

## RESEARCH ARTICLE

# Is the National Guideline Clearinghouse a Trustworthy Source of Practice Guidelines for Child and Youth Anxiety and Depression?

Stephanie Duda MSc<sup>1</sup>; Christine Fahim PhD(c)<sup>2</sup>; Peter Szatmari MD<sup>3</sup>; Kathryn Bennett PhD<sup>4</sup>

### Abstract

**Objective:** Innovative strategies that facilitate the use of high quality practice guidelines (PG) are needed. Accordingly, repositories designed to simplify access to PGs have been proposed as a critical component of the network of linked interventions needed to drive increased PG implementation. The National Guideline Clearinghouse (NGC) is a free, international online repository. We investigated whether it is a trustworthy source of child and youth anxiety and depression PGs. **Method:** English language PGs published between January 2009 and February 2016 relevant to anxiety or depression in children and adolescents ( $\leq 18$  years of age) were eligible. Two trained raters assessed PG quality using Appraisal of Guidelines for Research and Evaluation (AGREE II). Scores on at least three AGREE II domains (stakeholder involvement, rigor of development, and editorial independence) were used to designate PGs as: i) minimum quality ( $\geq 50\%$ ); and ii) high quality ( $\geq 70\%$ ). **Results:** Eight eligible PGs were identified (depression,  $n=6$ ; anxiety and depression,  $n=1$ ; social anxiety disorder,  $n=1$ ). Four of eight PGs met minimum quality criteria; three of four met high quality criteria. **Conclusions:** At present, NGC users without the time and special skills required to evaluate PG quality may unknowingly choose flawed PGs to guide decisions about child and youth anxiety and depression. The recent NGC decision to explore the inclusion of PG quality profiles based on Institute of Medicine standards provides needed leadership that can strengthen PG repositories, prevent harm and wasted resources, and build PG developer capacity.

**Key Words:** *practice guidelines, anxiety disorders, depressive disorder, child, adolescent*

### Résumé

**Objectif:** Les stratégies innovatrices qui facilitent l'utilisation de lignes directrices de la pratique (LDP) de grande qualité sont nécessaires. Par conséquent, il a été proposé que les dépositaires d'information destinés à simplifier l'accès aux LDP soient une composante essentielle du réseau des interventions liées nécessaires pour exécuter la mise en œuvre accrue des LDP. Le *National Guideline Clearinghouse* (NGC) est un dépositaire gratuit, international en ligne. Nous avons recherché si c'est une source fiable de LDP sur l'anxiété et la dépression chez les enfants et les adolescents. **Méthode:** Étaient admissibles les LDP en anglais publiées entre janvier 2009 et février 2016 et traitant de l'anxiété et la dépression chez les enfants et les adolescents ( $\leq 18$  ans). Deux évaluateurs compétents ont estimé la qualité des LDP à l'aide de la Grille d'évaluation de la qualité des recommandations pour la pratique clinique (AGREE II). Les scores à au moins 3 domaines d'AGREE II (participation des groupes concernés, rigueur d'élaboration, et indépendance éditoriale) ont été utilisés pour désigner les LDP comme étant : i) de qualité minimum ( $\geq 50\%$ ); et ii) de qualité élevée ( $\geq 70\%$ ).

<sup>1</sup>Research Coordinator, Department of Health Research Methods, Evidence, and Impact (formerly the Department of Clinical Epidemiology and Biostatistics), McMaster University, Hamilton, Ontario

<sup>2</sup>PhD candidate, Health Research Methodology program, Department of Health Research Methods, Evidence, and Impact (formerly the Department of Clinical Epidemiology and Biostatistics), McMaster University, Hamilton, Ontario

<sup>3</sup>Professor and Director of the Division of Child and Youth Mental Health, Department of Psychiatry, University of Toronto; Chief, Child and Youth Mental Health Collaborative, Hospital for Sick Children and Centre for Addiction and Mental Health, Toronto, Ontario

<sup>4</sup>Professor, Department of Health Research Methods, Evidence, and Impact (formerly the Department of Clinical Epidemiology and Biostatistics), McMaster University, Hamilton, Ontario

Corresponding E-Mail: [dudasj@mcmaster.ca](mailto:dudasj@mcmaster.ca)

Submitted: September 27, 2016; Accepted: February 10, 2017

**Résultats:** Huit LDP admissibles ont été identifiées (dépression, n = 6; anxiété et dépression, n = 1; trouble d'anxiété sociale, n = 1). Quatre LDP sur 8 satisfaisaient au critère de qualité minimum, et 3 sur 4 au critère de qualité élevée.

**Conclusions:** À l'heure actuelle, les utilisateurs du NGC qui n'ont pas le temps ni les compétences spéciales pour évaluer la qualité des LDP peuvent choisir à leur insu des LDP erronées pour guider leurs décisions relatives à l'anxiété et la dépression chez les enfants et les adolescents. La décision récente du NGC d'explorer l'inclusion des profils de qualité des LDP selon les normes de l'Institute of Medicine offre un leadership nécessaire qui peut renforcer les dépositaires de LDP, prévenir les dommages et le gaspillage des ressources, et accroître la capacité des élaborateurs de LDP.

**Mots clés:** lignes directrices de la pratique, troubles anxieux, trouble dépressif, enfant, adolescent

## Introduction

Many children and youth are affected by mental health problems (Polanczyk, Salum, Sugaya, Caye, & Rohde, 2015). Yet, although effective interventions are available, even in resource intensive settings fewer than one-third of those who need mental healthcare services receive them (Ford, 2008; Merikangas et al., 2011), and evidence-based practice is not the norm (Hoagwood, Burns, Kiser, Ringeisen, & Schoenwald, 2001; Kazak et al., 2010; Kazdin, 2010; Novins, Green, Legha, & Aarons, 2013). Innovations that increase access to effective child and youth mental health (CYMH) services are urgently needed. High quality practice guidelines (PGs) can address this gap, and are widely recognized as a foundation for health service excellence (Graham et al., 2011; Mental Health Commission of Canada, 2012). The recommendations they contain are intended to guide decisions about assessment, prevention and treatment, assist in the identification of gaps in the provision of evidence informed care, and drive needed practice change (Brouwers et al., 2015; Graham et al., 2011). PGs can contribute to practitioner and policy-maker capacity building, and enable users to understand the strengths (and limitations) of available evidence in the decision-making process. Numerous government organizations, professional societies, and others produce increasing numbers of PGs relevant to CYMH problems (Bennett, Gorman, Duda, Brouwers, & Szatmari, 2016). At the same time, influential organizations including the World Health Organization and the Mental Health Commission of Canada actively promote increased PG use in mental health service planning and point-of-care decision-making (Dua et al., 2011; Mental Health Commission of Canada, 2012; World Health Organization, 2012). However, despite the widespread international commitment to producing and promoting the value of their use, PG impact on health care processes and patient outcomes remains modest at best (Graham et al., 2011; Pronovost, 2013). Further research evaluating PG impact when paired with effective implementation strategies is needed.

