Journal of Evaluation in Clinical Practice
International Journal of Public Health Policy and Health Services Research

Journal of Evaluation in Clinical Practice ISSN 1365-2753

Rhetoric and argumentation: how clinical practice guidelines think

Jonathan Fuller BMSc

MD/PhD Student, Faculty of Medicine, University of Toronto, Ontario, Canada

Keywords
argumentation, clinical practice guidelines, elderly, generalization, logic, rhetoric

Correspondence
Mr Jonathan Fuller
Faculty of Medicine
University of Toronto
Room E349B 2075 Bayview Avenue
Toronto
Ontario
Canada M4N 3M5
E-mail: jonathan.fuller@mail.utoronto.ca

Accepted for publication: 18 March 2013

doi:10.1111/jep.12037

Abstract

Introduction Clinical practice guidelines (CPGs) are an important source of justification for clinical decisions in modern evidence-based practice. Yet, we have given little attention to how they argue their evidence. In particular, how do CPGs argue for treatment with long-term medications that are increasingly prescribed to older patients?

Approach and rationale I selected six disease-specific guidelines recommending treatment with five of the medication classes most commonly prescribed for seniors in Ontario, Canada. I considered the stated aims of these CPGs and the techniques employed towards those aims. Finally, I reconstructed and logically analysed the arguments supporting recommendations for pharmacotherapy.

Analysis The primary function of CPGs is rhetorical, or persuasive, and their means of persuasion include both a display of their credibility and their argumentation. Arguments supporting pharmacotherapy recommendations for the target population follow a common inductive pattern: statistical generalization from randomized controlled trial (RCT) and meta-analysis evidence. Two of the CPGs also argue their treatment recommendations for older patients in this style, while three fail to justify pharmacotherapy specifically for the older population.

Discussion The arguments analysed lack the auxiliary assumptions that would warrant making a generalization about the clinical effectiveness of medications for the older population. Guidelines reason using simple induction, while ignoring important inferential gaps. Future guidelines should aspire to be well-reasoned rather than simply evidence-based; argue from a plurality of evidence; be wary of hasty inductions; appropriately limit the scope of their recommendations; and avoid making law-like, prescriptive generalizations.

Introduction: prescribing in the 'age of evidence'

The coming decades will witness an increase in the number of older patients – a fact that has not gone totally unnoticed [1,2]. Population ageing has substantial economic implications as well as implications for quality of care in medicine. Pharmacotherapy is central to the medical management of chronic disease, and health care providers often prescribe multiple long-term medications for their older patients with multiple chronic diseases [3–7]. Multiple prescribing is expensive for countries with publicly funded medication programmes for seniors. The vast potential for harm resulting from polypharmacy due to adverse drug reactions [7,8], including drug–drug interactions, is of even greater concern. Yet, increased prescribing to seniors has been reported in several countries, including England [9], the United States [10] and Canada [11]. These potent concerns over costs and risks demand that the prescribing of medications to older patients is explicitly justified. But just what kinds of arguments count in medicine's modern context of justification?

The medical community finds itself in the 'age of evidence', where therapeutic decisions, especially those involving pharmaceutical treatment, must be based upon clinical research evidence, preferably from randomized controlled trials (RCTs). To find the justification for prescribing demanded above, one might thus turn to the medical literature, the rapidly expanding evidence base for clinical practice. The popular *Users' Guides to the Medical Literature* endorses the use of pre-appraised evidence, generated through a systematic review and summary of all literature relevant to a particular question, as long as the pre-appraised evidence is synthesized according to the principles of critical appraisal that are associated with evidence-based practice [12]. Evidence-based

clinical practice guidelines (CPGs) are one such source of pre-appraised evidence and are becoming more numerous. Beyond reviewing and summarizing evidence, guidelines provide recommendations for all domains of medical practice, but particularly for disease prevention and treatment. Quality assurance initiatives encourage adherence to guidelines [7,13], which help define the standard of care for the medical profession. Practice guidelines are thus a plausible source of justification for some of the more common, costly and consequential medical decisions made in routine practice.

Guidelines are produced by national societies and other organizations; the process involves a review of the literature, the forming of an expert panel to generate recommendations, and usually the convening of a consensus conference at which the recommendations are discussed and revised before a consensus is declared. A guideline is then published and, increasingly, efforts are undertaken to ensure its dissemination among doctors.

Given the effort devoted to the guideline-development process and the resources invested, it is natural to wonder about the primary function of the process. The consensus conference model has its origins in the National Institutes of Health (NIH) Consensus Development Conference (CDC) Programme in the United States [14]. Solomon argues that the role of the NIH CDC Programme and similar programmes is rhetorical. Solomon's use of the term 'rhetorical' is not to be understood pejoratively, but as identifying a primary persuasive function for consensus conferences; they aim to persuade doctors to adopt practices consistent with the available 'best evidence' [15]. If CPGs, as a product of many consensus conferences, function rhetorically, then the rationale supporting their recommendations could be analysed as argumentation, which some have considered to be an essential component of rhetoric since the time of Aristotle [16]. A logical analysis would permit an interrogation of some of the assumptions and inferences governing modern evidence-based care, especially in the realm of treatment.

Much has been written about the 'logic of medicine' and 'how doctors think' [17–19], but in an age of evidence, clinical reasoning on the part of the individual clinician is often supplemented by 'evidence summaries', including CPGs, that do some of the cognitive work for them. The use of CPGs or other rigorously prepared summaries is considered by evidence-based medicine (EBM) proponents to be one particular mode of practicing EBM (the 'using mode') [20]. Despite a preponderance of guidelines and the endorsement of EBM by undergraduate and postgraduate medical training programmes, countless associations and institutions, and major medical journals, insufficient attention has been paid to how the normative EBM model is realized through guideline-directed care. The present article analyses the logic of practice guidelines and provides an answer to the question: *how do CPGs think?*

In what follows, I demonstrate that CPGs are primarily rhetorical texts that argue their evidence for treatment recommendations. I then reconstruct and logically analyse arguments for drug therapy contained in chronic disease guidelines. Given the centrality of drug therapy to chronic disease management, as well as the significance of increased prescribing to the elderly seen in several countries, I focus on arguments justifying treatment with commonly prescribed long-term medications, including treatment for the older patient population. I conclude with some suggestions for how practice guidelines could be more 'conscientious' and 'judicious' in their evidence-based recommendations, consistent with the first principles of EBM [21]. This analysis ultimately sketches a method for describing and evaluating clinical reasoning within a common and influential context in modern practice.

