis and ongoing treatment. For example, a grid or flow diagram that sequenced the steps in the evaluation process, informed by the professionals consulted, to

guide the psychologist or physician on how to execute an efficient and complete diagnostic evaluation and treatment plan. Further, the guideline did not adequately address the high potential public and private cost, or potential economic impact on the patient, of conducting a multidisciplinary evaluation. Also missing were the monitoring objectives or auditing criteria that would trigger a guideline update.

The Domain 6 (editorial independence) criteria were not met and thus received a quality score of 0%. There was no mention of whether or not the views of the funding body (in this case the OPQ) had or had not influenced the content of the guideline. There was also no mention of the presence or absence of competing interests.

Reference and publication dates. The *Guidelines for Autism Spectrum Disorder* had the fewest number of listed references (*n*=19) of the five guidelines. The *mean* difference between the CPG's year of publication and the year of the references was 6.9 years (see Table 4), and 12.9 years difference relative to 2018.

***Guidelines for the Assessment of Mental Retardation* (2007)**

Inter-rater reliability on the individual AGREE items was excellent with a *mean* intra-class coefficient (2, 1) of .89 (*SD* = .01).

AGREE II assessment. For Domain 1 (scope and purpose), all raters agreed that the scope and purpose of the guideline were well explained, as reflected in all scores: the *consensus* score (6.7 on 7), and the AGREE Domain score (91%). Full points were not given because the stated rational for the guide was deemed too broad.

The Domain 2 (stakeholder involvement) had a low scaled domain score of 44% (group *consensus* = 3.7 on 7). The guideline development group was comprised of professors, researchers and psychologists. Points were deducted for not having included a broader number of professions on the development committee such as educators, social workers, counsellors,

methodology experts (such as systematic review experts), in addition to not seeking the views and preferences of the patient population or their caregivers.

Domain 3 (rigour of development) received a low quality Domain score of 19% (*consensus* = 2.1 on 7). The guideline did not explicitly state if a review was conducted or how, nor did it describe the selection criteria for the research evidence. Some though not all of the recommendations were cited, thus it was not clear if certain recommendations were drawn from empirical studies, or consensus or majority opinion. Clarifying this would have earned more points. According to the raters, there was also insufficient detail about if or how external reviews were conducted, and the procedure for updating the guideline was not provided.

Domain 4 (clarity of presentation) had a quality score of 81% (*consensus* score, 6 on 7). The recommendations in this guideline were specific and unambiguous. Several options were presented with their corresponding level of appropriateness for specific patient contexts, in the comorbidity section. The rest of the criteria were also met. Graphs and flow charts are part of the AGREE criteria and their absence prevented the raters from awarding full points in this section.

Domain 5 (applicability) had the second lowest scores among all the Domains with a quality score of 13% (*consensus*, 1.5 on 7). Missing were the facilitators and barriers to the applications of the recommendations, and the potential cost or resource implications of applying the recommendations. Monitoring of auditing criteria for future updates was also absent.

Domain 6 (editorial independence) scored the lowest (3%) among all six Domains. There was no specific mention that the views of the funding bodies did not influence the content of the guideline as required by AGREE II. There was also no statement of competing interests of the development group members.

Reference and publication dates. The *Guidelines for the Assessment of Mental Retardation* had 75 listed references. The *mean* difference between the CPG's year of publication and the year of the references was 12.6 years, and a 23.6 years difference relative to 2018.

Guidelines for the Assessment of a Child in Connection with a Request for Derogation to the Age of School Admission (2006)

Inter-rater reliability on the individual AGREE items was excellent with a *mean* intra-class coefficient (2,1) of .82 (*SD* = .12).

AGREE II assessment. For Domain 1 (scope and purpose), the scope and purpose of the guideline were well explained (scaled score 81%; *consensus* score was 6.7 on 7). This guideline's stated objective was to serve as a guide rather than a formal standard or best practice. A strong and clear statement of purpose in the first lines of the guideline would have led to a higher score.

Domain 2 (Stakeholder Involvement) quality score was 39% (*consensus*, 3 on 7). The guideline development committee included a document specialist, one legal advisor, several psychologists, however other relevant professionals group were missing, such as methodological specialists, educators, and psychoeducators, and members of the school board, parent or patient/client representatives. The score was low due to the fact that the development committee members were not listed with their proper professional roles and titles, or their respective organizations. Further, neither the targeted users, nor the intended users of the guideline were listed or defined in the body of the text, as required by the AGREE criteria.

Domain 3 (rigour of development) received a low AGREE quality score of 17% and a *consensus* score of 1.9 on 7. Although it appeared that much attention was paid to the process of gathering the research evidence, the following AGREE criteria were missing: 1) The systematic

method used to search the literature was not indicated; 2) The strengths and limitations of the body of evidence; 3) Explicit links between the recommendations and supporting evidence; 4) The mention of having subjected the guideline to a full external review; and 6) There was no procedure stated for updating the guideline in the future (see Tables 2 & 3).

For Domain 4 (clarity of presentation) the item scores were moderate with a quality score of 67% (*consensus*, of 4.7 on 7). Overall, the recommendations were specific, though the guideline required the users to use their own clinical judgment to self-assess for a given test competency. Key recommendations were easily identified. Though most of the criteria were met for this Domain, the guideline should have included examples of special cases or 'troubleshooting' examples for ambiguous cases, to earn full points.

The Domain 5 (applicability) applicability Domain scored the second-lowest among all the Domains for this guideline, with a quality score of 5% (*consensus* 1.3 on 7). This guideline did not provide possible barriers and facilitators to implementation. The potential resources implication of applying the recommendations were not discussed, nor were the monitoring and auditing criteria for when the guideline should be updated, presented.

Domain 6 (editorial independence) was rated at 0%. As required by the AGREE II instrument, the guideline did not explicitly state that the views of the funding bodies had not influenced the content of the guideline. Nor was there an explicit statement about competing interests influence been addressed.

Reference and publication dates. The *Guidelines for the Assessment of a Child in Connection with a Request for Derogation to the Age of School Admission* had 31 listed references. The *mean* difference between the CPG's year of publication and the year of the references was 12.2 years, with a 24.2 years difference relative to 2018.

***Guidelines for Expert Assessment Concerning Child Custody and Access Rights* (2006)**

Inter-rater reliability on the individual AGREE items was excellent with a *mean* intra-class coefficient (2, 1) of .90 (*SD* = .03).

AGREE II assessment. Domain 1 (scope and purpose) received a score of 65% (consensus, 4.7 on 7). The raters reported that it was possible to determine the topic and the objective of the guideline based on the information in the first few pages. However, it was missing a clear and official statement of objectives; therefore item 1 earned a *mean* score and *consensus* score of 4 on 7.

The domain 2 (stakeholder involvement) quality score was 22% (*consensus*, 2.2 on 7). No authors were listed; therefore it is unknown if relevant professionals were part of the development team, including legal counsel and parent groups. It is hence not possible to evaluate who the stakeholders were, or if any were involved. There was no evidence in this guideline that the views and preferences of the target population were sought.

Domain 3 (rigour of development) received an AGREE scaled score of 3% (*consensus* score of 1 on 7). No systematic search was mentioned, and no systematic method was used in the development process. The criteria for selecting the evidence were not described; neither were the strengths and limitations of the body of evidence. There was only one reference pertaining to the main source of this information inspired by another text developed in a separate jurisdiction, from which this present manual was adapted for Quebec (OPQ, 2006[b], p.3). There were no explicit links between the recommendations and the supporting evidence, and no supporting evidence on 'best practices' for the assessment of custody cases. Further, there was no indication that the guideline had been externally reviewed by experts prior to publication, and no procedure was provided for updating the guideline, as required by AGREE.

The Domain 4 (clarity of presentation) score had a low scaled score of 41% (*consensus* 2.7 on 7). For item 15, although recommendations were considered to be specific and unambiguous,

it received an item score 4 on 7 because it did not cite any supporting material or sources. The guideline satisfied most of the criteria for item 17, as key recommendations were easily identifiable, although some formatting could be improved with the incorporation of,summary boxes, bold title fonts, and flow charts.

Domain 5 (applicability) received a quality score of 5%. The guideline did not describe the facilitators and barriers for recommendations, or provide advice on how to put the recommendations into practice (item 18) (see Table 2). No supporting materials were present for users. The guideline did not address the potential resource and cost implications for the process of determining child access and custody rights. This guideline did not present any instructions for future updates.

Domain 6 (editorial independence) received a low scaled score (0%). For editorial independence, it was not stipulated if the views of the funding body had influenced the content of the guideline, nor was there any explicit statement regarding competing interest related to the development group members.

Reference and publication dates. The *Guidelines for Expert Assessment Concerning Child Custody and Access Rights* did not list any references.

Discussion

All of the guidelines presented the scope and purpose in sufficient detail, identifying the specific clinical problem, patient population, and expected health benefits, as demonstrated by the generally strong quality scores in Domain 1 (scope and purpose). They also performed relatively well in Domain 4 (clarity of presentation). In contrast, Domains, 2, 3, 5, and 6 did not receive strong quality scores.

For Domain 2, quality scores ranged widely across guidelines, from 22% to 59% (*consensus*: 2.7 to 4.7). For most of the guidelines, there was a homogeneous composition of

professionals in the development groups, to the exclusion of other important stakeholders. This is not to suggest that there was a conscious intent on the part of developers to exclude parents and patients, as well as other professionals from complementary disciplines, from contributing to the guidelines. It may, however, suggest an assumption of unique authority on the topic of the guideline, at the expense of otherwise enriching perspectives and contributions of other stakeholders (Armstrong, Mullins, Gronseth, & Gagliardi, 2018; Serrano-Aguilar et al., 2016). The expansion of stakeholders in guideline development groups is recommended in guideline development manuals such as the NICE guidelines (García et al., 2014) and SIGN50 (2001 & 2015), and is therefore strongly recommended for future guideline updates.

Possibly the greatest added value of the AGREE II instrument is its emphasis on the rigour of a guideline's development. Domain 3 is the largest of the six Domains and comprises 35% (8 of 23) of all the AGREE items highlighting the importance and complexity of methodological rigour in guideline development. Our study showed that there is a need for significant methodological improvements in all five of the OPQ guidelines. Among the trends identified were brief mentions, at best, of a literature search or method; information about the methods used was simply inadequate. More information is required pertaining to the search methods employed (e.g., key words, databases utilized, the timeframe of the search, and the inclusion and exclusion criteria), as well as the criteria for selecting the evidence on which to base the recommendations. Further, if no relevant empirical studies were available due to the nature of the guideline or the novelty of the search topic, a statement to that effect would have been necessary. All five guidelines failed to meet the criteria for item nine (9) "the strength and limitation of the body of evidence are clearly described," and item 14 "a procedure for updating the guideline is provided." With the exception of the evaluation of dyslexia guideline, the recommendations were poorly or inconsistently cited with any form of supporting evidence. This does not mean that the

recommendations found in the OPQ guidelines are not potentially useful to clinicians, although such a lack of rigour may call into question some of the recommendations contained in the guideline. Indeed, the results relate only to the lack of rigour and transparency with regards to the processes used to research and formulate the recommendations; any improvements made in both the methods and the transparency with which the methods are employed inevitably adds to the quality and trustworthiness of a guideline. Further, rigorous citing and referencing allows for more precise identification of outdated research, law, or public policies that may have been used to support a recommendation, thus rendering it invalid because of newer evidence or new standards of practice. Check-up tools have been developed to support this process (Vernooij, Alonso-Coello, Brouwers, & Martinez Garcia, 2017). Guidelines should be updated every two to five years (Shekelle, 2014). If not, doubts as to the guideline's validity may rise as it ages.

All five guidelines received low scores in the applicability Domain (5), with the exception of the *Guidelines for Autism Spectrum Disorder* (2012), which briefly addressed the barriers and facilitators to its application in a clinical setting. Developers must openly consider guideline users who work in diverse settings, perhaps with limited resources, and may serve a clientele who has limited financial means. These represent barriers to timely access to services that ought to be addressed in a guideline. Another important item in Domain 5 is the post-implementation monitoring plan. Once a guideline is produced, developers and publishers need to know whether or not it is having a positive impact on the services it was intended to improve. Unfortunately, this was not addressed in any of the OPQ guidelines, which concurs with the OPQ's feedback to our research team that no post-implementation studies were initiated for any of the guidelines. Therefore, we do not know what the impact of the guidelines are, in terms of health benefits, improved services, and costs savings. This is important to know given the cost of developing guidelines. It is also important to note that the funds used to develop guidelines also come from,

at least in part, from the annual fees collected by the OPQ from its members, who in turn deserve a quality product. Closer attention to 'applicability' is therefore needed in future guideline updates.

Finally, the guidelines received poor ratings in Domain 6 (editorial independence) (range, 0% to 56%), as they did not explicitly state whether the funding body had an influence on the recommendations, as required by AGREE II. Given that the OPQ is the funding body in this case, it is important that these above points be clarified in the guidelines. The guidelines were also required to provide a statement regarding the presence or absence of competing interests among the developers, with the exception of the guidelines on Dyslexia.

Given that some of the guidelines based their recommendations on research in fields that continuously produce newer studies, we analyzed the *mean* age of the references, and found that the average reference in four of the five guidelines ranged between 12.9 and 24.2 years relative to today. One guideline did not have a reference section. This is directly related to the requirement for well-described updating procedures for a guideline (Domain 3), in case there are changes to the diagnostic criteria, as with the DSM-V, or in the case where newer studies may render older ones invalid. For example, research in the area of Autism Spectrum Disorder (ASD) has been highly productive over the past years, and since there is a new DSM-V, the guidelines on ASD that are currently available to psychologists should be updated with new recommendations and supporting references. The same is very likely true for all the guidelines investigated in this study.

Finally, it is surprising that the OPQ did not have a record of costs and budgets available for analysis, except for the guidelines on Dyslexia which reportedly cost upward of $67,000. Developing guidelines requires an important investment in time and money. It would hence be expected that such investments of an institution's financial resources, which are drawn in large

part from the membership dues paid by psychologists, be the result of meticulous thought and planning. It would be equally reasonable to expect guideline developers to know how much their guidelines cost, and to conduct a study of the impact of their guidelines on practice in light of the amounts invested.

Conclusion

The present study went beyond the protocols provided but the AGREE Trust for applying the AGREE II instrument. The rater training meetings were held to ensure that the item criteria were interpreted similarly among the raters. It is not uncommon for professionals in the same field to assign slightly divergent item scores when evaluating the same guideline, due in part to interpretation differences of the criteria among the raters. Although the AGREE Trust provides explanations for each of the criteria in the AGREE II scoring manual, the rater consensus practice revealed some slightly dissimilar interpretations of some items found in Domains 1, 2, and 5. For this reason, it was decided that the harmonization of item interpretations would increase the reliability and validity of our findings. This extra step was viewed as a strength of the study.

Our findings inform both the developers and the end-users as to the strengths and weaknesses of the OPQ guidelines, and provide a framework from which to guide future updates and development initiatives. We credit the OPQ for being one of the few professional bodies in Canada to develop practice guidelines for psychologists. Their effort to support the psychologists in their jurisdiction is laudable. However, significant improvements in the methods used to develop guidelines have to be made, as the methods they use are deficient in many ways. Until such improvements are made, the guidelines produced will fail to meet the scientific rigour expected by psychologists and reduce their impact. We also recommend a full review of the OPQ guidelines currently available to psychologists.

References

AGREE Collaboration. (2003). Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: The AGREE project. *Quality & Safety in Health Care, 12*, 18-23. doi: 10.1136/qhc.12.1.18

AGREE Next Steps Consortium. (2017). *The AGREE II instrument.* [Electronic version]. Retrieved from https://www.agreetrust.org/wp-content/uploads/2017/12/AGREE-II-Users-Manual-and-23-item-Instrument-2009-Update-2017.pdf

Al-Ansary, L. A., Tricco, A. C., Adi, Y., Bawazeer, G., Perrier, L., Al-Ghonaim, M., . . . Straus, S. E. (2013). A systematic review of recent clinical practice guidelines on the diagnosis, assessment and management of hypertension. *PLoS ONE, 8*, e53744. doi:10.1371/journal.pone.0053744

Alonso-Coello, P., Irfan, A., Solà, I., Gich, I., Delgado-Noguera, M., Rigau, D., . . . Schunemann, H. (2010). The quality of clinical practice guidelines over the last two decades: A systematic review of guideline appraisal studies. *Quality and Safety in Health Care, 19*, e58. doi:10.1136/qshc.2010.042077

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5*. Washington, DC: Author.

Ansari, S., & Rashidian, A. (2012). Guidelines for guidelines: Are they up to the task? A comparative assessment of clinical practice guideline development handbooks. *PLoS ONE, 7*, e49864. doi:10.1371/journal.pone.0049864

Armstrong, M. J., Mullins, C. D., Gronseth, G. S., & Gagliardi, A. R. (2018). Impact of patient involvement on clinical practice guideline development: A parallel group study. *Implementation Science, 13*, 55-59. doi:10.1186/s13012-018-0745-6

Balogh, Z., Caldwell, E., Heetveld, M., D'Amours, S., Schlaphoff, G., Harris, I., & Sugrue, M. (2005). Institutional practice guidelines on management of pelvic fracture-related hemodynamic instability: Do they make a difference? *Journal of Trauma and Acute Care Surgery, 58*, 778-782. doi: 10.1097/01.ta.0000158251.40760.b2

Barriocanal, A. M., Lopez, A., Monreal, M., & Montané, E. (2016). Quality assessment of peripheral artery disease clinical guidelines. *Journal of Vascular Surgery, 63*, 1091-1098. doi:10.1016/j.jvs.2015.12.040

Beauchamp, S., Duplantie, J. P., & Mercier, C. (2011). *Development of practice guidelines in social care: A position of the INESSS social care practice guideline committee* [Policy statement]. Retrieved from Regroupement provinciale des comités usagers website: http://www.rpcu.qc.ca/pdf/english/INESSS_Practice_Guidelines_Social_Care.pdf

Bennett, K., Courtney, D., Duda, S., Henderson, J., & Szatmari, P. (2018). An appraisal of the trustworthiness of practice guidelines for depression and anxiety in children and youth. *Depression and Anxiety, 35*, 330-340. doi: 10.1002/da.22752

Bernard, H., Burlingame, G., Flores, P., Greene, L., Joyce, A., Kobos, J. C., . . . McEneaney, A. M. S. (2008). Clinical practice guidelines for group psychotherapy. *International Journal of Group Psychotherapy, 58*, 455-542. doi:1 0.1521/ijgp.2008.58.4.455

Brouwers, M. C., Kho, M. E., Browman, G. P., Burgers, J. S., Cluzeau, F., Feder, G., . . . The AGREE Next Step Consoritum (2010). AGREE II: Advancing guideline development, reporting and evaluation in health care. *Canadian Medical Association Journal, 182*, E839-E842. doi:10.1503/cmaj.090449

Burgers, J. S., Cluzeau, F. A., Hanna, S. E., Hunt, C., & Grol, R. (2003). Characteristics of high-quality guidelines. *International Journal of Technology Assessment in Health Care, 19*, 148-157. doi: 10.1017/S026646230300014X

Chang, S. G., Kim, D. I., Shin, E. S., Jang, J. E., Yeon, J. Y., & Lee, Y. S. (2016). Methodological quality appraisal of 27 Korean guidelines using a scoring guide based on the AGREE II instrument and a web-based evaluation. *Journal of Korean Medical Science, 31*, 682-687. doi:10.3346/jkms.2016.31.5.682

García, L. M., McFarlane, E., Barnes, S., Sanabria, A. J., Alonso-Coello, P., & Alderson, P. (2014). Updated recommendations: An assessment of NICE clinical guidelines. *Implementation Science, 9*, 72-91. doi: 10.1186/1748-5908-9-72

Gordon, J., & Cooper, B. (2010). Talking knowledge — practising knowledge: A critical best practice approach to how social workers understand and use knowledge in practice. *Practice, 22*, 245-257. doi: 10.1080/09503153.2010.495768

Graham, I. D., Beardall, S., Carter, A., Tetroe, J., & Davies, B. (2003). The state of the science and art of practice guidelines development, dissemination and evaluation in Canada. *Journal of Evaluation in Clinical Practice, 9*, 195-202.

Graham, I. D., Mancher, M., Wolman, D. M., Greenfield, S., & Steinberg, E. (2011). *Clinical practice guidelines we can trust*. Washington, DC: Institue of Medicine National Academies Press.

Graham, I. D., Evans, W. K., Logan, D., O'Connor, A., Palda, V., McAuley, L., . . . Harrison, M. B. (2000). Canadian oncologists and clinical practice guidelines: A national survey of attitudes and reported use. *Oncology, 59*, 283-290. doi: 10.1046/j.1365-2753.2003.00385.x

Hasenfeld, R., & Shekelle, P. G. (2003). Is the methodological quality of guidelines declining in the US? Comparison of the quality of US Agency for Health Care Policy and Research (AHCPR) guidelines with those published subsequently. *Quality and Safety in Health Care, 12*, 428-434. doi: 10.1136/qhc.12.6.428

Hershenberg, R., Drabick, D. A., & Vivian, D. (2012). An opportunity to bridge the gap between clinical research and clinical practice: Implications for clinical training. *Psychotherapy, 49*, 123-131.

Hollon, S. D., Areán, P. A., Craske, M. G., Crawford, K. A., Kivlahan, D. R., Magnavita, J. J., . . . Bufka, L. F. (2014). Development of clinical practice guidelines. *Annual Review of Clinical Psychology*, *10*, 213-241. doi: 10.1146/annurev-clinpsy-050212-185529

Hudson, P., Quinn, K., O'Hanlon, B., & Aranda, S. (2008). Family meetings in palliative care: Multidisciplinary clinical practice guidelines. *BMC Palliative Care, 7*, 1-6. doi: 10.1186/1472-684X-7-12

Kalies, H., Schöttmer, R., Simon, S. T., Voltz, R., Crispin, A., & Bausewein, C. (2017). Critical attitudes and beliefs towards guidelines amongst palliative care professionals – Results from a national survey. *BMC Palliative Care, 16*, 20-24. doi: 10.1186/s12904-017-0187-y

Kendall, T., Taylor, E., Perez, A., & Taylor, C. (2008). Guidelines: Diagnosis and management of attention-deficit/hyperactivity disorder in children, young people, and adults: Summary of NICE guidance. *British Medical Journal, 337*, 751-753.