Among the numerous barriers to PG use identified, finding high quality, up-to-date PGs presents a key challenge, particularly for busy clinicians (Francke, Smit, de Veer, & Mistiaen, 2008; Graham et al., 2011). This is because time and specialized skills that PG users may not have are required to locate and quality appraise potentially useful PGs. Accordingly, providing easy access to trustworthy PGs

is proposed as an essential component of the network of linked interventions needed to drive PG implementation and accordingly, strengthen CYMH services (Bennett et al., 2016; Graham et al., 2011). In fact, the availability of high quality PGs is a non-negotiable pre-requisite to implementing PG recommendations given the risk of harm and wasted resources that may be associated with the use of poor quality PGs. In CYMH, concern about the quality of PGs is of particular importance. This is because unlike numerous other healthcare specialties where PG quality has been documented regularly and found to be uneven (Alonso-Coello et al., 2010; Shaneyfelt, Mayo-Smith, & Rothwangl, 1999) little is currently known about CYMH PG quality. A recent review we conducted to appraise the rigor of CYMH PG development methods recommended by those who create them revealed that up to 70% of available CYMH PGs may have been developed using weak methods, highlighting the urgent need to determine the quality of available CYMH PGs (Bennett et al., 2016).

Therefore, we evaluated the trustworthiness of PGs for child and youth anxiety and depression contained in the National Guideline Clearinghouse (NGC) using systematic review methodology and the Appraisal of Guidelines for Research and Evaluation (AGREE II) PG quality assessment tool. The NGC is an online PG repository created by the United States Department of Health and Human Services' Agency for Healthcare Research and Quality (AHRQ) to enable easy access to PGs (Agency for Healthcare Research and Quality, 2015). While a number of PG repositories exist [for example (Australian Government, National Health and Medical Research Council, 2016; Canadian Medical Association, n.d.; Guidelines International Network, 2015; Guia Salud, n.d.)], we selected the NGC because it is internationally known, is publically available at no charge, and includes a large index of PGs related to mental health disorders (over 200 PGs are contained in its Mental Disorders index)(Agency for Healthcare Research and Quality, n.d.). In addition, the NGC has been increasingly concerned with PG quality following the 2011 Institute of Medicine (IOM) recommendation that the NGC include information about PG trustworthiness (Graham et al., 2011). The NGC subsequently modified its PG inclusion criteria and now requires PGs to meet specific methodologic standards, for example the use of a systematic review to derive PG recommendations is now required (The National Guideline Clearinghouse, 2014b). The NGC also decided to develop a PG

quality appraisal tool derived from IOM quality standards, and include the resulting PG appraisals on their website, but this commitment is still in the planning stages and not yet implemented (The National Guideline Clearinghouse, 2016). We focused on anxiety and depression since these are two of the most prevalent CYMH disorders, affecting about 6.5% (anxiety) and 2.6% (depression) of those  $\leq 18$  years of age at any given point in time (Polanczyk et al., 2015). The associated impairments are often profound and chronic, and include disrupted relationships with others, failure to graduate high school or obtain post-secondary education (Kessler, Foster, Saunders, & Stang, 1995; Woodward & Fergusson, 2001), an increased risk of teenage pregnancy (Kessler et al., 1997), an increased risk of lifelong persistent psychiatric disorder (Pine, Cohen, Gurley, Brook, & Ma, 1998; Woodward & Fergusson, 2001), higher risk for marital discord (Kessler, Walters, & Forthofer, 1998), and an increased risk of suicide related behavior (Fleischmann, Bertolote, Belfer, & Beautrais, 2005; Woodward & Fergusson, 2001).

## Methods

### Search Strategy and Selection Criteria

The NGC was electronically searched to identify eligible PGs published since 2009 (January 2009 – February 2016). Two searches were conducted using the NGC's 'advanced search' function for the following terms: (i) keyword = anxiety; target population = child (2 to 12 years), adolescent (13 to 18 years); clinical specialty = psychiatry, psychology; and (ii) keyword = depression; target population = child (2 to 12 years), adolescent (13 to 18 years); clinical specialty = psychiatry, psychology.

PGs meeting the following inclusion criteria were eligible: i) PG as defined by the *1997 Criteria for Inclusion of Clinical Practice Guidelines in NGC* (The National Guideline Clearinghouse, 2014a); ii) published in the English language; iii) focused on the prevention or treatment of anxiety or depression [excluding PGs singularly focused on obsessive compulsive disorder or post-traumatic stress disorder, as these disorders are no longer classified as anxiety disorders in DSM-5 (American Psychiatric Association, 2013) or bipolar disorder, as this disorder is in a distinctly different diagnostic category than depressive disorders (American Psychiatric Association, 2013)]; iv) focused on children or youth 18 years of age or younger (may include adults, but must have distinct recommendations for children or youth); and, v) published, revised, updated, or reaffirmed since 2009 (i.e., January 2009 – February 2016).

### Screening

One reviewer (SD) screened titles and summaries of all records identified by the search and applied inclusion/exclusion criteria.

### Quality Assessment Tool: AGREE II

The AGREE II tool was used to assess the quality of eligible PGs (AGREE Next Steps Consortium, 2013; Brouwers et al., 2010). AGREE II consists of 23 items grouped in six domains: scope and purpose; stakeholder involvement; rigor of development; clarity of presentation; applicability; and editorial independence (AGREE Next Steps Consortium, 2013; Brouwers et al., 2010). Items are rated on a 7-point Likert scale ranging from '1 = strongly disagree' to '7 = strongly agree' (AGREE Next Steps Consortium, 2013; Brouwers et al., 2010). Two additional items assess: i) overall quality ('1 = lowest possible quality' to '7 = highest possible quality') and ii) whether the PG should be recommended for use (yes, yes with modifications, or no) (AGREE Next Steps Consortium, 2013; Brouwers et al., 2010).

### Quality Assessment Process

First, two reviewers (SD and CF) participated in a training exercise which involved completion of the online AGREE II Overview Tutorial (The AGREE Research Trust, n.d.) followed by a practice assessment of a PG. Reviewers met with a senior investigator (KB) to discuss disagreements and 'lessons learned' regarding use of the tool. Following training, two reviewers independently applied AGREE II criteria to each eligible PG to arrive at their initial item scores. When applying AGREE II criteria, reviewers ensured that any companion documents for a given PG (e.g., tools and resources to aid PG implementation, technical reports, health economic analyses, PG evaluation tools, etc.) were considered in addition to the main PG document. Next, inter-rater differences on initial item scores were discussed. Reviewers then revised their item scores as they deemed appropriate; however, no attempt was made to reach consensus. Thus, any scoring differences that remained following discussion represent differences in judgment rather than errors due to failure to detect specific pieces of information within PG documents. These item scores were then considered final for each reviewer. Inter-rater agreement was assessed using the intraclass correlation coefficient (ICC) (Fleiss & Cohen, 1973). The degree of agreement was classified as per Fleiss' categories: poor (0-0.40), fair to good (0.41-0.75), and excellent (>0.75) (Fleiss, 1981). Analyses were performed using the statistical software SPSS, version 23. Final item scores were then aggregated into six domain scores by summing each reviewer's scores for all items within a given domain and standardizing as a percentage of the maximum possible score (ranging from 0-100%) using the formula described in the AGREE II User's Manual (AGREE Next Steps Consortium, 2013). The overall quality item score was also calculated as a percentage of the maximum possible score.