Approach and rationale

Although a trend of increased drug prescribing has been reported in several countries, recent and detailed data on prescribing to older patients in primary care are generally lacking. Bajcar and colleagues analysed the trend of increased prescription drug claims for seniors (those aged 65+) seen in primary care in Ontario, Canada, according to the unique medication classes prescribed [11]. Using these data, Mutasingwa and colleagues mapped the medication classes corresponding to the most frequent prescription drug claims onto the most common chronic diseases managed by primary care doctors [22]. Generally, Mutasingwa and colleagues found a one-to-one correspondence between a medication class and a chronic disease for which it is prescribed. Thus, in the following analysis, I sought the argument supporting the use of a particular class of medications in the disease-specific CPG onto which the class was mapped.

I used the list of CPGs cited by Mutasingwa and colleagues [22], which consisted of guidelines commonly used by primary care doctors in Ontario, after updating the list to include the most recent guidelines published. I chose this particular collection of guidelines for the sake of consistency; the trends in prescribing by individual medication class are known from Ontario data. I selected five guidelines that could be mapped with the four most commonly claimed medication classes. From greatest to fewest number of claims, these classes were: cardiovascular medications (for the treatment of hypertension [23] and heart failure [24]), psychotropics (for the treatment of depression [25]), gastrointestinal medications (for the treatment of gastro-oesophageal reflux disease [26]), and lipid-lowering medications (for the treatment of dyslipidemia and the prevention of cardiovascular disease [27]). I selected a sixth guideline, corresponding to the class of medications that had the highest-fold increase in number of prescription claims from 1997 to 2006: osteoporosis medications (for the treatment of osteoporosis [28]).

To determine the function of CPGs and particularly whether they contain rhetoric, I searched these six guidelines, as well as the websites of the societies that produced them [29,30], for the stated aims of the guideline-development process. Kuypers and King defined rhetoric as the 'strategic use of communication, oral or written, to achieve specifiable goals' [31], including the use of the written word to persuade an audience to act in a certain way. If CPGs aim to persuade doctors, their audience, to adopt certain practices, then they can be analysed as rhetorical texts or artefacts. Rhetorical analysis or criticism, the understanding and evaluation of rhetorical artefacts, has its own flourishing traditions from various perspectives in the social sciences and humanities. My purpose was not a thorough rhetorical critique of CPGs, but I engaged selectively with the methodology of 'traditional' criticism descendent from Aristotle's *Rhetoric* [32].

For Aristotle, a rhetorician has three technical means of persuasion at their disposal: *pathos*, *ethos and logos* [16]. *Pathos* is the rousing of passion within the audience to win sympathy for the

interlocutor's thesis. In the second technique, *ethos*, the interlocutor portrays him- or herself as credible in order to gain the confidence of their audience. The final technique, *logos*, consists of the arguments made in order to convince the audience. Arguments were the central concern of the present study.

For the analysis of argumentation, I examined all consensus statements that were based, in part or whole, on the clinical effectiveness of a pharmaceutical intervention. I define *clinical effectiveness* as the capacity of a drug to cause a particular clinical outcome for some members in a population. While many recommendations compared the relative effectiveness of different drugs, I chose those recommendations that are based, at least in part, on clinical effectiveness because the belief that the drug might work is a necessary condition for justified treatment with that drug. A doctor might choose to prescribe a drug that is less effective than another drug or alternative intervention due to other considerations, such as cost or harms, but prescribing cannot be adequately justified without the belief that the drug is reasonably effective. I divided the chosen consensus statements into those that make recommendations for the target population of the guideline generally, and those that make recommendations for older patients specifically. Where no consensus statements pertaining to older patients could be found, I considered the recommendations for the target population as extending to older patients, as long as the target population was inclusive of older adults.

I analysed the consensus statements and the justification supporting each as argumentation, using techniques of informal logic to reconstruct arguments [33]. In Walton's argument theory [33–35], arguments are dialectical; they take place within a dialogue between two agents. Applying Walton's theory, clinicians reading CPGs are engaged in a one-sided conversation, with the guideline as speaker and the clinician as audience member. Arguments in rhetorical contexts are often probative; that is, the premises (starting points) support the conclusion (end point). One would expect that arguments for clinical effectiveness possess this structure, as it is incumbent that *evidence*-based guidelines support effectiveness claims (end points) with *evidence* (starting points). Therefore, I reconstructed probative arguments by considering recommendations to be conclusions and statements interpreting evidence to be premises. I then looked for patterns common to token guideline arguments and constructed abstract argument types. Finally, I suggested one type of argument that might be used to justify prescribing for the elderly in cases where justification was absent. I represented argument types in ordinary language, rather than using logical symbolism or argument diagramming.

Subsequently, I evaluated the reasoning in the abstract argument types. Again taking a cue from Walton, reasoning is the process of inferring conclusions from premises by means of warrants, which are rules or frames for inference [35]. When arguments are deductive (the conclusion necessarily follows from the premises), they can be evaluated as valid (correct) or invalid (incorrect). When arguments are inductive (the premises provide grounds for accepting the conclusion but do not guarantee the conclusion), they can be evaluated as good/strong or bad/weak. Inductive arguments are often judged to be weak if they commit a fallacy, or some stereotyped violation of the accepted warrants for reasoning. After characterizing each argument type as deductive or inductive, I discussed potential weaknesses and suggested additional premises that would hypothetically strengthen the inference.

Analysis

Function of CPGs and techniques used

To determine the function of CPGs, I looked to the aims of the guideline-development process, as stated within the CPGs and by the national associations supporting them. The osteoporosis CPG mentioned the need to address the 'care gap' between 'appropriate assessment or treatment' and actual assessment or treatment of patients [28]. Similarly, the website of the Canadian Cardiovascular Society, which supports the development of the guidelines for dyslipidemia and heart failure, stated that their aim 'is to have a process that supports the development of key resources and tools for those dedicated to closing the gap between "what we know" and "what we do" ' [29]. This discourse around closing gaps between knowledge and practice is typical of the field of *knowledge translation* (KT) [36,37], also known as *knowledge transfer and exchange* [38]. The aims of CPGs suggest that they are an instrument for KT. In fact, the programme responsible for developing the hypertension guidelines describes itself as a 'knowledge translation programme' [30].