Lenzer, J. (2013). Why we can't trust clinical guidelines. *British Medical Journal, 346*, 1-5. doi:0.1136/bmj.f3830

Mental Health Commission of Canada. (2012). *Changing directions, changing lives: The mental health strategy for Canada* [Strategy report]. Calgary, Canada: Author. Retrieved from https://www.mentalhealthcommission.ca/sites/default/files/MHStrategy_Strategy_ENG.pdf

Middleton, J. C., Kalogeropoulos, C., Middleton, J. A., & Drapeau, M. (2018). Assessing the methodological quality of Canadian Psychiatric Association's anxiety and depression

clinical guidelines. *Journal of Evaluation in Clinical Practice, 2,* 1-9. doi:10.1111/jep.13026

Miville, M. L., Duan, C., Nutt, R. L., Waehler, C. A., Suzuki, L., Pistole, M. C., . . . Corpus, M. (2009). Integrating practice guidelines into professional training: Implications for diversity competence. *The Counseling Psychologist, 37*, 519-563. doi: 10.1177/0011000008323651

Ordre des psychologues du Québec (OPQ). (2012). *Guidelines for autism spectrum disorder: Clinical evaluation*. Montreal, Canada: Author. Retrieved from https://www.ordrepsy.qc.ca/documents/26707/63191/spectre-de-lautisme/868d8894-44f1-471d-941d-94a5f571395c

Ordre des psychologues du Québec (OPQ). (2006). *Guidelines for the assessment of a child in connection with a request for derogation to the age of school admission*. Montreal, Canada: Author. Retrieved from https://www.ordrepsy.qc.ca/documents/26707/0/Les+lignes+directrices+pour+l%27%C3%A9valuation+d%27un+enfant+en+vue+d%27une+demande+de+d%C3%A9rogation+%C3%A0+l%27%C3%A2ge+d%27admission+%C3%A0+l%27%C3%A9cole/b2c2e606-7fad-4b18-8ec8-883fb69e572e

Ordre des psychologues du Québec (OPQ). (2006). *Guidelines for expert assessment concerning child custody and access rights.* Montreal, Canada: Author. Retrieved from https://www.ordrepsy.qc.ca/documents/26707/6319/Lignes+directrices+pour+l%27expertise+en+mati%C3%A8re+de+garde+d%27enfants+et+des+droits+d%27acc%C3%A8s/d0c94c08-c04d-49ae-8fea-b37bd0ed83b8

Ordre des psychologues du Québec (OPQ). (2007). *Guidelines for the assessment of mental retardation*. Montreal, Canada: Author. Retrieved from

https://www.ordrepsy.qc.ca/documents/26707/63191/Lignes+directrices+pour+l%27%C3%A9valuation+du+retard+mental/f20a2002-4535-4070-8a03-15889a9d5450

Ordre des psychologues du Québec (OPQ). (2014). *Guidelines for the evaluation of dyslexia in children*. Montreal, Canada: Author. Retrieved, from http://www.ordrepsy.qc.ca/documents/26707/63191/Lignes+directrices+pour+l%27%C3%A9valuation+de+la+dyslexie+chez+les+enfants/139edf81-9163-4060-9c66-06f02eee0ed8

Pilling, S. (2008). History, context, process, and rationale for the development of clinical guidelines. *Psychology and Psychotherapy: Theory, Research and Practice, 81*, 331-350. doi:10.1348/147608308X324923

Ruszczyński, M., Horvath, A., Dziechciarz, P., & Szajewska, H. (2016). Cow's milk allergy guidelines: A quality appraisal with the AGREE II instrument. *Clinical & Experimental Allergy, 46*, 1236-1241. doi:10.1111/cea.12784

Rutten, G. M., Degen, S., Hendriks, E. J., Braspenning, J. C., Harting, J., & Oostendorp, R. A. (2016). Adherence to clinical practice guidelines for low back pain in physical therapy: Do patients benefit? *Physical Therapy, 90*, 1111-1122. doi: 10.2522/ptj.20090173

Serrano-Aguilar, P., Trujillo-Martin, M. M., del Pino-Sedeño, T., Pérez de la Rosa, A., de Pascual y Medina, A. M., Perestelo-Pérez, L., . . . Sarría Santamera, A. (2016). Patient participation in the development of a clinical guideline for inherited retinal dystrophies. *Expert Opinion on Orphan Drugs, 2*, 691-697. doi:10.1080/21678707.2016.1182907

Shaner, R. (2001). Practical psychotherapy: How practice guidelines can rescue psychotherapy in public systems. *Psychiatric Services, 52*, 1021-1025. doi:10.1176/appi.ps.52.8.1021

Shekelle. (2014). Updating practice guidelines. *Journal of the American Medical Association, 311*, 2072-2074. doi:10.1001/jama.2014.4950

Shiroiwa, T., Fukuda, T., Ikeda, S., Takura, T., & Moriwaki, K. (2016). Development of an official guideline for the economic evaluation of drugs/medical devices in Japan. *Value in Health. 20*, 372-378. doi 10.1016/j.jval.2016.08.726

Scottish Intercollegiate Guidelines Network (SIGN). (2001). *SIGN 50: A guideline developer's handbook*. Edinburgh, Scottland: Author

Scottish Intercollegiate Guidelines Network (SIGN). (2015). *SIGN50: A guideline developer's handbook* [Electronic version]. Retrieved from https://sign.ac.uk/assets/sign50_2015.pdf

Stamoulos, C., Reyes, A., Trepanier, L., & Drapeau, M. (2014). The methodological quality of clinical practice guidelines in psychology: A pilot study of the guidelines developed by the Ordre des Psychologues du Québec. *Canadian Psychology/Psychologie canadienne, 55*, 177-186. doi:10.1037/a0037287

The ADAPTE Collaboration (2009). *The ADAPTE process: Resource toolkit for guideline adaptation* (Version 2.0). Retrieved from the Guidelines International Network (G-I-N) website http://wwwg-i-n.net.

Trepanier, L., Stamoulos, C., & Reyes, A. (2017). Assessing the quality of the OPQ's guidelines for the evaluation of dyslexia in children using the Appraisal of Guidelines for Research and Evaluation (AGREE) II: A brief report. *Canadian Psychology/Psychologie canadienne, 58,* 211-217. doi:10.1037/cap0000121

van Dijk, M. K., Oosterbaan, D. B., Verbraak, M. J., & van Balkom, A. J. (2013). The effectiveness of adhering to clinical-practice guidelines for anxiety disorders in secondary mental health care: the results of a cohort study in the Netherlands. *Journal of Evaluation in Clinical Practice, 19*, 791-797. doi: 10.1111/j.1365-2753.2012.01851.x

Vernooij, R. W., Alonso-Coello, P., Brouwers, M., & Martinez Garcia, L. (2017). Reporting items for updated clinical guidelines: Checklist for the reporting of updated guidelines (CheckUp). *PLoS Med, 14*, e1002207. doi:10.1371/journal.pmed.1002207

World Health Organization (WHO). (2010). *WHO handbook for guideline development.* Geneva, Switzerland: Author. Retrieved from http://www.who.int/hiv/topics/mtct/grc_handbook_mar2010_1.pdf

Wolf, J. S., Hubbard, H., Faraday, M. M., & Forrest, J. B. (2011). Clinical practice guidelines to inform evidence-based clinical practice. *World Journal of Urology, 29*, 303-309. doi: 10.1007/s00345-011-0656-5

Table1.

Inter-rater agreement of guideline evaluation using AGREE II

	Mean ICC	*SD*	Range	Min	Max
Guidelines for the Evaluation of Children with Dyslexia (2014).	**0.92**	0.03	0.05	0.90	0.95
Autism Spectrum Disorder – The Clinical Evaluation (2012).	**0.90**	0.04	0.07	0.88	0.95
Guidelines Mental Retardation Assessment (2007).	**0.89**	0.01	0.02	0.88	0.90
Guidelines for the Assessment of a Child in Connection with a Request for Derogation to the Age of School Admission (2006).	**0.82**	0.12	0.21	0.74	0.95
Guidelines for Expert Assessment Concerning Child custody and Access Rights (2006).	**0.90**	0.03	0.05	0.88	0.93

Table 2.

Mean Consensus and Domain Scores from AGREE II Assessment of Five OPQ Guidelines

	Child Custody & Access Rights		Derogation for Age of School Admission		Mental Retardation		Autism		Dyslexia*	
Domain Groups	*Mean Cons.* by Domain	AGREE Domain Score	*Mean Cons.* by Domain	AGREE Domain Score	*Mean Cons.* by Domain	AGREE Domain Score	*Mean Cons.* by Domain	AGREE Domain Score	*Mean Cons.* by Domain	AGREE Domain Score
1. Scope & Purpose	4.7	65%	6.7	81%	6.7	91%	7.0	100%	7.0	100%
2. Stakeholder Involvement	2.2	22%	3.0	39%	3.7	44%	3.3	43%	4.7	59%
3. Rigour of Development	1.0	3%	1.9	17%	2.1	19%	1.5	10%	3.4	51%
4. Clarity of Presentation	2.7	41%	4.7	67%	6.0	81%	4.0	67%	5.0	65%
5. Applicability	1.3	5%	1.3	5%	1.5	13%	2.0	18%	2.3	28%
6. Editorial Independence	1.0	0%	1.0	0%	1.0	3%	1.0	0%	4.5	56%

Note. From Trepanier, Stamoulos, & Reyes, 2017.

Table 3.

AGREE II item mean *scores and consensus scores for the five OPQ practice guidelines*

		Child Access & Custody Rights		Early Derogation for Age of Adm.		Mental Retardation		Autism Guideline		Dyslexia Guideline*	
Domain Groups	AGREE II Items	*Mean* Score	Group Cons.	*Mean* Score	Group Cons.	*Mean* Score	Group Cons.	*Mean* Score	Group Cons.	*Mean* Score	Group Cons.
1. Scope & Purpose	1. The overall objectives are specifically described.	4	4	6	6	7	7	7	7	7	7
	2. The health questions covered are specifically described.	6	5	5	7	7	7	7	7	7	7
	3. The population to whom the guideline is meat to apply is specifically described.	5	5	7	7	6	6	7	7	7	7
2. Stakeholder Involvement	4. The development group includes individuals from all relevant professional groups.	1	2	4	4	4	3	2	2	6	6
	5. The views and preferences of the target population were sought.	1	1	1	1	1	1	1	1	1	1
	6. The target users are clearly defined.	5	5	5	4	6	7	6	7	7	7

3. Rigour of Development	7. Systematic methods were used to search for evidence.	1	1	2	2	2	2	1	2	6	6
	8. The criteria for selecting the evidence are clearly described.	1	1	1	1	1	1	1	1	4	2
	9. The strengths and limitations of the body of evidence are clearly described.	1	1	1	1	1	1	1	1	1	1
	10. The method for formulating the recommendations are clearly described.	1	1	1	1	1	1	1	1	3	3
	11. The health benefits, side effects, and risks have been considered in formulating the recommendations.	2	1	6	6	4	5	3	3	3	4
	12. There is an explicit link between the recommendations and the supporting evidence.	1	1	2	1	4	5	1	1	7	6
	13. The guideline has been externally reviewed by experts prior to its publication.	1	1	3	2	2	1	2	2	6	4

	14. A procedure for updating the guideline is provided.	1	1	1	1	1	1	1	1	1	1
4. Clarity of Presentation	15. The recommendations are specific and unambiguous.	4	4	6	5	6	5	5	4	6	7
	16. The different options for management of the condition or health issue are clearly presented.	1	1	3	3	6	6	4	2	2	1
	17. Key recommendations are easily identifiable.	6	5	6	6	6	7	6	6	7	7
5. Applicability	18. The guideline describes facilitators and barriers to its application.	1	1	1	1	1	1	3	3	1	1
	19. The guideline provides advice and/or tools on how the recommendations can be put into practice.	1	1	2	2	3	3	2	3	6	6
	20. The potential resource implications of applying the recommendations have	2	2	1	1	1	1	2	1	1	1

	been considered. 21. The guideline presents monitoring and/or auditing criteria.	1	1	1	1	2	1	1	1	2	1
6. Editorial Independence	22. The views of the funding body have not influenced the content of the guideline. 23. Competing interests of the development group members have been recorded and addressed.	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	2 6	2 7
Overall Score		2	2	3	3	4	3	4	3	5	5

*Note**. From Trepanier, Stamoulos, & Reyes, 2017.

Table 4.

Mean age of listed references of five OPQ practice guidelines relative to the year of the guideline's publication and to 2018.

	Child Access & Custody Rights	Early Derogation for Age of Adm.	Assessment of mental retardation	Autism spectrum disorder-clinical evaluation	Evaluation of dyslexia in children
Year of CPG Publication	2006	2006	2007	2012	2014
Number of References	0	31	75	19	150*
Mean difference between CPG yr. and the yr. of the references	-	12.2	12.6	6.9	20.1
Mean difference between 2018 and the references	-	24.2	23.6	12.9	24.1
Earliest publication yr. / most recent publication yr.	-	1980 / 2004	1996 / 2007	1996 / 2011	1865 / 2013
Range	-	24	11	15	148

Note. *Includes seminal works that span from 1865 across the 20th century (50% were published before the year 2000).

Linking Manuscripts 1 and 2

In study 1, we evaluated five OPQ guidelines that were published between 2006 and 2014 using the AGREE II assessment tool. These five guidelines represented all of the clinical practice guidelines available to psychologists (guidelines intended for the diagnosis, assessment, and treatment of a disorder); each was downloaded from the OPQ website (www.ordrepsy.qc.ca). While in the process of thoroughly examining these guidelines, the evaluators noted that the OPQ did not present the author information (e.g. name, title/designation, occupation, employer) uniformly. Some guidelines provided all of the necessary information, while others provided almost no information on the developers. This drew our attention to the composition of the guideline development committees. Our research team began to ponder possible relationships between the poor overall quality of the guidelines (see study 1), and the development groups' composition, such that the AGREE domains that performed the poorest across all of the guideline were Rigour of Development (Domain 3) and Applicability (Domain 5). The items in Domain 3 were centered on assessing the quality of the systematic research methodology, and items in Domain 5 on the ease of use, implementation, and on planning impact studies and updates. Thus, researchers would be ideally suited to ensure that the guidelines are sound methodologically, plan for future impact studies, and inform fellow committee members on the latest research for potential updates.

Therefore, our team opted to examine the group composition of the OPQ guideline development committees. Next, we wished to further identify the number of committee members who were also researcher-experts on the topic of the guideline that they developed, by examining their publication productivity as measured by peer-reviewed publication counts, and the *h*-index. Our findings may explain the assessed quality of the guidelines.

Chapter 3

Manuscript 2

(To be submitted as a brief report)

What is an Expert? : Publication Productivity as a Complementary Indicator of Expertise of Guideline Development Committee Members

Trepanier, Lyane[1], Reyes, Andrea[1], Stamoulos, Constantina[1], Beauchamp, Sylvie [1, 2], & Drapeau, Martin[1, 3]

Department of Counselling Psychology, McGill University[1]

Montreal West Island Integrated University Health and Social Services Center[2]

Department of Psychiatry, McGill University[3]

Author note:

Address correspondence to Dr. Martin Drapeau, MPPRG – McGill University, 3700 McTavish,Montreal, Quebec, H3A 1Y2, Canada. E-mail: martin.drapeau@mcgill.ca

The work of Lyane Trepanier was supported by the SSHRC Joseph-Bombardier Canada Graduate Scholarship Doctoral Award. The work of Martin Drapeau was supported by SSHRC Research Grant. The work of Constantina Stamoulos was supported by the SSHRC Doctoral Award and FQRSC Doctoral Research Training Scholarship, and Andrea Reyes was supported by the Fonds de Recherche sur la Société et la Culture du Québec (FRQSC). No funding body had any input into any aspect of this research.

Abstract

The development and implementation of clinical practice guidelines (CPGs) has flourished over the past two decades. Unfortunately, many studies have found that the quality of such guidelines were highly variable (Alonso-Coello et al., 2010; MacQueen et al., 2016). More specifically, research in this area suggests that some of the CPGs used in psychology were developed using poor methods for guideline development (Bennett, Courtney, Duda, Henderson, & Szatmari, 2018; Trepanier, Stamoulos, & Reyes, 2017). While there remains a dearth of research in this area, typically it is guidelines themselves that are examined by researchers and too little attention is paid to the developers, specifically how the guideline development groups are composed and the nature of the expertise of those involved in guideline development. In light of this, this study examined group composition as well as the expertise of guideline development committee members at the Order of Psychologists of Quebec (OPQ), as defined by academic research productivity. We analyzed the peer-reviewed publication productivity of committee members using PsycINFO and MEDLINE, and retrieved their *h*-index from Scopus and from Web of Knowledge. Results show that there is a clear imbalance between clinical and research expertise, with only a small percentage of researchers represented on these committees. Our findings highlight the need for improved group composition for future guideline development.

Keywords: Guideline development, clinical practice, publication productivity, expertize, psychology guidelines, mental health guidelines, guideline development group, *h*-index.

What is an Expert? : Publication Productivity as a Complementary Indicator of Expertise of Guideline Development Committee Members

Guideline development committees, otherwise known as guideline steering groups or working groups, involve a team of experts who select, review, and rate the scientific evidence that leads to and supports one or many recommendations for practice, and are responsible for the final formulation of the recommendations in a guideline (Brouwers et al., 2010; Fretheim, Schunemann, & Oxman, 2006; Oxman, Fretheim, & Schunemann, 2006; SIGN50, 2015). These committees are generally comprised of experts from various disciplines and specializations that are relevant to the topic of the guideline that is to be developed. If a committee member is not an expert on the topic of a guideline per se, the expertise of this person may lie in his or her understanding of best practice in knowledge synthesis or in guideline development (Fretheim et al., 2006; Gordon & Cooper, 2010; Shiffman, 2016).

Best practices for guideline development have been developed by international organizations such as NICE (www.nice.org.uk) and SIGN (www.sign.ac.uk). The AGREE Trust (www.agreetrust.org) developed the Appraisal of Guidelines for Research and Evaluation (AGREE) II (AGREE, 2017; Brouwers, et al., 2010). All three of these organizations have developed documents that provide criteria for good development methodology, content, and ease of use. More specifically, they detail rigorous development methods required for best practices. For example, one of the methodological requirements presented in these 'guidelines for guidelines' is the need to solicit input from a variety of experts; these may include clinicians, policy makers, service managers, service users, as well as researchers. Another recommendation is the establishment of a formal procedure for selecting and analyzing the highest quality studies on which to base guideline recommendations (AGREE, 2017; Brouwers, et al., 2010). More specifically, effective, high-quality guidelines depend on good systematic reviews (Beauchamp,

Drapeau, Dionne, et al., 2015). Systematic reviews are conducted by focusing on a specific research question, and then executing searches that follow explicit scientific methods established by the developers, to identify, select, assess, and summarize the findings in a transparent and well-documented manner (Graham, Mancher, Wolman, Greenfield, & Steinberg, 2011). Graham and colleagues (2011) also suggest that for a systematic review to support the development of a CPG, it should be conducted by a team of topic experts and experts in systematic reviews. Furthermore, the methods used, the processes followed, and a critical summary of the findings of individual studies, should be compiled into a *systematic review report* for good record keeping (Graham, et al. 2011).

High performing and experienced researchers are well suited to the task of selecting, reviewing, and rating scientific evidence, which should be the foundation upon which any guideline should be built (Graham et al., 2011; Kunz et al., 2012; SIGN50, 2015; WHO, 2014). The question remains however to what extent guideline committee members have relevant expertise, more specifically research expertise, for the development of a guideline, and what the nature of this research expertise is. The aim of this study was thus to examine the extent to which researchers are involved in guideline development committees, and their level of research productivity as measured by publication count and the *h*-index. Our focus was on five practice guidelines published by the Order of Psychologists of Quebec (OPQ). These guidelines represent all of the available clinical practice guidelines for treating or evaluating disorders, developed by the OPQ at the time of this writing. These guidelines were selected specifically because they were published by the OPQ, as this organization is the regulatory body (College) for all 8 900 psychologists in Quebec, which represents half of all psychologists in Canada. The OPQ is also the most prolific publisher of practice guidelines for psychologists in Canada. More importantly, because the OPQ is a regulatory body, unlike psychology associations, it has both moral and legal

authority over practitioners. It therefore handles complaints from the public, conducts investigations into the conduct of Quebec psychologists and takes action against psychologists who display unethical behaviors, misconducts, or fail to respect the OPQ Code of Ethics, which states that psychologists "must practice according to generally recognized scientific and professional principles, in keeping with good practice in psychology" (Code of ethics of psychologists, c. C-26, r. 212, a. 5). The OPQ considers its guidelines to reflect these scientific and professional principles and good practices in psychology (e.g., Desjardins, 2010).

Method

Participants

The participants were comprised of committee members who participated in the development of one or more of the five clinical practice guidelines produced by the OPQ: (1) the *Guidelines for the Evaluation of Dyslexia in Children* (2014)*;* (2) the *Guidelines for Autism Spectrum Disorder - Clinical Evaluation* (2012; jointly developed with the Quebec College of Physicians); (3) the *Guidelines for the Assessment of Mental Retardation* (2007); (4) the *Guidelines for the Assessment of a Child in Connection with a Request for Derogation to the Age of School Admission* (2006); and (5) the *Guidelines for Expert Assessment Concerning Child Custody and Access Rights* (2006; jointly developed with the Order of Social Workers and Family Therapists of Quebec).

For the purpose of this study, we focused on individuals who were involved in the development of the guidelines (as a member of the guideline "Working Group", "Development Committee" or "Consultative Committee") and who as such had a determining influence on the methodology used to design the guideline or on the final recommendations. This study did not include "external reviewers" because they were not specifically identified in the guidelines. For example, the *Guidelines for Autism Spectrum Disorder* reported that eight committee members

participated in the development of the guideline, four psychiatrists and four psychologists. However, it was also reported that feedback from over 40 psychologists and psychiatrists was considered upon determining the guideline's final recommendations; the names of these professionals were not provided and their involvement was limited to providing feedback. For these reasons, the latter were not included in this study.

We identified guideline development committee members by referring to the guidelines themselves, and searched for committee roles and biographical information. Any statements within the guidelines that described the criteria for selecting committee members were also noted.

Metrics

We first examined how the committees were balanced in relation to each member's expertise on the topic of the guideline (e.g., clinical expertise, research expertise, legal expertise, and patient/care-giver representatives), and whether there was a stated expert in guideline development within the committee's composition. We then focused on the researchers involved in the development of the guidelines. Research productivity, as measured by publication and citation counts, are often used as objective metrics for evaluating a researcher's performance (Alonso, Cabrerizo, Herrera-Viedma, & Herrera, 2009; Carleton, Parkerson, & Horswill, 2012; Ng, 2018). This study therefore focused on two metrics: publication count and *h*-index. Searches were conducted to assess the publication productivity of each committee member using publicly available data provided by MEDLINE, PsychINFO, Scopus, and Web of Knowledge/Web of Science. The publication searches were completed using the committee members' full names. Disambiguation was done manually for each publication by cross-referencing with discipline (e.g., psychology, psychiatry, social work, etc.), and cross-referencing the member's associated institution as listed on the publication. Search results belonging to different authors with the same or similar names were therefore removed.