### **PG Quality**

The AGREE II User's Manual does not provide criteria to designate PGs as high or low quality (AGREE Next Steps Consortium, 2013). Thus, the interpretation of domain scores and overall PG quality assessments are determined by the user (AGREE Next Steps Consortium, 2013). We used scores on three AGREE II domains - stakeholder involvement, rigor of development, and editorial independence - to designate PGs as: i) minimum quality (domain score  $\geq 50\%$ ); and ii) high quality (domain score  $\geq 70\%$ ). These three domains were selected because they address the extent to which risk of bias was minimized in the identification and interpretation of the evidence used to derive the guideline. The remaining three domains, although important, do not evaluate the validity of the PG; rather they focus on the problem statement, clarity of presentation, and implementation.

### **Data Extraction**

Characteristics of eligible PGs (e.g., topic, developer, date of publication, country of origin) were extracted by a single reviewer (SD).

## **Results**

### **Search & Screening**

The search identified 125 unduplicated records (see Figure 1). Following screening, eight PGs met inclusion criteria (Beyondblue, 2010; Cincinnati Children's Hospital Medical Center, 2010a; Cincinnati Children's Hospital Medical Center, 2010b; Guidelines and Protocols Advisory Committee, 2010; National Institute for Health and Care Excellence, 2013e; National Institute for Health and Care Excellence, 2015g; U.S Preventive Services Task Force, 2009; Working group of the Clinical Practice Guideline on the Management of Major Depression in Childhood and Adolescence, 2009). Supplementary File 1, available online, contains a list of excluded PGs with reasons.

### **PG Characteristics**

Characteristics of the eight included PGs are presented in Table 1. Six PGs focused on depression, one PG addressed both anxiety and depression, and one PG focused solely on social anxiety disorder. The USA was the most frequent country of origin ( $n=3$ ), followed by the United Kingdom ( $n=2$ ), Australia ( $n=1$ ), Canada ( $n=1$ ), and Spain ( $n=1$ ). Six of eight PGs addressed both screening/diagnosis and treatment; one PG focused solely on screening/diagnosis while another addressed treatment only. Six of eight PGs had companion documents; two PGs were stand-alone documents.

### **Inter-rater Agreement**

Overall inter-rater agreement on the AGREE II item scores was excellent (ICC = 0.93, 95% CI = 0.91-0.95). Good to excellent inter-rater agreement was observed for each AGREE II domain (scope and purpose: ICC = 0.81, 95% CI = 0.61-0.91; stakeholder involvement: ICC = 0.95, 95% CI = 0.88-0.98; rigor of development: ICC = 0.93, 95% CI = 0.89-0.96; clarity of presentation: ICC = 0.72, 95% CI = 0.45-0.87; applicability: ICC = 0.90, 95% CI = 0.81-0.95; editorial independence: ICC = 0.96, 95% CI = 0.88-0.99).

### **PG Quality Ratings**

Table 2 shows the AGREE II ratings (% maximum possible score) for each domain. AGREE II item scores on which the domain scores are based are available from the author upon request. Four of eight PGs met our definition of minimum quality (i.e., scored  $\geq 50\%$  on stakeholder involvement, rigor of development, and editorial independence domains), namely the National Institute for Health and Care Excellence (NICE) depression PG (National Institute for Health and Care Excellence, 2015g), NICE social anxiety PG (National Institute for Health and Care Excellence, 2013e), Beyondblue depression PG (Beyondblue, 2010), and the Spanish Ministry of Health and Social Policy depression PG (Working group of the Clinical Practice Guideline on the Management of Major Depression in Childhood and Adolescence, 2009). When all six domains were considered, three out of the four PGs scored  $\geq 50\%$  on all six. Four PGs did not meet our definition of minimum quality. Specifically, the British Columbia Medical Services Commission PG on anxiety and depression (Guidelines and Protocols Advisory Committee, 2010), the Cincinnati Children's Hospital Medical Center depression screening PG (Cincinnati Children's Hospital Medical Center, 2010a), and the Cincinnati Children's Hospital Medical Center depression treatment PG (Cincinnati Children's Hospital Medical Center, 2010b) failed to score  $\geq 50\%$  on all three domains (i.e., stakeholder involvement, rigor of development, and editorial independence). The U.S Preventive Services Task Force (USPSTF) depression PG (U.S Preventive Services Task Force, 2009) scored  $\geq 50\%$  on two of these three key domains, namely rigor of development and editorial independence domains. When all six domains were considered, three out of the four scored  $\geq 50\%$  on two of the six domains, while one scored  $\geq 50\%$  on four of the six domains.

Three of eight PGs met our definition of high quality (i.e., scored  $\geq 70\%$  on stakeholder involvement, rigor of development, and editorial independence domains) - the NICE depression PG (National Institute for Health and Care Excellence, 2015g), the NICE social anxiety PG (National Institute for Health and Care Excellence, 2013e), and the Beyondblue depression PG (Beyondblue, 2010). When all six domains were considered, these three PGs (Beyondblue, 2010; National Institute for Health and Care Excellence, 2013e; National Institute for Health and Care Excellence,



**Table 1. Characteristics of eligible practice guidelines**

Guideline reference	Author/ developer	Last updated	Country of origin	Disorder			Purpose			Companion documents	
				Anxiety	Depression	Other	Prevention	Treatment	Other		
(National Institute for Health and Care Excellence, 2015g)	National Institute for Health and Care Excellence (NICE)	March 2015	United Kingdom		✓				✓		<ul style="list-style-type: none"> <li>• Full Guideline (National Institute for Health and Care Excellence, 2005)</li> <li>• Addendum (National Institute for Health and Care Excellence, 2015a)</li> <li>• Appendices P-V (National Institute for Health and Care Excellence, 2015b; National Institute for Health and Care Excellence, 2015c; National Institute for Health and Care Excellence, 2015d; National Institute for Health and Care Excellence, 2015e)</li> <li>• Evidence Update (National Institute for Health and Care Excellence, 2013a)</li> <li>• Tools &amp; Resources (National Institute for Health and Care Excellence, 2015f)</li> <li>• NICE Charter (National Institute for Health and Care Excellence, 2013b)</li> <li>• NICE Guidelines Manual (National Institute for Health and Care Excellence, 2016)</li> </ul>
(National Institute for Health and Care Excellence, 2013e)	National Institute for Health and Care Excellence (NICE)	May 2013	United Kingdom	✓					✓		<ul style="list-style-type: none"> <li>• Appendices 6-26 (National Institute for Health and Care Excellence, 2013c)</li> <li>• Tools &amp; Resources (National Institute for Health and Care Excellence, 2013d)</li> <li>• NICE Charter (National Institute for Health and Care Excellence, 2013b)</li> <li>• NICE Guidelines Manual (National Institute for Health and Care Excellence, 2016)</li> </ul>
(Beyondblue, 2010)	Beyondblue	February 2011	Australia		✓			✓	✓		<ul style="list-style-type: none"> <li>• NHMRC Standards and Procedures (National Health and Medical Research Council, 2007)</li> <li>N/A</li> </ul>
(Working group of the Clinical Practice Guideline on the Management of Major Depression in Childhood and Adolescence, 2009)	Spanish Ministry of Health and Social Policy	July 2009	Spain		✓				✓		