The KT field arose out of the recognition that making primary research evidence accessible to doctors is not enough to change their practice [36–38]. Doctors have to be *persuaded* to make changes. Practice guidelines are seen here to be one rhetorical device used to accomplish this objective. All six CPGs contained consensus recommendations worded in strongly prescriptive language: in such-and-such circumstances, patients 'should be' managed in such-and-such ways [23–28]. Each recommendation is then followed, in square brackets, by the 'level' or 'grade' of evidence upon which the recommendation is based. Where high-grade evidence is available, its invocation strengthens the rhetorical force of the recommendation.

As a KT initiative, the primary function of CPGs is rhetorical. But what are their means of persuasion? Guidelines do not use emotive language, nor do they appeal to the sympathies of clinicians (*pathos*), likely because their goals are to persuade doctors to adopt certain practices among several alternatives, where all possible actions share the ultimate *telos* of healing or helping the patient.

I will consider in more depth whether CPGs use the rhetorical technique of *ethos*: persuading the audience of the speaker's credibility in order to add credence to their thesis. Guideline development takes place under the banner of a respected national organization. Of the six CPGs reviewed, four mention the involvement of national or international specialists in the consensus panel [23,25,27,28], four mention its multidisciplinary make-up [24–26,28], and all six use the term 'expert' to describe its members [23–28]. The objectivity of the decisions rendered by the group is promoted in four guidelines through mention of the processes in place to reveal potential conflicts of interest among panel members [23,25–27]. A perception of objectivity is also created by the mechanism of group judgement used in all instances, culminating in the reaching of consensus on the final product. The consensus panel is thus portrayed as competent and unbiased and the CPG as trustworthy, which may engender faith among care providers in the positions taken by the panel and published in the CPG. Practice guidelines establish their own credibility through these techniques.

I now turn to the substance of the guidelines. Do the consensus statements carry the power to persuade on their own, if detached from the imprimatur of their authors? Can one make out coherent argumentation (*logos*) defending these statements? The answer, to each is *yes*. Whenever evidence from experimental or observational clinical studies is available, it is used to argue for a recommendation. The structure and direction of the argument is suggested by the simple fact that evidence-based recommendations are recommendations *supported by* evidence. Each consensus recommendation can be viewed as a conclusion, proceeding from the supportive evidence (premises) summarized in the text above or below. This finding allows us to reconstruct the justification for treatment with commonly prescribed drugs as probative arguments and then expose the arguments to logical analysis.

Arguments for the effectiveness of drugs for the target population

All six guidelines reviewed refer extensively to the clinical literature [23–28], citing relevant studies and summarizing the results of several of these studies. All CPGs also contain general statements endorsing treatment with a certain family of medications, 'medications Y'. Most commonly, these statements suggest that the drugs 'should be used', although some suggest that they 'are recommended' or 'can be used'. The generality of the recommendations implies that they are applicable to the entire target population of the guideline, 'patient population X'. In most cases, X consists of adults. Recommendations have the general form of the proposition: 'medications Y should be used to achieve outcome Z in patient population X [Grade Φ]'. The grade or level of the recommendation is based on the highest grade or level of evidence from which the proposition is inferred; high-grade recommendations are based on high-grade evidence. Although each CPG uses a slightly different grading system, all agree that the highest grade evidence comes from RCTs or meta-analysis of RCTs [23–28].

The evidence to which Φ refers is briefly summarized in the rationale that follows the recommendation. For propositions recommending drug therapy for the target population, the evidence described usually consists *exclusively* of RCTs and/or meta-analyses of RCTs. In fact, three of the CPGs included only RCTs and systematic reviews of RCTs in their literature search to update their evidence base for therapeutic recommendations [23,27,28]. In all six guidelines, the authors claim that one or more trials show that the drugs achieved outcome Z, or that 'medications Y were clinically effective with respect to outcome Z'. Their evidence is an increased frequency of a positive response or a decreased frequency of a negative event in the group receiving the medication compared with the group receiving a placebo.

I found that all token arguments recommending treatment for the target population with a certain family of medications were of the same type (Table 1). This structure is a *serial argument*, in which proposition two (ii) serves as a link between two sub-arguments; (ii) is both a conclusion inferred from premise one (i) above and a premise from which the consensus statement below is inferred. The general structure is instantiated by several token arguments taken from the CPGs for osteoporosis (X = women over age 50, Y = osteoporosis medications, Z = prevention of fractures); hypertension (X = adults, Y = antihypertensive medications, Z = prevention of cardiovascular disease events); and the prevention of cardiovascular disease (X = high-risk adults, Y = statins, Z = prevention of cardiovascular disease events) (Table 2).

Table 1 Argument type supporting pharmacotherapy for the target population within disease-specific clinical practice guidelines

(i) In one or more RCTs or meta-analyses of RCTs involving samples of patient population X, medications Y were clinically effective with respect to outcome Z.
(ii) Therefore, medications Y are clinically effective in patient population X with respect to outcome Z.
Therefore, medications Y should be used to achieve outcome Z in patient population X [Grade Φ].

RCT, randomized controlled trial.

'Extrapolating' or making 'in general' inferences (Table 2) involves generalizing from particular trial populations (i) to the wider target population (ii). The inference from (i) to (ii) is inductive, as the conclusion (ii) does not follow with necessity from the premise (i). Given the truth of (i), the truth of (ii) is not guaranteed. Premise (i) only establishes the effectiveness of the drug in the *trial population*; logical consistency is preserved even if (ii) happens to be false, or the medication is not clinically effective in the *wider target population*. The inference from (i) to (ii) is also more specifically a *statistical generalization*, in which we infer a population statistic from a sample statistic. In the argument type I abstracted, the generalization proceeds from an average treatment effect statistic, measured in particular samples of patient population X, to a general conclusion about the average treatment effect in X.

