International Political Sociology (2023) **17**, 1–22

Methods Regimes in Global Governance: The Politics of Evidence-Making in Global Health

Annabelle Littoz-Monnet and Juanita Uribe
Graduate Institute, Geneva, Switzerland

This article opens up the blackbox through which evidence is selected and assessed in the making of guidelines and recommendations in global governance, through an exploration of "methods regimes." Methods regimes are a special kind of sociomaterial arrangement, which govern the production and validation of knowledge, by establishing a clear hierachy between alternative forms of research designs. When such regimes become inscribed in processes of global governance, they shape and control what knowledge is deemed valid and thus relevant for policy. We shed light that through a mode of operation that relies on a discourse of procedurality, a dispersed but powerful network of epistemic operators, and a dense web of infrastructures, methods regimes constitute and police the making of "policy-relevant knowledge" in global governance. Through an examination of the case of "GRADE" (Grading of Recommendations, Assessment, Development, and Evaluation), a standardized system that evaluates and grades the quality of evidence in global health, we show that its dominance has worked to the effect of empowering a new cast of methodologists, seen as more objective and portable across domains, sidelining certain forms of evidence that do not conform with its own methodological criteria of scientificity, and "clinicalizing" research in medicine and beyond.

Cet article ouvre la « boîte noire » de la sélection et de l'évaluation des connaissances existantes lors de l'élaboration de lignes directrices et de recommandations en gouvernance mondiale, en s'intéressant aux « régimes de méthodes ». Les régimes de méthodes, des agencements sociomatériels avec leurs caractéristiques propres, gouvernent la production et la validation des connaissances en établissant une échelle hiérarchique très claire entre différents types de méthodes. Lorsque ces régimes deviennent ancrés dans les mécanismes de gouvernance mondiale, ils façonnent et déterminent les connaissances considérées comme valides, et donc pertinentes pour l'élaboration d'une politique. Nous mettons en lumière que, par l'intermédiaire d'un mode opératoire fondé sur un discours procédural, un réseau diffus, mais influent, d'opérateurs épistémiques et un maillage dense d'infrastructures, les régimes de méthodes déterminent et régissent l'élaboration des « connaissances politiques pertinentes » en gouvernance mondiale. Nous étudions le cas de « GRADE » (Grading of Recommendations, Assessment, Development and Evaluation) un système normalisé qui évalue et note les études scientifiques existantes dans le domaine de la santé. Nous montrons que sa prépondérance a permis l'émergence et le renforcement d'un nouveau groupe de méthodologistes, considérés plus objectifs et polyvalents, la mise à l'écart de certains types de données non conformes à ses propres critères méthodologiques de qualification scientifique et de

Littoz-Monnet, Annabelle, and Juanita Uribe. (2023) Methods Regimes in Global Governance: The Politics of Evidence-Making in Global Health. *International Political Sociology*, https://doi.org/10.1093/ips/olad005
Corresponding author e-mail: annabelle.littoz-monnet@graduateinstitute.ch

« cliniciser » la recherche, en médecine mais également dans d'autres domaines.

Este artículo abre la caja negra a través de la cual se seleccionan y evalúan las pruebas para la elaboración de directrices y recomendaciones en materia de gobernanza mundial, mediante una exploración de los «regímenes metodológicos». Los regímenes metodológicos son un tipo especial de arreglo socio-material, relacionado con la producción y validación de conocimiento, sobre la base de su diseño de investigación. Cuando estos regímenes se integran en los procesos de gobernanza mundial, configuran y controlan el conocimiento que se considera válido y, por tanto, relevante para la política. A través de un modo de operar que se sustenta en un discurso procedimental, en una red dispersa pero poderosa de operadores epistémicos y en una densa red de infraestructuras, los regímenes metodológicos constituyen y supervisan la creación de «conocimiento relevante para las políticas» en la gobernanza global. A través de un examen del caso de «GRADE» (Clasificación de Recomendaciones, Valoración, Desarrollo y Evaluación de Pruebas), un sistema estandarizado que clasifica la calidad de las pruebas en el ámbito de la salud mundial, mostramos que su predominio ha servido para empoderar a un nuevo elenco de metodólogos, considerados más objetivos y transferibles entre dominios, dejando de lado ciertos tipos de pruebas que no se ajustan a sus propios criterios metodológicos de cientificidad, y «clinicalizando» la investigación en medicina y fuera de ella.

The production of guidelines, recommendations, best practices, and other forms of guidance makes for the daily business of global governance. International organizations (IOs) produce a plethora of documents that aim at guiding political action in domains as diverse as education, development, and health, to name a few. Such guidelines, the word goes, are always based on "the best available evidence" and "systematic" reviews of existing research. We know little, however, about what evidence is considered to be "best" and what existing criteria of validity and scientificity are adopted when such evidence is selected, assembled, and assessed in global governance fora.

The politics of methods are at the heart of the processes through which certain forms of evidence come to be seen as valid, scientific, and relevant, while others are dismissed. Methods seem mechanical, procedural, and somehow neutral. Yet, they embody specific epistemological assumptions and judgments about what constitutes "science." Largely discussed in the sociology of knowledge literature (Knorr Cetina 2007; Shapin and Schaffer 2011), such situated ways of delineating what is scientific can also become "methods regimes" when they are enacted, inscribed, and materialized in the institutions, processes, and practices of governance. Specific conceptions of methods have indeed become hegemonic in global governance, such as experimental methods in international development, risk assessment for food policy, or auditing practices (Timmermans and Angell 2001; Jatteau 2013; Donovan 2018). At the same time, a new cast of professionals, the self-labeled "methodologists" of global health, but also the auditors, or the risk analysts, of other fields, has emerged.

Methods regimes are a special kind of sociomaterial arrangement, which have to do with the *production* and *validation* of knowledge, rather than knowledge claims that directly delineate the way specific objects of global governance are constituted. By designating which methods of knowledge production are "right," they directly regulate and control the kind of evidence that is deemed to be accurate and relevant for the governance of global problems. Methods regimes, thus, are at the core of the politics of "evidence-based" policy-making. Yet, their role in the making of "policy-relevant evidence" and, ultimately, global recommendations, guidelines,

or best practices has gone largely unnoticed. Making their presence and role in global governance more explicit and visible is, thus, a necessary intervention that brings new insights into how global recommendations and standards of all sorts are delineated. We ask, thus, two interrelated questions. What makes it possible for methods regimes to operate and gain traction in global governance? And what are their specific effects on the politics of "evidence," and knowledge-making, in given governance domains?