*continued*

**Table 1. continued**

Guideline reference	Author/ developer	Last updated	Country of origin	Disorder			Purpose			Companion documents
				Anxiety	Depression	Other	Screening/ diagnosis	Treatment	Other	
(Guidelines and Protocols Advisory Committee, 2010)	British Columbia Ministry of Health Services	January 2010	Canada	✓	✓		✓	✓		N/A
(Cincinnati Children's Hospital Medical Center, 2010a)	Cincinnati Children's Hospital Medical Center	March 2010	USA		✓		✓			<ul style="list-style-type: none"> <li>• Table of Evidence Levels (Cincinnati Children's Hospital Medical Center, 2009c)</li> <li>• Grading Body of Evidence Table (Cincinnati Children's Hospital Medical Center, 2009a)</li> <li>• Judging the Strength of a Recommendation (Cincinnati Children's Hospital Medical Center, 2009b)</li> </ul>
(Cincinnati Children's Hospital Medical Center, 2010b)	Cincinnati Children's Hospital Medical Center	January 2010	USA		✓		✓			<ul style="list-style-type: none"> <li>• Table of Evidence Levels (Cincinnati Children's Hospital Medical Center, 2009c)</li> <li>• Grading Body of Evidence Table (Cincinnati Children's Hospital Medical Center, 2009a)</li> <li>• Judging the Strength of a Recommendation (Cincinnati Children's Hospital Medical Center, 2009b)</li> </ul>
(U.S. Preventive Services Task Force, 2009)	U.S. Preventive Services Task Force (USPSTF)	April 2009	USA		✓		✓	✓		<ul style="list-style-type: none"> <li>• Evidence Review (Williams, O'Connor, Eder, &amp; Whitlock, 2009b)</li> <li>• Full Evidence Synthesis (Williams, O'Connor, Eder, &amp; Whitlock, 2009a)</li> <li>• Methods of the USPSTF (Harris et al., 2001)</li> </ul>

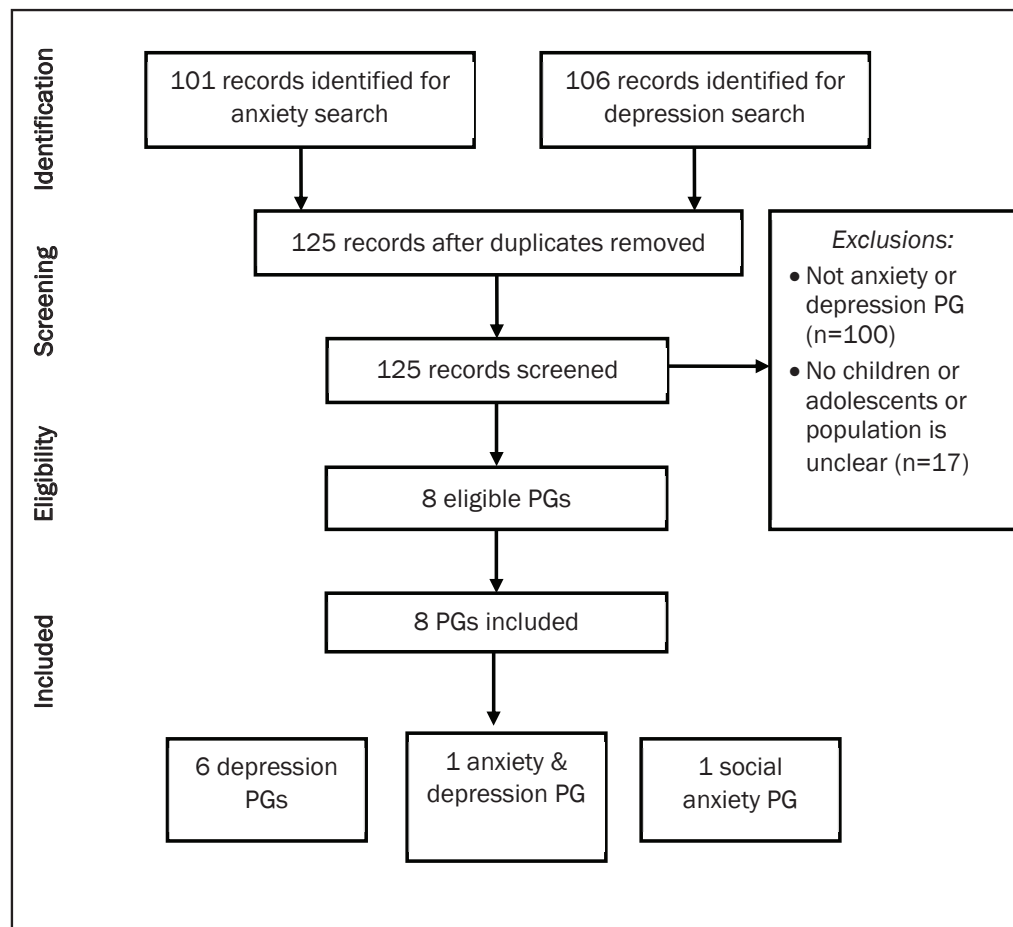
\*accessing care; N/A = Not Applicable

2015g) also scored  $\geq 70\%$  on all six. Among the five PGs that did not meet our high quality definition, the USPSTF depression PG scored  $\geq 70\%$  on two of the three key domains (U.S. Preventive Services Task Force, 2009), the Spanish Ministry of Health and Social Policy depression PG scored  $\geq 70\%$  on one of three key domains (Working group of the Clinical Practice Guideline on the Management of Major Depression in Childhood and Adolescence, 2009), and the three remaining PGs failed to score  $\geq 70\%$  on any of the key domains (Cincinnati Children's Hospital Medical Center, 2010a; Cincinnati Children's Hospital Medical Center, 2010b; Guidelines and Protocols Advisory Committee, 2010). When all six domains were considered, the Spanish depression PG (Working group of the Clinical Practice Guideline on the Management of Major Depression in Childhood and Adolescence, 2009) and the USPSTF depression PG (U.S. Preventive Services Task Force, 2009) scored  $\geq 70\%$  on three of six domains. The remaining three PGs (Cincinnati Children's Hospital Medical Center, 2010a; Cincinnati Children's Hospital Medical Center, 2010b; Guidelines and Protocols Advisory Committee, 2010) scored  $\geq 70\%$  on less than half of the domains.

## Discussion

### Main Findings

Easy access to high quality PGs through user-friendly web-based repositories is recommended as a PG dissemination strategy that may increase awareness and understanding of PGs, and facilitate their implementation (Graham et al., 2011). Since the NGC is free, internationally available online, and houses a large number of mental health PGs, we sought to determine its usefulness as a source of

**Figure 1. Eligibility of Practice Guidelines**

trustworthy PGs for decision-makers concerned with two high prevalence CYMH disorders – anxiety and depression. PG quality was of particular concern because to date such ratings are not available for PGs that address CYMH problems. Despite the rapid rise in PG production over the past 20 years (Alonso-Coello et al., 2010) we found only eight PGs relevant to child and youth anxiety or depression in the NGC. Their quality based on AGREE II assessments was variable with only four of the eight achieving minimum quality standards; three of the four met criteria for high quality.