Arguments for the effectiveness of drugs for seniors

I now consider consensus statements recommending drug therapy *specifically* for older patients, which are present in only three of the guidelines reviewed, namely those for depression, hypertension and heart failure. Most of the arguments justifying these treatment recommendations are of the kind previously identified (Table 3). Notably, the depression CPG treats seniors as its target population and contains an additional recommendation specifically for older patients with medical co-morbidities (Table 3).

The hypertension CPG recommends treatment with antihypertensive medications in older hypertensive patients, with the caveat that RCT results cannot be extrapolated to the frail elderly (Table 3). Interestingly, conflicting non-RCT evidence is presented in the background for this recommendation: 'Observational studies raise concern that blood pressure lowering in patients 80 years of age or older may be poorly tolerated and associated with increased mortality' [40]. Curiously, after the results of a single RCT involving treatment with one particular class of antihypertensive agents in patients that were 'healthier than the general Canadian population 80 years of age and older' are described, no further consideration is given to the discrepant 'observational studies', as if the mere existence of higher-grade evidence invalidates the lower-grade evidence.

In all CPGs in which statements explicitly recommending drug therapy for seniors were not found [26–28], the target population

Table 2 Examples of arguments supporting pharmacotherapy for the target population within disease-specific clinical practice guidelines

Guidelines for 'the diagnosis and management of osteoporosis' [28]	(i) 'There is consistent evidence from randomized clinical trials that all therapies currently available in Canada reduce the risk of vertebral fractures for menopausal women with osteoporosis . . . There is also evidence that some interventions prevent non-vertebral and/or hip fractures'. (ii) 'In general, pharmacotherapy reduces the risk of vertebral fracture by 30–70%, depending on the agent and level of adherence. The effect on non-vertebral fractures is lower and varies by fracture site'. 'For menopausal women requiring treatment of osteoporosis, alendronate, risedronate, zoledronic acid and denosumab can be used as first-line therapies for prevention of hip, non-vertebral and vertebral fractures [grade A]'.
Guidelines for 'the management of hypertension' [39]	(i) '20. Psaty BP, Lumley T, Furberg CD, Schellenbaum G, Pahor M, Alderman MH, Weiss NH. Health outcomes associated with various antihypertensive therapies used as first-line agents: A network meta-analysis. JAMA 2003;289:2534–44'. (ii) 'There is a large and compelling body of literature demonstrating that antihypertensive drug therapy is associated with a 20% to 25% reduction in cardiovascular events and a 10% reduction in mortality (20) [referenced above] . . . These studies form the foundation for the grades A and B monotherapy recommendations'. 'Initial therapy should consist of monotherapy with a thiazide diuretic (grade A), a beta-blocker (in patients younger than 60 years of age, grade B), an angiotensin-converting-enzyme (ACE) inhibitor (in non-Black patients, grade B), a long-acting calcium channel blocker (CCB) (grade B) or an angiotensin receptor blocker (ARB) (grade B)'.
Guidelines for 'the diagnosis and treatment of dyslipidemia and prevention of cardiovascular disease' [27]	(i) 'The Cholesterol Treatment Trialists (CCT) meta-analysis (17) of 14 statin trials showed a dose-dependent relative reduction in CVD [cardiovascular disease] with LDL-C lowering'. (ii) 'Extrapolating from the available data, a 2.0 mmol/L absolute reduction or a 50% relative reduction in LDL-C provides optimal benefit in terms of CVD reduction (52)'. 'Thus, for high-risk subjects, the target levels should be an LDL-C of less than 2.0 mmol/L, or a 50% or greater reduction from baseline LDL-C (class I, level A). In the majority of patients, this is achievable with statin monotherapy . . . In high-risk individuals, treatment should be started immediately'.

Arguments are organized according to the argument type in Table 1.

Table 3 Examples of arguments supporting pharmacotherapy for seniors within disease-specific clinical practice guidelines

Guidelines for 'the assessment and treatment of depression' [25]	(i) 'There have been several meta-analyses of treatment in people over age 60, showing similar levels of response compared with younger patients . . . The response rate for both groups is generally 30% with placebo and 60% with treatment. (ii) 'Pharmacological treatment of non-psychotic major depressive episodes using antidepressants has been shown beneficial in the elderly'. 'Older patients have a response rate with antidepressant therapy similar to younger adults. Clinicians should approach elderly depressed individuals with therapeutic optimism [A]'. (i) 'A Cochrane meta-analysis of [antidepressant] treatment in adults of all ages with a variety of physical illnesses (e.g. diabetes, cancer, HIV, Parkinson disease, myocardial infarction) found a NNT [number needed to treat] of 4, similar to treatment in a general older population'. (ii) '[T]his data indicates that antidepressant treatment can be efficacious for depressive disorders in the medically ill'. 'Antidepressants should be used when indicated, even in patients with multiple co-morbidities and serious illnesses, as they have similar efficacy rates compared with use in well elderly . . . [B]'.
Guidelines for 'the management of hypertension' [40]	(i) 'This year, the landmark HYVET randomized placebo-controlled trial (3) examined the efficacy and safety of antihypertensive therapy in this very elderly population . . . After a median 1.8 years of follow-up, the trial was terminated early with a 30% RR [relative risk] reduction in the primary end point, fatal or nonfatal stroke (95% CI–1% to 51%; $P = 0.06$), a 39% reduction in fatal stroke (95% CI 1% to 62%; $P = 0.05$) and a 21% reduction in the total mortality (95% CI 4% to 35%; $P = 0.02$) in the indapamide group compared with the placebo group'. (ii) 'HYVET clearly demonstrates the beneficial effects of antihypertensive therapy in reducing the risk of stroke and death in very elderly patients; these results form the basis of the new recommendation to prescribe antihypertensive therapy regardless of age'. 'Antihypertensive therapy should be considered in all patients meeting the above indications [for treatment] regardless of age (grade B). Caution should be exercised in elderly patients who are frail'.

Arguments are organized according to the argument type in Table 1.