Total number of peer-reviewed publications. To count each committee member's contributions to the scientific literature, we conducted a search of all peer-reviewed research in MEDLINE and PsycINFO. We excluded from our search conference proceedings, books, book chapters, and abstracts. The MEDLINE search was completed, between September 5th and September 14th, 2014, from the year 1946 (when the database was created) to one year after the publication date of a given guideline. Likewise, the search in PsycINFO was completed from the year 1967 (when the database was created), to one year after the publication date of the guideline. Data retrieved from both databases were merged, and duplicates were removed. The search parameters were set to determine the publication productivity of each committee member at the time of a guideline's publication. The additional year in the search parameters was to account for the extra time needed for submitted or in-press articles to be published.

A second search was conducted using the method described above, focusing exclusively on the total number of topic-related peer-reviewed publications (i.e. on the specific topic addressed in a given guideline), using topic-relevant search terms (see Table 1). The purpose was to examine the difference between committee members who were generally productive researchers, and productive researchers who were also experts on the topic of addressed in each guideline.

***h*-index.** The *h*-index, also known as the Hirsch-index, is a metric that measures a researcher's productivity and the citation impact of his or her publications (Alonso et al., 2009; Carleton et al., 2012; Hirsch, 2005). The index quantifies an individual's scientific research impact based on the number of cited papers produced, and on the number of citations received in other people's publications. More specifically, it is based on the list of publications by an author, ranking in descending order by the times it was cited, where the value *h* is equal to the number of articles (*N*) in the list that have *N* or more citations (see Hirsch, 2005). This metric is used here as an objective and complementary measure upon which to assess a guideline development

committee member's research impact. Though it cannot be used to isolate research output on a given topic or single area of expertise, the *h*-index remains very widely used and is considered to be objective.

Two index scores were retrieved from the Scopus and Web of Knowledge databases using the procedures provided by the University of Illinoi Libraries[1]. The search in Scopus was completed from the year 1960 (when the database was created) to one year after the publication date of the guideline. This extra year after the publication date is to account for the extra time typically need for submitted or in-press articles to be published. The search in Web of Knowledge was also set from the earliest possible year, 1900, to one year after the publication date of the guideline.

Results

No committee members were listed for the *Guidelines for Expert Assessment Concerning Child Custody and Access Rights* (2006); this guideline was therefore excluded from the analyses.

The *Guidelines for the Evaluation of Dyslexia in Children* (2014) had a total of 14 committee members (43% female). Nine of the 14 committee members had at least one publication. Based on the data retrieved from MEDLINE and PsycINFO, six of the 14 (42.8%) had 10 or more lifetime publications. Of the 14 committee members, three conducted research on the topic addressed in the guideline. One had 41 topic-related publications, another had 11, and a third committee member had two topic-related publications, at the time this guideline was being developed.

Overall, the Scopus *h*-index ranged from zero to 28 (0-28) and the Web of Knowledge *h*-index ranged from zero to 13 (0-13) (see Table 2). For those who had conducted research on the

[1] See: https://researchguides.uic.edu/c.php?g=252299&p=1683205.

topic addressed in the guideline, the Scopus *h*-index scores ranged from five to 17, and the Web of Knowledge scores ranged from two to 13 (see Table 2).

The *Guidelines for Autism Spectrum Disorder - Clinical Evaluation* (2012) had a total of eight committee members of whom two had at least one publication, although only one of the eight (12.5%) was a researcher in the field, as evidenced by the 113 topic-related peer-reviewed publication found, with a high Scopus *h*-index of 37, and a Web of Knowledge index score of 19 (see Table 2).

For the *Guidelines for the Assessment of Mental Retardation* (2007), five of a total of 10 committee members had at least one publication. Only one member of the development committee was a published researcher on the topic of the guideline (see Table 2). This committee member had a total publication count of 59, and a topic-related publication count of 26. The Scopus *h*-index ranged from zero to13 (0-13) and the Web of Knowledge *h*-index ranged from zero to 4 (0-4), among the published committee members.

Although *Guidelines for the Assessment of a Child in Connection with a Request for Derogation to the Age of School Admission* (2006) listed 11 committee members, we did not find any peer-reviewed publications over the lifetime of the committee members, or on the topic or related to the topic of the guideline for any of the group members. No *h*-index was found for these authors within the time parameters of our search.

Discussion

Overall, our findings suggest that there were some researchers on most of the guideline committees, which is congruent with recommendations for guideline development (Fretheim et al., 2006; Kunz et al., 2012; WHO, 2012). With some notable exceptions, the research productivity of these published committee members was however low, with the total number of lifetime publications ranging between one and 143 (M = 30.94, SD = 44.65). Furthermore, the

impact of these committee members also appeared to be low, with the Scopus *h*-indices ranging between zero and 37 (M = 7.88, SD = 10.94), and the Web of Knowledge *h*-indices ranging between zero and 19 (M = 3.81, SD = 5.67). More importantly perhaps, a closer examination indicates that few of the committee members were research experts on the topic of the guideline; three of the four guidelines that listed its committee members had as few as one researcher who was a research expert on the topic of the guideline at the time of its development.

More specifically, the *Guidelines for the Evaluation of Dyslexia in Children,* had the highest number of published committee members, with a total of nine members who had published at least one paper in peer-reviewed journals. While most had very few publications, three had over 30 lifetime publications. Furthermore, six of the 9 had not published on a topic related to the guideline, and three had published on that topic. While these numbers appear low, it is difficult to determine if such research productivity reflects true expertise on a given topic. A study by Carleton and colleagues (2012) can however shed light here. These researchers provided normative data surrounding the publication productivity of CPA-accredited Canadian clinical psychology professors using publically available data from all universities across Canada. These data were for all CPA accredited programs, including clinical programs (as opposed to programs offered in research intensive universities). The researchers reported that psychology professors published between zero and four articles annually. When professors were grouped by rank (assistant, associate, and full professors) and sex, they found that the Full Professors had a *mean* publication count of 32.05 (9.63 *h*-index) for women, and 47.33 (15.10 *h*-index) for men. This study also showed that Assistant Professors in psychology had approximately half the mean publication counts of Full Professors. In the case of the OPQ guidelines on dyslexia, it appears that most committee members have research outputs that are well below the normative data provided by Carleton and colleagues (2012). Moreover, according to Drapeau (2019), two of the

experts who contributed to these OPQ guidelines had asked that their names be removed from the document, a request which was denied, and several members of the committee expressed dissatisfaction about the content of the guideline and the methods used for its development. It is unclear if those who expressed such concerns were amongst the most, or the least productive researchers in the group.

The *Guidelines for Autism Spectrum Disorder - Clinical Evaluation* had a total of eight committee members of which two had at least one publication, and one (12.5%) was a highly productive research expert in the field. The latter had 124 lifetime publications and 113 topic-related peer-reviewed articles, which exceeded the mean publication count of the average Full Professors of psychology in Canadian universities (Carleton et al., 2012). However, this guideline offered little information regarding the research methods used to find and assess the evidence, or who was responsible for gathering, assessing, and selecting the research evidence.

The *Guidelines for the Assessment of Mental Retardation* had 10 committee members in total, and over half had at least one publication, yet only one committee member (10%) was a productive researcher (Carleton et al., 2012), with 59 total publications and 26 topic-related research articles. Therefore, like the guidelines for the evaluation of autism spectrum disorders, this guideline development committee did not have a strong representation of expert researchers in the topic addressed in the guideline. Like all of the guidelines in this study, little is known about the development methods followed to produce this guideline, which does not reflect 'best practices' in guideline development (Graham et al., 2011, 2010, 2012).

The *Guidelines for the Assessment of a Child in Connection with a Request for Derogation to the Age of School Admission* did not have any committee members with peer-reviewed publications, nor did the guideline indicate that a person adept at conducting systematic reviews was on the committee or provided services as a consultant. Furthermore, the description of

professional credentials, which was limited to one sentence placed at the top of the list of committee members that read 'Psychologists and Experts in the field,' was not in keeping with best practices in guideline development (Brouwers et al., 2010).

Unfortunately, the *Guidelines for Expert Assessment Concerning Child Custody and Access Rights*, did not list the names of those who participated in the development of these guidelines. Instead, the authors explain that the content of the guideline was "inspired by" a document published by the Association of Family and Conciliation Courts (AFCC). This association describe itself as being "the premier interdisciplinary and international association of professionals dedicated to the resolution of family conflict" (https://www.afccnet.org). According to their website, AFCC members include practitioners, researchers, teachers and policymakers in the family court arena, and have a fee-based membership. Irrespective of exactly how much of this guideline was "inspired" by information produced by the AFCC, the person(s) tasked with vetting the information of that guideline remain unidentified, which is contrary to best practice manuals in guideline development (Schünemann et al., 2014; SIGN50, 2015). Since the development of the *Guidelines for Expert Assessment Concerning Child Custody and Access Rights*, the ADAPTE Collaboration has published a tool kit to guide guideline adaptations to a local context (ADAPTE-Collaboration, 2009). This toolkit should be considered for any future revisions of this guideline to ensure that the adaptation to the local context is adequate.

Due to a paucity of information with regards to methods used to develop all of the guidelines in this study, our research team reached out to the OPQ[2] and requested: 1) All of the documentation pertaining to the development procedures used that were provided to persons on the guideline development committees; 2) Any supplemental material that may not have been

[2] Electronic correspondence requesting information through Quebec's Access to Information Act on April 4, 2018.

included in the guideline that describe the methods used to develop the guidelines; and 3) Whether a systematic literature search was completed, and if so, if they had any information on the method used. The OPQ responded by stating that they did not have information on the development process of any of their guidelines, aside from what was already reported in each guideline. They did however produce an email addressed to psychologists requesting their feedback about the *Guidelines for Autism Spectrum Disorder*. The representative from the OPQ confirmed that they had received feedback from over 40 psychologists for this guideline. Unfortunately, it remains unknown if feedback providers included researchers, how that feedback was considered, and nothing more pertaining to the development methods was provided. For some guidelines, information about the methodology used for their design was indeed reported in the guidelines themselves. However, previous studies have shown that the methodology used in most OPQ guidelines do not follow recommended guideline development methods (Stamoulos, Reyes, Trepanier, & Drapeau, 2014; Trepanier et al., 2017).

While this study highlights important limitations in how these guidelines were developed, including a potentially insufficient use of science and of researchers, it is important to note that this study has a number of limitations. For example, publication productivity is not the only measure of a committee member's contribution value. Many crucial insights in the assessment and care of the client or patient are obtained from strong clinical experience with a particular patient population, while some clinicians also keep up-to-date with the newest research in their field. Furthermore, Hirsch (2005), the developer of the *h*-index, acknowledged that seeking a method to quantify a researcher's impact was "potentially distasteful," but he also argued that it was needed for evaluation and comparison purposes. Furthermore, there are some known disadvantages to the *h*-index. For example, the total number of papers does not account for the quality of the publications, and the total number of citations can be disproportionately weighted

by having authored a few publications in high-impact journals, having many publications with few citations each, and it does not exclude self-citations (Sekercioglu, 2008; Zhang, 2009). Likewise, the application of PsycINFO and MEDLINE exclude potentially valuable, non-peer reviewed works, such as abstracts, chapters, and books.

While we value the participation of clinicians, and other stakeholders, in guideline development and the rich knowledge they bring to development committees, it is an enormous task to develop guidelines that reach the most individuals possible. But guidelines must be grounded in science (Beauchamp, Drapeau, & Dionne, 2015; Brouwers et al., 2010), and our study highlight a number of important problems in that regard, as some guidelines were developed without the input of researchers, and the researchers with a topic-specific expertise were often not highly productive. High publication productivity suggests that researchers are well positioned to evaluate the quality of studies in their field, enjoy frequent collaborations with other researchers, attend scientific conferences, and are therefore potentially privy to the newest findings prior to their publication. Thus, highly active researchers can offer valuable contributions on guideline development committees. Recommendations for future studies include examining the search and selection process for assembling a guideline development committee, develop a definition for the term “expert” committee member, determined how best to diversify expertise, examine how COIs are managed, and examine how to expand end-user involvement.

References

AGREE Next Steps Consortium. (2017). *The AGREE II instrument.* [Electronic version]. Retrieved from https://www.agreetrust.org/wp-content/uploads/2017/12/AGREE-II-Users-Manual-and-23-item-Instrument-2009-Update-2017.pdf

Alonso-Coello, P., Irfan, A., Solà, I., Gich, I., Delgado-Noguera, M., Rigau, D., . . . Schunemann, H. (2010). The quality of clinical practice guidelines over the last two decades: A systematic review of guideline appraisal studies. *Quality and Safety in Health Care, 19,* e58. doi:10.1136/qshc.2010.042077

Beauchamp, S., Drapeau, M., & Dionne, C. (2015). The development of practice guidelines in the social and human sciences. *Canadian Psychology/Psychologie canadienne, 56,* 357-367. doi:10.1037/a0038811

Beauchamp, S., Drapeau, M., Dionne, C., & Duplantie, J. P. (2015). *Institut national d'excellence en santé et en services sociaux (INESSS): Cadre d'élaboration des guides de pratique dans le secteur des services sociaux* [Policy statement]. Montreal, Canada: Author. Retrieved from INESSS website: https://www.inesss.qc.ca/fileadmin/doc/INESSS/Rapports/ServicesSociaux/INESSS_Cadre_elaboration_guides_pratique_servicessociaux.pdf

Bennett, K., Courtney, D., Duda, S., Henderson, J., & Szatmari, P. (2018). An appraisal of the trustworthiness of practice guidelines for depression and anxiety in children and youth. *Depression and Anxiety, 35,* 330-340. doi:10.1002/da.22752

Brouwers, M. C., Kho, M. E., Browman, G. P., Burgers, J. S., Cluzeau, F., Feder, G., . . . The AGREE Next Step Consoritum (2010). AGREE II: Advancing guideline development, reporting and evaluation in health care. *Canadian Medical Association Journal, 182,* E839-E842. doi:10.1503/cmaj.090449

Carleton, R. N., Parkerson, H. A., & Horswill, S. C. (2012). Assessing the publication productivity of clinical psychology professors in Canadian Psychological Association-accredited Canadian psychology departments. *Canadian Psychology/Psychologie canadienne, 53*, 226-237. doi:10.1037/a0027731

Drapeu, M. (2019, March 7). The OPQ guidelines on dyslexia: Guidelines for the assessment of dyslexia published by the Order of Psychologists of Quebec. [Web log comment]. Retrieved from http://www.mpprg.mcgill.ca/blog_OPQGuidelinesDyslexia.html

Fretheim, A., Schunemann, H. J., & Oxman, A. D. (2006). Improving the use of research evidence in guideline development: 5. group processes. *Health Resources and Policy Systems, 4*, 17-35. doi:10.1186/1478-4505-4-17

Gordon, J., & Cooper, B. (2010). Talking knowledge — practising knowledge: A critical best practice approach to how social workers understand and use knowledge in practice. *Practice, 22*, 245-257. doi: 10.1080/09503153.2010.495768

Graham, I. D., Mancher, M., Wolman, D. M., Greenfield, S., & Steinberg, E. (2011). *Clinical practice guidelines we can trust*. Washington, DC: Institue of Medicine National Academies Press.

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *PNAS, 102*, 16569-16572. doi: 10.1073/pnas.0507655102

Kunz, R., Fretheim, A., Cluzeau, F., Wilt, T. J., Qaseem, A., Lelgemann, M., . . . Coordinating Efforts in, C. G. D. (2012). Guideline group composition and group processes, article 3 in Integrating and coordinating efforts in COPD guideline development - An official ATS/ERS workshop report. *Proceedings of the American Thoractic Society, 9*, 229-233. doi:10.1513/pats.201208-056ST

MacQueen, G., Santaguida, P., Keshavarz, H., Jaworska, N., Levine, M., Beyene, J., & Raina, P. (2016). Systematic review of clinical practice guidelines for failed antidepressant treatment response in major depressive disorder, dysthymia, and subthreshold depression in adults. *The Canadian Journal of Psychiatry*, *62*, 11–23. doi:10.1177/0706743716664885

Ng, C. F. (2018). Research productivity and impact of professors in Canadian undergraduate psychology programmes. *Canadian Psychology/Psychologie canadienne, 59*, 211-218. doi: /10.1037/cap0000137

Office of Professions of Quebec. (2018). *Code of ethics for psychologists* (c. C26, r. 212, a. 5) [Online reference]. Retrieved from www.legisquebec.gouv.qc.ca/

Ordre des psychologues du Québec (OPQ). (2012). *Guidelines for autism spectrum disorder: Clinical evaluation*. Montreal, Canada: Author. Retrieved from https://www.ordrepsy.qc.ca/documents/26707/63191/spectre-de-lautisme/868d8894-44f1-471d-941d-94a5f571395c

Ordre des psychologues du Québec (OPQ). (2006). *Guidelines for the assessment of a child in connection with a request for derogation to the age of school admission*. Montreal, Canada: Author. Retrieved from https://www.ordrepsy.qc.ca/documents/26707/0/Les+lignes+directrices+pour+l%27%C3%A9valuation+d%27un+enfant+en+vue+d%27une+demande+de+d%C3%A9rogation+%C3%A0+l%27%C3%A2ge+d%27admission+%C3%A0+l%27%C3%A9cole/b2c2e606-7fad-4b18-8ec8-883fb69e572e

Ordre des psychologues du Québec (OPQ). (2006). *Guidelines for expert assessment concerning child custody and access rights.* Montreal, Canada: Author. Retrieved from https://www.ordrepsy.qc.ca/documents/26707/6319/Lignes+directri

ces+pour+l%27expertise+en+mati%C3%A8re+de+garde+d%27enfants+et+des+droits+d%27acc%C3%A8s/d0c94c08-c04d-49ae-8fea-b37bd0ed83b8

Ordre des psychologues du Québec (OPQ). (2007). *Guidelines for the assessment of mental retardation*. Montreal, Canada: Author. Retrieved from https://www.ordrepsy.qc.ca/documents/26707/63191/Lignes+directrices+pour+l%27%C3%A9valuation+du+retard+mental/f20a2002-4535-4070-8a03-15889a9d5450

Ordre des psychologues du Québec (OPQ). (2014). *Guidelines for the evaluation of dyslexia in children*. Montreal, Canada: Author. Retrieved, from http://www.ordrepsy.qc.ca/documents/26707/63191/Lignes+directrices+pour+l%27%C3%A9valuation+de+la+dyslexie+chez+les+enfants/139edf81-9163-4060-9c66-06f02eee0ed8

Oxman, A. D., Fretheim, A., & Schunemann, H. J. (2006). Improving the use of research evidence in guideline development: Introduction. *Health Resources and Policy Systems, 4*, 12-18. doi:10.1186/1478-4505-4-12

Schünemann, H. J., Wiercioch, W., Etxeandia, I., Falavigna, M., Santesso, N., Mustafa, R., . . . al., e. (2014). Guidelines 2.0: Systematic development of a comprehensive checklist for a successful guideline enterprise. *Canadian Medical Association Journal, 186*, E123-E142. doi:10.1503/cmaj.131237

Sekercioglu, C. H. (2008). Quantifying coauthor contributions. *Science, 322*, 371-371. doi: 10.1126/science.322.5900.371a

Shiffman, R. N. (2016). Best practices for implementation of clinical decision support. *Clinical Decision Support Systems, 2,* 99-109. doi: 10.1007/978-3-319-31913-1_6

Scottish Intercollegiate Guidelines Network (SIGN). (2015). *SIGN50: A guideline developer's handbook* [Electronic version]. Retrieved from https://sign.ac.uk/assets/sign50_2015.pdf

Stamoulos, C., Reyes, A., Trepanier, L., & Drapeau, M. (2014). The methodological quality of clinical practice guidelines in psychology: A pilot study of the guidelines developed by the Ordre des Psychologues du Québec. *Canadian Psychology/Psychologie canadienne, 55*, 177-186. doi:10.1037/a0037287

The ADAPTE Collaboration (2009). *The ADAPTE process: Resource toolkit for guideline adaptation* (Version 2.0). Retrieved from the Guidelines International Network (G-I-N) website http://wwwg-i-n.net.

Trepanier, L., Stamoulos, C., & Reyes, A. (2017). Assessing the quality of the OPQ's guidelines for the evaluation of dyslexia in children using the Appraisal of Guidelines for Research and Evaluation (AGREE) II: A brief report. *Canadian Psychology/Psychologie canadienne, 58*, 211-217. doi:10.1037/cap0000121

World Health Organization (WHO). (2010). *WHO handbook for guideline development.* Geneva, Switzerland: Author. Retrieved from http://www.who.int/hiv/topics/mtct/grc_handbook_mar2010_1.pdf

Wold Health Organization (WHO). (2012). *WHO handbook for guideline development*. Geneva, Switzerland: Author. Retrieved from https://apps.who.int/iris/bitstream/handle/10665/75146/9789241548441_eng.pdf?sequence=1

Zhang, C. (2009). A proposal for calculating weighted citations based on author rank. *EMBO Reports, 10*, 416–417. doi:10.1038/embor.2009.74

Table 1.

Search words used to find committee member publications on the topic of the guideline

Guideline	Search terms in English / French
Dyslexia Guideline (2014)	Dyslexia / Dyslexie
	Learning disabilities / Troubles d'apprentissage
	Reading disabilities / Troubles de lecture
Autism Guideline (2012)	Asperger / Asperger
	Autism, Autistic / Autisme, Trouble autistique
	Autism spectrum disorders (ASD) / Désordre du spectre autistique (DSA)
	Developmental delays / Retard de developpement
	Pervasive developmental disorder (PDD) / Troubles envahissants du développement (TED)
	Rett syndrome / Le syndrome de Rett
Mental Retardation Guideline (2007)	Assessment intellectual functioning / Evaluation fonctionnement intellectuel
	Developmental intellectual delays / Déficience intellectuelle
	Intellectually handicapped / Handicap intellectuel
	Intellectual disabilities / Troubles d'apprentissage
	Mental Retardation / Retard mental
Derogation for Age of School Admission Guideline (2006)	Committee members were not published in peer-reviewed journals
Child Access & Custody Rights Guideline (2006)	Committee members were not listed.

Table 2.

Publication productivity and h-index of guideline development committee members.