We explore these questions through an in-depth study of the case of "GRADE" (Grading of Recommendations, Assessment, Development, and Evaluation), a formal and standardized system that grades the quality of evidence and the strength of recommendations in global health.[1] International relations' (IR) conventional disciplinary view of health as a secondary matter of international politics has limited its engagement with this domain, often limiting its focus to questions of effectiveness (Beisheim and Liese 2014). While critical scholars in IR are today paying more attention to global health (Elbe 2014; Hanrieder 2019), they have not focused on the politics of evidence-based policy, or expertise (but see Littoz-Monnet 2022). GRADE has been adopted by about 110 organizations worldwide, and most notably by the World Health Organization (WHO), as the framework for developing its guidelines and recommendations. In the words of the WHO, which develops guidelines on topics as diverse as pharmacological interventions, sugar intake, or vaccine introduction, its own "legitimacy and technical authority lie in its rigorous adherence to the systematic use of evidence as the basis for all policies" (WHO 2014, 1). The way evidence is *produced*, *evaluated*, and *synthetized* thus plays a crucial role in the final content and shape of recommendations, and constitutes the field of global health as well as related scientific inquiry.

We argue that GRADE acts as a powerful methods regime through its specific characteristics and mode of operation. GRADE, first, is a (seemingly) *procedural* regime, in the sense that it sets and stabilizes hierarchies among different forms of knowledge through a pre-established evidence scale, which ranks evidence on the basis of the procedures and study design used for its production. Although GRADE embodies clear epistemological claims about what forms of evidence are more valuable than others, it seems to be just a matter of "mechanical steps" to be followed. The apparent procedurality of GRADE renders it natural and harder to contest. GRADE is also enacted and sustained through a highly *dispersed* network of epistemic operators, professional methodologists, and their knowledge sites, which not only produce and perpetuate the regime's claims, but also act to control, verify, and validate what forms of evidence are relevant for the purpose of policy. GRADE as a regime is also *embedded*, or materially inscribed in global governance processes through a complex web of material objects, such as evaluation procedures, learning modules, a standardized "GRADE CV," or yet computer programs that embody and reproduce its core assumptions. This mode of operation makes it possible for GRADE to shape the production of policy-relevant knowledge, and the field of global health more broadly, in a highly effective way.

Our observations point to three kinds of effects. We argue, to begin, that GRADE has contributed to a reconfiguration of hierarchies in global health governance. It has, in particular, sustained a displacement of epistemic authority away from what global health professionals refer to as "content experts"—in general doctors with experience in their area of specialization—toward an emerging transnational cast of methodologists, seen as more objective and capable of working across domains. In addition, through its formalized standards, GRADE has worked to the effect of hierarchizing knowledge in global health. Those forms of knowledge that do not rank high according to GRADE, such as observational studies or case reports, are

[1] Although there are a number of tools that help support guideline development, we focus on GRADE because of its dominance worldwide and its adoption by the WHO.

often disregarded or seen as anecdotal. Finally, GRADE has fostered the clinicalization of research in medicine, but also beyond, spreading out a mode of reasoning that only accounts for the efficacy of given interventions, while broader questions are barely addressed.

Methods regimes in global governance act as a form of police that authorizes certain forms of evidence and discredits others. Given their "epistemic" nature, which has to do with the production and validation of evidence, methods regimes operate in a distinct way. Through claims that are seemingly procedural, void of content, and neutral, a dispersed but powerful network of epistemic operators, and a dense web of material inscriptions that perpetuate the regime and give it a life of its own, methods regimes act as a special kind of sociomaterial arrangement, which powerfully shapes the *production* and *warranting* of knowledge in global governance. As such, they act as epistemological ways of governing, in the sense that they delineate how problems can be known and thus, ultimately, acted upon.

The article relies on an in-depth case study work. Through immersion with the details of the case, we explored GRADE's assumptions and rhetoric, its network of operators, and its material inscriptions. This was done, first, through an extensive textual analysis of documents from the WHO, as well as a number of scientific articles by prominent GRADE methodologists in scientific and medical journals such as *The Lancet*, *The British Medical Journal*, and the *Journal of Clinical Epidemiology*. Documents from the WHO included Chapters 9 and 10 from the 2014 WHO Handbook for Guideline Development, which provide instructions on evidence assessment and the methodological standards for the development of recommendations. Second, through an exploration of the websites of those centers strongly linked with the GRADE Working Group, such as McMaster University, the Cochrane centers, and the WHO's Guidelines Review Committee (GRC), we mapped the epistemic operators of GRADE and traced relationships between them, as well as their ties with other actors or sites that, although not sitting at the core of the network, also promote GRADE. Third, we observed the materiality of GRADE, looking at its teaching modules, the "GRADE CV," or the MagicApp, examining both their content and form. Third, we conducted sixteen semi-structured interviews between October 2020 and May 2022 with WHO officials, including methodologists and members of the GRC, medical practitioners, the founders of GRADE, and the members of both the GRADE Working Group and Cochrane.

Methods always involve selecting, assembling, and interpreting facts (Ruppert and Scheel 2019), and so do our own. With our own observations, we do not aim to make truth claims, but rather expand existing methodological repertoires, so that policy knowledge, as well as academic practice, moves beyond the narrowness of scope of the current clinical "dogma" not only in global health, but also in other domains such as development for instance. We do not see objectivity as a sort of capacity to "remove bias," but as an attitude toward research, which is capable of resorting to heterogeneous knowledge sources and recognizes multiple meanings (Leander 2016). In unpacking, and making more explicit, the politics at play in the inclusion or exclusion of certain forms of knowledge in global governance, we aim to enhance understandings of expert discourses and open up possibilities for methodological pluralism, more diverse forms of expertise, and a more inclusive governance architecture.

Expert Knowledge and Global Governance

Claims about the "evidence-based" nature of global agendas and interventions abound in global governance. Whether in health, climate, education, or development aid, global policies are presented and legitimized by reference to their reliance on the "best" available evidence. Although the role of evidence in governance has been widely debated in the public policy literature

(Davies and Nutley 2000; Parkhurst 2017), scholarship in IR has not engaged much with the politics associated with claims to evidence-based policy. While scholars have focused on the construction or assembling of certain forms of expertise (Bueger 2015; Sending 2015; Littoz-Monnet 2020; Aue 2021), we do not know much about how certain types of knowledge come to be seen as evidence and are assembled and translated into policy recommendations. The role of methods in such processes, although central to the establishment of an evidence-based style of governing in global governance, has also not been explored.

Partly, this is due to an entrenched preoccupation, in the field of IR, with the way scientists or experts influence policy-makers (Haas 1992). If IOs rely on solid scientific evidence, the word goes, their agendas and programs will also be more sound (Haas and Stevens 2011). As a result, this approach has been leaving aside the question of how such evidence is assembled and evaluated in the first place. Other accounts in IR have focused on how IOs mobilize certain forms of expertise as a way to assert their authority, revealing that by "emphasizing the objective nature of their knowledge, staff of IOs are able to present themselves as technocrats whose advice is unaffected by partisan squabbles" (Barnett and Finnemore 2004, 24; Liese et al. 2021). Yet, this body of work concentrates on the politics of knowledge mobilization, rather than knowledge production. In addition, it focuses exclusively on IOs, missing on the way IOs are embedded in ecosystems of knowledge production that involve a complex web of actors, sites, and infrastructures.