Table 4 Hypothetical argument type supporting pharmacotherapy for a sub-population, based on randomized controlled trial (RCT) evidence involving samples of the wider target population

(i)	In one or more RCTs or meta-analyses of RCTs involving samples of patient population X, medications Y were clinically effective with respect to outcome Z.
(ii)	Therefore, medications Y are clinically effective in patient population X with respect to outcome Z.
(iii)	Therefore, medications Y are clinically effective in sub-population X_o with respect to outcome Z.
Therefore,	medications Y should be used to achieve outcome Z in sub-population X_o.

of the guideline (most commonly, adults) included older individuals. If medications Y should be used in all adults, it necessarily follows that medications Y should be used in older adults because older adults are included in the extension of the class 'adults'. Thus, recommendations for patient population X could be extended to older adults, sub-population X_o. None of the guidelines that omitted specific recommendations for older adults considered whether the evidence presented is generalizable to this important sub-population [26–28].

I present a second argument type as one possible justification for prescribing in older patients, which follows the style of reasoning used throughout the CPGs and, in fact, builds upon the arguments made explicitly within them (Table 4). In this hypothetical argument, the inference from (ii) to (iii) is 'the reverse' of the inference from (i) to (ii). While the inference from (i) to (ii) is an induction from a sample statistic to a population statistic, the inference from (ii) to (iii) is an induction from a population statistic to a sub-population statistic. Both inferences require auxiliary assumptions if they are to be warranted, as will be discussed.

Discussion

I have suggested that the primary function of CPGs is rhetorical, consistent with the function Solomon identifies for the NIH CDC Programme and other programmes following the consensus conference model [15]. Guidelines are a tool for knowledge translation, for persuading practicing physicians to change their practices consistent with the highest quality evidence available. Given the lack of sophistication in the arguments that CPGs make, their persuasiveness may depend on the confidence inspired in their audience due to the perceived competence of the developers, the objectivity of the process and the resulting trustworthiness of the recommendations. This technique is not new. In Aristotle's *Rhetoric*, the first of three persuasive argument techniques is conveying the personal attributes possessed by the interlocutor to make him or her seem credible [16].

The importance of this rhetorical technique for the function of guidelines is illustrated by the fact that scientific papers containing reasons for the recommendations are only one way in which the recommendations are communicated to the medical community. A doctor reading a CPG summary, which often simply lists the recommendations, must trust in the experts and the guideline if they are to be persuaded to adhere to it. Nonetheless, expert pronouncements are not enough for a process governed by the principles of EBM. Arguments are indispensable for the full-length CPGs because the conclusions of the consensus panel must be premised on the current best evidence.

Many have criticized EBM for eschewing traditional forms of authority while sanctioning new forms [41–45], including evidence-based guidelines that codify practice. The present analysis reveals that guidelines do indeed pronounce normative ideals for medicine and may derive their authority as much from the collective expertise of their authors as from their authenticity as legitimate EBM products. Putting the criticisms of EBM's authoritarianism aside, we might trustingly accept the recommendations made by CPGs, given the expertise of their consensus panel members and the rigor of their methodology. However, to do so uncritically would be to ignore important concerns concerning the explicit reasoning within these guidelines, discussed below.

The question of what grounds clinical studies provide for decision making has been largely neglected, despite its significance [45–47], especially in the context of long-term preventative therapies, in which the individual practitioner's clinical experience provides less guidance. In practice guidelines, the logic of generalizing from the results of clinical studies to various patient populations is shown to be simple or crude induction, ultimately drawing on evidence from RCTs or meta-analyses of RCTs. Guidelines use statistical generalizations from these studies to make conclusions about target populations. In this argument type, 'clinical effectiveness' means not merely that the treatment caused the outcome in some patients, but that it did so in a clinically significant percentage of the population, reported as some measure of the average treatment effect.

Statistical generalizations from average treatment effects rest on the assumption that the sample (trial) population is representative of the sampled (target) population. A common statistical method used to achieve representativeness is random sampling. Although RCTs utilize 'randomization', they do not sample randomly from the target population. Instead, they use strict inclusion and exclusion criteria to enrol a trial population that is usually not representative of the target population [42]. Thus, extrapolating or generalizing from a proposition about the trial population to a conclusion about the target population risks a fallacy in inductive logic, namely *hasty generalization* or the *fallacy of biased statistics* [34].

Because effectiveness generalizations concern causal claims, an added requirement is that the causally relevant factors are fixed between the trial and target populations. In other words, the conditions of clinical practice must not differ systematically in relevant ways from the trial conditions under which the effect was investigated. Unfortunately, trial results do not reveal which external factors are important, and consideration to these factors was not routinely given in the guidelines reviewed, even though trial conditions are markedly different from the conditions of the clinic. In an RCT, patients may be monitored and cared for more intensely, may follow more stringent dietary restrictions, may show greater adherence to therapy, and perhaps most critical for the older demographic, are typically on fewer concurrent medications. As the trial population and trial conditions are not representative of the target population and practice conditions, the general argument type in Table 1 lacks the assumptions that would strengthen the generalization. Yet, this type of argument is commonly used to justify pharmacotherapy for several populations, including the elderly, within the guidelines for clinical practice.

If we wish to argue more convincingly, we must first establish what a claim for the clinical effectiveness of drugs literally asserts. According to Ashcroft, clinical effectiveness is a causal power, intrinsic to the treatment, for some outcome [48]. I would add that effectiveness is *effectiveness in certain patients* in a given population. RCT evidence does not allow us to predict with confidence which particular patients will experience benefit. As Cartwright has shown, the only kind of causal conclusion that an RCT provides is an 'it-works-somewhere' claim [49,50]. Getting from 'it-works-somewhere' to 'it-will-work-for-us' requires strong auxiliary assumptions not found within the argument types proposed in Table 1 or Table 4. When these assumptions are missing, as in an inference about the effectiveness of drugs in older patients that is based on effectiveness in trials enrolling adults, hasty induction results.

One strong premise that, if true, would strengthen the above inference is that the older population is representative of the adult population in all ways relevant to a particular outcome. This premise is difficult to sustain, given the unique physiology of older individuals and their altered pharmacokinetic and pharmacodynamic drug responses compared with younger individuals [51,52]. Aside from these physiological changes associated with healthy ageing, older patients have a higher burden of co-morbidity, cognitive impairment and functional impairment [6,53,54], which may similarly affect their response to treatment.

As Upshur noted, when we attend narrowly to RCT evidence that has poor external validity on its own, the result is a privileging of inferential gaps in treatment decision making [47]. The lack of respect given to these gaps within evidence-based guidelines is not surprising, considering that their methodology is concerned with grading evidence rather than grading arguments. A myopic view to the internal validity of study designs forgets that evidence carries no special power to make inferences on its own [47]. Rather, it falls to the arguer to cross the inferential gap between the evidence and the conclusion.