These findings are noteworthy for a number of reasons. First, the criteria we used to define minimum quality were lenient – a score of  $\geq 50\%$  on three of the six AGREE II domains (stakeholder involvement, rigor of development and editorial independence). Even with this relatively forgiving threshold, only four of eight PGs were considered worthy of consideration for use. Second, the four PGs that met our definition of minimum quality scored consistently high on all six AGREE II domains, while three of the four PGs that did not achieve a minimum quality rating scored poorly on all six domains. The consistency we observed across domains suggests that PGs may fall into two quality categories – strong or weak – and accordingly that judgments by

users about which PGs to consider and which to avoid may be relatively straightforward. The consistency is likely explained by the fact that PG developers either choose to align their methods with international quality standards ensuring rigor in all phases of the development process, or choose to use less demanding methods in each step of PG development. Third, the areas of weakness among low quality PGs (i.e., stakeholder involvement, rigor of development, applicability, and editorial independence domains) are all standards that require significant time, expertise and financial resources to conduct the methodologically rigorous activities needed to achieve a high quality score. For example, the rigor of development domain, arguably the most influential domain with respect to PG quality, requires PG developers to engage in the following resource intensive activities: i) conduct a systematic review of the evidence (i.e., a comprehensive literature search using clear inclusion/exclusion criteria and including assessment of risk of bias across the body of evidence); ii) convene a panel of stakeholders to formulate recommendations that are clearly linked to the supporting evidence and the risks and benefits identified in the systematic review; iii) conduct an external expert review of the PG; and iv) commit to a plan for PG updating (AGREE Next Steps Consortium, 2013). It is possible,

as discussed recently by Classen and Mermel (2015), that when PG quality is low this is explained by lack of the substantial financial resources needed to engage in the type of activities required to develop rigorous PGs. Finally, it is important to note that the PG weaknesses we identified are not unique to child and youth anxiety and depression PGs. For example, in a review of PGs focused primarily on adult physical health conditions, Alonso-Coello *et al* (2010) reported that rigor of development, stakeholder involvement, editorial independence, and applicability were amongst the weakest rated domains.

## Implications

At the present time NGC users do not have access to quality ratings based on AGREE II or other international standards. Even though the NGC has recently revised their PG inclusion criteria (The National Guideline Clearinghouse, 2014b) it is unclear whether the new criteria will over time achieve their intended goal – to exclude PGs deemed to be weak using quality appraisal tools such as AGREE II or the IOM based tool they intend to develop. Consequently, at present NGC users may unknowingly choose weak or flawed PGs to guide decisions about child and youth anxiety and depression. Although some users may have the skills needed to evaluate PG quality and find the time needed to apply them to PGs of potential interest, it is unrealistic to expect the majority of potential PG users to do this. The baseline knowledge of research methodology and time needed to develop PG quality appraisal skills combined with the significant time and effort required to locate and evaluate the relevant PG documents (estimated to be between 2 and 4 hours per PG) is not likely available to many PG users.

It is important to recognize that the absence of information about PG quality is not a problem that is unique to the NGC. When we investigated three other prominent English language repositories [Guidelines International Network (G-I-N), the Canadian Medical Association CPG Infobase, and the Australian Clinical Guidelines Portal (Australian Government, National Health and Medical Research Council, 2016; Canadian Medical Association, n.d.; Guidelines International Network, 2015)] we were unable to locate quality ratings in any of them (although the Australian Portal uses a modified version of the NGC inclusion criteria to determine whether a PG will be available through their portal). In contrast, AHRQ and the NGC are exemplary given their recent attention to the importance of information about PG quality. In so doing, they have taken a leadership role compared to other repositories. The NGC proposal to develop a PG quality appraisal tool, and include the resulting quality assessments on their website provides needed direction to the international PG community in two critical areas. First, NGC plans regarding PG quality appraisal can stimulate other repositories to adopt similar policies and methods, increasing the availability of crucial information PG users need to make choices between PGs based on quality/

trustworthiness. Second, making PG quality ratings available can contribute to building the capacity of PG developers, providing a call to action regarding the need to raise the PG methodologic rigor bar. Capacity building can be accomplished through initiatives that teach guideline developers how to: prepare systematic reviews using Cochrane Collaboration (Higgins & Green, 2011) and PRISMA standards (Moher, Liberati, Tetzlaff, Altman, & The PRISMA Group, 2009); assemble multi-disciplinary PG development groups; apply explicit and transparent methods to address financial and intellectual conflict of interest; and increase the attention paid to applicability and implementation tools relevant to barriers and facilitators that may hinder PG use. Increased attention to PG quality can also stimulate breaking down barriers associated with the substantial resources needed to develop rigorous PGs. For example, through increased collaboration and partnerships among CYMH PG developers available resources could be pooled making more effective use of the time and effort currently invested in PG development. The result could be fewer PGs of higher quality, an outcome that should be regarded as desirable by PG users and developers alike.

Our findings also point to potential gaps in the CYMH problems addressed by currently available PGs. For example, no PGs that meet our quality criteria were found that address generalized anxiety disorder, separation anxiety disorder, panic disorder or specific phobias. The value of creating specialty specific PG repositories is worthy of consideration. For example, a repository of CYMH PGs that includes PG quality ratings, similar to the PG repository developed for cancer (The Canadian Partnership Against Cancer, 2013), would allow those concerned with CYMH to quickly identify high quality PGs relevant to their needs. Such a repository could raise the profile of PGs within the CYMH community, helping to create a common language and knowledge base regarding assessment, prevention and treatment. CYMH clinician trainees and their mentors may find such a resource invaluable, particularly given competency-based training is increasingly embraced by healthcare professional organizations (Milne, Tierney, & Doig, 2016). Youth and families may also benefit from the availability of such a repository, helping them to understand what to expect when they seek mental healthcare, and facilitating patient engagement in mental health decision-making (Légaré *et al.*, 2011).

Finally, the development and use of PG quality assessment tools, and the creation of strategies to facilitate PG choice based on quality raises several challenging issues. For example, what criteria should be used to evaluate quality and how good is good enough? Who is ultimately responsible for judgments regarding whether or not a PG should be used? That is, are decisions about whether a PG is good enough to guide decisions the responsibility of organizations/professional bodies, or individual PG users (e.g., clinicians, policy-makers, youth and families) informed by