IR scholars have, in recent years, examined in greater detail the processes and practices of knowledge-making in global governance (Bueger 2018; Leander and Wæver 2018). Such research has looked into the fabric of expertise (Best 2014; Bueger 2015, 2018) and the relationships that connect together actors, actions, and material objects (Latour 2005; Law 2008). Taking inspiration from the fields of sociology and Science and Technology Studies (STS), this body of work has shed light on the epistemic and also sociomaterial components that embody and perpetuate certain ways of knowing things. Scholars have examined how certain orthodoxies or discourses, embodied in sociomaterial arrangements, constitute subjects (Lakoff and Collier 2008; Towghi and Vora 2014) and objects of governance (Aue 2021). Recent scholarship in IR has built upon the concept of "assemblage" (Bueger 2018; Ruppert and Scheel 2019) to qualify those patterned arrangements that connect discursive and nondiscursive elements. Assemblage thinking focuses on how constituencies, concepts, techniques, and material objects hold together, forming arrangements that not only enact certain realities but also produce actors, objects, and power relationships (Law 2004; Bueger 2015; Leander and Wæver 2018; Ruppert and Scheel 2019).

Our approach not only builds upon these insights, but also departs from them in two ways. First, we shift away from a focus on the fluidity and ever-changing character of knowledge arrangements. In reinstating the notion of "regime" rather than that of assemblage, for instance, we aim to better capture the stability and resilience of sociomaterial arrangements. Recent practice and assemblage scholarship often leave aside questions of resources and hierarchies; when hierarchy is conceptualized, it is in relation to—and by means of—other actors (Latour 2005) and thus also amenable to fluctuations. However, resources and hierarchies delineate what kind of assembling or ordering of expertise is possible and make them more stable and self-perpetuating than often assumed. In turn, sociomaterial arrangements also produce specific forms of politics and hierarchies. In reintroducing hierarchies and resources into the picture, we shed light on the conditions of possibility of *specific* forms of assemblage, as well as of their stability.

Second, in focusing on methods in global governance, we make a shift away from studying knowledge arrangements to studying the *epistemic conditions that shape the production of knowledge*. We examine a special kind of arrangement that operates at the level of the production, selection, and assembling of "policy evidence"—a

form of knowledge that in turn makes possible the emergence of certain knowledge truths and not others. Like the "paradigms," "research programs," or yet "epistemic cultures" described in the STS literature on knowledge production (Kuhn 1962; Lakatos 1978; Knorr Cetina 1999), methods regimes are epistemic in that they have to do with the validation of knowledge. The politics of methods have been brought to the fore in IR, as critical accounts have made it possible to see that supposedly neutral methods in fact embody political visions and knowledge claims (Aradau and Huysmans 2014; Ruppert and Scheel 2019). Yet, methods have effects beyond their role in specific academic disciplines or laboratories. They might also take public authority and get inscribed, enacted, and materialized in global governance processes. While scholars have started to engage with the dominance of certain epistemological assumptions and their associated knowledge techniques in global governance, such as an overreliance on quantitative forms of data, like measurements and rankings (Hansen and Porter 2012; Cooley 2015; Aue 2021), digital techniques of data production (Ruppert and Scheel 2019), and randomized controlled trials (RCTs) (Kelly and McGoey 2018), we know little about the kinds of categorizations and hierarchies that inform the production and selection of the evidence used to formulate global policy recommendations. In focusing on methods regimes, we shed light on the political work that shapes the delineation of the standards and recommendations of all sorts, which abound in global fora and are instrumental to global ways of governing (Broome and Seabrooke 2012; Hearson 2018; Doshi, Kelley, and Simmons 2019).

Methods Regimes

Methods regimes are a special kind of social and material machinery. They are epistemic, as they have to do with the *production* and *validation* of knowledge; they consist of the seemingly neutral procedures, networks of operators, and material objects that shape the selection and assembling of the evidence that counts in global governance. Because methods regimes act at an epistemic level, they have specific characteristics and a special mode of operation. They are *procedural*, in the sense that they work through a prespecified set of criteria that determine "how to" produce and evaluate knowledge, thus not making explicit their more substantial knowledge claims. They are also *dispersed*, in that they are enacted and sustained by a network of methodologists and their knowledge centers located throughout the globe. They are *embedded*, in that they are inscribed in global governance processes through a web of material inscriptions, which embody and stabilize their core assumptions. This mode of operation makes it possible for methods regimes to gain traction in governance and, as a result, to have constitutive and policing effects on the politics of evidence-making in global health governance.

Procedural: Unlike other sociomaterial arrangements, or knowledge regimes, a methods regime does not rely on substantial knowledge claims, such as how to address poverty or unemployment, for instance; instead, it consists of prespecified sets of criteria, "how-to" technical guidelines, and methodological techniques. It seems empty of any substantial principles and relies on a discourse of detachment toward what it fashions as "content knowledge." Like other "technologies of truth," methods regimes have an appearance of impersonality (Porter 1995), yet they serve to set clear and fixed hierarchies among different forms of knowledge—on the basis of the study design through which these have been produced. Methods regimes are procedural in that they distinguish "high-quality" research that produces reliable evidence from weak designs that produce "poor" quality evidence. Acting as a "policing technology" (Leander 2020), they signal what procedures of knowledge production are "right," with the effect that knowledge not matching certain design criteria tends to be dismissed as anecdotal or biased. In doing so, and

despite their seemingly mechanical functioning, methods regimes embody strong knowledge claims. Beyond their centrality in the delineation of policy-relevant knowledge, methods regimes also indirectly shape the conduct of scientific research in certain academic fields, by enacting markers of what is scientific and what is not. Researchers either internalize the regime's principles or comply with them in order to make their studies publishable and relevant. Those who do not run the risk of not being considered as scientists at all (Adams 2002). Such principles in turn constrain the kind of questions that can be asked. In delineating what counts as valuable and relevant evidence for policy, methods regimes also shape what kind of evidence is produced at all.