A limitation of the present analysis is that the justification for the recommendations within CPGs may not equate with any reasoning that supported these recommendations during the guideline-development process. Consistent with the persuasive role of argumentation in the guideline context, authors may have chosen the arguments because they are simple and palatable, and thus more convincing to the average clinical reader. Perhaps stronger reasoning, based on a plurality of kinds of evidence, takes place at the consensus conferences. Two findings raise doubts about the use of pluralistic evidential reasoning regarding effectiveness. First, half of the CPGs limited their literature searches to RCTs and systematic reviews of RCTs [23,27,28]. Second, the hypertension CPG mentioned evidence from 'observational studies' that contradicted a treatment recommendation, only to set it aside in the light of higher quality evidence from a placebo-controlled RCT [40].

A second limitation of this analysis is that it only considered certain classes of medications and the Canadian guidelines onto which these classes could be mapped. However, the medications considered here are standard treatment for the management of chronic conditions that are common globally [2], and the medications comprise a significant portion of total prescriptions for the elderly [11]. The lack of variation in how chronic disease guidelines reasoned suggests that their basic argument may not depend on the medication class. Further, because guideline panels review the same international evidence base and use similar established mechanisms for grading and selecting evidence, it is also expected that the pattern of argumentation does not differ substantially across countries, although future analysis would be needed to verify this prediction. The arguments typified by these CPGs may thus provide a glimpse into a statistical induction style of reasoning prevalent in evidence-based guidelines and, more broadly, evidence-based practice.

It is notable that none of the CPGs reviewed were generated through the GRADE (Grades of Recommendation, Assessment, Development and Evaluation) approach to rating quality of evidence and grading strength of recommendations [55]. In this approach, rather than evaluating evidence according to a strict hierarchy of research designs in which RCTs always trump observational studies, the rater partly determines quality of evidence via other important considerations, including the degree of similarity between the study population and the target population, as well as the similarity between the study intervention and the intervention in practice [56]. Of the CPGs analysed, only two considered whether a recommendation was an extrapolation based on a dissimilar trial population as part of their grading schemes [23,25]. Future studies could examine the influence of more sophisticated approaches to appraising evidence on the argumentation in guidelines.

I have not ruled out the possibility that effectiveness claims in the elderly might be justified by means of more complete arguments than those provided in guidelines. Yet, even if these individual prescribing decisions can be justified singularly, the justification required is more demanding in cases of multiple prescribing due to the resulting medication regimen complexity and attendant problems [57]. Complex regimens have been associated with poor adherence [58–60], unmanageable treatment burden for the patient [61,62] and harms due to treatment conflicts when multiple CPGs are followed at once [7,8]. Despite these concerns and those raised by the preceding analysis regarding the effectiveness of several classes of drugs in the elderly, in one Canadian province these drugs are increasingly prescribed to older patients [11], and prescribing to the elderly is also on the rise in England [9] and the United States [10].

Practice guidelines reason about patient populations, not individual patients. Thus, there is an inherent propensity to overgeneralize, with the trust that individual prescribers will not apply recommendations indiscriminately and will use their own careful judgement. We need more data on how guideline users reason from guideline recommendations and from the evidence base upon which CPGs are founded. A qualitative study of primary care clinicians' experiences with treatment decision making for older persons with multiple chronic conditions in the United States revealed variability in beliefs among participants regarding the benefits of guideline-directed care and the relevance of the evidence base for older patients [63].

In the meantime, in order to encourage the conscientious and judicious use of clinical evidence, CPGs should:

- Aspire to be well-reasoned rather than simply evidence-based. Simple inductions from 'high-quality' evidence may be persuasive, but they make for a weak argument. Causal claims based on RCT average treatment effects hold for the trial population only. Further assumptions are needed to assert that a treatment is

clinically effective in the target population or a sub-population of the target. More consideration should be given to inferential gaps in guideline-driven, evidence-based care.

• Include a plurality of evidence, where available, in their reasoning. While the RCT may be less prone to various kinds of bias, groups studied via other observational and experimental designs may be more representative of the general population, an important factor for inferences of statistical generalization.

• Be wary of making hasty inductions. It is incumbent on the experts to model good clinical reasoning for the rest of the profession. Meanwhile, the philosophy of medicine stands to make important contributions by revealing weaknesses in common argument types.

• Limit the scope of recommendations to populations for which a reasoned argument for the clinical effectiveness and safety of treatment can be made. Guidelines may tend to over-generalize to quite inclusive populations when their evidence base consists of more narrow subsets of these populations. Older patients differ from the general population in ways that may affect their response to treatment.

• Generate consensus statements that are contingent and descriptive rather than law-like and prescriptive. Treatment decisions should always be contingent on the facts of the individual case at hand. While expert consensus statements on the evidence may be useful, when phrased prescriptively, they are susceptible to abuse. The creators of CPGs must accept some of the responsibility for how these guides are used and abused by regulatory forces and individual clinicians. Prescriptive, law-like recommendations are easily adopted by quality assurance schemes, potentially encouraging undue standardization in the way various patient populations are managed. Sweeping generalizations that certain drugs 'should be used' in a large target population are a prescription for over-prescribing.

Acknowledgements

I am thankful for the comments and suggestions from two anonymous reviewers. I am grateful for the support from the McLaughlin Centre, the Canadian Institutes of Health Research and the Wilson Centre for Research in Health Professions Education. I have no financial interests to disclose.