Table 2. AGREE II Domain Scores (%) for Eligible Practice Guidelines		Guideline reference						
AGREE II Domain	(National Institute for Health and Care Excellence, 2015g)	(National Institute for Health and Care Excellence, 2013e)	(Beyondblue, 2010)	(Working group of the Clinical Practice Guideline on the Management of Major Depression in Childhood and Adolescence, 2009)	(Guidelines and Protocols Advisory Committee, 2010)	(Cincinnati Children's Hospital Medical Center, 2010a)	(Cincinnati Children's Hospital Medical Center, 2010b)	(U.S Preventive Services Task Force, 2009)
Scope & Purpose	91.7	100.0	94.4	94.4	52.8	66.7	83.3	94.4
Stakeholder Involvement	86.1	86.1	91.7	63.9	2.8	16.7	13.9	36.1
Rigor of Development	85.4	80.2	81.3	68.8	7.3	49.0	42.7	78.1
Clarity of Presentation	94.4	86.1	100.0	88.9	63.9	72.2	80.6	66.7
Applicability	72.9	72.9	79.2	43.8	41.7	39.6	31.3	22.9
Editorial Independence	79.2	95.8	83.3	79.2	0.0	4.2	4.2	79.2
Overall Quality Score	91.7	75.0	91.7	83.3	0.0	50.0	41.7	58.3
PG Meets Minimum Quality Definition*	✓ (3/3)	✓ (3/3)	✓ (3/3)	✓ (3/3)	✗ (0/3)	✗ (0/3)	✗ (0/3)	✗ (2/3)
Total PG Domains ≥ 50%, #	6/6	6/6	6/6	5/6	2/6	2/6	2/6	4/6
PG Meets High Quality Definition**	✓ (3/3)	✓ (3/3)	✓ (3/3)	✗ (1/3)	✗ (0/3)	✗ (0/3)	✗ (0/3)	✗ (2/3)
Total PG Domains ≥ 70 %, #	6/6	6/6	6/6	3/6	0/6	1/6	2/6	3/6
* Scored ≥ 50% on three AGREE II domains - stakeholder involvement, rigor of development and editorial independence								
** Scored ≥ 70% on three AGREE II domains - stakeholder involvement, rigor of development and editorial independence								
Abbreviations: PG = Practice Guideline								

their individual needs, circumstances, and preferences and values?

Enormous strides in the development and uptake of quality standards for research about the assessment, prevention and treatment of health problems have occurred in the past few decades. For example, the Cochrane Collaboration and Grading of Recommendations Assessment, Development and Evaluation (GRADE) methods are increasingly recognized internationally as key resources that can guide healthcare decisions, with the overall aim of increasing the use of effective interventions and eliminating the use of ineffective or harmful ones (Guyatt et al., 2011; Higgins & Green, 2011). Strategies that help users sift through the continuously increasing stock of PGs and associated recommendations for care are a logical extension of the work of bodies such as Cochrane and GRADE. However, consistent with the tenets of evidence-informed healthcare, mental healthcare decisions should be made taking into account not only research based knowledge but also i) professional/personal knowledge and ii) patient values and preferences (Sackett, Rosenberg, Gray, Haynes, & Richardson, 1996). Accordingly, PGs contribute to decision-making, but do not prescribe the outcome. Regarding the question ‘How good is good enough?’, it is our view that recommendations based on a rigorous systematic review, informed by the needs and preferences of the key stakeholders who can benefit from the PG, and protected from the influence of academic and financial conflicts of interest represent minimum, non-negotiable standards. Regarding who should decide whether a PG is worthy of use, this is ultimately up to the individual decision-maker. However, it is our belief that rigorous and credible tools, such as PG quality appraisal tools, are needed to assist PG users in the decision-making process, are welcomed by many if not most of them, and can ultimately contribute to raising the bar on the use of research and PGs in mental healthcare decisions.

## Strengths and Limitations

The strengths of our study include the following. First, we used systematic review methods to search, screen, and critically appraise eligible PGs identified in a large, well-known international PG repository. Second, two trained raters (SD and CF) applied AGREE II quality criteria to each PG and achieved excellent inter-rater agreement. Third, we set criteria for minimum and high quality providing a simple yardstick to aid understanding of where each PG is located on a scale from high to low quality defined using AGREE II domain items. However, our study is not without potential limitations. First, validated cut-points defining minimum and high quality AGREE II scores are not currently available (AGREE Next Steps Consortium, 2013). The ones we used were based on the professional judgment of our team and are intended to help readers understand the AGREE II scores. Other stakeholders may have different views.

Although a number of published studies utilizing AGREE II have applied cut-points to meet their research objectives, no consensus has emerged. For example, Bekkering *et al* (2014) deemed PGs high quality if the rigor of development domain score was > 50%. In contrast, Lee, Yamada, Kyo-lolo, Shorkey and Stevens (2014) recommended PGs for use if they received overall quality scores of 6 or 7, while Sanclemente, Acosta, Tamayo, Bonfill and Alonso-Coello (2014) applied a cut-point of 60% to each domain. Second, PG quality ratings are influenced by the completeness of reporting of the PG development process which may differ from the actual methods applied by the developers. It is possible that PGs with low quality ratings actually adhered to a specific standard but did not report sufficient detail to receive a score that aligned with our definitions of minimum and high PG quality.

## Conclusion

At the present time, the quality of child and youth anxiety and depression PGs contained in the NGC is variable. Innovative strategies including PG repositories that allow users to quickly access high quality PGs are needed to increase the use of high quality PGs, avoid the application of flawed or harmful PGs, prevent the associated harm and wasted resources that may result, and facilitate capacity building among those who develop CYMH PGs. The recent NGC decision to explore the inclusion of PG quality profiles based on IOM standards provides needed leadership that can strengthen PG repositories and build PG developer capacity.

## Acknowledgements / Conflict of Interests

The authors have no financial relationships to disclose. No funding was received to support this work.

## References

- Agency for Healthcare Research and Quality (n.d.). National guideline clearinghouse: Guidelines by topic. Retrieved from <http://www.guideline.gov/browse/by-topic-detail.aspx?id=39068&ct=1>
- Agency for Healthcare Research and Quality. (2015). National guideline clearinghouse: About. Retrieved from <http://www.guideline.gov/about/index.aspx>
- AGREE Next Steps Consortium. (2013). AGREE-II user's manual. Retrieved from <http://www.agreetrust.org>
- Alonso-Coello, P., Irfan, A., Sola, 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 & Safety in Health Care*, 19(6):e58.
- American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders, fifth edition. Arlington, VA: American Psychiatric Association.
- Australian Government, National Health and Medical Research Council. (2016). Clinical practice guidelines portal. Retrieved from <https://www.clinicalguidelines.gov.au/portal>
- Bekkering, G. E., Aertgeerts, B., Asueta-Lorente, J. -, Atrique, M., Goossens, M., Smets, K.,...for the ADAPTE-youth project group. (2014). Practitioner review: Evidence-based practice guidelines on