Dispersed: A methods regime is enacted not *only* by public institutions, such as IOs, but also by a highly dispersed, yet connected, network of "epistemic operators," typically composed of knowledge sites, such as research clusters, knowledge centers, professional organizations, and high-profile journals and experts, which all together produce and reproduce the ideas of the regime in global governance. Epistemic operators can act as knowledge producers, or perform "rituals of verification" (Power 1999), through which they "inspect" or "scrutinize" what is objective or valid for the purpose of policy. Methods regimes are upheld in knowledge sites, where the regime's professionals work, but that also sustain the regime's ideas, as well as a specific set of relations, beyond the role of specific actors. In finance, for example, such sites can consist of rating agencies, which rate a country's economic "worth," the analyst departments of banks and investment companies, or accounting firms (Knorr Cetina 2007). In global health, they consist of influential research centers, which act as the epicenter of the regime, expert committees, networks of professional organizations, or even scientific journals (Littoz-Monnet 2022). There are hierarchies within the regime, so that powerful knowledge centers, endowed with financial, social, and epistemic resources, sit at its core. Knowledge centers serve as powerful "mediation sites" where knowledge "has to pass" (Bueger 2015) and link together a dense network of spatially dispersed partner organizations, which play a central role in perpetuating and reproducing the regime. Like scientific "laboratories" (Latour and Woolgar 1986), or "centers of calculation" (Latour 2005), such centers play a central role in not only producing the knowledge of the regime, but also standardizing and validating some of its "products," such as its courses, certifications, or codebooks—those inscriptions, or devices, which will be discussed below. Journals can also act as central epistemic operators; they not only are some sort of "outsprings" of methods regimes but also amplify their effects, in that they actively filter, validate, and circulate knowledge. More peripheral sites, often accredited by those sitting at the core of the regime, also fulfill a central function in sustaining the ideas of the regime in a more decentralized and dispersed fashion. Professionals also play a central role in producing the ideas of methods regimes and transporting these across locations. In such regimes, the "experts" are professionals endowed with "skills" rather than issue-specific knowledge. They are the self-labeled "methodologists" of global health, or the "risk-assessment analysts," the "auditors," we encounter elsewhere. Their skills, which typically have to do with the possession of a toolkit of procedures, or evaluation techniques, seem easily and endlessly transferable, from one location to the next. This dense web of epistemic operators, with its own experts, its prestigious and well-endowed knowledge centers, and its networks, but also strong links with high-profile journals, makes methods regimes powerful and effective.

Embedded: A methods regime is also embedded, in the sense that it is inscribed into a set of material objects that make its effects wide-ranging, entrenched, and self-perpetuating. In contrast to the regime's epistemic operators, which consist of people, networks, those sites where they meet, and their relationships, those objects

have a *material* dimension. They may consist of standardized courses, sets of guidelines, computer programs, documents, or yet apps. These "inscriptions," Latour and Woolgar tell us, translate the untidy work and discussions taking place in the knowledge centers of the regime into "written documents" (Latour and Woolgar 1986). Through their "techne" (Walters 2002), they embody, stabilize, and perpetuate certain ideas and meanings, while also having their own regularities and effects (Star 1999; Knorr Cetina 2007). A material object has its own form of existence that is "nested in and arises from its use" (Knorr Cetina 2007). They may of course be translated or transformed. Yet, their existence and portability are a central element of a regime's stability, in that material objects delineate what can be seen in a way that is hard to notice and endlessly reproducible. The inscriptions of a methods regime, whether its codebooks, or technological devices, are all "certified" by the regime's core operators and thus strongly embody the regime's ideas. At the same time, inscriptions function as "durable, mobile traces which can be transported between locales" (Walters 2002, 91). The codes, guidelines, routinely produced documents, courses, apps, and computer programs can be picked up and operated by anyone, anywhere, without any need for particular instructions. Thus, material objects allow actors or knowledge sites "to exert their power through things that don't sleep and associations that don't break down" (Sayes 2014, 140). When a regime is inscribed into written documents, syllabi, or devices, this helps the regime keep going, even when its ideas are contested.

GRADE as a Methods Regime

Brief Genealogy of GRADE

GRADE is a standardized system that categorizes and hierarchizes the quality or certainty of evidence and determines the strength of recommendations for healthcare interventions.[2] It consists of a scale for the evaluation of the quality and certainty of evidence, as well as a system for translating evidence into recommendations, called the "evidence to decision framework," which lists social criteria that are important to consider when making a decision, such as benefits and harms, feasibility, among others (Alonso-Coello et al. 2016). In the context of guideline development, the quality of the evidence is defined by the GRADE Working Group as "the extent to which one can be confident that an estimate of the effect or association is correct" (Balshem et al. 2011). The emergence of GRADE is directly enmeshed with evidence-based medicine (EBM). Not only is GRADE based on the same core ideational assumptions, but it has also been developed by those people and organizations that were the strongest proponents of EBM. GRADE strongly echoes EBM's empiricism—a belief that evidence speaks for itself—and the worth and efficacy of biostatistical techniques. Like EBM, GRADE also assumes that observer bias can be removed from research through good statistical techniques and effective randomization. The way theory and observation are intertwined, and even rigorous RCTs are shaped by human judgment at every stage of their design, and interpretation, is not much considered (Grossman and Mackenzie 2005). As a result, the diverse blends of criticism raised against EBM and RCTs, its golden tool, are not integrated into the making of evidence for global health (Adams 2002; Timmermans and Berg 2010; Hawe 2015).

Initially, the EBM movement emerged as a critique of the largely unchallenged authority of doctors and clinicians, in a context in which clinical epidemiology was establishing itself as a discipline throughout the 1980s (Landzelius 2006). A number of medical schools were starting to include courses in clinical epidemiology, with a

[2] GRADE classifies recommendations made in guidelines as either strong or weak. The strength of a recommendation reflects the extent to which guideline developers can be confident that the desirable effects of an intervention outweigh the undesirable effect (Guyatt et al. 2008).

strong focus on statistics and mathematics in their curricula. The British epistemologist Archie Cochrane, in his book *Effectiveness and Efficiency*, wrote a plea for the use of a new form of trials, which introduced the use of control groups to patient trials (what we know as RCTs) as well as systematic reviews of research results produced through RCTs on given medical issues (Daly 2005, 131–32). Here too, the claim was that access to what was fashioned as high-quality evidence would diminish doctors' reliance on personal intuition and experience.

It is not surprising, thus, that the GRADE Working Group, which produced the GRADE scale, emerged in 2000 at the Department of Clinical Epidemiology and Biostatistics at McMaster University, the alma mater of EBM. It was indeed at McMaster in Canada that a team of researchers played a crucial role in establishing courses on the systematic assessment of evidence in medical schools' curricula (Guyatt et al. 1992). The International Clinical Epidemiology Network (INCLEN) had by then started to create Clinical Epidemiology Units, designed to push for "the application of quantitative measurement principles … in the development of clinical and health care policy" (Hanemaayer 2016, 464). McMaster included INCLEN's training programs in the methods of clinical epidemiology; it was thus not accidentally that the first formal formulations of EBM appeared at McMaster and that GRADE also emerged there (Hanemaayer 2016). Gordon Guyatt, Andrew Oxman, and Holger Schu¨nemann, the then professors of medicine and/or clinical epidemiology at McMaster, created the GRADE Working Group, in order to develop what they saw as a standardized rating system to assess the quality of evidence for the purpose of making clinical practice recommendations.