References

1. Centers for Disease Control and Prevention (2003) Trends in aging – United States and worldwide. *Morbidity and Mortality Weekly Report*, 52 (6), 101–104, 106.
2. World Health Organization (2011) Global Status Report on Noncommunicable Diseases 2010. Geneva: WHO.
3. Health Council Canada (2007) Why Health Care Renewal Matters: Learning From Canadians with Chronic Health Conditions. Toronto: Health Council Canada.
4. Denton, F. T. & Spencer, B. G. (2010) Chronic health conditions: changing prevalence in an aging population and some implications for the delivery of health care services. *Canadian Journal of Aging*, 29 (1), 11–21.
5. Snaedal, J. (2011) Person centred medicine for the older patient, with specific reference to the person with dementia. *International Journal of Person Centered Medicine*, 1 (1), 53–55.
6. Upshur, R. E. & Tracy, S. (2008) Chronicity and complexity: is what's good for the diseases always good for the patients? *Canadian Family Physician*, 54 (12), 1655–1658.
7. Boyd, C. M., Darer, J., Boult, C., Fried, L. P., Boult, L. & Wu, A. W. (2005) Clinical practice guidelines and quality of care for older patients with multiple comorbid diseases: implications for pay for performance. *Journal of the American Medical Association*, 294 (6), 716–724.
8. Caughey, G. E., Roughead, E. E., Shakib, S., Vitry, A. I. & Gilbert, A. L. (2011) Co-morbidity and potential treatment conflicts in elderly heart failure patients: a retrospective, cross-sectional study of administrative claims data. *Drugs & Aging*, 28 (7), 575–581.
9. Cole, A. (2008) Over 60s′ use of prescription drugs has doubled in past decade in England. *British Medical Journal*, 337, a1132.
10. Qato, D. M., Alexander, G. C., Conti, R. M., Johnson, M., Schumm, P. & Lindau, S. T. (2008) Use of prescription and over-the-counter medications and dietary supplements among older adults in the United States. *Journal of the American Medical Association*, 300 (24), 2867–2878.
11. Bajcar, J. M., Wang, L., Moineddin, R., Nie, J. X., Tracy, C. S. & Upshur, R. E. (2010) From pharmaco-therapy to pharmaco-prevention: trends in prescribing to older adults in Ontario, Canada, 1997–2006. *BMC Family Practice*, 11, 75.
12. Guyatt, G., Rennie, D., Meade, M. O. & Cook, D. J. (2008) Users' Guides to the Medical Literature: Essentials of Evidence-Based Clinical Practice, 2nd edn. New York: McGraw-Hill Medical.
13. Tinetti, M. E., Bogardus, S. T. Jr & Agostini, J. V. (2004) Potential pitfalls of specific disease guidelines for patients with multiple conditions. *New England Journal of Medicine*, 351 (27), 2870–2874.
14. National Institutes of Health (last updated 7 February 2013) *About the consensus development program*. Available at: http://prevention.nih.gov/cdp/about.aspx (last accessed 30 April 2012).
15. Solomon, M. (2011) Group judgment and the medical consensus conference. In Handbook of the Philosophy of Science. Volume 16: Philosophy of Medicine (eds D. M. Gabbay, P. Thagard & J. Woods), pp. 239–254. Amsterdam: Elsevier BV.
16. Christof, R. (2010) *Aristotle's rhetoric*. Available at: http://plato.stanford.edu/entries/aristotle-rhetoric/#means (last accessed 28 March 2012).
17. Murphy, E. A. (1997) The Logic of Medicine, 2nd edn. Baltimore, MD: Johns Hopkins University Press.
18. Groopman, J. (2007) How Doctors Think. Boston, MA: Houghton Mifflin Co.
19. Montgomery, K. (2006) How Doctors Think. New York: Oxford University Press.
20. Straus, S. E., Glasziou, P., Richardson, W. S. & Haynes, R. B. (2011) Evidence-Based Medicine: How to Practice and Teach It, 4th edn. New York: Churchill-Livingstone.
21. Sackett, D. L., Rosenberg, W. M., Gray, J. A., Haynes, R. B. & Richardson, W. S. (1996) Evidence-based medicine: what it is and what it isn't. *British Medical Journal*, 312, 71–72.
22. Mutasingwa, D. R., Ge, H. & Upshur, R. E. (2011) How applicable are clinical practice guidelines to elderly patients with comorbidities? *Canadian Family Physician*, 57 (7), e253–e262.
23. Rabi, D. M., Daskalopoulou, S. S., Padwal, R. S., *et al*. (2011) The 2011 Canadian Hypertension Education Program recommendations for the management of hypertension: blood pressure measurement, diagnosis, assessment of risk, and therapy. *Canadian Journal of Cardiology*, 27 (4), 415–433.
24. Arnold, J. M. O., Liu, P., Demers, C., *et al*. (2006) Canadian Cardiovascular Society consensus conference recommendations on heart failure 2006: diagnosis and management. *Canadian Journal of Cardiology*, 22 (1), 23–45.