Guidelines	Committee Members *n* (%)	MEDLINE / PsycINFO Total no. of Publications*	MEDLINE / PsycINFO No. of Publications on Guideline Topic**	*h*-index*** Scopus	*h*-index*** WoK
Dyslexia Guideline (2014)					
Total Committee Members	14				
Members with at least 1 pub.	9 (64.3%)				
Committee member 1		16	0	3	2
Committee member 2		143	0	28	11
Committee member 3		36	11	10	6
Committee member 4		10	0	5	1
Committee member 5		59	41	17	13
Committee member 6		2	0	0	0
Committee member 7		27	2	5	2
Committee member 8		4	0	3	1
Committee member 9		3	0	2	1
Range		2-143	0 - 41	0 -28	0-13

Autism Guideline (2012)					
Total Committee Members	8				
Members with at least 1 pub.	2 (25%)				
Committee member 1		124	113	37	19
Committee member 2		3	3	1	0
Range		3-124	3-113	1-37	0-19
Mental Retardation Guideline (2007)					
Total Committee Members	10				
Members with at least 1 pub.	5 (50%)				
Committee member 1		1	0	2	0
Committee member 2		3	1	0	0
Committee member 3		59	26	13	4
Committee member 4		4	0	0	0
Committee member 5		1	0	0	0
Range		1-59	26-0	0-13	0-4
Derogation for Age of School Admission (2006)					
Total Committee Members	11				

Members with at least 1 pub.	0 (0%)	0	0	0	0
Child Access and Custody Rights (2006)					
Total Committee Members	--				
Members with at least 1 pub.	--	--	--	--	--

Note. * Includes all peer-reviewed articles published up to one year after the publication of the guideline.

**Includes all peer-reviewed articles on the topic of the guideline, published up to one year after the publication of the guideline.

*** The *h*-index search parameters were limited to one year after the publication of the guidelines, and by default of the index's search algorithms, includes peer-reviewed and non peer-reviewed works.

Linking Manuscripts 2 and 3

Study 1 pointed to specific methodological development flaws in the five guidelines studied, and study 2 showed that the committees were not composed of enough researchers who were experts on the topic of the guideline. This led to our next investigative leap. We endeavored to survey the committee members and ask them specific questions about: 1) Their familiarity with established development methods; 2) If they felt free to express their views; 3) If they believed their opinions were duly considered; 4) If there were procedures for dealing with divergent views during the development process, or whether they were vague or lacking; and 4) If records were kept when disagreements occurred when views differed on which recommendations would be best. Given that the AGREE II domain 6 addressed conflicts of interest (COI) which was also poorly addressed in the guidelines, we inquired whether the participants viewed the COIs as having been adequately reported. Finally, it was important to know if the developers believed that the guideline they worked on ought to be updated and this point, and if they endorsed their guideline; as we will see this was not always the case.

We then proceeded to invite each person who worked on one of the OPQ guidelines to participate in our survey. Unfortunately, the responses were too low (n=9) to interpret meaningfully, so we opted to widen our participant pool to other guideline developers in allied disciplines, such as social workers, psychoeducators, counselors, occupational therapists, and physicians, who participated on guideline development committees for their respective Orders. The following provides insight into the views, knowledge, and inner workings of guideline development committees in the social sciences and medicine.

Chapter 4

Manuscript 3

(Paper to be submitted for publication)

Practice Guideline Developers Share their Views and Experiences as Members of a Guideline Development Committee within the Social Sciences

Trepanier, Lyane[1], Stamoulos, Constantina[1], Beauchamp, Sylvie[1,2], Larivée, Serge[3] & Drapeau, Martin[1,4]

Author Note

Department of Counselling Psychology, McGill University[1], Montreal West Island Integrated University, Health and Social Services Center[2], School of Psychoeducation, University of Montreal[3], Department of Psychiatry, McGill University[4]

Lyane Trepanier's was supported by the SSHRC Joseph-Bombardier Canada Graduate Scholarship Doctoral Award. The work of Martin Drapeau was supported by SSHRC Research Grant. The work of Constantina Stamoulos was supported by the SSHRC Doctoral Award and FQRSC Doctoral Research Training Scholarship, and Andrea Reyes was supported by the Fonds de Recherche sur la Société et la Culture du Québec (FRQSC). No funding body had any input into any aspect of this research.

Correspondence address: Dr. Martin Drapeau, MPPRG – McGill University, 3700 McTavish, Montreal, Quebec, H3A 1Y2, Canada. E-mail: martin.drapeau@mcgill.ca

Abstract

Clinical practice guidelines (CPGs) have proliferated in medicine and more recently, in the social sciences. Although the quality of CPGs available to practitioners has received increasing attention over the past decade, little attention has been directed toward examining the perspectives and processes of guideline development groups in the social sciences. In this study, we surveyed 40 CPG development committee members who worked on one of the 17 guidelines published by six Quebec regulatory bodies in the social sciences. We examined their knowledge about guideline development methodology and solicited their views on the quality of the guidelines they developed. Results show that the developers' familiarity with established development methods must be improved; group composition must be broadened; procedures for dealing with divergent views during the development process were vague or lacking; conflicts of interest were inadequately reported; and that the guidelines, although still currently available to practitioners, should be updated. This study provides new insights on the inner working of guideline development groups in the social sciences.

Keywords: Clinical practice guidelines, guideline development, guideline quality, survey, psychology guidelines, social work guidelines, psychoeducation guidelines, counselling guidelines.

Practice Guideline Developers Share their Views and Experiences as Members of a Guideline Development Committee within the Social Sciences

The purpose of clinical practice guidelines (CPGs) is to enhance practitioner knowledge, facilitate the application of scientifically supported recommendations, improve outcomes, and lower costs (Graham, Mancher, Wolman, Greenfield, & Steinberg, 2011; Hollon et al., 2014; Rutten et al., 2016; van Dijk, Oosterbaan, Verbraak, & van Balkom, 2013; Wolf, Hubbard, Faraday, & Forrest, 2011). Guidelines are also known to reduce variation in treatment delivery (Schuh et al., 2017) and discourage outdated or ineffective clinical approaches (Fervers et al., 2011; Shaner, 2001). For these reasons, CPGs are appreciated within a vast number of healthcare professions or settings (Bernard et al., 2008; Gordon & Cooper, 2010; Hudson, Quinn, O'Hanlon, & Aranda, 2008; Kendall, Taylor, Perez, & Taylor, 2008; OPQ, 2007; van Dijk et al., 2013).

However, useful guidelines are not easy to produce. They require specific processes and considerations. In 2011, the *Institute of Medicine* (IOM; now the *National Academy of Medicine*) outlined formal standards of trustworthy guidelines that specify procedures for producing scientifically valid, transparent, and reproducible results and recommendations (Graham et al. 2011). The IOM formal standards for trustworthy guidelines are comprised of eight key criteria: 1) Transparency of process; 2) Fully disclosed conflicts of interest; 3) Guideline development group composition, whereby development groups should be composed of methods experts, clinicians, representatives of stakeholders, and affected populations; 4) Systematic review, a process by which evidence from multiple studies is collected and analyzed using predetermined and rigorous methods; 5) The quality and strength of the evidence on which the recommendations are based have been thoroughly examined; 6) Well articulated recommendations, such as how

and when they should be used; 7) External reviews, conducted by independent experts; and 8) Guidelines are kept up to date.

To date, the eight criteria listed above have guided the formal assessment of practice guidelines with special attention paid to methodological rigour, transparency of the data collection methods, and the usability of the final product, by applying the *Appraisal of Guidelines for Research and Evaluation* (AGREE II; Brouwers et al., 2010; Graham et al., 2011; Hollon et al., 2014; Pilling, 2012). While considerable research has been conducted on these key features of guidelines (Stamoulos, Reyes, Trepanier, & Drapeau, 2014; Trepanier, Stamoulos, & Reyes, 2017; Tudor, Kozina, & Marušić, 2013; Watine & Bunting, 2008), little research has examined the composition of guideline development groups or explored the views and perspectives of committee members who participated in the guideline development process. Like all group endeavors, developing practice guidelines is an interpersonal process (Eccles, Grimshaw, Shekelle, Schünemann, & Woolf, 2012; Richter Sundberg, Garvare, & Nyström, 2017). For example, although the specific features of a good guideline are well established (Graham, et al., 2011), little is known about how decisions are made, more specifically, how consensus among developers is reached, what happens when there is disagreement within critical developmental activities, or whether in the end developers fully endorse the final product. An investigation into the views of guideline development committee members would provide additional information on which aspects of the development processes work well, and what could be improved to make the process smoother and more transparent and efficient for other development projects. Given that the production of practice guidelines has accelerated among regulatory bodies, governmental agencies, and professional associations in the past few years, a closer examination of the guideline development process would inform future development initiatives.

The scope of this study was limited to guidelines that were produced by Quebec Colleges (referred to as "Orders" in Quebec) that operate as the professional regulatory and licensing bodies in Quebec. They are responsible for protecting the public and service users and imposing rules and standards for their professionals to follow, and bear the responsibility to exercise legal and moral authority over their members. They must therefore be held to a high standard when they publish documents that guide the professionals under their authority. Thus, we surveyed the guidelines development committee members who worked on the guidelines published by six Quebec regulatory bodies: the *Collège des médecins du Québec* (CMQ: College of Physicians of Quebec); the *Ordre des psychologues du Québec* (OPQ; College of psychologists of Quebec); the *Ordre des travailleurs sociaux et des thérapeutes conjugaux et familiaux du Québec* (OTSTCFQ: College of Social Workers and Family Therapists of Quebec); the *Ordre des psychoéducateurs et psychoéducatrices du Québec* (OPPOQ: College of Psychoeducators of Quebec); the *Ordre des conseillers et conseillères d'orientation du Québec* (OCCOQ: College of Guidance Counsellors of Quebec); and the *Ordre des ergothérapeutes du Québec* (OEQ: College of Occupational Therapists of Quebec). We examined their prior knowledge of guidelines development methodology, their experience as members of a development committee, and solicited their views on the quality of the guidelines they helped to develop.

Methodology

This study received ethical approval from the McGill University Research Ethics Board II (REB # 409-0415).

Participants

The participants were comprised of the guideline development committee members, also known as 'working groups,' identified in one or more of the 17 CPGs developed by the

professional regulatory bodies listed in Table 1. We identified 125 committee members listed in the guidelines; 11 of the 125 committee members had been involved in the development of more than one guideline, bringing the total number of unique guideline developers and potential study participants to 114. Of these, we were able to contact 96 committee members via email to solicit their participation in this study; 48 agreed to participate. However, the number of completed surveys was further reduced to 40, as eight individuals abandoned the survey once started. Thus, the total number of participants included in this study was 40.

We also examined the role participants had on their respective guideline development committees. Of the 40 participants, 26 indicated that they were clinicians, and therefore brought a clinical perspective to the group, six were full-time researchers, five participants indicated that their role was primarily administrative, and three were group representatives (e.g., service user).

Measure

A five-point Likert scale survey with 43 statements was administered in French via the Qualtrics.com platform (see Table 3 for the full survey). The survey addressed several key areas of guideline development at the committee level, including the following: expertise on the topic of the guideline; the collaborative experience; whether rules and procedures were clearly established, including what to do in case of disagreement among members; the research methods used; the consensus process; if they viewed the guideline as out of date; familiarity with development manuals; and finally, whether the respondent endorses the guideline. Response choices were assigned the following scores: *Strongly Disagree* = 1, *Disagree* = 2, *Neither Agree nor Disagree* = 3, *Agree* = 4, and *Strongly Agree* = 5. The responses *I don't know* or *no response* were not included in the calculation of the means, medians, and standard deviations reported below. When relevant, we compared the responses between clinicians and researchers.

The survey involved some deception found in a section that assessed the respondent's knowledge about guidelines development manuals. This information was to assess the value of the responses and social desirability (Ferrando, 2008; O'Connor, 2006).

Procedures

Once the guideline developers were identified, their contact information was collected by contacting the Colleges to which they belonged, the organizations listed in the guideline, or via an Internet search. Each person received a survey that referred to the guideline(s) they were involved in. If a guideline developer was involved in more than one guideline, that person received an email invitation for each guideline they were involved with, to provide an opportunity to report on each development experience separately. Participants could enter a draw to win an Apple IPad Mini once they had completed the survey. Some Colleges are more prolific than others for developing practice guidelines, which is why we included as few as two guidelines from some, while others have as many as 11 guidelines in this study.

Data analysis

The demographic characteristics were documented, and descriptive statistics were calculated. We combined and tallied the "*Agree* and *Strongly agree*" items and presented them with their corresponding percentages. A Mann-Whitey *U* test was also performed to examine the response differences between the clinicians and researchers, categorized by their reported role on the development committees, with Bonferroni corrections when needed. Finally, we examined the responses by year of publication using frequency tables and scatterplot graphs to identify a trend in responses across time.

Results

Participants

Forty (40) surveys were completed online. Participants were comprised of representatives from six professional orders and self-identified as having one of several principal occupations (see Table 2 for full demographic data). Eight surveys were partially completed and were consequently excluded from our analysis. All eight abandoned surveys were for the OPQ guidelines (see Table 1). One person belonging to one of the OPQ guideline development groups informed our researchers via email that he/she would not participate due to the age of the guideline, as he or she could not accurately recall the details of the development process.

Survey Response Results

Questions one to three addressed the respondent's overall familiarity with various aspects of the guideline topic. Respondents *Agreed* or *Strongly agreed* that they were familiar with the research (Q.1, 83%) and clinical work (Q. 2, 90%), as well as with policies and regulations related to the guideline topic (Q.3, 80%) (see Table 3 for all response scores).

The developer's overall impression of the collaborative atmosphere was captured in questions four, seven, eight, and nine. The combined *Agree* and *Strongly Agree* scores ranged from 84% to 95% for these items (see Table 3). Thus, the results show that the majority of committee members believed that their expertise was fully utilized, their contributions were invited and duly considered, and that all members had an equal voice.

Questions 10 through 14 addressed how disagreements were resolved. Most of the respondents *Agreed* or *Strongly agreed* that there was room for disagreement in their respective development groups (Question 10, 85%), but few experienced disagreements (Question 11, 19%) (see Table 3). Question 12 asked about the transparency of the rules and procedures for reconciling disagreements regarding how research findings should be interpreted. A total of 45% either *Agreed* or *Strongly agreed* that they were aware of such a procedure. When asked about

whether there was a procedure to reconcile disagreements between committee members about the content of the guidelines recommendations (Question 13), slightly fewer agreed (39%). Some reported incidences of actual conflicts between committee members (Question 14; 16%).

Questions 15 to 17 queried the formality of the decision-making processes, and how dissenting opinions, incongruent with the final guideline recommendations, were documented (see Table 3). More respondents *Agreed* or *Strongly Agreed* that the decision-making processes were informal (Question 15; 75%), rather than formal (Question 16; 7%). For Question 17, only 27 of 40 (66%) respondents provided a response to this item, which may suggest that many did not know what a "minority report" was, or were not sure if one was kept during the development process. Of the 27 responses, only two responded that they *Agreed* or *Strongly Agreed* that a minority report was created (Question 17, 8%). A total of 68% *Agreed* or *Strongly Agreed* that the guideline's recommendations were based on the highest quality scientific evidence (Question 18). Likewise, 70% (Question 6) agreed that the guideline was developed using best practices in guideline development.

Questions 19, 25, 32, and 33 were related to how much consideration was given to the guideline's applicability in a clinical context. Only 25% of respondents *Agreed* or *Strongly Agreed* that barriers, such as cost, and the professional resources needed to implement the guidelines were considered (Question 25). The applicability of the scientific evidence to the clinical context was discussed in the development meetings (Question 32) according to 46% of the respondents, while 53% reported that the health benefits, risks, and side-effects of the guideline recommendations were also considered (Question 33). Finally, 50% of respondents *Agreed* or *Strongly Agreed* that the guidelines should be updated to reflect newer findings (Question 19).

Questions 21 to 24 related to the intended purpose of the guideline; 90% believed that the guideline's purpose was to make suggestions that guided and informed professionals (Question 22). The statement with the lowest agreement score in this set was for Question 24 (68%), which stated that the purpose of the guideline was to outline and provide recommendations that are evidence-based.

The lowest *Agree* and *Strongly Agree* scores were found in the cluster of questions that addressed the guideline development and research methodology. The percentage scores for Questions 27 to 31 ranged from 30% to 37%. Only 32% reported that a comprehensive plan was established to conduct a systematic search for the evidence (Question 27), and 35% of respondents for Question 28 agreed that this plan was adhered to. Slightly fewer (30%) agreed that explicit criteria for including or excluding research evidence was adhered to (Question 29), while 33% of respondents agreed that the strengths and limitations of the studies retrieved were duly considered (Question 30).

This survey also included one question that addressed potential conflicts of interest (COIs) for guidelines. About half (55%) reported that they agreed that COI risks for members of the guideline development committee was evaluated and reported (see Question 26). However, nine individuals who completed the survey did not respond to this question.

Finally, to assess whether guideline committee members were familiar with guideline development manuals, such as the ones listed in Questions 36 to 43, respondents were asked to rate how familiar there were with each one on a scale of 1 to 5. The responses ranged between 0% and 15% of respondents who were *Familiar* or *Very Familiar* with the development manuals listed). Two false options (Questions 38 and 42) were presented to identify possible social desirability responses, of which all responded with *Very Unfamiliar and Unfamiliar*.

Question 35 asked the respondent if they would still endorse their guideline and 74% *Agreed* and *Strongly Agreed* that they would.

Comparing Researchers to Clinicians

As the data were skewed, a Mann-Whitney *U* test with a Bonferroni adjustment (for $\alpha = .05$) was performed to compare the researcher group ($n = 6$) with the clinician group ($n = 26$), for each survey question. The test indicated that the researcher group gave lower scores to all four significant results. For Question 16, more clinicians (*Mdn*= 2.0) agreed that a "*Formal consensus was the method used to formulate the final recommendations in this guideline,*" than the researchers (*Mdn* = 1.0), $U = 9.00$, $p = .00$, $r = -0.69$. For Question 30, clinicians agreed more (*Mdn*= 3.0) that "*The strengths and limitation of the studies retrieved were assessed by examining their design, and methodology*" than did the researchers (*Mdn* = 1.0), $U = 15.0$, $p = .00$, $r = -0.55$. More clinicians (*Mdn*= 4.0) believed that "*the health benefits, side effects, and risks, were considered when formulating the recommendations*," than researchers (*Mdn* = 2.0), $U = 15.0$, $p = .00$, $r = -0.55$ (Question 33). Finally, for Question 34, more clinicians agreed (*Mdn*= 4.0) that "*The quality of the guidelines development process serves as an excellent example for future guideline development*" than did the researchers (*Mdn* = 1.5), $U = 23.00$, $p = .00$, $r = -.48$. Other comparisons were performed and were not found to be significant[1].

Discussion

Little is known about how guideline development committees are formed or how they operate. This study was one of the first to examine the experience, views, and methodological knowledge of guideline development committee members in the social sciences. The researchers

[1]An examination of responses across time by year of publication revealed no trend for any of the statements provided. Scatterplots showed that the data was generally flat for each statement, with occasional outliers not consistent with any trend.

found that there was a relatively high incidence (8 of 48) of abandoned surveys with four of the 17 clinical practice guidelines in this study, all of which belonged to the OPQ. The reason why is unclear since the OPPQ and the OTSTCFQ had a similar number of guidelines and respondents in this study. Another observation was that the composition of the guideline development committees was disproportionately composed of clinical practitioners and members of the organizations that produced the guideline. Organizations such as the World Health Organization (WHO) and the IOM recommend that the composition of development committees include other stakeholders such as researchers, method experts, as well as affected populations or their caregivers, & Oxman, 2006; Graham et al., 2011).

The committee members in this study reported that they were familiar with the clinical work, research, and policies on the topic of the guideline they developed. This was expected since these individuals were invited to be part of the development committee due to their presumed expertise on the topic of the guideline.

When asked about the quality of the collaboration between committee members, the vast majority of the respondents agreed that they had an equal voice (85%), and were encouraged to contribute to the group (95%). Overall, the respondents reported that the atmosphere within their respective groups was positive and collaborative. Most described the decision-making process as informal (75%), such that most of the decisions were reached by group discussion, rather than through more formal means like a Delphi (Questions 16 and 17). Group differences between clinicians and researchers revealed that the researchers were more likely to describe the development process as informal.

We also examined whether or not there was a clear procedure in place that guided the development process as a whole, and one that guided the search and selection of the research

evidence on which the recommendations were based. It is here that the respondents reported the lowest agreement with the statements provided. When asked if there was a comprehensive plan established to guide a systematic search for scientific evidence, only 10 (32%) agreed or strongly agreed that such a plan was in place, and adhered to (Questions 27 and 28). Less than a third of respondents also agreed that there were explicit inclusion and exclusion criteria for the literature search, and that the design, methodology, and consistency of the studies (direction of results) were sufficiently scrutinized. Interestingly, 28 (70%) respondents agreed that their guideline was developed using best practices in the guideline development (Question 6), and 25 (68%) of respondents reported that the guideline they developed was evidence-based (Question 18). The responses to Question 27 and Question 28 are incompatible with the responses to Question 6 and Question 18. The absence of a comprehensive plan for the development process and literature search indicates that best practices were likely not used. Thus, it is unclear what the development committee respondents consider as "best-practices" or "evidence-based." This is unfortunate, since users would reasonably expect that all of the recommendations contained in a guideline would be based on rigorously assessed research.

Our previous research, which examined the quality of five OPQ guidelines published between 2006 and 2014 using *The Appraisal of Guidelines for Research & Evaluation* (AGREE) II appraisal instrument, showed that the guidelines had the lowest scores for items related to the transparency of the method used, and for poorly cited guideline recommendations (Stamoulos et al., 2014; Trepanier et al., 2017). The other guidelines in this study, to our knowledge, have not been objectively appraised, therefore we cannot draw any conclusion about their quality based on our findings. However, approximately two-thirds (2/3) of the respondents reported that there was no formal research method on which the recommendations were based (as addressed in Questions

27 to 32), which is noteworthy since this is an important component of the development process for trustworthy guidelines (e.g., Graham et al., 2011). When the guidelines were analyzed by year of publication, there was no appreciable difference in responses across time. A group difference between researchers and clinicians was found for Question 30, *the strengths and limitations of the studies were assessed by examining their design and methodology*, where researchers disagreed with this statement significantly more. A reasonable conjecture for this finding is that researchers had more stringent criteria for what constitutes a thorough examination of a study's quality.