Members of the GRADE working group actively participated in the introduction of GRADE within the WHO, as well as other health organizations.[3] Before the adoption of the system, so-called content experts, most often doctors specialized in a medical domain, were making recommendations based on their experience and clinical training. Doctors benefited from a high degree of epistemic authority, but their large autonomy. It was in this context that the WHO brought in Andrew Oxman, co-founder of GRADE, to reflect on how it could be more systematic in its use of evidence.[4] Oxman was proactive and soon published a piece in *The Lancet* where he accused the WHO of relying too heavily on "expert opinion," instead of evidence (Oxman, Lavis, and Fretheim 2007). It was in reaction to these charges that the WHO convened, in 2009, the GRC, tasked to ensure that WHO recommendations are "based on the best available evidence" (WHO 2022). The GRC was strongly enmeshed with the same people who had produced GRADE and had also acted as major proponents of EBM. Gordon Guyatt, Andrew Oxman, and Holger Schünemann, who were all part of the GRADE Working Group, played a central role in putting together the GRC and developing the Handbook for Guideline Development.[5] The GRADE system is now fully established within the WHO and directly shapes global health policy knowledge.

GRADE's Mode of Operation

Procedural: Hierarchizing and Filtering Evidence

GRADE has a procedural mode of operation in that it defines what is considered as relevant and valid knowledge through an evidence scale, its foundational discursive basis. The scale establishes a clear and stable hierarchy among different forms of evidence, based on the methods through which they have been produced, but it does so in a fashion that seems a matter of just following certain "procedures" and "technical steps" (Balshem et al. 2011). GRADE conceives

[3] Interview with former WHO staff, November 3, 2020

[4] Interview with GRADE methodologist, December 4, 2020.

[5] Interview with a member of the GRADE Working Group, April 6, 2021.

"epistemological correctness" (Mercer 2008) as achieved through a scrupulous compliance with a specific set of methodological rules, all standardized through handbooks, guidelines, and "tick the box" procedures. The scale, as introduced initially by the GRADE working group, acts as the ideational foundation of the regime and embodies specific epistemological assumptions about what evidence ranks as "high," "moderate," "low," or "very low" quality (Balshem et al. 2011). In this scale, RCTs rank at the top as high-quality evidence, while observational studies and expert opinions rank at the bottom, as low-quality evidence. There might be circumstances when methodologists can upgrade observational studies, or reversely downgrade those RCTs that they feel suffer from a number of limitations, such as risk of bias, imprecision, or inconsistency, among others (Balshem et al. 2011). Despite its flexibility, GRADE is mainly responsive to "limitations" that come from study design. Other types of biases, for instance, when the study design ignores local circumstances and the socioeconomic specifics of a given study population, cannot be captured (Luke et al. 2022). Also, GRADE's point of departure is that randomized types of evidence are inherently superior, in line with the core assumptions of EBM. Nonrandomized types of evidence are considered, from the outset, to be of poorer quality than RCTs, the unquestionable "gold standard." In the words of a WHO official, what makes GRADE so authoritative is the widespread perception that GRADE is just an impartial method "supposed to take the arbitrary away."[6] While GRADE, like all systems of evidence production and evaluation, relies on specific assumptions, it indeed makes these quasi-invisible; in fashioning certain forms of evidence as more impartial, it makes the adoption of certain methods as just a question of "procedures" and "steps," which need to be properly followed.

The GRC, which reviews all WHO guidelines, gives traction to the system by verifying that the WHO's recommendations adhere to GRADE's methodological rules. The GRC indeed relies on procedures now highly codified in the WHO Handbook for Guideline Development (WHO 2014)—partly drafted by some of the founders of GRADE—which sets all steps that need to be taken when assessing evidence and translating it into recommendations. In that sense, the GRC acts as a filtering and policing mechanism, which controls that the WHO uses the GRADE scale to select and assess which forms of evidence make it to policy. The GRADE system thus establishes the confidence that the WHO can place in specific studies and how much emphasis it should give to these findings. However, beyond the central filtering role of the GRC, GRADE as a methods regime is also enacted through a dispersed web of epistemic operators and material inscriptions or devices that sustain the regime.

Dispersed: The Epistemic Operators of GRADE

GRADE as a regime operates through its powerful networks of methodologists and knowledge sites. Beyond the role of the GRC at the WHO, it is enacted by a dispersed web of epistemic operators, mainly a myriad of well-established academic centers and networks across the world, staffed with a new cast of methodologists who typically have a degree in epidemiology and/or medicine, and who produce and disseminate the ideas of GRADE. GRADE now has nineteen academic centers in North America, Europe, and Asia, as well as powerful organizations, such as the Cochrane network, professional associations, journals, and civil society organizations, such as the European Stroke Organization or the Robert Koch Institute, which also act as GRADE partners.[7]

McMaster University and Cochrane act as the regime's core knowledge centers, which produce its knowledge and standardize it through the production of a

[6] Interview with a WHO former staff member, November 9, 2020.

[7] A list of 110 organizations that have endorsed GRADE or are using it can be found at https://www.gradeworkinggroup.org/.

number of derivative products, such as certified GRADE trainings, graduate courses, and technological devices. As detailed above, McMaster University is the alma mater of GRADE, and where the GRADE Working Group still has its headquarters. The Guidelines International Network (GIN), which advertises itself as "the connector in the guideline world" and claims to be providing the best way to "connect with the most influential people in the guideline world," is also located at McMaster University's Department of Health Research Methods, Evidence and Impact (GIN 2021). GIN and McMaster have officially partnered to offer the International Guideline Credentialing & Certification Program (INguide), targeted toward organizations that develop not only guidelines, but also trainings in guideline development methods. INGuide is chaired by Holger Schünemann, co-founder of GRADE, and describes itself as "the product of a partnership between GIN and world-renowned experts in guideline research, development, and implementation" (GIN 2021).