- alcohol and drug misuse among adolescents: A systematic review. *Journal of Child Psychology and Psychiatry*, 55(1), 3-21.
- Bennett, K., Gorman, D. A., Duda, S., Brouwers, M., & Szatmari, P. (2016). Practitioner review: On the trustworthiness of clinical practice guidelines - A systematic review of the quality of methods used to develop guidelines in child and youth mental health. *Journal of Child Psychology and Psychiatry*, 57(6), 662-673.
- Beyondblue. (2010). Clinical practice guidelines: Depression in adolescents and young adults. Melbourne: Beyondblue: The national depression initiative.
- Brouwers, M. C., Kho, M. E., Browman, G. P., Burgers, J. S., Cluzeau, F., Feder, G.,...for the AGREE Next Steps Consortium. (2010). AGREE-II: Advancing guideline development, reporting and evaluation in health care. *Canadian Medical Association Journal*, 182(18), e839-e842.
- Brouwers, M. C., Makarski, J., Kastner, M., Hayden, L., Bhattacharyya, O., & and the GUIDE-M Research Team. (2015). The guideline implementability decision excellence model (GUIDE-M): A mixed methods approach to create an international resource to advance the practice guideline field. *Implementation Science*, 10(36). doi:10.1186/s13012-015-0225-1
- Canadian Medical Association. (n.d.). CPG infobase: Clinical practice guidelines database. Retrieved from <https://www.cma.ca/En/Pages/clinical-practice-guidelines.aspx>
- Cincinnati Children's Hospital Medical Center. (2009a). Grading a body of evidence to answer a clinical question. Retrieved from <http://www.cincinnatichildrens.org/assets/0/78/1067/2709/2777/2793/9200/bd6f4eea-825c-49c3-a0e5-3e66c54dc066.pdf>
- Cincinnati Children's Hospital Medical Center. (2009b). Judging the strength of a recommendation. Retrieved from <http://www.cincinnatichildrens.org/assets/0/78/1067/2709/2777/2793/9200/d7344329-03d0-45f3-b6ca-02c746a472ec.pdf>
- Cincinnati Children's Hospital Medical Center. (2009c). Table of evidence levels - levels of individual studies by domain, study design & quality. Retrieved from <http://www.cincinnatichildrens.org/assets/0/78/1067/2709/2777/2793/9200/5ce396bf-fdcb-4c65-a9f2-1b9888d4fc7e.pdf>
- Cincinnati Children's Hospital Medical Center. (2010a). Best evidence statement (BEST): Screening of children and adolescents for major depressive disorder (MDD). Cincinnati: Cincinnati Children's Hospital Medical Center.
- Cincinnati Children's Hospital Medical Center. (2010b). Best evidence statement (BEST): Treatment of children and adolescents with major depressive disorder (MDD) during the acute phase. Cincinnati: Cincinnati Children's Hospital Medical Center.
- Classen, D. C., & Mermel, L. A. (2015). Specialty society clinical practice guidelines: Time for evolution or revolution? *Journal of the American Medical Association*, 314(9), 871-872.
- Dua, T., Barbui, C., Clark, N., Fleischmann, A., Poznyak, V., van Ommeren, M.,...Saxena, S. (2011). Evidence-based guidelines for mental, neurological, and substance use disorders in low- and middle-income countries: Summary of WHO recommendations. *PLoS Medicine*, 8(11), e1001122.
- Fleischmann, A., Bertolote, J. M., Belfer, M., & Beautrais, A. (2005). Completed suicide and psychiatric diagnoses in young people: A critical examination of the evidence. *American Journal of Orthopsychiatry*, 75(4), 676-683.
- Fleiss, J. L. (1981). Statistical methods for rates and proportions (2nd ed.). New York: John Wiley.
- Fleiss, J. L., & Cohen, J. (1973). The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educational and Psychological Measurement*, 33, 613-619.
- Ford, T. (2008). Practitioner review: How can epidemiology help us plan and deliver effective child and adolescent mental health services. *Journal of Child Psychology and Psychiatry*, 49(9), 900-914.
- Francke, A. L., Smit, M. C., de Veer, A. J. E., & Mistiaen, P. (2008). Factors influencing the implementation of clinical guidelines for health care professionals: A systematic meta-review. *BMC Medical Informatics and Decision Making*, 8(38). doi: 10.1186/1472-6947-8-38.
- Graham, R., Mancher, M., Miller Wolman, D., Greenfield, S., Steinberg, E., & Committee on Standards for Developing Trustworthy Clinical Practice Guidelines, Institute of Medicine. (2011). In Graham R., Mancher M., Miller Wolman D., Greenfield S. and Steinberg E. (Eds.), Clinical practice guidelines we can trust. Washington, DC: National Academies Press.
- Guia Salud. (n.d.) Biblioteca de Guías de Práctica Clínica del Sistema Nacional de Salud. Retrieved from <http://portal.guiasalud.es/web/guest/home>
- Guidelines and Protocols Advisory Committee. (2010). Anxiety and depression in children and youth - diagnosis and treatment. Victoria, BC: British Columbia Ministry of Health.
- Guidelines International Network. (2015). Resources/international guideline library. Retrieved from <http://www.g-i-n.net/library/international-guidelines-library>
- Guyatt, G., Oxman, A. D., Akl, E. A., Kunz, R., Vist, G., Brozek, J.,... Schünemann, H. J. (2011). GRADE guidelines: 1. Introduction—GRADE evidence profiles and summary of findings tables. *Journal of Clinical Epidemiology*, 64(4), 383-394.
- Harris, R. P., Helfand, M., Woolf, S. H., Lohr, K. N., Mulrow, C. D., Teutsch, S. M.,...for the Methods Work Group, Third U.S. Preventive Services Task Force. (2001). Current methods of the U.S. preventive services task force: A review of the process. *American Journal of Preventive Medicine*, 20(3S), 21-35.
- Higgins, J. P. T., & Green, S. (2011). Cochrane handbook for systematic reviews of interventions version 5.1.0. Retrieved from [www.cochrane-handbook.org](http://www.cochrane-handbook.org)
- Hoagwood, K., Burns, B. J., Kiser, L., Ringeisen, H., & Schoenwald, S. K. (2001). Evidence-based practice in child and adolescent mental health services. *Psychiatric Services*, 52(9), 1179-1189.
- Kazak, A. E., Hoagwood, K., Weisz, J. R., Hood, K., Kratochwill, T. R., Vargas, L. A., & Banez, G. (2010). A meta-systems approach to evidence-based practice for children and adolescents. *American Psychologist*, 62(2), 85-97.
- Kazdin, A. E. (2010). Improving the quality of care and reducing the burden of clinical dysfunction. *Administration and Policy in Mental Health and Mental Health Services Research*, 37, 160-166.
- Kessler, R., Walters, E., & Forthofer, M. (1998). The social consequences of psychiatric disorders, III: Probability of marital stability. *American Journal of Psychiatry*, 155(8), 1092-1096.
- Kessler, R. C., Berglund, P. A., Foster, C. L., Saunders, W. B., Stang, P. E., & Walters, E. E. (1997). Social consequences of psychiatric disorders, II: Teenage parenthood. *American Journal of Psychiatry*, 154(10), 1405-1411.
- Kessler, R. C., Foster, C. L., Saunders, W. B., & Stang, P. E. (1995). Social consequences of psychiatric disorders, I: Educational attainment. *American Journal of Psychiatry*, 152(7), 1026-1032.
- Lee, G. Y., Yamada, J., Kyololo, O., Shorkey, A., & Stevens, B. (2014). Pediatric clinical practice guidelines for acute procedural pain: A systematic review. *Pediatrics*, 133(3), 500-515.
- 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(6), E45-74.
- Mental Health Commission of Canada. (2012). Changing directions, changing lives: The mental health strategy for Canada. Calgary, AB: Mental Health Commission of Canada.
- Merikangas, K. R., He, J., Burstein, M., Swendsen, J., Avenevoli, S., Case, B., . . . Olfson, M. (2011). Service utilization for lifetime mental disorders in U.S. adolescents: Results of the National Comorbidity Survey—Adolescent supplement (NCS-A). *Journal of the American Academy of Child & Adolescent Psychiatry*, 50(1), 32-45.
- Milne, V., Tierney, M. & Doig, C. (2016). Excitement and apprehension over shift to competency-based medical education.