How disagreements between committee members were handled is also of great importance. Seven (19%) respondents reported that there were discussions that resulted in disagreements, and six (16%) of them reported that they felt the disagreements remained unresolved. Most did not respond to the 'minority report' question (Question 17). Less than half (45%) reported that there was a transparent system in place to reconcile disagreements over how the scientific data should be interpreted, and 39% reported that there was a system in place to reconcile differences of opinion about the content of the references in the guideline. Disagreement between committee members can be expected; however, it is also important to document when one or more members disagree or partially disagree with a recommendation in the guideline (Beauchamp, Drapeau, & Dionne, 2015; Beauchamp, Duplantie, & Mercier, 2011; Fretheim et al., 2006; Kunz et al., 2012). This is typically documented in what is called a 'minority report.' When respondents were asked if a minority report was kept, 13 respondents either did not answer or responded with "I don't know"; of those who did respond, only two (7%) agreed that one was maintained by their committee. It is important to keep a record of this since a person's participation on a guideline development committee is publically available information, and any disagreement with a

recommendation found in a guideline has professional implications for those mentioned in the guideline's credits, and therefore divergent views should be publically documented.

Another important, yet seemingly overlooked requirement for transparent guideline development, is the reporting of conflicts of interest (COIs) in many healthcare guidelines (Morciano et al., 2016; Wang et al., 2017). In the present study, nine individuals responded "I don't know" to question 26, which pertained to the evaluation and reporting of COIs among committee members, whereas 17 (55%) respondents agreed that these were examined and reported. Given that most of the guidelines had committee members who where employed by the regulatory bodies that funded and produced the guideline, such is a disclosure is relevant to the end-user (Graham et al., 2011).

At a minimum, a good guideline is up to date (Graham et al., 2011; Shekelle, 2014; Shekelle, Woolf, Grimshaw, Schünemann, & Eccles, 2012). A surprisingly high number of respondents, 18 (50%), agreed that the guidelines they developed should be updated at this point. Guidelines should be updated every three to five years (Ahmadzai et al., 2013; Alderson, Alderson, & Tan, 2014; Shekelle, 2014). None of the guidelines in our study provided a specific time or date for which the guidelines aught to be reviewed, beyond the suggestion that it should be updated when newer evidence became available.

Only 23 (59%) respondents reported that the quality of the development process serves as an excellent example for future guideline development initiatives. Clinicians were shown to be significantly more likely to respond in the positive about the development process, than researchers. This may suggest that researchers may expect to have a more formal and rigorous approach to guideline development, or are more knowledgeable in this area.

There were no clinician/researcher group differences when asked if they endorsed the guideline, as 28 (74%) endorsed their guidelines, and seven responded that they disagreed or strongly disagreed with this statement. Three individuals who responded in the negative stated that they were dissatisfied with the how the development process unfolded, how conflicts were managed, and the committee's leadership.

Finally, Questions 36 to 43 listed guideline development manuals, with two false options to test for responses influenced by 'social desirability.' The percentage of those who were familiar with guideline development manuals was very low, ranging from 0% to 15%, and two people responded that they were familiar and very familiar with one of the false options. This is revealing because most of the manuals and evaluation instruments listed in this series are very useful tools for guideline development and further lend support for having at least one guideline methodologist on a development committee. One of the earliest guideline development tools listed in the series of statements at the end of the survey (Question 36 to Question 43) was the SIGN50 which was published in 2001, followed by the AGREE I in 2003. This is noteworthy since the earliest guidelines in our study were published in 2006, with the vast majority of the guidelines published after 2010. Therefore at least some of these tools were available at the time the guidelines were being developed; it is therefore reasonable to expect a certain familiarity with these tools as guideline developers. However, it is often the case that we do not know what we don't know, which suggests a need for at least one committee member who is knowledgeable on guideline development methods to ensure that best practices in guideline development are being followed.

In sum, this study was the first to explore the views and experiences of clinical practice guideline developers and allowed us to peak behind this closed-door process. A strength of this

study was that it provided never-before investigated insight into the subjective experiences of guideline committee the members, as well as the technical procedures and development processes followed. Moreover, the committee members' responses further supported our previous findings which showed that established methods for guideline development are not consistently adhered to (Trepanier et al., 2017). A weakness of this study was its retrospective design and the potential for inaccurate recall. Some guidelines were produced five years before this investigation began, while others were produced up to 12 years prior. We also obtained group imbalances between clinicians and researchers, which greatly limited our statistical comparisons. It was unfortunate that so few of those who identified themselves as researchers responded to our survey. However, our previous research has also demonstrated that some of the practice guidelines in this study, such as the ones developed by the OPQ, were mostly developed by clinicians; this may also be the case for other guidelines in this study (see Study 2 in this thesis).

Future research should focus on obtaining more detailed information on the committee selection procedures, the guideline development processes from beginning to end, and perhaps an *in Vivo* observation of how the committee meetings are conducted. Such research would alert guideline producers of potential problems in these areas, and possibly improve the quality and trustworthiness of the final product for end-users.

References

AGREE Collaboration. (2003). Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: The AGREE project. *Quality & Safety in Health Care, 12*, 18-23. doi: 10.1136/qhc.12.1.18

AGREE Next Steps Consortium. (2017). *The AGREE II instrument.* [Electronic version]. Retrieved from https://www.agreetrust.org/wp-content/uploads/2017/12/AGREE-II-Users-Manual-and-23-item-Instrument-2009-Update-2017.pdf

Ahmadzai, N., Newberry, S. J., Maglione, M. A., Tsertsvadze, A., Ansari, M. T., Hempel, S., . . . Shanman, R. (2013). A surveillance system to assess the need for updating systematic reviews. *Systematic Revie*ws, *2*, 1-4. doi: 10.1186/2046-4053-2-104

Alderson, L. J., Alderson, P., & Tan, T. (2014). Median life span of a cohort of National Institute for Health and Care Excellence clinical guidelines was about 60 months. *Journal of Clinical Epidemiology*, *67*, 52-55. doi: 10.1016/j.jclinepi.2013.07.012

Beauchamp, S., Drapeau, M., & Dionne, C. (2015). The development of practice guidelines in the social and human sciences. *Canadian Psychology/Psychologie canadienne, 56*, 357-367. doi:10.1037/a0038811

Beauchamp, S., Duplantie, J. P., & Mercier, C. (2011). *Development of practice guidelines in social care: A position of the INESSS social care practice guideline committee* [Policy statement]. Retrieved from Regroupement provinciale des comités usagers website: http://www.rpcu.qc.ca/pdf/english/INESSS_Practice_Guidelines_Social_Care.pdf

Bernard, H., Burlingame, G., Flores, P., Greene, L., Joyce, A., Kobos, J. C., . . . McEneaney, A. M. S. (2008). Clinical practice guidelines for group psychotherapy. *International Journal of Group Psychotherapy, 58,* 455-542. doi:10.1521/ijgp.2008.58.4.455

Brouwers, M. C., Kho, M. E., Browman, G. P., Burgers, J. S., Cluzeau, F., Feder, G., . . . The AGREE Next Step Consoritum (2010). AGREE II: Advancing guideline development, reporting and evaluation in health care. *Canadian Medical Association Journal, 182*, E839-E842. doi:10.1503/cmaj.090449

Eccles, M. P., Grimshaw, J. M., Shekelle, P., Schünemann, H. J., & Woolf, S. (2012). Developing clinical practice guidelines: Target audiences, identifying topics for guidelines, guideline group composition and functioning and conflicts of interest. *Implementation Science, 7*, 1-8. doi:10.1037/a0035767

Ferrando, P. J. (2008). The impact of social desirability bias on the EPQ-R item scores: An item response theory analysis. *Personality and Individual Differences, 44*, 1784-1794. doi:10.1016/j.paid.2008.02.005

Fervers, B., Burgers, J. S., Voellinger, R., Brouwers, M., Browman, G. P., Graham, I. D., . . . ADAPTE Collaboration*, T. A. (2011). Guideline adaptation: An approach to enhance efficiency in guideline development and improve utilisation. *BMJ Quality & Safety, 20*, 228-236. doi:10.1136/bmjqs.2010.043257

Fretheim, A., Schunemann, H. J., & Oxman, A. D. (2006). Improving the use of research evidence in guideline development: 5. group processes. *Health Resources and Policy Systems, 4*, 17-35. doi:10.1186/1478-4505-4-17

Gordon, J., & Cooper, B. (2010). Talking knowledge — practising knowledge: A critical best practice approach to how social workers understand and use knowledge in practice. *Practice, 22*, 245-257. doi: 10.1080/09503153.2010.495768

Graham, I. D., Mancher, M., Wolman, D. M., Greenfield, S., & Steinberg, E. (2011). *Clinical practice guidelines we can trust*. Washington, DC: Institue of Medicine National Academies Press.

Hollon, S. D., Areán, P. A., Craske, M. G., Crawford, K. A., Kivlahan, D. R., Magnavita, J. J., . . . Bufka, L. F. (2014). Development of clinical practice guidelines. *Annual Review of Clinical Psychology, 10,* 213-241. doi: 10.1146/annurev-clinpsy-050212-185529

Hudson, P., Quinn, K., O'Hanlon, B., & Aranda, S. (2008). Family meetings in palliative care: Multidisciplinary clinical practice guidelines. *BMC Palliative Care, 7,* 1-6. doi: 10.1186/1472-684X-7-12

Kendall, T., Taylor, E., Perez, A., & Taylor, C. (2008). Guidelines: Diagnosis and management of attention-deficit/hyperactivity disorder in children, young people, and adults: Summary of NICE guidance. *British Medical Journal, 337,* 751-753.

Kunz, R., Fretheim, A., Cluzeau, F., Wilt, T. J., Qaseem, A., Lelgemann, M., . . . Coordinating Efforts in, C. G. D. (2012). Guideline group composition and group processes, article 3 in Integrating and coordinating efforts in COPD guideline development - An official ATS/ERS workshop report. *Proceedings of the American Thoractic Society, 9,* 229-233. doi:10.1513/pats.201208-056ST

Morciano, C., Basevi, V., Faralli, C., Hilton Boon, M., Tonon, S., & Taruscio, D. (2016). Policies on conflicts of interest in health care guideline development: A cross-sectional analysis. *PLoS ONE, 11,* e0166485. doi:10.1371/journal.pone.0166485

O'Connor, B. P. (2006). Social Desirability Measures and the Validity of Self-Reports: A Comprehensive Search for Moderated Relationships in Five-Factor Model Space. In A.

Columbus (Ed.), *Advances in psychology research* (pp. 39-73). Hauppauge, NY, US: Nova Science Publishers.

Pilling, S. (2012). *Developing clinical guidelines for adults: Handbook of Evidence-Based Practice in Clinical Psychology*. Hobogen, New Jersey: John Wiley & Sons, Inc.

Richter, L., Garvare, R., & Nyström, M. E. (2017). Reaching beyond the review of research evidence: A qualitative study of decision making during the development of clinical practice guidelines for disease prevention in healthcare. *BMC Health Services Research, 17,* 344-457. doi:10.1186/s12913-017-2277-1

Rutten, G. M., Degen, S., Hendriks, E. J., Braspenning, J. C., Harting, J., & Oostendorp, R. A. (2016). Adherence to clinical practice guidelines for low back pain in physical therapy: Do patients benefit? *Physical Therapy, 90*, 1111-1122. doi: 10.2522/ptj.20090173

Schuh, S., Babl, F. E., Dalziel, S. R., Freedman, S. B., Macias, C. G., Stephens, D., . . . Plint, A. C. (2017). Practice variation in acute bronchiolitis: A pediatric emergency research networks study. *Pediatrics,* e20170842. doi: 10.1542/peds.2017-0842

Shaner, R. (2001). Practical psychotherapy: How practice guidelines can rescue psychotherapy in public systems. *Psychiatric Services, 52*, 1021-1025. doi:10.1176/appi.ps.52.8.1021

Shekelle. (2014). Updating practice guidelines. *Journal of the American Medical Association, 311*, 2072-2074. doi:10.1001/jama.2014.4950

Shekelle, Woolf, S., Grimshaw, J. M., Schünemann, H. J., & Eccles, M. P. (2012). Developing clinical practice guidelines: Reviewing, reporting, and publishing guidelines. *Implementation Science, 7,* 62-67. doi: 10.1186/1748-5908-7-62

Stamoulos, C., Reyes, A., Trepanier, L., & Drapeau, M. (2014). The methodological quality of clinical practice guidelines in psychology: A pilot study of the guidelines developed by

the Ordre des Psychologues du Québec. *Canadian Psychology/Psychologie canadienne, 55*, 177-186. doi:10.1037/a0037287

Trepanier, L., Stamoulos, C., & Reyes, A. (2017). Assessing the quality of the OPQ's guidelines for the evaluation of dyslexia in children using the Appraisal of Guidelines for Research and Evaluation (AGREE) II: A brief report. *Canadian Psychology/Psychologie canadienne, 58,* 211-217. doi:10.1037/cap0000121

Tudor, K. I., Kozina, P. N., & Marušić, A. (2013). Methodological rigour and transparency of clinical practice guidelines developed by neurology professional societies in Croatia. *PLoS ONE, 8*, e69877. doi:10.1371/journal.pone.0069877

van Dijk, M. K., Oosterbaan, D. B., Verbraak, M. J., & van Balkom, A. J. (2013). The effectiveness of adhering to clinical-practice guidelines for anxiety disorders in secondary mental health care: the results of a cohort study in the Netherlands. *Journal of Evaluation in Clinical Practice, 19*, 791-797. doi: 10.1111/j.1365-2753.2012.01851.x

Wang, X., Chen, Y., Yao, L., Zhou, Q., Wu, Q., Estill, J., . . . Norris, S. L. (2017). Declarations and conflicts of interest in WHO guidelines. *Journal of Clinical Epidemiology*. *98,* 1-8. doi: 10.1016 /j.jclinepi.2017.12.021

Watine, J. C., & Bunting, P. S. (2008). Mass colorectal cancer screening: Methodological quality of practice guidelines is not related to their content validity. *Clinical Biochemistry, 41*, 459-466. doi:10.1016/j.clinbiochem.2007.12.020

Wolf, J. S., Hubbard, H., Faraday, M. M., & Forrest, J. B. (2011). Clinical practice guidelines to inform evidence-based clinical practice. *World Journal of Urology, 29*, 303-309. doi: 10.1007/s00345-011-0656-5

Table 1.

Descriptive information on guideline committee members and survey responders

Prof. Orders	Guidelines	Total Committee members	Committee Members Contacted	Survey Responders	Abandoned surveys	Completed surveys
CMQ	Le médecin, la télémédecine et les technologies de l'information et de la communication / The Physician, Telemedicine and Information, and Communications Technologies (2015)	11	10	4	0	4
CMQ	L'évaluation médicale de l'aptitude à conduire un véhicule automobile- Guide d'exercice / Practice Guideline for the Medical Evaluation of Aptitude for Driving an Automobile (2007)	6	4	3	0	3
CMQ & OPQ	Les troubles du spectre de l'autisme - L'évaluation Clinique / Guidelines for Autism Spectrum Disorder - Clinical Evaluation (2012)	8	8	8	4	4
OPQ	Lignes directrices pour l'évaluation de la dyslexie chez les enfants / Guidelines for the Evaluation of Dyslexia in Children (2014).	14	12	7	2	5
OPQ	Lignes directrices pour l'évaluation du retard mental / Guidelines for Mental Retardation Assessment (2007)	6	6	2	1	1
OPQ	Lignes directrices pour l'évaluation d'un enfant en vue d'une demande de derogation à l'âge d'admission à l'école / Guidelines for the assessment of a child in connection with a request for derogation to the age of school admission (2006)	11	7	1	1	0

OCCOQ	Guide de pratique – Orientation en formation générale des jeunes / Guidelines for the Orientation and General Formation of Young People (2010)	7	3	2	0	2
OTSTCFQ	Décider de l'utilisation des mesures de contention et d'isolement dans le cadre de la Loi sur les services de santé et des services sociaux pour les autochtones cris / Decision on the Utilization of Restraint and Isolation Method Within the Law of Health and Social Services and Indigenous Populations (2011)	6	5	2	0	2
OTSTCFQ	Évaluer une personne ayant un trouble mental ou neuropsychologique attesté par un diagnostic par une évaluation effectuée par un professionnel habilité / Evaluating a Person with a Mental or Neuropsychological Disorder as Attested by a Trained Professional (2011)	7	5	3	0	3
OTSTCFQ	Déterminer un plan d'intervention pour une personne atteinte d'un trouble mental ou présentant un risque suicidaire qui est hébergée dans une installation d'un établissement qui exploite un centre de réadaptation pour les jeunes en difficulté d'adaptation (2013)	6	3	1	0	1
OTSTCFQ	Évaluation psychosociale dans le contexte des régimes de protection, du mandat donné en prévision de l'inaptitude et des autres mesures de protection au majeur / (2011)	5	3	1	0	1
OPPQ	Évaluer un enfant qui n'est pas encore admissible à l'éducation préscolaire et qui présente des indices de retard de développement dans le but de déterminer des services de réadaptation et d'adaptation répondant à ses besoins / Evaluating a Child Not Yet Admissible to Pre-School but Who Presents with Developmental Delays to Determine the Needs	8	4	1	0	1

	for Therapeutic and Guide Intergration Services (2011)					
OPPQ	Lignes directrices sur l'évaluation d'un adolescent dans le cadre d'une décision du tribunal en application de la Loi sur le système de justice pénale pour les adolescents / Guidelines for the Evaluation of an Adolescent in Relation to a Tribunal Decision when Applying the Law Respecting the Penal Justice System for Adolescents (2014)	6	5	2	0	2
OPPQ	Lignes directrices sur l'évaluation psychoéducative de la personne en difficulté d'adaptation / Practice Guidelines for Persons with Adaptation Difficulties (2014)	4	4	3	0	3
OPPQ	Lignes directrices sur l'évaluation du retard de développement / Guidelines Mental Retardation Assessment (2013)	6	5	5	0	5
OEQ	Lignes directrices sur le plan d'intervention pour une personne hébergée en centre jeunesse / Practice Guidelines to Plan an Interventions for Persons Residing in Youth Centres (2014)	5	5	1	0	1
OEQ	Interventions relatives à l'utilisation d'un véhicule routier / Interventions for the utilisation of a vehicle (2008)	9	7	2	0	2
Total		**125**	**96**	**48**	**8**	**40**

Note. CMQ: College of Physicians of Quebec; OPQ: College of psychologists of Quebec; OTSTCFQ: College of Social Workers and Family Therapists of Quebec; OPPOQ: College of Psychoeducators of Quebec; OCCOQ: College of Guidance Counsellors of Quebec; OEQ: College of Occupational Therapists of Quebec.

Table 2.

Demographic Information (N=40)

Variable	*N* (%)
Female gender	26 (65%)
Highest diploma earned	
Bachelors	5 (12.5%)
Masters	14 (35%)
PhD	10 (25%)
Post-Doctoral	10 (25%)
Missing*	1 (2%)
Participants by Order	
CMQ	8 (20%)
OCCOQ	2 (5%)
OEQ	2 (5%)
OPQ	9 (22.5%)
OPPQ	11 (27.5%)
OTSTCFQ	8 (20%)
Role on committee	
Clinician	26 (65%)
Researcher	6 (15%)
Coordinator/Administrator	5 (12.5%)
Group Representative	3 (7.5%)
Principal occupation	
Administrator/Manager	3 (7.5%)
Counsellor	2 (5%)

Lawyer	1 (2.5%)
Neuropsychologist	5 (12.5%)
Occupational Therapist	2 (5%)
Paediatrician	1 (2.5%)
Physician	4 (10%)
Professor	1 (2.5%)
Psychoeducator – Clinician	7 (17.5%)
Psychoeducator – Researcher	1 (2.5%)
Psychologist – Clinician	3 (7.5%)
Psychologist – Researcher	1 (2.5.%)
Researcher	1 (2.5%)
Social Worker	6 (15%)
“Specialist” on the topic of the guideline	2 (5%)
Principal place of work	
Collège des médecins du Québec (CMQ)	8 (20%)
Government Agency	10 (25%)
Hospital	10 (25%)
Private Practice	4 (10%)
School or Cegep	1 (2.5%)
University	7 (17.5%)

Note. *Response to question left blank.

Table 3.