Cochrane, a nongovernmental organization created in 1992, also acts as a central epistemic operator of the GRADE regime through its dispersed network of centers. Established to promote epidemiological methods, Cochrane is the largest global network of organizations that perform and collect systematic reviews. In fact, most of the WHO guidelines include Cochrane reviews (Cochrane 2022). Cochrane centers "act as clearing houses for EBM certified information" (Mercer 2008, 412) and have also developed a complex "science" of systematic reviews, for which it has adopted the GRADE scale as a method (Cochrane 2022). Cochrane thus acts as a central operator in promoting GRADE, through its practice of doing systematic reviews, but also its trainings, handbooks, and publications available via its platforms. As put by a member of the Grade Working Group:

> We try to harmonize guidance within Cochrane and GRADE. The Cochrane handbook has entire chapters on how to use GRADE and Cochrane is one of the groups that helped refine the GRADE approach and it continually does so. Cochrane is a key partner in the same enterprise.[8]

There is in fact a strong degree of circularity within the GRADE–Cochrane networks with several GRADE members holding positions within Cochrane and vice versa.[9] Several WHO staff members, in particular the methodologists from the GRC, belong to Cochrane's review groups.[10] At the same time, the decentralized mode of existence of Cochrane, which relies on multiple Cochrane-accredited centers throughout the globe, makes it possible for GRADE to operate anywhere.

Moreover, GRADE strongly relies on the collaboration of leading journals in the field of medicine, which play a central role in shaping the production of scientific research in global health. The *Journal of Clinical Epidemiology*, widely perceived as "belonging" to GRADE, has a series entirely devoted to GRADE, the "GRADE series" (JCE 2022).[11] Members of the GRADE Working Group also hold positions in the editorial boards of these journals. Gordon Guyatt, co-founder and co-chair of GRADE, is currently a member of the *Journal of Clinical Epidemiology*'s editorial board and participates in the *British Medical Journal*'s Open editorial advisory board (BMJ Open 2022). Such journals filter what knowledge can be published and endow with prestige the studies that they accept.

The WHO's GRC, GRADE's academic clusters, Cochrane, and its multiple centers around the world act as the central epistemic operators of GRADE, which not only produce and stabilize the epistemological assumptions of the regime, but also perpetuate them by providing standardized trainings and certification programs to methodologists, who then get hired by the WHO and other organizations. They

[8] Interview with a GRADE methodologist, April 6, 2020.
[9] Interview with a GRADE methodologist, October 14, 2020.
[10] Interview with a WHO methodologist, October 12, 2020.
[11] Interview with a public health scholar, June 1, 2022.

also entertain a close relationship with high-profile scientific journals, which consolidate the dominance of GRADE, as will be examined below. The epistemological operators of GRADE are, thus, actively "protecting its brand and methodological power."[12]

Embedded: The Material Inscriptions of GRADE

GRADE is also sustained through a dense web of material techniques, documents, and technological devices that diffuse and give the regime's ideas an autonomous existence. GRADE is materially inscribed into a plethora of detailed handbooks and guidelines, certification programs, a standardized "GRADE CV," and technological devices such as apps and computer programs. Although one can find alternative methodologies for developing clinical guidelines and assessing the quality of evidence, GRADE dominates bureaucratic life, health journals, and academic training.[13] To begin, the GRADE Working Group has been consolidating GRADE's foundational literature, guidelines, and methodology through the production and diffusion of about seventy publications in scientific journals. These procedures and publications are easily accessible and can be applied by any practitioner, health agency, or organization, thus giving the regime a self-sustaining capacity, or a life of its own. In 2008, members of the GRADE Working Group initiated an introductory series of articles in the *British Medicine Journal*, which detailed the system at great length and gave specific instructions for users. This literature inscribes GRADE and its epistemological assumptions, while endowing them with further scientificity.

GRADE also has its own courses, trainings, certificates and even a standardized GRADE CV, to the extent that GRADE has also become a profitable business, as people are paying some thousands of dollars to get accredited.[14] Possessing some sort of GRADE training has now become a normal expectation and a door opener to work as a methodologist at the WHO or other health organizations that seek greater "methodological rigor and transparency" (Sultan et al. 2020). The WHO Handbook for Guideline Development even explicitly encourages those WHO officials who perform systematic reviews in-house to have an in-depth knowledge of GRADE and take their training modules (WHO 2014). As a matter of fact, McMaster offers GRADE online learning modules specifically targeted toward WHO guideline developers (McMaster University 2022). Recently, authors involved with GRADE have published a framework designed to "enable guideline-producing organizations to identify guideline methodologists with the relevant and appropriate level of knowledge and skills to lead guidelines" (Sultan et al. 2020, 561–62). The skills required, no need to say, are those set by GRADE. The GRADE CV has itself become highly standardized. Susan Norris, member of the GRADE Working Group, has described the main characteristics of a standardized "GRADE CV" in the *Journal of Clinical Epidemiology* (Norris et al. 2016), where she explains how to assess methodologists' expertise and dresses up a list of a "set of minimum skills and experience" for methodologists to claim that they "can do GRADE." They must have attended GRADE's trainings or webinars, read GRADE's publications, and used GRADE web-based tools for creating evidence profiles and summary of tables, among others (Norris et al. 2016, 151). Through its material inscriptions, GRADE operates in a quasi-autonomous fashion, beyond the role of those individuals who founded GRADE, and exerts a significant control over the legitimate practices in the field.

Finally, GRADE as a methods regime is embodied into technological objects, such as GRADE's MAGICapp, a web-based tool that guides professionals through the process of writing a guideline, and GRADEpro, a computer program that creates summaries of tables. A number of health ministries, universities, national research

[12] Interview with a public health scholar, June 1, 2022.

[13] Interviews with WHO methodologists, between October 2020 and December 2020.

[14] Interview with a GRADE methodologist, December 18, 2020.

councils, and institutions including the European Commission and the WHO now use GRADEpro (GRADEpro 2021). Similarly, the MAGICapp, launched in 2013 and initially conceived by the members of the GRADE Working Group, was developed to support the application of GRADE's methodology. MAGIC, which stands for "making GRADE the irresistible choice" (MAGIC 2020a) is solely devoted to "improving the application or advancing evidence-based medicine and GRADE" (MAGIC 2020b). The nonprofit foundation that sustains the app (also called MAGIC) has now established partnerships with the *British Medical Journal*'s rapid recommendation project, Cochrane, and the WHO (MAGIC 2020b). GRADE relies on a dense and powerful web of operators and material inscriptions, which are strongly tied to those of EBM, making it difficult for alternative systems to compete, as will be discussed in greater detail below.

GRADE's mode of operation is procedural, dispersed, and embedded. For it to operate at an epistemic level, GRADE filters and polices knowledge production through apparently simple "procedures." This apparent procedurality, together with the regime's dense but dispersed networks of epistemic operators, as well as its multiple material products, from the standardized courses to its technological devices, gives the regime traction, but in a way that is quasi-mechanical. Given these characteristics, the way the regime delineates the contours of knowledge production in global health is hard to notice, but ubiquitous.