- Retrieved from <http://healthydebate.ca/2016/04/topic/competency-based-medical-education>
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Journal of Clinical Epidemiology*, *62*, 1006-1012.
- National Health and Medical Research Council. (2007). NHMRC standards and procedures for externally developed guidelines. Canberra: National Health and Medical Research Council.
- National Institute for Clinical Excellence. (2005). Depression in children and young people: Identification and management in primary, community and secondary care. London: The British Psychological Society and the Royal College of Psychiatrists.
- National Institute for Health and Care Excellence. (2013a). Depression in children and young people: Evidence update June 2013. Manchester: National Institute for Health and Care Excellence.
- National Institute for Health and Care Excellence. (2013b). NICE charter. Retrieved from [https://www.nice.org.uk/Media/Default/About/Who-we-are/NICE\\_Charter.pdf](https://www.nice.org.uk/Media/Default/About/Who-we-are/NICE_Charter.pdf)
- National Institute for Health and Care Excellence. (2013c). Social anxiety disorder: Recognition, assessment and treatment - appendices. Retrieved from <https://www.nice.org.uk/guidance/cg159/evidence>
- National Institute for Health and Care Excellence. (2013d). Social anxiety disorder: Recognition, assessment and treatment - tools and resources. Retrieved from <https://www.nice.org.uk/guidance/cg159/resources>
- National Institute for Health and Care Excellence. (2013e). Social anxiety disorder: The NICE guideline on recognition, assessment and treatment. London: The British Psychological Society and the Royal College of Psychiatrists.
- National Institute for Health and Care Excellence. (2015a). Addendum to clinical guideline 28, depression in children and young people. London: National Institute for Health and Care Excellence.
- National Institute for Health and Care Excellence. (2015b). Appendices Q-T. Retrieved from <https://www.nice.org.uk/guidance/cg28/evidence>
- National Institute for Health and Care Excellence. (2015c). Appendix P: Evidence profile tables. Retrieved from <https://www.nice.org.uk/guidance/cg28/evidence>
- National Institute for Health and Care Excellence. (2015d). Appendix U: Psychological interventions - forest plots. Retrieved from <https://www.nice.org.uk/guidance/cg28/evidence>
- National Institute for Health and Care Excellence. (2015e). Appendix V: Antidepressant drug treatment - forest plots. Retrieved from <https://www.nice.org.uk/guidance/cg28/evidence>
- National Institute for Health and Care Excellence. (2015f). Depression in children and young people: Identification and management - tools and resources. Retrieved from <https://www.nice.org.uk/guidance/cg28/resources>
- National Institute for Health and Care Excellence. (2015g). Depression in children and young people: Identification and management (CG 28). London: National Institute for Health and Care Excellence.
- National Institute for Health and Care Excellence. (2016). Developing NICE guidelines: The manual. Retrieved from <https://www.nice.org.uk/article/pmg20/resources/non-guidance-developing-nice-guidelines-the-manual-pdf>
- Novins, D. K., Green, A. E., Legha, R. K., & Aarons, G. A. (2013). Dissemination and implementation of evidence-based practices for child and adolescent mental health: A systematic review. *Journal of the American Academy of Child & Adolescent Psychiatry*, *52*(10), 1009-1025.e18.
- Pine, D. S., Cohen, P., Gurley, D., Brook, J., & Ma, Y. (1998). The risk for early-adulthood anxiety and depressive disorders in adolescents with anxiety and depressive disorders. *Archives of General Psychiatry*, *55*(1), 56-64.
- Polanczyk, G. V., Salum, G. A., Sugaya, L. S., Caye, A., & Rohde, L. A. (2015). Annual research review: A meta-analysis of the worldwide prevalence of mental disorders in children and adolescents. *The Journal of Child Psychology and Psychiatry*, *56*(3), 345-365.
- Pronovost, P. J. (2013). Enhancing physicians' use of clinical guidelines. *Journal of the American Medical Association*, *310*(23), 2501-2502.
- 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. *BMJ*, *312*, 71. doi: <http://dx.doi.org/10.1136/bmj.312.7023.71>.
- Sancllemente, G., Acosta, J., Tamayo, M., Bonfill, X., & Alonso-Coello, P. (2014). Clinical practice guidelines for treatment of acne vulgaris: A critical appraisal using the AGREE II instrument. *Archives of Dermatological Research*, *306*(3), 269-277.
- 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*(20), 1900-1905.
- The AGREE Research Trust. (n.d.). AGREE II overview tutorial. Retrieved from <http://www.agreetrust.org/resource-centre/agree-ii-training-tools/>
- The Canadian Partnership Against Cancer. (2013). Standards and guidelines evidence (SAGE) reports and publications. Retrieved from [http://www.cancerview.ca/cv/portal/Home/TreatmentAndSupport/TSPProfessionals/ClinicalGuidelines/GRCMain/GRCSAGE/GRCSAGEReports?\\_afLoop=2632968906403239&lang=en&\\_afWindowMode=0&\\_adf.ctrl-state=1cf2xby7i6\\_603](http://www.cancerview.ca/cv/portal/Home/TreatmentAndSupport/TSPProfessionals/ClinicalGuidelines/GRCMain/GRCSAGE/GRCSAGEReports?_afLoop=2632968906403239&lang=en&_afWindowMode=0&_adf.ctrl-state=1cf2xby7i6_603)
- The National Guideline Clearinghouse. (2014a). 1997 criteria for inclusion of clinical practice guidelines in NGC. Retrieved from <http://www.guideline.gov/about/inclusion-criteria.aspx>
- The National Guideline Clearinghouse. (2014b). 2013 (revised) criteria for inclusion of clinical practice guidelines in NGC. Retrieved from <http://www.guideline.gov/about/inclusion-criteria.aspx>
- The National Guideline Clearinghouse. (2016). Determining extent adherence to the IOM standards in NGC. Retrieved from <https://guideline.gov/help-and-about/summaries/determining-extent-adherence-to-the-iom-standards-in-ngc>
- U.S Preventive Services Task Force. (2009). Screening and treatment for major depressive disorder in children and adolescents: US preventive services task force recommendation statement. *Pediatrics*, *123*(4), 1223-1228.
- Williams, S. B., O'Connor, E., Eder, M., & Whitlock, E. (2009a). Screening for child and adolescent depression in primary care settings: A systematic evidence review for the U.S. Preventive Services Task Force. (Evidence Synthesis No. 69. AHRQ Publication No. 09-05130-EF-1). Rockville, Maryland: Agency for Healthcare Research and Quality.
- Williams, S. B., O'Connor, E. A., Eder, M., & Whitlock, E. P. (2009b). Screening for child and adolescent depression in primary care settings: A systematic evidence review for the US Preventive Services Task Force. *Pediatrics*, *123*(4), e716-e735.
- Woodward, L. J., & Fergusson, D. M. (2001). Life course outcomes of young people with anxiety disorders in adolescence. *Journal of the American Academy of Child and Adolescent Psychiatry*, *40*(9), 1086-1093.
- Working group of the Clinical Practice Guideline on the Management of Major Depression in Childhood and Adolescence. (2009). Clinical practice guideline on major depression in childhood and adolescence. quality plan for the national health system of the ministry of health and social policy. Spain: Axencia de Avaliacion de Tecnoloxias Sanitarias de Galicia (avalial-t).
- World Health Organization. (2012). WHO handbook for guideline development. Geneva: WHO Press