25. Buchanan, D., Tourigny-Rivard, M. -F., Cappeliez, P., *et al*. (2006) National guidelines for seniors' mental health: the assessment and treatment of depression. *Canadian Journal of Geriatrics*, 9 (Suppl 2), S52–S58.
26. Armstrong, D., Marshall, J. K., Chiba, N., *et al*. (2005) Canadian consensus conference on the management of gastroesophageal reflux disease in adults – update 2004. *Canadian Journal of Gastroenterology*, 19 (1), 15–35.
27. Genest, J., McPherson, R., Frohlich, J., *et al*. (2009) 2009 Canadian Cardiovascular Society/Canadian guidelines for the diagnosis and treatment of dyslipidemia and prevention of cardiovascular disease in the adult – 2009 recommendations. *Canadian Journal of Cardiology*, 25 (10), 567–579.
28. Papaioannou, A., Morin, S., Cheung, A. M., *et al*. (2010) 2010 clinical practice guidelines for the diagnosis and management of osteoporosis in Canada: summary. *Canadian Medical Association Journal*, 182 (17), 1864–1873.
29. Canadian Cardiovascular Society (2012) *About CCS*. Available at: http://www.ccs.ca/about_ccs/index_e.aspx (last accessed 28 March 2012).
30. Hypertension Canada (2012) *Recommendations*. Available at: http://www.hypertension.ca/chep-recommendations (last accessed 28 March 2012).
31. Kuypers, J. A. & King, A. (2009) What is rhetoric? In Rhetorical Criticism: Perspectives in Action (ed. J. A. Kuypers), pp. 1–12. Lanham, MD: Lexington Books.
32. Hill, F. I. (2009) The 'traditional' perspective. In Rhetorical Criticism: Perspectives in Action (ed. J. A. Kuypers), pp. 39–61. Lanham, MD: Lexington Books.
33. Walton, D. (1996) Argument Structure: A Pragmatic Theory. Toronto: University of Toronto Press.
34. Walton, D. (1989) Informal Logic: A Handbook for Critical Argumentation. New York: Cambridge University Press.
35. Walton, D. (1990) What is reasoning? What is an argument? *Journal of Philosophy*, 87, 399–419.
36. Canadian Institutes of Health Research (2011) *More about knowledge translation at CIHR*. Available at: http://www.cihr-irsc.gc.ca/e/39033.html (last accessed 30 April 2012).
37. World Health Organization (2005) Bridging the 'Know–Do' Gap: Meeting on Knowledge Translation in Global Health. Geneva: World Health Organization.
38. Mitton, C., Adair, C. E., McKenzie, E., Patten, S. B. & Waye Perry, B. (2007) Knowledge transfer and exchange: review and synthesis of the literature. *The Milbank Quarterly*, 85 (4), 729–768.
39. Khan, N. A., Hemmelgarn, B., Herman, R. J., *et al*. (2008) The 2008 Canadian Hypertension Education Program recommendations for the management of hypertension: part 2 – therapy. *Canadian Journal of Cardiology*, 24 (6), 465–475.
40. Khan, N. A., Hemmelgarn, B., Herman, R. J., *et al*. (2009) The 2009 Canadian Hypertension Education Program recommendations for the management of hypertension: part 2 – therapy. *Canadian Journal of Cardiology*, 25 (5), 287–298.
41. Upshur, R. E. G. & Tracy, C. S. (2004) Legitimacy, authority, and hierarchy: critical challenges for Evidence-based medicine. *Brief Treatment and Crisis Intervention*, 4, 197–204.
42. Bluhm, R. & Borgerson, K. (2011) Evidence-based medicine. In Handbook of the Philosophy of Science. Volume 16: Philosophy of Medicine (eds D. M. Gabbay, P. Thagard & J. Woods), pp. 203–238. Amsterdam: Elsevier BV.
43. Goldenberg, M. J. (2009) Iconoclast or creed? Objectivism, pragmatism, and the hierarchy of evidence. *Perspectives in Biology and Medicine*, 52 (2), 168–187.
44. Rosenfeld, J. A. (2004) The view of evidence-based medicine from the trenches: liberating or authoritarian? *Journal of Evaluation in Clinical Practice*, 10 (2), 153–155.
45. Feinstein, A. R. & Horwitz, R. I. (1997) Problems in the 'evidence' of 'evidence-based medicine'. *American Journal of Medicine*, 103, 529–535.
46. Horton, R. (2000) Common sense and figures: the rhetoric of validity in medicine (Bradford Hill Memorial Lecture 1999). *Statistics in Medicine*, 19, 3149–3164.
47. Upshur, R. E. (2005) Looking for rules in a world of exceptions: reflections on evidence-based practice. *Perspectives in Biology and Medicine*, 48 (4), 477–489.
48. Ashcroft, R. (2002) What is clinical effectiveness? *Studies in History and Philosophy of Biological and Biomedical Sciences*, 33 (2), 219–233.
49. Cartwright, N. (2011) A philosopher's view of the long road from RCTs to effectiveness. *Lancet*, 377, 1400–1401.
50. Cartwright, N. & Munro, E. (2010) The limitations of randomized controlled trials in predicting effectiveness. *Journal of Evaluation in Clinical Practice*, 16 (2), 260–266.
51. Hilmer, S. N., McLachlan, A. J. & Le Couteur, D. G. (2007) Clinical pharmacology in the geriatric patient. *Fundamental and Clinical Pharmacology*, 21 (3), 217–230.
52. ElDesoky, E. S. (2007) Pharmacokinetic-pharmacodynamic crisis in the elderly. *American Journal of Therapeutics*, 14 (5), 488–498.
53. Jadad, A. R., To, M. J., Emara, M. & Jones, J. (2011) Consideration of multiple chronic diseases in randomized controlled trials. *Journal of the American Medical Association*, 306 (24), 2670–2672.
54. Fusco, D., Lattanzio, F., Tosato, M., Corsonello, A., Cherubini, A., Volpato, S., Maraldi, C., Ruggiero, C. & Onder, G. (2009) Development of criteria to assess appropriate medication use among elderly complex patients (CRIME) project. *Drugs and Aging*, 26 (Suppl 1), 3–13.
55. Guyatt, G. H., Oxman, A. D., Schünemann, H. J., Tugwell, P. & Knottnerus, A. (2011) GRADE guidelines: a new series of articles in the Journal of Clinical Epidemiology. *Journal of Clinical Epidemiology*, 64 (4), 380–382.
56. Guyatt, G. H., Oxman, A. D., Kunz, R., *et al*. (2011) GRADE guidelines: 8. Rating the quality of evidence –indirectness. *Journal of Clinical Epidemiology*, 64 (12), 1303–1310.
57. Fuller, J. & Upshur, R. E. (2011) Medication regimen complexity and the care of the chronically ill patient. *International Journal of Person Centered Medicine*, 1 (4), 719–725.
58. Corsonello, A., Pedone, C., Lattanzio, F., Lucchetti, M., Garasto, S., Carbone, C., Greco, C., Fabbietti, P. & Incalzi, R. A. (2009) Regimen complexity and medication nonadherence in elderly patients. *Journal of Therapeutics and Clinical Risk Management*, 5 (1), 209–216.
59. Ingersoll, K. S. & Cohen, J. (2008) The impact of medication regimen factors on adherence to chronic treatment: a review of literature. *Journal of Behavioral Medicine*, 31 (3), 213–224.
60. Osterberg, L. & Blaschke, T. (2005) Adherence to medication. *New England Journal of Medicine*, 353 (5), 487–497.
61. May, C., Montori, V. M. & Mair, F. S. (2009) We need minimally disruptive medicine. *British Medical Journal*, 339, b2803.
62. Gallacher, K., May, C. R., Montori, V. M. & Mair, F. S. (2011) Understanding patients' experiences of treatment burden in chronic heart failure using normalization process theory. *Annals of Family Medicine*, 9 (3), 235–243.
63. Fried, T. R., Tinetti, M. E. & Iannone, L. (2011) Primary care clinicians' experiences with treatment decision making for older persons with multiple conditions. *Archives of Internal Medicine*, 171 (1), 75–80.