Means and frequencies for each item, and combined scores for 'Strongly Agree' *and* 'Agree' *responses.*

Survey Questions	*n* responses	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree	*Mean*	*Std. Dev.*	Strongly Agree and Agree *n* (%)
Q.1 I was familiar with the research on the topic of the guideline.	40	3	2	2	13	20	4.13	1.20	33 (83%)
Q.2 I was familiar with clinical practice related to the topic of the guideline.	40	3	1	0	8	28	4.43	1.15	36 (90%)
Q.3 I was familiar with policies and regulations related to the topic of the guideline.	40	1	4	3	7	25	4.28	1.13	32 (80%)
Q.4 My expertise in the topic of the guideline was fully utilized by the guideline development committee when the guidelines were being developed.	39	2	1	2	9	25	4.38	1.07	34 (85%)
Q.5 The rules and procedures that guided the work of the guideline development committee were clear to me throughout the process.	39	2	4	5	13	15	3.90	1.19	28 (72%)
Q.6. The guideline was developed using 'best practices in guideline development'.	36	3	2	3	17	11	3.86	1.17	28 (70%)
Q.7 I was encouraged to contribute to the discussion when the guidelines were being developed.	40	1	0	1	11	27	4.58	0.78	38 (95%)
Q.8 The guideline development committee considered my comments and suggestions fairly.	40	2	3	1	14	20	4.18	1.13	34 (85%)
Q.9 All members of the guideline development committee had an opportunity to express their ideas and opinions and I had an equal voice.	38	1	2	3	12	20	4.26	1.00	32 (84%)

Q.10 There was room for disagreement in the guideline development committee.	40	2	2	2	13	21	4.23	1.10	34 (85%)
Q.11 There were discussions that resulted in conflicts between members of the guideline development committee.	37	12	13	5	3	4	2.30	1.31	7 (19%)
Q.12 There was a transparent system or process put in place to reconcile disagreements between committee members on the guideline development committee as to how the scientific data should be interpreted.	31	5	4	8	7	7	3.23	1.38	14 (45%)
Q.13 There was a transparent system or process put in place to reconcile disagreements between committee members on the guideline development committee about the content of the *recommendations* to be published in this guideline.	31	5	4	10	5	7	3.16	1.37	12 (39%)
Q.14 There were conflicts between members of the guideline development committee that remained unresolved.	37	19	9	3	4	2	1.95	1.25	6 (16%)
Q.15 *Informal consensus* (e.g., by simple discussion) was the method used to make final decisions on the recommendations contained in this guideline.	39	3	3	4	17	12	3.82	1.19	29 (75%)
Q.16 *Formal consensus* (e.g. using methods like the Delphi or Glaser techniques) was the method used to formulate the final recommendations in this guideline.	31	13	13	3	1	1	1.84	0.97	2 (7%)
Q.17 A procedure (often referred to as a "minority report") was in place to document, report and make available to those who so request it, the opinions of committee members that are not congruent with the final content of the	27	14	8	3	2	0	1.74	0.94	2 (8%)

guideline.									
Q.18 The guideline recommendations were based on the highest quality scientific evidence when the guideline was produced.	38	3	4	5	7	19	3.92	1.34	26 (68%)
Q.19 The guideline should be updated at this point in time to reflect newer findings.	36	7	6	5	12	6	3.11	1.41	18 (50%)
Q.20 The financial resources available for the development of this guideline seemed sufficient to accomplish the task satisfactorily.	31	1	4	4	11	11	3.87	1.15	22 (71%)
Q.21 The intended purpose of this guideline was to protect the public.	39	2	1	6	14	16	4.05	1.07	30 (77%)
Q.22 The intended purpose of these guidelines was to make suggestions to guide and inform professionals.	39	2	2	0	9	26	4.41	1.09	35 (90%)
Q.23 The intended purpose of these guidelines was to make suggestions to guide and inform professionals.	39	1	4	6	11	17	4.00	1.12	28 (72%)
Q.24 The intended purpose of these guidelines was to outline and provide recommendations that are evidence-based.	37	2	5	5	11	14	3.81	1.24	25 (68%)
Q.25 The barriers (such as cost, professional resources needed) to implementing the recommendations in the guideline, were considered during development.	33	2	8	9	12	2	3.12	1.05	14 (25%)
Q.26 The risk of conflicts of interest for members of the guideline development committee was evaluated and reported.	31	3	9	2	9	8	3.32	1.40	17 (55%)
Q.27 A comprehensive plan was established to guide a systematic search for evidence, such as defining the terms used, the sources consulted	32	8	10	4	6	4	2.63	1.39	10 (32%)

(e.g. electronic databases, hand-searching journals, reviewing conference proceedings, and other guidelines), etc.									
Q.28 The plan that guided the systematic search for evidence was adhered to.	29	6	5	8	6	4	2.90	1.35	10 (35%)
Q.29 Explicit criteria for including/excluding studies and evidence identified by the search were established and clear.	33	8	10	5	6	4	2.64	1.37	10 (30%)
Q.30 The strengths and limitations of the studies retrieved were assessed by examining their design, the methodology.	34	8	10	5	6	5	2.71	1.40	11 (33%)
Q.31 The consistency of the evidence across studies and the direction of results across studies was assessed when developing the guideline.	35	8	7	7	8	5	2.86	1.40	13 (37%)
Q.32 The applicability of the scientific evidence to practice and context was assessed when developing the guideline.	33	7	6	5	9	6	3.03	1.45	15 (46%)
Q.33 Health benefits, side effects, and risks were considered when formulating the recommendations.	34	4	3	9	9	9	3.47	1.31	18 (53%)
Q.34 The quality of the guideline development process serves as an excellent example for future guideline development.	39	5	4	7	15	8	3.44	1.29	23 (59%)
Q.35 I endorse this guideline.	38	5	2	3	13	15	3.82	1.37	28 (74%)
To what extent are you familiar with the following instruments for guideline development? (1 – unfamiliar to 5 extremely familiar) with:									
Q.36. AGREE I	35	31	1	0	1	2	1.34	1.06	3 (9%)

Q.37 AGREE II	35	32	1	0	1	1	1.23	0.84	2 (6%)
Q.38 AGREE III	36	33	1	0	1	1	1.22	0.83	2 (6%)
Q.39 GRADE	34	30	1	1	0	2	1.32	1.01	2 (6%)
Q.40 SIGN50	37	35	2	0	0	0	1.05	0.23	0 (0%)
Q.41 NICE	35	26	4	1	1	3	1.60	1.24	4 (11%)
Q.42 NGN	38	36	2	0	0	0	1.05	0.23	0 (0%)
Q.43 INESSS 2012 and 2015	34	23	3	3	2	3	1.79	1.34	5 (15%)

Chapter 6

General Discussion

Summary of Main Findings

The aim of the first study was to evaluate five guidelines from the Ordre des psychologues du Québec (OPQ): the *Guidelines for the Evaluation of Dyslexia in Children* (2014), the *Guidelines for Autism Spectrum Disorder - Clinical Evaluation* (2012), the *Guidelines for the Assessment of Mental Retardation* (2007), the *Guidelines for the Assessment of a Child in Connection with a Request for Derogation to the Age of School Admission* (2006), and the *Guidelines for Expert Assessment Concerning Child Custody and Access Rights* (2006), to assess their quality. These guidelines were selected because they represented all of the available guidelines offered by the OPQ at the time of this study, and are intended to guide clinical evaluations and treatments for specific disorders. Since their purpose is to provide empirically supported guidance in the areas of assessment, diagnosis, general functioning, treatment, and clinical decision-making, psychologists often assume that they are of good quality, with scientifically supported recommendations that are up-to-date. However, there have been no impact studies or follow-up studies to evaluate how these guidelines have been received by psychologists. Therefore, the aim of this study was to evaluate the quality of these guidelines using the Appraisal of Guidelines for Research and Evaluation II (AGREE II) guideline evaluation instrument.

Our results from Study 1 showed that although there were some modest improvements in the quality of the OPQ guidelines over time, we identified some important methodological problems in all five CPGs, and suggested areas of improvement for future updates. Common to all of the guidelines in this study was the need for greater methodological rigour and transparency of the developmental process.

There was very little information in most of the guidelines as to how the recommendations were selected and few recommendations provided citations, as required by established standards (Brouwers et al., 2010; García et al., 2014; WHO, 2010, 2012). For example, more information is required pertaining to the search methods employed as well as the criteria for selecting the evidence on which the recommendations are based. We also found that not all stakeholders, including those from related disciplines, (e.g. special educators, integration aids, social workers, and patient groups etc.) were invited to provide their input into the final product. We recommended that developers give more opportunities to all stakeholders to provide their input during the development process and ensure that the feedback provided is reflected in the final version of the guideline. This is supported by other studies that have shown the importance of considering patient preferences during treatment planning to improve the rate of adherence to the recommendations provided (Boivin et al., 2009; Légaré et al., 2011). Finally, guidelines should be updated every two to five years (Shekelle, 2014). The guidelines in this study ranged from five to 13 years old. We proposed that the above be addressed in future guideline updates.

The objective of the second study was to examine the research expertise of the OPQ guideline development committee members, on the topic of the guideline. To accomplish this, we analyzed the peer-reviewed publication productivity of committee members using PsycINFO and MEDLINE, and retrieved their *h*-index from Scopus and from Web of Knowledge. Overall, we found that there were at least some researchers on most of the guideline committees, which is congruent with recommendations for guideline development (Fretheim, Schunemann, & Oxman, 2006; Kunz et al., 2012; WHO, 2012). However, our results also showed that there was a clear imbalance between clinical and research expertise, with only a small percentage of researchers represented on these committees. Furthermore, with some notable exceptions, the research productivity of the published committee members was generally low, with the total number of

lifetime publications ranging between one and 143 (M = 30.94, SD = 44.65). Likewise, we found that the research impact of these committee members appeared to be low as compared to the mean national productivity scores for Canadian psychology professors (Carleton, Parkerson, & Horswill, 2012).

Study 2 also revealed that few committee members were research experts on the topic of the guideline. An examination of individual guidelines revealed that the *Guidelines for the Evaluation of Dyslexia in Children,* had the highest number of published committee members, with a total of nine members who had published at least one paper in peer-reviewed journals, yet only three had published on a topic related to the guideline. The *Guidelines for Autism Spectrum Disorder - Clinical Evaluation* had only one committee member of eight (12.5%) who was a productive researcher in the field as compared to the national average of full psychology professors in Canadian universities (Carleton et al., 2012). Likewise, the *Guidelines for the Assessment of Mental Retardation* had only one committee member (10%) who was a productive researcher on the topic of the guideline. Unfortunately, the *Guidelines for the Assessment of a Child in Connection with a Request for Derogation to the Age of School Admission* did not have any committee members with peer-reviewed publications, nor did the guideline indicate whether a person adept at conducting systematic reviews was on the committee or provided services as a consultant. Furthermore, the description of professional credentials, which was limited to one sentence placed at the top of the list of committee members stated ‘Psychologists and Experts in the field’ (French to English translation), which was not in keeping with best practices in guideline development (Brouwers et al., 2010). Therefore, most of the above guidelines had a poor representation of expert researchers in the field of the guideline they were tasked to develop. Finally, the *Guidelines for Expert Assessment Concerning Child Custody and Access Rights* did not list the names of those who participated in the development of these guidelines. Instead, the

authors explain that the content of the guideline was "inspired by" a document published by the Association of Family and Conciliation Courts (AFCC), on in keeping with guideline development best practices (Schünemann et al., 2014; SIGN50, 2015; WHO, 2012). We recommended that future updates and revisions of this guideline include a systematic review of the literature, and apply the ADAPTE Collaboration's tool kit designed to guide guideline adaptations to a local context (ADAPTE-Collaboration, 2009).

It is interesting to note that the guideline that performed the best on the AGREE II instrument in Domain 3 "Rigour of Development," was the Guidelines for Dyslexia, which had the most researchers who were topic experts. This appears to support the assertion that topic experts who are well published on the topic of the guideline, are indispensable contributors to development committees.

The third study sought the views and experience of 40 CPG development committee members who worked on one of the 17 guidelines published by six Quebec regulatory bodies in the social sciences: the *Collège des médecins du Québec* (CMQ: College of Physicians of Quebec); the *Ordre des psychologues du Québec* (OPQ; College of psychologists of Quebec); the *Ordre des travailleurs sociaux et des thérapeutes conjugaux et familiaux du Québec* (OTSTCFQ: College of Social Workers and Family Therapists of Quebec); the *Ordre des psychoéducateurs et psychoéducatrices du Québec* (OPPOQ: College of Psychoeducators of Quebec); the *Ordre des conseillers et conseillères d'orientation du Québec* (OCCOQ: College of Guidance Counsellors of Quebec); and the *Ordre des ergothérapeutes du Québec* (OEQ: College of Occupational Therapists of Quebec). There were a relatively high number of abandoned surveys (8 of 48), all of which were four of the five guidelines that belonged to the OPQ. The reason for this was unclear.

In study 3, we examined committee members' knowledge about guideline development methodology and solicited their views on the quality of the guidelines they developed. Developers reported that they did not have a clear procedure to follow, or a comprehensive plan to guide the search and selection of the research evidence on which the recommendations were to be based. It was interesting to find that 70% respondents felt that their guideline was developed using 'best practices' in guideline development. However the reported absence of a comprehensive plan for the development process and literature search indicates that best practices were likely not used (e .g., Graham et al., 2011). Therefore it is unclear what the respondents considered as 'best-practices' or 'evidence-based.' Furthermore, most respondents reported that there were no formal research methods that guided their search of the literature and selection criteria (e.g., inclusion and exclusion criteria). This is noteworthy since well-defined research methods are an important component of the development process for trustworthy guidelines (Graham et al., 2001). We also found that researchers where significantly less likely than clinicians to agree with the statement "strengths and limitations of the studies were assessed by examining their design and methodology." A reasonable conjecture for this finding is that researchers had more stringent criteria for what constituted a thorough examination of a study's quality.

The above is congruent with the low percentage of respondents who reported familiarity with guideline development manuals and tools (0% -15%). This was noteworthy since development manuals and tools were available to developers at the time the guidelines were being developed (AGREE, 2003, SIGN50, 2015). Furthermore, expertise on the topic of the guideline does not necessarily translate into expertise on knowledge transfer and guideline development. Taken together, these factors supports the notion that there is a need for at least

one committee member who is knowledgeable on guideline development methods to ensure that best practices in guideline development are being followed.

Study 3 also showed that the procedures for dealing with divergent views during the development process were vague or lacking, and conflicts of interest were inadequately reported. A relatively small set of respondents (19%) reported that there were discussions that resulted in disagreements, while six of the 40 (15%) reported that they felt the disagreements remained unresolved. Disagreement between committee members can be expected. However, they must be documented, since a person's participation on a committee is publically available information, and any disagreement with a recommendation found in a guideline should also be duly stated in the document (Beauchamp, Drapeau, & Dionne, 2015; Beauchamp, Duplantie, & Mercier, 2011; Fretheim et al., 2006; Kunz et al., 2012). It is also good to document divergent views for the benefit of the user who is presumed to trust the contents of the guidelines, while using their own professional judgment. This is typically documented in what is called a 'minority report,' for which only two (7%) respondents agreed that one was maintained by their committee. It is important to keep a record of this since it makes the development process more transparent, and documents unresolved disagreements between committee members.

When asked about the current applicability of the guidelines, a high number of respondents, 18 (45%), agreed that their guidelines should be updated at this point. Indeed, development experts have reported that guidelines should be updated every three to five years (Ahmadzai et al., 2013; Alderson, Alderson, & Tan, 2014; Shekelle, 2014). Given that this survey was conducted in early 2018, and 11 of the 17 guidelines were published in 2013 or before, the majority of these guidelines should be refreshed.

Finally, given that there were reported disagreements among committee members, our research team was very interested to know if the respondents endorsed the guideline they

developed. Only 28 (74%) respondents reported that they endorsed their guidelines, and seven (17.5%) reported that they did not endorse their guideline; five (12.5%) did not respond to this question. It is worth noting that three individuals who responded in the negative stated that they were dissatisfied with the how the development process unfolded, how conflicts were managed, and with the committee's leadership. Study 3 provided new insights on the inner working of guideline development groups in the social sciences.

Implications of Findings and Direction for Future Research

We examined the guidelines developed by the OPQ from three difference perspectives: 1) three independent raters assessed the guidelines' quality using the AGREE II instrument, a 'gold standard' assessment tool used internationally by guideline developers; 2) we examined the expertise of the researchers, based on topic-related publications, for each of the guideline development committees members; and 3) we solicited the views and experiences of those who participated on guideline development committees. For the latter, we expanded our sample to include five other Quebec regulatory bodies (Colleges) to increase power due to the low number of total committee members who developed the OPQ guidelines and who responded to our survey. In total nine members from the OPQ responded to the survey.

A critical factor for a guideline's trustworthiness is the validity of the recommendations it provides. Study 1 showed that the scores for the 'Rigour of Development' domain were generally poor among all five of the OPQ guidelines. Future guideline development projects and updates should begin with a clear development plan and employ a systematic search method (e.g., keywords, databases utilized, the timeframe of the search, and the inclusion and exclusion criteria), as well as the criteria for selecting the evidence on which to base the recommendations. If no relevant empirical studies were available due to the nature of the guideline or the novelty of the search topic, a statement to that effect should be provided in the guideline. Each guideline

recommendation ought to cite the supporting evidence within the body of the text. The findings from the second study showed that the guidelines developed by more than one researcher identified as a topic expert, appeared to be developed with more rigour and earned higher quality scores (study 1). Study 3 showed that this problem is not specific to the OPQ, since a high percentage of committee members acknowledged that decision-making methods to establish guideline recommendations were generally informal and made via group discussions. The researchers were more inclined than the clinicians to reply in the negative when responding to whether rigorous development methods were followed. The results from the Study 1 and 2, demonstrate that there is a need for the OPQ to expand the development committees to be composed of all stakeholders (e.g., relevant professionals and representatives of affected populations).

The importance of clinicians has not been overlooked in this study. We acknowledge that clinicians are in the best position to ensure that guidelines are applicable for psychologists who work in a clinical setting. Using the guideline development tools that are currently available to developers, clinicians are needed to contribute their knowledge and skills in order to improve the applicability of the OPQ guidelines in future development projects and updates. As we have presented in the first study, the 'Applicability' scores were one of the lowest of the AGREE II Domain scores. For example, it was found that developers must carefully consider the types of barriers faced by guidelines users who work in diverse settings, often with limited resources, and may serve a clientele who have limited financial means. Such barriers ought to be addressed in a guideline. Further, given the high cost of producing guidelines, future research should include post-implementation monitoring to see if and how psychologists utilize the guidelines, identify the cost-benefit to the public system, and most importantly, to learn whether or not they contributed to improved patient outcomes.

We also found that four out of the five (75%) OPQ guidelines should be reviewed, and in total, 11 of the 17 guidelines (65%) should be reviewed (Study 3). This is an important finding since it suggests that many social service and health professional may not be using up-to-date guidelines. We recommend that each of the older guidelines undergo a new systematic review of the literature bring their recommendations up to date.

Another AGREE quality domain that did not have high scores overall was Domain 6 (Editorial Independence) (Study 1). Most guidelines did not explicitly state whether or not the funding body had an influence on the recommendations, as required by conventional guideline development methods (Graham, et al. 2011, WHO, 2012). To implement these findings, future updates of all guidelines, should expressly state all COIs for members on the development committees.

Limitations

A number of limitations should be considered when interpreting the findings of the present research. For example, in Study 1, the AGREE Trust recommends that the guidelines be distributed to evaluators by an administrator via the AGREEtrust.org website. The evaluators are typically blind to one another and do not meet at any time to discuss their scores. According to the Agreetrust.org website, it is common for guideline quality score to vary widely between evaluators for various reasons, including differences in the interpretation of the criteria, as well as differences in the interpretation of the guideline. It is the role of the administrator to compile the results and take note of all the scores and comments. In an effort to maintain the integrity and consistency of the evaluation, our research team held a consensus meeting after each independent rating, to discuss the item scores and to ensure that each was deemed a fair reflection of the AGREE item scoring criteria. It is for this reason that we presented both the *mean* item scores

obtained independently, and the consensus scores, though little variation between the two was found.

While Study 2 highlighted the insufficient inclusion of highly productive topic-related researchers, we also acknowledged that publication productivity is not the only measure of a committee member's knowledge and contribution value. Many crucial insights in the assessment and care of clients are gained from strong clinical experience with a particular patient population, or from the experience of being the service recipient. Furthermore, the application of PsycINFO and MEDLINE excludes valuable, non-peer reviewed works, such as abstracts, chapters, and books. We also appreciate that there are some known disadvantages to the *h*-index. Books, chapters, conference proceedings, and other non-peer reviewed material were excluded from the publication counts, but could not be excluded from the *h*-indices, since the algorithm automatically captures all academic activities regardless of the type or topic (Sekercioglu, 2008; Zhang, 2009). Therefore, awareness and caution is to be used when referring to these numbers.

Additionally, our knowledge of the guideline committee members was limited to publically available information. Although Curriculum Vitaes are generally publically posted by university professors, they typically only highlight their most important work and publications, and clinicians generally do not have their complete academic and clinical experience publically posted. This made quantifying clinical experience a highly complex task that was beyond the scope of this study. Therefore, it was decided that the data for Study 2 would include only data that could be objectively and publically accessed and quantified.

A weakness of Study 3 was its retrospective design and the potential for inaccurate recall by the participants. Some guidelines were produced five years before this investigation started, while others were produced up to 12 years prior. We also obtained group imbalances between clinicians and researchers, which greatly limited our statistical comparisons. It was unfortunate

that so few of those who identified themselves as researchers responded to our survey. However, Study 2 has also demonstrated that some of the practice guidelines in this study, such as the ones developed by the OPQ, were mostly developed by clinicians; this may also be the case for guidelines developed by other Orders.

Conclusion

Good quality guidelines have been shown to be beneficial for many health and psychological needs in the areas of evaluation, diagnosis, and treatment planning (Köhler, Hoffmann, Unger, Steinacher, & Fydrich, 2012; Martínez, Reyes, Lorenzo, & Menéndez, 2009; Middleton, Kalogeropoulos, Middleton, & Drapeau, 2018; van Dijk, Oosterbaan, Verbraak, & van Balkom, 2013). However, some health professionals have reported that they do not trust the guidelines that were produced for them; many have expressed concern surrounding their quality (Ansari & Rashidian, 2012; Bindslev, Schroll, Gotzsche, & Lundh, 2013; Kalies et al., 2017; Lenzer, 2013; Stamoulos, Reyes, Trepanier, & Drapeau, 2014; Trepanier, Stamoulos, & Reyes, 2017).

As treatments become more varied and complex over time, it is increasingly difficult to stay abreast of the new research and treatment approaches unless clinicians spend vast amounts of time pouring over the research in their field every day (Davidoff, Haynes, Sackett, & Smith, 1995). This is not always feasible, but clinicians are nevertheless expected to maintain their skills and their competencies. Good quality practice guidelines, that are up to date, offer clinicians scientifically supported interventions that are recommended by experts in the field, and ought to save practitioners time, with the ultimate goal of benefitting end-users (Brouwers, et al. 2010; Freitheim et al., 2006). Therefore, it is important for guideline producers to maintain high standards of development practices in order to earn and maintain the trust of clinicians who use the practice guidelines.

The research presented in Studies 1, 2, and 3, points to areas that must improve, specifically in the areas of planning, searching, developing, reporting, dissemination, and follow-up studies to measure the impact of a practice guideline. The thoughtful planning of group composition, stakeholder involvement, and a well-document external review process, are important to developing applicable and trustworthy guidelines. The present thesis attempted to provide concrete suggestions to guide developers, and support the development of quality clinical practice guidelines for psychologists to have at their fingertips.