The Effects of GRADE on the Politics of Knowledge-Making in Global Health

Methods regimes have significant effects on the way global problems are being governed. GRADE has worked to the effect of empowering a new cast of methodologists, seen as more objective and portable across domains, demoting certain forms of evidence that do not conform with its own methodological criteria of scientificity, and "clinicalizing" research in medicine and beyond.

The Reconfiguration of Epistemic Authority in Global Health: The Rise of Methods Experts

GRADE has contributed to the reconfiguration of hierarchies in global health governance, resulting in what an interviewee calls a drastic "cultural change" in global health.[15] GRADE has indeed sustained a displacement of epistemic authority away from the so-called content experts—in general, doctors with experience in their area of specialization—toward a transnational cast of methods experts, known as "Graders," who have come to play a central function in the production of global health guidelines. As GRADE seems to be "just" a matter of methods, methodologists were able to fashion themselves as objective, neutral, and also highly mobile experts. Although their authority faced some contestation, as will be discussed below, the networks, knowledge centers, journals, and technological devices that support GRADE and its ideas contributed to the expansion of the authority of GRADE, as well as that of the regime's experts, the methodologists.

Within the WHO, the power of the GRC signals the triumph of methods experts and of "Graders" in particular.[16] The GRC is endowed with a high degree of epistemic authority, as "there is still this fake idea (within the WHO) that the GRC is a science lab."[17] Methodologists trained with GRADE, who benefit from the prestige associated with GRADE's influential academic and knowledge centers and their training programs, have gained prominence within the WHO and want to impose its clinical paradigm, no matter the type of intervention examined.[18] A WHO official says that "GRADE members define standards (…) in a few years that if you don't

[15] Interview with a WHO methodologist, November 3, 2020.
[16] Interview with a GRADE methodologist, December 3, 2020.
[17] Interview with a WHO staff member, December 3, 2020.
[18] Interview with a former WHO staff member, November 9, 2020.

have this kind of GRADE certifications you will not be able to work as a methodologist."[19] The WHO is also quasi-systematically resorting to Cochrane, a central epistemic operator of GRADE, to do its systematic reviews, which have come to be seen as the new standard. As put by a global health official, the WHO is contracting out external "Graders," in order to "plan the scope, the key questions, find someone to do systematic reviews, make sure that the systematic reviews meet methodological standards, help to frame evidence to decision framework, and help the group formulate recommendations based on the evidence."[20] As GRADE's knowledge networks have become so dominant, and "Graders" have also largely penetrated the WHO, the use of methodologists trained with GRADE increasingly appears as unavoidable.

Unlike content experts, who derive their authority from in-depth knowledge in a specific area and their experience as physicians, the knowledge of methods experts focuses on how to follow a technique or apply a specific set of rules. Because of the seemingly procedural nature of the tasks they perform, they are perceived as objective and disinterested, in contrast to content experts, portrayed as prone to bias and conflicts of interests. Thus, methodologists see themselves as "helping content experts to understand the merits of being transparent, comprehensive and the risk of minimizing bias."[21] As expressed by a GRC member, "thanks to GRADE, guidelines are not based on opinions anymore."[22] This illusion of objectivity, no need to say, is central to the authority of methodologists. In addition, methodologists seem mobile and easily portable. In contrast to content experts, they are not tied to any particular field of expertise and can circulate freely across issue areas and organizations. In health and beyond, they are seen as attractive for organizations, as they can easily transfer their skills across topics, and from one place to the next.[23]

Some health practitioners see this radical shift as a threat to their well-established authority and worry that content experts be replaced by "a new cast of experts characterized by absence of formal training, certification and clinical experience in the subject of guidelines" (Grannis 2009). For a WHO official, "no one wants a generalist telling you how to do a guideline, this is offensive in many ways."[24] According to another WHO official, "the GRC should become obsolete, it is just delaying the process."[25] For content experts, the GRC focuses too exclusively on randomized data while "real-life evidence" and "repeated experiences" stemming from observations are being dismissed even if they would be highly relevant.[26] Thus, some WHO staff members have contested the GRC by trying to circumvent the standard review process. Some of them "do not call a document guideline so they do not have to submit it to the GRC for review."[27] The Strategic Advisory Group of Experts on Immunization has, for instance, contested the increasing clinicalization of guideline development and asked for exceptions and adjustments to GRADE. According to a former WHO staff, GRADE did not "do justice to the overall body of evidence, but only to the study design of the data. Without having clinical data, but only surveillance data, you see that certain vaccines such as measles work, but according to GRADE this real life data would be low quality, and we needed to find ways of saying this is actually very good data."[28] The immunization expert group

[19] Interview with a former WHO methodologist, November 3, 2020.
[20] Interview with a former WHO methodologist, November 3, 2020.
[21] Interview with a former WHO methodologist, November 3, 2020.
[22] Interview with a WHO former staff member, November 3, 2020.
[23] Interview with WHO methodologist, October 12, 2020.
[24] Interview with a former WHO methodologist, November 3, 2020.
[25] Interview with Senior Health Advisor at the WHO, May 23, 2022.
[26] Interview with Senior Health Advisor at the WHO, May 23, 2022.
[27] Interview with a WHO consultant, October 27, 2020.
[28] Interview with former staff at the WHO, November 9, 2020; interview with a public health scholar, June 1, 2022.

was able to make minor adjustments to GRADE, but "we all ended up using it as a framework because it was over powering everything."[29] Despite internal tensions, the perceived objectivity and portability of methodologists, as well as their strong networks of support—Cochrane and its multiple centers, GRADE's academic clusters, and the "GRADE journals"—have contributed to the expansion of their epistemic authority, within the WHO, which is crucial for the making of global health recommendations, but also beyond. UpToDate, an online resource for authors and practitioners, with one of the founders of GRADE as executive editor, has for instance also become a tool used by doctors "to make appropriate care decisions and drive better outcomes" (Wolters Kluwer 2022). This technical device presents doctors with a synthesis of "the available clinical evidence and best clinical practices" and in effect gives traction to GRADE and its ideas within the medical community. On its website, UpToDate claims that "more than 100 research studies demonstrate that widespread use of UpToDate is associated with improved outcomes and hospital performance" (Wolters Kluwer 2022). As stated by a public health practitioner when asked about the possibility for using alternative frameworks within the WHO or beyond: "GRADE sets the standards, it is the default, if you contest, you still need to explain why you do not use GRADE, maybe I have just ended up accepting it the way it is."[30]