References

AGREE Collaboration. (2003). Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: The AGREE project. *Quality & Safety in Health Care, 12,* 18-23. doi: 10.1136/qhc.12.1.18

AGREE Next Steps Consortium. (2017). *The AGREE II instrument.* [Electronic version]. Retrieved from https://www.agreetrust.org/wp-content/uploads/2017/12/AGREE-II-Users-Manual-and-23-item-Instrument-2009-Update-2017.pdf

Ahmadzai, N., Newberry, S. J., Maglione, M. A., Tsertsvadze, A., Ansari, M. T., Hempel, S., . . . Shanman, R. (2013). A surveillance system to assess the need for updating systematic reviews. *Systematic Revie*ws, *2*, 1-4. doi: 10.1186/2046-4053-2-104

Alderson, L. J., Alderson, P., & Tan, T. (2014). Median life span of a cohort of National Institute for Health and Care Excellence clinical guidelines was about 60 months. *Journal of Clinical Epidemiology*, *67*, 52-55. doi: 10.1016/j.jclinepi.2013.07.012

Ansari, S., & Rashidian, A. (2012). Guidelines for guidelines: Are they up to the task? A comparative assessment of clinical practice guideline development handbooks. *PLoS ONE, 7*, e49864. doi:10.1371/journal.pone.0049864

Beauchamp, S., Drapeau, M., & Dionne, C. (2015). The development of practice guidelines in the social and human sciences. *Canadian Psychology/Psychologie canadienne, 56,* 357-367. doi:10.1037/a0038811

Beauchamp, S., Duplantie, J. P., & Mercier, C. (2011). *Development of practice guidelines in social care: A position of the INESSS social care practice guideline committee* [Policy statement]. Retrieved from Regroupement provinciale des comités usagers website: http://www.rpcu.qc.ca/pdf/english/INESSS_Practice_Guidelines_Social_Care.pdf

Bindslev, J. B. B., Schroll, J., Gotzsche, P. C., & Lundh, A. (2013). Underreporting of conflicts of interest in clinical practice guidelines: cross sectional study. *BMC Medical Ethics, 14*, 19-24. doi: 10.1186/1472-6939-14-19

Boivin, A., Green, J., van der Meulen, J., Légaré, F., & Nolte, E. (2009). Why consider patients' preferences?: A discourse analysis of clinical practice guideline developers. *Medical care, 47*, 908-915. doi:10.1097/MLR.0b013e3181a81158

Brouwers, M. C., Kho, M. E., Browman, G. P., Burgers, J. S., Cluzeau, F., Feder, G., . . . The AGREE Next Step Consoritum (2010). AGREE II: Advancing guideline development, reporting and evaluation in health care. *Canadian Medical Association Journal, 182*, E839-E842. doi:10.1503/cmaj.090449

Carleton, R. N., Parkerson, H. A., & Horswill, S. C. (2012). Assessing the publication productivity of clinical psychology professors in Canadian Psychological Association-accredited Canadian psychology departments. *Canadian Psychology/Psychologie canadienne, 53*, 226-237. doi:10.1037/a0027731

Davidoff, F., Haynes, B., Sackett, D., & Smith, R. (1995). Evidence based medicine. *British Medical Journal, 310*, 1085-1093.

Fretheim, A., Schunemann, H. J., & Oxman, A. D. (2006). Improving the use of research evidence in guideline development: 5. group processes. *Health Resources and Policy Systems, 4*, 17-35. doi:10.1186/1478-4505-4-17

García, L. M., McFarlane, E., Barnes, S., Sanabria, A. J., Alonso-Coello, P., & Alderson, P. (2014). Updated recommendations: An assessment of NICE clinical guidelines. *Implementation Science, 9*, 72-91. doi: 10.1186/1748-5908-9-72

Kalies, H., Schöttmer, R., Simon, S. T., Voltz, R., Crispin, A., & Bausewein, C. (2017). Critical attitudes and beliefs towards guidelines amongst palliative care professionals – Results from a national survey. *BMC Palliative Care, 16*, 20-24. doi: 10.1186/s12904-017-0187-y

Köhler, S., Hoffmann, S., Unger, T., Steinacher, B., & Fydrich, T. (2012). Adherence to guidelines and effectiveness of inpatient treatment for unipolar depression. *International Journal of Psychiatry in Clinical Practice, 16*, 103-112. doi: 10.3109/13651501.2011.638142

Kunz, R., Fretheim, A., Cluzeau, F., Wilt, T. J., Qaseem, A., Lelgemann, M., . . . Coordinating Efforts in, C. G. D. (2012). Guideline group composition and group processes, article 3 in Integrating and coordinating efforts in COPD guideline development - An official ATS/ERS workshop report. *Proceedings of the American Thoractic Society, 9*, 229-233. doi:10.1513/pats.201208-056ST

Légaré, F., Boivin, A., van der Weijden, T., Pakenham, C., Burgers, J., Légaré, J., … Gagnon, S. (2011). Patient and public involvement in clinical practice guidelines: A knowledge synthesis of existing programs. *Medical Decision Making*, *31*, 45–74. doi:10.1177/0272989X11424401

Lenzer, J. (2013). Why we can't trust clinical guidelines. *British Medical Journal, 346*, 1-5. doi: 0.1136/bmj.f3830

Martínez, R., Reyes, S., Lorenzo, M. J., & Menéndez, R. (2009, April). *Impact of guidelines on outcome: The evidence.* Paper presented at the Seminars in Respiratory and Critical Care Medicine, Valencia, Spain. Retrieved from https://www.thiemeconnect.com/products/ejournals/html/10.1055/s-0029-1202936

Middleton, J. C., Kalogeropoulos, C., Middleton, J. A., & Drapeau, M. (2018). Assessing the methodological quality of Canadian Psychiatric Association's anxiety and depression clinical guidelines. *Journal of Evaluation in Clinical Practice, 2,* 1-9. doi:10.1111/jep.13026

Schünemann, H. J., Wiercioch, W., Etxeandia, I., Falavigna, M., Santesso, N., Mustafa, R., . . . al., e. (2014). Guidelines 2.0: Systematic development of a comprehensive checklist for a successful guideline enterprise. *Canadian Medical Association Journal, 186*, E123-E142. doi:10.1503/cmaj.131237

Sekercioglu, C. H. (2008). Quantifying coauthor contributions. *Science, 322*, 371-371. doi: 10.1126/science.322.5900.371a

Shekelle. (2014). Updating practice guidelines. *Journal of the American Medical Association, 311*, 2072-2074. doi:10.1001/jama.2014.4950

Scottish Intercollegiate Guidelines Network (SIGN). (2015). *SIGN50: A guideline developer's handbook* [Electronic version]. Retrieved from https://sign.ac.uk/assets/sign50_2015.pdf

Stamoulos, C., Reyes, A., Trepanier, L., & Drapeau, M. (2014). The methodological quality of clinical practice guidelines in psychology: A pilot study of the guidelines developed by the Ordre des Psychologues du Québec. *Canadian Psychology/Psychologie canadienne, 55*, 177-186. doi: 10.1037/a0037287

The ADAPTE Collaboration (2009). *The ADAPTE process: Resource toolkit for guideline adaptation* (Version 2.0). Retrieved from the Guidelines International Network (G-I-N) website http://wwwg-i-n.net.

Trepanier, L., Stamoulos, C., & Reyes, A. (2017). Assessing the quality of the OPQ's guidelines for the evaluation of dyslexia in children using the Appraisal of Guidelines for Research

and Evaluation (AGREE) II: A brief report. *Canadian Psychology/Psychologie canadienne, 58,* 211-217. doi:10.1037/cap0000121

van Dijk, M. K., Oosterbaan, D. B., Verbraak, M. J., & van Balkom, A. J. (2013). The effectiveness of adhering to clinical-practice guidelines for anxiety disorders in secondary mental health care: the results of a cohort study in the Netherlands. *Journal of Evaluation in Clinical Practice, 19,* 791-797. doi: 10.1111/j.1365-2753.2012.01851.x

World Health Organization (WHO). (2010). *WHO handbook for guideline development.* Geneva, Switzerland: Author. Retrieved from http://www.who.int/hiv/topics/mtct/grc_handbook_mar2010_1.pdf

Wold Health Organization (WHO). (2012). *WHO handbook for guideline development*. Geneva, Switzerland: Author. Retrieved from https://apps.who.int/iris/bitstream/handle/10665/75146/9789241548441_eng.pdf?sequence=1

Zhang, C. (2009). A proposal for calculating weighted citations based on author rank. *EMBO Reports, 10,* 416–417. doi: 10.1038/embor.2009.74

Master Reference List

AGREE Collaboration. (2003). Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: The AGREE project. *Quality & Safety in Health Care, 12,* 18-23. doi: 10.1136/qhc.12.1.18

AGREE Next Steps Consortium. (2017). *The AGREE II instrument.* [Electronic version]. Retrieved from https://www.agreetrust.org/wp-content/uploads/2017/12/AGREE-II-Users-Manual-and-23-item-Instrument-2009-Update-2017.pdf

Ahmadzai, N., Newberry, S. J., Maglione, M. A., Tsertsvadze, A., Ansari, M. T., Hempel, S., . . . Shanman, R. (2013). A surveillance system to assess the need for updating systematic reviews. *Systematic Revie*ws, *2*, 1-4. doi: 10.1186/2046-4053-2-104

Al-Ansary, L. A., Tricco, A. C., Adi, Y., Bawazeer, G., Perrier, L., Al-Ghonaim, M., . . . Straus, S. E. (2013). A systematic review of recent clinical practice guidelines on the diagnosis, assessment and management of hypertension. *PLoS ONE, 8*, e53744. doi:10.1371/journal.pone.0053744

Alderson, L. J., Alderson, P., & Tan, T. (2014). Median life span of a cohort of National Institute for Health and Care Excellence clinical guidelines was about 60 months. *Journal of Clinical Epidemiology*, *67*, 52-55. doi: 10.1016/j.jclinepi.2013.07.012

Alonso-Coello, P., Irfan, A., Solà, I., Gich, I., Delgado-Noguera, M., Rigau, D., . . . Schunemann, H. (2010). The quality of clinical practice guidelines over the last two decades: A systematic review of guideline appraisal studies. *Quality and Safety in Health Care, 19*, e58. doi:10.1136/qshc.2010.042077

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5*. Washington, DC: Author.

Ansari, S., & Rashidian, A. (2012). Guidelines for guidelines: Are they up to the task? A comparative assessment of clinical practice guideline development handbooks. *PLoS ONE, 7,* e49864. doi:10.1371/journal.pone.0049864

Armstrong, M. J., Mullins, C. D., Gronseth, G. S., & Gagliardi, A. R. (2018). Impact of patient involvement on clinical practice guideline development: A parallel group study. *Implementation Science, 13,* 55-59. doi:10.1186/s13012-018-0745-6

Babione, J. M. (2010). Evidence-based practice in psychology: An ethical framework for graduate education, clinical training, and maintaining professional competence. *Ethics & Behavior, 20*, 443-453. doi:10.1080/10508422.2010.521446

Balogh, Z., Caldwell, E., Heetveld, M., D'Amours, S., Schlaphoff, G., Harris, I., & Sugrue, M. (2005). Institutional practice guidelines on management of pelvic fracture-related hemodynamic instability: Do they make a difference? *Journal of Trauma and Acute Care Surgery, 58*, 778-782. doi: 10.1097/01.ta.0000158251.40760.b2

Barriocanal, A. M., Lopez, A., Monreal, M., & Montané, E. (2016). Quality assessment of peripheral artery disease clinical guidelines. *Journal of Vascular Surgery, 63*, 1091-1098. doi:10.1016/j.jvs.2015.12.040

Beauchamp, S., Drapeau, M., & Dionne, C. (2015). The development of practice guidelines in the social and human sciences. *Canadian Psychology/Psychologie canadienne, 56,* 357-367. doi:10.1037/a0038811

Beauchamp, S., Drapeau, M., Dionne, C., & Duplantie, J. P. (2015). *Institut national d'excellence en santé et en services sociaux (INESSS): Cadre d'élaboration des guides de pratique dans le secteur des services sociaux* [Policy statement]. Montreal, Canada: Author. Retrieved from INESSS website:

https://www.inesss.qc.ca/fileadmin/doc/INESSS/Rapports/ServicesSociaux/INESSS_Cadre_elaboration_guides_pratique_servicessociaux.pdf

Beauchamp, S., Duplantie, J. P., & Mercier, C. (2011). *Development of practice guidelines in social care: A position of the INESSS social care practice guideline committee* [Policy statement]. Retrieved from Regroupement provinciale des comités usagers website: http://www.rpcu.qc.ca/pdf/english/INESSS_Practice_Guidelines_Social_Care.pdf

Bennett, K., Courtney, D., Duda, S., Henderson, J., & Szatmari, P. (2018). An appraisal of the trustworthiness of practice guidelines for depression and anxiety in children and youth. *Depression and Anxiety, 35*, 330-340. doi:10.1002/da.22752

Bergman, D. A. (1999). Evidence-based guidelines and critical pathways for quality improvement. *Pediatrics, 103*, 225-232.

Bernard, H., Burlingame, G., Flores, P., Greene, L., Joyce, A., Kobos, J. C., . . . McEneaney, A. M. S. (2008). Clinical practice guidelines for group psychotherapy. *International Journal of Group Psychotherapy, 58*, 455-542. doi:10.1521/ijgp.2008.58.4.455

Berrigan, L., Marshall, S., McCullagh, S., Velikonja, D., & Bayley, M. (2011). Quality of clinical practice guidelines for persons who have sustained mild traumatic brain injury. *Brain Injury, 25*, 742-751. doi:10.3109/02699052.2011.580317

Bindslev, J. B. B., Schroll, J., Gotzsche, P. C., & Lundh, A. (2013). Underreporting of conflicts of interest in clinical practice guidelines: cross sectional study. *BMC Medical Ethics, 14*, 19-24. doi: 10.1186/1472-6939-14-19

Blozik, E., Nothacker, M., Bunk, T., Szecsenyi, J., Ollenschläger, G., & Scherer, M. (2012). Simultaneous development of guidelines and quality indicators – how do guideline groups act? *International Journal of Health Care Quality Assurance, 25*, 712-729. doi:10.1108 /09526861211270659

Blum, D. J. (1996). The evolution of physician credentialing into managed care selective contracting. *American Journal of Law and Medicine, 22*, 173-179.

Boivin, A., Green, J., van der Meulen, J., Légaré, F., & Nolte, E. (2009). Why consider patients' preferences?: A discourse analysis of clinical practice guideline developers. *Medical Care, 47*, 908-915. doi:10.1097/MLR.0b013e3181a81158

Brouwers, M. C., Kho, M. E., Browman, G. P., Burgers, J. S., Cluzeau, F., Feder, G., . . . The AGREE Next Step Consoritum (2010). AGREE II: Advancing guideline development, reporting and evaluation in health care. *Canadian Medical Association Journal, 182*, E839-E842. doi:10.1503/cmaj.090449

Burgers, J. S., Anzueto, A., Black, P. N., Cruz, A. A., Fervers, B., Graham, I. D., . . . Yawn, B. P. (2012). Adaptation, evaluation, and updating of guidelines. *Proceedings of the American Thoracic Society, 9*, 304-310. doi:10.1513/pats.201208-067ST

Burgers, J. S., Cluzeau, F. A., Hanna, S. E., Hunt, C., & Grol, R. (2003). Characteristics of high-quality guidelines. *International Journal of Technology Assessment in Health Care, 19*, 148-157. doi: 10.1017/S026646230300014X

Burgers, J. S., Grol, R., Klazinga, N. S., Makela, M., & Zaat, J. (2003). Towards evidence-based clinical practice: An international survey of 18 clinical guideline programs. *International Journal for Quality in Health Care, 15*, 31-45. doi: 10.1093/intqhc/15.1.31

Cabana, M. D. (1999). Why don't physicians follow clinical practice guidelines?: A framework for improvement. *Journal of the American Medical Association, 282*, 1458-1465. doi: 10.1177/2333794X19847026

Cahill, N. E., & Heyland, D. K. (2010). Bridging the guideline–practice gap in critical care nutrition: A review of guideline implementation studies. *Journal of Parenteral and Enteral Nutrition, 34*, 653-659. doi:10.1177/0148607110361907

Carleton, R. N., Parkerson, H. A., & Horswill, S. C. (2012). Assessing the publication productivity of clinical psychology professors in Canadian Psychological Association-accredited Canadian psychology departments. *Canadian Psychology/Psychologie canadienne, 53,* 226-237. doi:10.1037/a0027731

Chang, S. G., Kim, D. I., Shin, E. S., Jang, J. E., Yeon, J. Y., & Lee, Y. S. (2016). Methodological quality appraisal of 27 Korean guidelines using a scoring guide based on the AGREE II instrument and a web-based evaluation. *Journal of Korean Medical Science, 31,* 682-687. doi:10.3346/jkms.2016.31.5.682

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

Davidoff, F., Haynes, B., Sackett, D., & Smith, R. (1995). Evidence based medicine. *British Medical Journal, 310,* 1085-1093.

Davis, Goldman, J., & Palda, V. (2007). A guide to Canadian Medical Association handbook on clinical practice guidelines. *Canadian Medical Association Journal, 177,* 1221-1226. doi:10.1503/cmaj.070880

Dozois, D. J. A. (2013). Psychological treatments: Putting evidence into practice and practice into evidence. *Canadian Psychology/Psychologie canadienne, 54,* 1-11. doi:10.1037/a0031125

Dozois, D. J. A., Mikail, S. F., Alden, L. E., Bieling, P. J., Bourgon, G., Clark, D., . . . Johnston, C. (2014). The CPA presidential task force on evidence-based practice of psychological treatments. *Canadian Psychology/Psychologie canadienne, 55,* 153-160. doi:10.1037/a0035767

Drapeau, M., & Hunsley, J. (2014). Where's the science? Introduction to a special issue of Canadian Psychology on science in psychology. *Canadian Psychology/Psychologie canadienne, 55*, 145-152. doi: 10.1037/a0037321

Eccles, M. P., Grimshaw, J. M., Shekelle, P., Schünemann, H. J., & Woolf, S. (2012). Developing clinical practice guidelines: Target audiences, identifying topics for guidelines, guideline group composition and functioning and conflicts of interest. *Implementation Science, 7*, 1-8. doi:10.1037/a0035767

Ellen, M., Brown, A., & Cockerill, R. (2009). Do clinical practice guidelines influence length of stay in ontario acute care hospitals? *Healthcare Quarterly, 12*, 64-71. doi:10.1080/10437797.2003.10779115

Ferrando, P. J. (2008). The impact of social desirability bias on the EPQ-R item scores: An item response theory analysis. *Personality and Individual Differences, 44*, 1784-1794. doi:10.1016/j.paid.2008.02.005

Fervers, B., Burgers, J. S., Voellinger, R., Brouwers, M., Browman, G. P., Graham, I. D., . . . ADAPTE Collaboration*, T. A. (2011). Guideline adaptation: An approach to enhance efficiency in guideline development and improve utilisation. *BMJ Quality & Safety, 20*, 228-236. doi:10.1136/bmjqs.2010.043257

Field, M. J., & Lohr, K. N. (Eds.). (1992). Guidelines for clinical practice: From development to use. Washington, DC: Institute of Medicine National Academies Press.

Flexner, A. (1910). *Medical education in the United States and Canada: A report to the Carnegie Foundation for the advancement of teaching* (Bulletin No. 4). New York, NY: The Merrymount Press. Retrieved from http://archive.carnegiefoundation.org/pdfs/elibrary/Carnegie_Flexner_Report.pdf

Fretheim, A., Schunemann, H. J., & Oxman, A. D. (2006). Improving the use of research evidence in guideline development: 5. group processes. *Health Resources and Policy Systems, 4,* 17-35. doi:10.1186/1478-4505-4-17

Gambrill, E. (2003). Evidence-based practice: Sea of change or the emperor's new clothes? *Journal of Social Work Education, 39,* 3-23. doi:10.1080/10437797.2003.10779115

García, L. M., McFarlane, E., Barnes, S., Sanabria, A. J., Alonso-Coello, P., & Alderson, P. (2014). Updated recommendations: An assessment of NICE clinical guidelines. *Implementation Science, 9,* 72-91. doi: 10.1186/1748-5908-9-72

Gordon, J., & Cooper, B. (2010). Talking knowledge — practising knowledge: A critical best practice approach to how social workers understand and use knowledge in practice. *Practice, 22,* 245-257. doi: 10.1080/09503153.2010.495768

Gould, N., & Kendall, T. (2007). Developing the NICE/SCIE guidelines for dementia care: The challenges of enhancing the evidence base for social and health care. *British Journal of Social Work, 37,* 475-490. doi: 1093/bjsw/bcm020

Graham, I. D., Beardall, S., Carter, A., Tetroe, J., & Davies, B. (2003). The state of the science and art of practice guidelines development, dissemination and evaluation in Canada. *Journal of Evaluation in Clinical Practice, 9,* 195-202.

Graham, I. D., Mancher, M., Wolman, D. M., Greenfield, S., & Steinberg, E. (2011). *Clinical practice guidelines we can trust*. Washington, DC: Institue of Medicine National Academies Press.

Graham, I. D., Evans, W. K., Logan, D., O'Connor, A., Palda, V., McAuley, L., . . . Harrison, M. B. (2000). Canadian oncologists and clinical practice guidelines: A national survey of attitudes and reported use. *Oncology, 59,* 283-290. doi:10.1046/j.1365-2753.2003.00385.x

Gray, M., Plath, D., Webb, S. A., & Webb, S. (2009). *Evidence-based social work: A critical stance.* London, England: Routledge.

Grimshaw, J., Eccles, M. P., Lavis, J. N., Hill, S. J., & Squires, J. E. (2012). Knowledge translation of research findings. *Implementation Science, 7*, 50-58. doi: 10.1186/1748-5908-7-50

Grimshaw, J., Freemantle, N., Wallace, S., Russell, I., Hurwitz, B., Watt, I., . . . Sheldon, T. (1995). Developing and implementing clinical practice guidelines. *Quality in Healthcare, 4*, 55-63.

Grol, R. (2001). Successes and failures in the implementation of evidence-based guidelines for clinical practice. *Medical care, 39*, II-46-II-54.

Harrison, M. B., Légaré, F., Graham, I. D., & Fervers, B. (2010). Adapting clinical practice guidelines to local context and assessing barriers to their use. *Canadian Medical Association Journal, 182*, E78-E84. doi: 10.1503/cmaj.081232

Hasenfeld, R., & Shekelle, P. G. (2003). Is the methodological quality of guidelines declining in the US? Comparison of the quality of US Agency for Health Care Policy and Research (AHCPR) guidelines with those published subsequently. *Quality and Safety in Health Care, 12*, 428-434. doi: 10.1136/qhc.12.6.428

Hayward, R. S., Guyatt, G. H., Moore, K.-A., McGibbon, K. A., & Carter, A. O. (1997). Canadian physicians' attitudes about and preferences regarding clinical practice guidelines. *Canadian Medical Association Journal, 156*, 1715-1723.

Hershenberg, R., Drabick, D. A., & Vivian, D. (2012). An opportunity to bridge the gap between clinical research and clinical practice: Implications for clinical training. *Psychotherapy, 49*, 123-131.

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *PNAS, 102*, 16569-16572. doi: 10.1073/pnas.0507655102

Hollon, S. D., Areán, P. A., Craske, M. G., Crawford, K. A., Kivlahan, D. R., Magnavita, J. J., . . . Bufka, L. F. (2014). Development of clinical practice guidelines. *Annual Review of Clinical Psychology, 10*, 213-241. doi: 10.1146/annurev-clinpsy-050212-185529

Hudson, P., Quinn, K., O'Hanlon, B., & Aranda, S. (2008). Family meetings in palliative care: Multidisciplinary clinical practice guidelines. *BMC Palliative Care, 7*, 1-6. doi: 10.1186/1472-684X-7-12

Hussain, T., Michel, G., & Shiffman, R. N. (2009). The Yale guideline recommendation corpus: A representative sample of the knowledge content of guidelines. *International Journal of Medical Informatics, 78*, 354-363. doi: 10.1016/j.ijmedinf.2008.11.001

Ionita, G., & Fitzpatrick, M. (2014). Bringing science to clinical practice: A Canadian survey of psychological practice and usage of progress monitoring measures. *Canadian Psychology/Psychologie canadienne, 55*, 187-196. doi: 10.1037/a0037355

Irani, S., Rashidian, A., Yousefi-Nooraie, R., & Soltani, A. (2011). Evaluating clinical practice guidelines developed for the management of thyroid nodules and thyroid cancers and assessing the reliability and validity of the AGREE instrument. *Journal of Evaluation in Clinical Practice, 17*, 729-736. doi: 10.1111/j.1365-2753.2010.01561.x

Kalies, H., Schöttmer, R., Simon, S. T., Voltz, R., Crispin, A., & Bausewein, C. (2017). Critical attitudes and beliefs towards guidelines amongst palliative care professionals – Results from a national survey. *BMC Palliative Care, 16*, 20-24. doi: 10.1186/s12904-017-0187-y

Kendall, T., Taylor, E., Perez, A., & Taylor, C. (2008). Guidelines: Diagnosis and management of attention-deficit/hyperactivity disorder in children, young people, and adults: Summary of NICE guidance. *British Medical Journal, 337*, 751-753.

doi: 10.1136/bmj.a1239

Köhler, S., Hoffmann, S., Unger, T., Steinacher, B., & Fydrich, T. (2012). Adherence to guidelines and effectiveness of inpatient treatment for unipolar depression. *International Journal of Psychiatry in Clinical Practice, 16,* 103-112. doi: 10.3109/13651501.2011.638142

Kryworuchko, J., Stacey, D., Bai, N., & Graham, I. D. (2009). Twelve years of clinical practice guideline development, dissemination and evaluation in Canada (1994 to 2005). *Implementation Science, 4,* 49-56. doi:10.1186/1748-5908-4-49

Kunz, R., Fretheim, A., Cluzeau, F., Wilt, T. J., Qaseem, A., Lelgemann, M., . . . Coordinating Efforts in, C. G. D. (2012). Guideline group composition and group processes, article 3 in Integrating and coordinating efforts in COPD guideline development - An official ATS/ERS workshop report. *Proceedings of the American Thoractic Society, 9,* 229-233. doi:10.1513/pats.201208-056ST

Légaré, F., Boivin, A., van der Weijden, T., Pakenham, C., Burgers, J., Légaré, J., … Gagnon, S. (2011). Patient and public involvement in clinical practice guidelines: A knowledge synthesis of existing programs. *Medical Decision Making*, *31*, 45–74. doi:10.1177/0272989X11424401

Lenzer, J. (2013). Why we can't trust clinical guidelines. *British Medical Journal, 346,* 1-5. doi: 0.1136/bmj.f3830

Lesho, E. P., Myers, C. P., Ott, M., Winslow, C., & Brown, J. E. (2005). Do clinical practice guidelines improve processes or outcomes in primary care? *Military Medicine, 170,* 243-246. doi: 10.7205/MILMED.170.3.243

MacQueen, G., Santaguida, P., Keshavarz, H., Jaworska, N., Levine, M., Beyene, J., & Raina, P. (2016). Systematic review of clinical practice guidelines for failed antidepressant treatment

response in major depressive disorder, dysthymia, and subthreshold depression in adults. *The Canadian Journal of Psychiatry, 62*, 11–23. doi:10.1177/0706743716664885

Margo, C. E. (2004). Quality care and practice variation: The roles of practice guidelines and public profiles. *Survey of ophthalmology, 49*, 359-371. doi: 10.1016/j.survophthal.2004.02.010

Martínez, R., Reyes, S., Lorenzo, M. J., & Menéndez, R. (2009, April). *Impact of guidelines on outcome: The evidence.* Paper presented at the Seminars in Respiratory and Critical Care Medicine, Valencia, Spain. Retrieved from https://www.thiemeconnect.com/products/ejournals/html/10.1055/s-0029-1202936

Menéndez, R., Ferrando, D., Vallés, J. M., & Vallterra, J. (2002). Influence of deviation from guidelines on the outcome of community-acquired pneumonia. *Chest, 122*, 612-617.

Mental Health Commission of Canada. (2012). *Changing directions, changing lives: The mental health strategy for Canada* [Strategy report]. Calgary, Canada: Author. Retrieved from https://www.mentalhealthcommission.ca/sites/default/files/MHStrategy_Strategy_ENG.pdf

Michie, S., & Johnston, M. (2004). Changing clinical behaviour by making guidelines specific. *British Medical Journal, 328*, 343-345.

Middleton, J. C., Kalogeropoulos, C., Middleton, J. A., & Drapeau, M. (2018). Assessing the methodological quality of Canadian Psychiatric Association's anxiety and depression clinical guidelines. *Journal of Evaluation in Clinical Practice, 2*, 1-9. doi:10.1111/jep.13026

Miville, M. L., Duan, C., Nutt, R. L., Waehler, C. A., Suzuki, L., Pistole, M. C., . . . Corpus, M. (2009). Integrating practice guidelines into professional training: Implications for diversity competence. *The Counseling Psychologist, 37*, 519-563. doi: 10.1177/0011000008323651

Morciano, C., Basevi, V., Faralli, C., Hilton Boon, M., Tonon, S., & Taruscio, D. (2016). Policies on conflicts of interest in health care guideline development: A cross-sectional analysis. *PLoS ONE, 11*, e0166485. doi:10.1371/journal.pone.0166485

Mosavianpour, M., Collet, J.-P., Sarmast, H., & Kissoon, N. (2016). Barriers to the implementation of sepsis guideline in a Canadian pediatric tertiary care centre. *Journal of Nursing Education and Practice, 6*, 34. doi:10.5430/jnep.v6n12p34

Nathan, P. E. (1998). Practice guidelines: Not yet ideal. *American Psychologist, 53*(3), 290. doi: 10.1037/0003-066X.53.3.290

Ng, C. F. (2018). Research productivity and impact of professors in Canadian undergraduate psychology programmes. *Canadian Psychology/Psychologie canadienne, 59*, 211-218. doi: /10.1037/cap0000137

Norris, S. L., Holmer, H. K., Ogden, L. A., & Burda, B. U. (2011). Conflict of interest in clinical practice guideline development: A systematic review. *PLoS ONE, 6*. doi:10.1371/journal .pone.0025153

Norris, S. L., Meerpohl, J. J., Akl, E. A., Schünemann, H. J., Gartlehner, G., Chen, Y., & Whittington, C. (2016). The skills and experience of GRADE methodologists can be assessed with a simple tool. *Journal of Clinical Epidemiology*, *79*, 150-158. doi:10.1016/j.jclinepi.2016.07.001

O'Connor, B. P. (2006). Social Desirability Measures and the Validity of Self-Reports: A Comprehensive Search for Moderated Relationships in Five-Factor Model Space. In A. Columbus (Ed.), *Advances in psychology research* (pp. 39-73). Hauppauge, NY, US: Nova Science Publishers.

Office of Professions of Quebec. (n.d.). *Code of ethics for psychologists* (c. C26, r. 212, a. 5) [Online reference]. Retrieved from www.legisquebec.gouv.qc.ca/

Ordre des psychologues du Québec (OPQ). (2012). *Guidelines for autism spectrum disorder: Clinical evaluation*. Montreal, Canada: Author. Retrieved from https://www.ordrepsy.qc.ca/documents/26707/63191/spectre-de-lautisme/868d8894-44f1-471d-941d-94a5f571395c

Ordre des psychologues du Québec (OPQ). (2006). *Guidelines for the assessment of a child in connection with a request for derogation to the age of school admission*. Montreal, Canada: Author. Retrieved from https://www.ordrepsy.qc.ca/documents/26707/0/Les+lignes+directrices+pour+l%27%C3%A9valuation+d%27un+enfant+en+vue+d%27une+demande+de+d%C3%A9rogation+%C3%A0+l%27%C3%A2ge+d%27admission+%C3%A0+l%27%C3%A9cole/b2c2e606-7fad-4b18-8ec8-883fb69e572e

Ordre des psychologues du Québec (OPQ). (2006). *Guidelines for expert assessment concerning child custody and access rights.* Montreal, Canada: Author. Retrieved from https://www.ordrepsy.qc.ca/documents/26707/6319/Lignes+directrices+pour+l%27expertise+en+mati%C3%A8re+de+garde+d%27enfants+et+des+droits+d%27acc%C3%A8s/d0c94c08-c04d-49ae-8fea-b37bd0ed83b8

Ordre des psychologues du Québec (OPQ). (2007). *Guidelines for the assessment of mental retardation*. Montreal, Canada: Author. Retrieved from https://www.ordrepsy.qc.ca/documents/26707/63191/Lignes+directrices+pour+l%27%C3%A9valuation+du+retard+mental/f20a2002-4535-4070-8a03-15889a9d5450

Ordre des psychologues du Québec (OPQ). (2014). *Guidelines for the evaluation of dyslexia in children*. Montreal, Canada: Author. Retrieved, from http://www.ordrepsy.qc.ca/documents/26707/63191/Lignes+directrices+pour+l%27%C3

%A9valuation+de+la+dyslexie+chez+les+enfants/139edf81-9163-4060-9c66-06f02eee0ed8

Oxman, A. D., Fretheim, A., & Schunemann, H. J. (2006). Improving the use of research evidence in guideline development: Introduction. *Health Resources and Policy Systems, 4*, 12-18. doi:10.1186/1478-4505-4-12

Palda, V. A., Davis, D., & Goldman, J. (2007). A guide to the Canadian Medical Association handbook on clinical practice guidelines. *Canadian Medical Association Journal, 177*, 1221-1226. doi:10.1503/cmaj.070880

Parry, G., Cape, J., & Pilling, S. (2003). Clinical practice guidelines in clinical psychology and psychotherapy. *Clinical Psychology & Psychotherapy, 10*, 337-351. doi:10.1002/cpp.381

Patel, A., Vieira, M. M. C., Abraham, J., Reid, N., Tran, T., Tomecsek, K., . . . Staton, C. A. (2016). Quality of the development of traumatic brain injury clinical practice guidelines: A systematic review. *PLoS ONE, 11*, e0161554. doi:10.1371/journal.pone.0161554

Pawson, R., Boaz, A., Grayson, L., Long, A., & Barnes, C. (2003). *Types and quality of knowledge in social care* (Report No. 3). London, England: Queen Mary University Press. Retrieved from www.kcl.ac.uk/sspp/departments/politicaleconomy/research/cep/pubs/papers/assets/wp18.pdf

Peterson, K., McDonagh, M. S., & Fu, R. (2011). Decisions to update comparative drug effectiveness reviews vary based on type of new evidence. *Journal of Clininical Epidemiology, 64*, 977-984. doi: 10.1016/j.jclinepi.2010.11.019

Pilling, S. (2008). History, context, process, and rationale for the development of clinical guidelines. *Psychology and Psychotherapy: Theory, Research and Practice, 81*, 331-350. doi:10.1348/147608308X324923

Pilling, S. (2009). Developing evidence-based guidance – implications for systemic interventions. *Journal of Family Therapy, 31,* 194-205. doi:10.1111/j.1467-6427.2009.00462.x

Pilling, S. (2012). *Developing clinical guidelines for adults: Handbook of Evidence-Based Practice in Clinical Psychology*. Hobogen, New Jersey: John Wiley & Sons, Inc.

Pincus, D., Kuhn, J. E., Sheth, U., Rizzone, K., Colbenson, K., Dwyer, T., … Wasserstein, D. (2017). A systematic review and appraisal of clinical practice guidelines for musculoskeletal soft tissue injuries and conditions. *The American Journal of Sports Medicine, 45,* 1458–1464. doi:10.1177/0363546516667903

Qaseem, A., Forland, F., Macbeth, F., Ollenschläger, G., Phillips, S., & van der Wees, P. (2012). Guidelines international network: Toward international standards for clinical practice guidelines. *Annals of Internal Medicine, 156,* 525-531. doi:10.7326/0003-4819-156-7-201204030-00009

Ransohoff, D. F., Pignone, M., & Sox, H. C. (2013). How to decide whether a clinical practice guideline is trustworthy. *Journal of the American Medical Association, 309,* 139-140. doi: 10.1001/jama.2012.156703

Rosenfeld, R. M., & Shiffman, R. N. (2009). Clinical practice guideline development manual: a quality-driven approach for translating evidence into action. *Otolaryngology--Head and Neck Surgery, 148,* 1-55. doi: 10.1016/j.otohns.2009.04.015

Richter, L., Garvare, R., & Nyström, M. E. (2017). Reaching beyond the review of research evidence: A qualitative study of decision making during the development of clinical practice guidelines for disease prevention in healthcare. *BMC Health Services Research, 17,* 344-457. doi:10.1186/s12913-017-2277-1

Ruszczyński, M., Horvath, A., Dziechciarz, P., & Szajewska, H. (2016). Cow's milk allergy guidelines: A quality appraisal with the AGREE II instrument. *Clinical & Experimental Allergy, 46,* 1236-1241. doi:10.1111/cea.12784

Rutten, G. M., Degen, S., Hendriks, E. J., Braspenning, J. C., Harting, J., & Oostendorp, R. A. (2016). Adherence to clinical practice guidelines for low back pain in physical therapy: Do patients benefit? *Physical Therapy, 90*, 1111-1122. doi:10.2522/ptj.20090173

Sackett, D. L., Rosenberg, W. M. C., Gray, J. A. M., Haynes, R. B., & Richardson, W. S. (1996). Evidence based medicine: What it is and what it isn't. *British Medical Journal, 312,* 70-72. doi:10.1136/bmj.312.7023.71

Serrano-Aguilar, P., Trujillo-Martin, M. M., del Pino-Sedeño, T., Pérez de la Rosa, A., de Pascual y Medina, A. M., Perestelo-Pérez, L., . . . Sarría Santamera, A. (2016). Patient participation in the development of a clinical guideline for inherited retinal dystrophies. *Expert Opinion on Orphan Drugs, 2,* 691-697. doi:10.1080/21678707.2016.1182907

Schuh, S., Babl, F. E., Dalziel, S. R., Freedman, S. B., Macias, C. G., Stephens, D., . . . Plint, A. C. (2017). Practice variation in acute bronchiolitis: A pediatric emergency research networks study. *Pediatrics*, e20170842. doi:10.1542/peds.2017-0842

Schünemann, H. J., Wiercioch, W., Etxeandia, I., Falavigna, M., Santesso, N., Mustafa, R., . . . al., e. (2014). Guidelines 2.0: Systematic development of a comprehensive checklist for a successful guideline enterprise. *Canadian Medical Association Journal, 186*, E123-E142. doi:10.1503/cmaj.131237

Sekercioglu, C. H. (2008). Quantifying coauthor contributions. *Science, 322*, 371-371. doi: 10.1126/science.322.5900.371a

Shaner, R. (2001). Practical psychotherapy: How practice guidelines can rescue psychotherapy in public systems. *Psychiatric Services, 52*, 1021-1025. doi:10.1176/appi.ps.52.8.1021

Shaneyfelt, T. M., Mayo-Smith, M. F., & Rothwangl, J. (1999). Are guidelines following guidelines?: The methodological quality of clinical practice guidelines in the peer-reviewed medical literature. *Journal of the American Medical Association, 281*, 1900-1905. doi:10.1001/jama.281.20.1900

Shekelle. (2014). Updating practice guidelines. *Journal of the American Medical Association, 311*, 2072-2074. doi:10.1001/jama.2014.4950

Shekelle, P. G., Ortiz, E., Rhodes, S., & et al. (2001). Validity of the agency for healthcare research and quality clinical practice guidelines: How quickly do guidelines become outdated? *Journal of the American Medical Association, 286*, 1461-1467. doi:10.1001/jama.286.12.1461

Shekelle, Woolf, S., Grimshaw, J. M., Schünemann, H. J., & Eccles, M. P. (2012). Developing clinical practice guidelines: Reviewing, reporting, and publishing guidelines. *Implementation Science, 7*, 62-67. doi:10.1186/1748-5908-7-62

Shiffman, R. N. (2016). Best practices for implementation of clinical decision support. *Clinical Decision Support Systems*, *2,* 99-109. doi: 10.1007/978-3-319-31913-1_6

Shiroiwa, T., Fukuda, T., Ikeda, S., Takura, T., & Moriwaki, K. (2016). Development of an official guideline for the economic evaluation of drugs/medical devices in Japan. *Value in Health. 20*, 372-378. doi 10.1016/j.jval.2016.08.726

Shnier, A., Lexchin, J., Romero, M., & Brown, K. (2016). Reporting of financial conflicts of interest in clinical practice guidelines: A case study analysis of guidelines from the Canadian Medical Association Infobase. *BMC Health Services Research, 16*, 1-8. doi:10.1186/s12913-016-1646-5

Scottish Intercollegiate Guidelines Network (SIGN). (2001). *SIGN 50: A guideline developer's handbook*. Edinburgh, Scottland: Author

Scottish Intercollegiate Guidelines Network (SIGN). (2015). *SIGN50: A guideline developer's handbook* [Electronic version]. Retrieved from https://sign.ac.uk/assets/sign50_2015.pdf

Stamoulos, C., Reyes, A., Trepanier, L., & Drapeau, M. (2014). The methodological quality of clinical practice guidelines in psychology: A pilot study of the guidelines developed by the Ordre des Psychologues du Québec. *Canadian Psychology/Psychologie canadienne, 55,* 177-186. doi:10.1037/a0037287

Steinert, T., Richter, D., & Bergk, J. (2010). Sources of bias in guideline development. *Psychiatric Services, 61,* 946. doi: 10.1176/appi.ps.61.9.946

Straus, S. E., Tetroe, J., & Graham, I. D. (2009). Defining knowledge translation. *Canadian Medical Association Journal, 181,* 165-168. doi:10.1503/cmaj.081229

The ADAPTE Collaboration (2009). *The ADAPTE process: Resource toolkit for guideline adaptation* (Version 2.0). Retrieved from the Guidelines International Network (G-I-N) website http://wwwg-i-n.net.

Timmermans, S., & Kolker, E. S. (2004). Evidence-based medicine and the reconfiguration of medical knowledge. *Journal of Health and Social Behaviour, 45,* 177-193.

Trepanier, L., Stamoulos, C., & Reyes, A. (2017). Assessing the quality of the OPQ's guidelines for the evaluation of dyslexia in children using the Appraisal of Guidelines for Research and Evaluation (AGREE) II: A brief report. *Canadian Psychology/Psychologie canadienne, 58,* 211-217. doi:10.1037/cap0000121

Tudor, K. I., Kozina, P. N., & Marušić, A. (2013). Methodological rigour and transparency of clinical practice guidelines developed by neurology professional societies in Croatia. *PLoS ONE, 8,* e69877. doi:10.1371/journal.pone.0069877

van Dijk, M. K., Oosterbaan, D. B., Verbraak, M. J., & van Balkom, A. J. (2013). The effectiveness of adhering to clinical-practice guidelines for anxiety disorders in secondary mental health care: the results of a cohort study in the Netherlands. *Journal of Evaluation in Clinical Practice, 19,* 791-797. doi:10.1111/j.1365-2753.2012.01851.x

Vernooij, R. W., Alonso-Coello, P., Brouwers, M., & Martinez Garcia, L. (2017). Reporting items for updated clinical guidelines: Checklist for the reporting of updated guidelines (CheckUp). *PLoS Med, 14*, e1002207. doi:10.1371/journal.pmed.1002207

Wang, X., Chen, Y., Yao, L., Zhou, Q., Wu, Q., Estill, J., . . . Norris, S. L. (2017). Declarations and conflicts of interest in WHO guidelines. *Journal of Clinical Epidemiology*. *98,* 1-8. doi:10.1016 /j.jclinepi.2017.12.021

Ward, J. E., & Grieco, V. (1996). Why we need guidelines for guidelines: a study of the quality of clinical practice guidelines in Australia. *Medical Journal of Australia, 165*, 574-576.

Watine, J. C., & Bunting, P. S. (2008). Mass colorectal cancer screening: Methodological quality of practice guidelines is not related to their content validity. *Clinical Biochemistry, 41*, 459-466. doi:10.1016/j.clinbiochem.2007.12.020

Weinberger, M. (2009). Seventeen years of asthma guidelines: Why hasn't the outcome improved for children? *The Journal of Pediatrics, 154,* 786-788. doi: 10.1016/j.peds,2009.01.003

Weisz, G., Cambrosio, A., Keating, P., Knaapen, L., Schlich, T., & Tournay, V. J. (2007). The emergence of clinical practice guidelines. *Milbank Quarterly, 85,* 691-727. doi:10.1111/j.1468-0009.2007.00505.x

World Health Organization (WHO). (2003). *Global programme on evidence for health policy: Guidelines for WHO guidelines.* Geneva, Switzerland: Author. Retrieved from

https://apps.who.int/iris/bitstream/handle/10665/68925/EIP_GPE_EQC_2003_1.pdf?sequence=1&isAllowed=y

World Health Organization (WHO). (2010). *WHO handbook for guideline development.* Geneva, Switzerland: Author. Retrieved from http://www.who.int/hiv/topics/mtct/grc_handbook_mar2010_1.pdf

Wold Health Organization (WHO). (2012). *WHO handbook for guideline development*. Geneva, Switzerland: Author. Retrieved from https://apps.who.int/iris/bitstream/handle/10665/75146/9789241548441_eng.pdf?sequence=1

Wolf, J. S., Hubbard, H., Faraday, M. M., & Forrest, J. B. (2011). Clinical practice guidelines to inform evidence-based clinical practice. *World Journal of Urology, 29*, 303-309. doi:10.1007/s00345-011-0656-5

Wollersheim, H., Burgers, J., & Grol, R. (2005). Clinical guidelines to improve patient care. *Netherlands Journal of Medicine, 63*, 188-192

Woolf, S. H., Grol, R., Hutchinson, A., Eccles, M., & Grimshaw, J. (1999). Clinical guidelines: Potential benefits, limitations, and harms of clinical guidelines. *British Medical Journal, 318*, 527-536. doi: 10.1136/bmj.318.7182.527

Ye, Z. K., Liu, Y., Cui, X., & Liu, L. (2016). Critical appraisal of the quality of clinical practice guidelines for stress ulcer prophylaxis. *PLoS ONE, 11*, e0155020. doi:10.1371/journal.pone.0155020

Zhang, C. (2009). A proposal for calculating weighted citations based on author rank. *EMBO Reports, 10*, 416–417. doi: 10.1038/embor.2009.74