The Anecdotalization of Qualitative Forms of Evidence

GRADE's hierarchization of knowledge makes it possible to work with the assumption that practitioners can rigorously find out "what works," should they rely on the right kind of evidence (Bédécarrats et al. 2020). Through its enmeshment with the WHO's GRC, its networks, and its embedding into a wide net of material inscriptions, the GRADE system has become the dominant way of assessing evidence in health governance. Health organizations, as well as practitioners worldwide, increasingly use GRADE in order to draft guidelines, design their studies, or assess the efficacy of competing treatment options before doing a prescription. Crucially with regard to the making of global health guidelines, the WHO "prioritizes certain types of evidence," while those forms of knowledge that do not "rank high" according to GRADE are seen as at risk of resulting in "prognostic imbalance" and thus undervalued.[31] The power of the GRC and its methodologists within the organization, and the way GRADE's technological devices have come to be seen as useful tools within the WHO, has made this shift possible (GRADEpro 2021). MAGICapp, one of GRADE's technological devices, is now used for the development of health guidelines, for instance, in relation to COVID-19 drug treatments (WHO 2020). The WHO also relies on GRADEpro, a device that "guides through the process of guideline development while *seamlessly* making sure it adheres to the GRADE methodology" (GRADEpro 2021). Such tools are far from neutral. The UK National Institute for Health and Care Excellence (NICE) reported that the initial GRADEpro software lacked the capacity to evaluate observational studies, forcing staff to find other ways of presenting such evidence (Thornton et al. 2013).[32] According to a WHO official:

> A hard-core grader will see thing black and white. For example in environmental health, graders can come in, but they do not have twenty years of experience with environmental data, and they want to impose a grade clinical model on the assessment on the association between some particle in the air and health outcomes. This does

[29] Interview with former staff at the WHO, November 9, 2020; interview with a public health scholar, June 1, 2022.
[30] Interview with global health expert, May 30, 2022.
[31] Interviews with WHO officials, December 2020.
[32] Interview with a GRADE methodologist, December 18, 2020.

> not work well. There needs to be a flexibility and step back from dogmatism to work with content experts and understand *their* data.[33]

The hierarchization, from the outset, of qualitative forms of evidence, such as observations and case reports at the bottom of the evidence scale, is problematic. First, like other existing research methods, RCTs also embody certain biases. It is obvious when trials, often with positive outcomes, are funded by biopharmaceutical companies that omit inconvenient data (Wahlberg and McGoey 2007). However, beyond this, and even when trials are publicly funded, a number of decisions are taken by researchers at every stage of *any* research design. Formulating the research question, selecting variables, assembling the sample before it is randomized, conducting the analysis, and interpreting the results all involve decisions that reflect certain theoretical presuppositions. To mention only one example, choosing when to end a trial and when to collect endline data directly affects the nature of the results and thus the kind of claim that can be made about the effects of a treatment or any social intervention (Timmermans and Berg 2010; Krauss 2018). Thus, assessing randomized data as objective and thus inherently superior can result in the omission of other valuable forms of data. For instance, global recommendations related to Covid-19 ignored non-randomized data drawn from observation and fieldwork, considered to be of "low quality."[34] A WHO official explains that a strict adherence to GRADE has resulted in the undervaluation of ecological evidence, and that, as a result, "we are having Covid recommendations without this type of data," although it would have been highly relevant for the problem at stake.[35] The application of GRADE's hierarchies of evidence to the making of COVID recommendations also resulted in giving strong emphasis to certain forms of data in relation to the efficacy of drugs, while dismissing others. For instance, the WHO conducted a systemic review of the existing studies on the effectiveness (and side effects) of corticosteroids for COVID-19. While a large number of observational studies concluded to the lack of effectiveness of this drug and an increase in side effects reported, the WHO relied, in an interim guidance document, on the *only* RCT study available, which concluded that corticosteroids may reduce mortality in moderate–severe acute respiratory distress syndrome, not necessarily COVID related (Ortolani and Pastorello 2020).

Second, while RCTs may be suited to measure the outcomes of a particular intervention, they are of little help when it comes to understanding how complex phenomena work.[36] GRADE does not work well for data about *how* things work, and still people "try to force it into that model."[37] For instance, the use of GRADE in the context of the WHO's occupational health guidelines "unfairly downgrades environmental evidence" (Verbeek, Heroux, and van Deventer 2017). This occurs despite the fact that environmental exposures are inherently complex and interrelated and, as such, must rely on observational, human, animal, and in vitro mechanistic studies (and not only on RCTs of a single chemical) (Bero, Norris, and Lawrence 2019). Examples of how GRADE has contributed to demote certain forms of knowledge extend beyond the domain of health. In the field of nutrition, critics contend that guidelines fail to take full account of available evidence because they rely on "*methods borrowed from other fields*" (Bero, Norris, and Lawrence 2019, emphasis added). The selection and evaluation of evidence for nutrition guidance are currently being driven by methodological criteria, rather than the questions that need to be answered. As a result, most of the nutrition guidelines and policy

[33] Interview with a former WHO methodologist, November 3, 2020.
[34] Interview with a WHO consultant, October 27, 2020.
[35] Interview with a WHO consultant, October 27, 2020.
[36] Interview with Public Health Scholar, June 1, 2022.
[37] Interview with a former WHO methodologist, November 3, 2020.

statements in the WHO relate to single-nutrient interventions, which can be evaluated through randomized trials (Bero, Norris, and Lawrence 2019). Holding RCTs as superior to all other forms of evidence, therefore, results in demoting or simply ignoring forms of evidence that could prove helpful (Rehfuess et al. 2019); it also absolves professionals from comparing and reconciling results across different forms of evidence. In addition, RCTs are almost always based on clinical protocols and universal diagnostic and disease label categories, making it illusory for non-Western forms of medicine with different conceptions of sickness to "prove" the worth of their therapies through such trials and get accredited as medicines in the West (Adams 2002). Although the GRADE Working Group has been reflecting on its own weaknesses,[38] with one of its subgroups currently working on the "confidence" that might be granted to findings from qualitative studies, its rationale and underlying assumptions remain the same.[39]

The Clinicalization of Academic Research

GRADE relies on a clinical frame, according to which the most important aspect of an intervention is its efficacy. A medical practitioner explains that "GRADE tells us that it does not matter *how* something works, the only thing that matters is that it works."[40] Not only does the clinical paradigm embodied into GRADE shapes the assessment of knowledge in health governance, but it also deeply alters academic practices in medicine and beyond. Through its embedding into dispersed objects such as its teaching modules, its vast body of literature, technological devices meant to "help" professionals, and its ties with prestigious scientific journals, GRADE works to the effect of clinicalizing academic research. For example, if medical students or practitioners want to publish articles related to guideline development in some of the most prestigious medical journals, they are encouraged to adopt GRADE's methodological criteria. As evoked by one interviewee, entire academic careers are built on GRADE.[41] For instance, the *British Medical Journal*, one of the most prominent journals in the medical field, recommends that authors who include statements on the quality of evidence and strength of a recommendation use GRADE (BMJ 2018). Medical students also turn toward GRADE, which operates within medical schools not only through GRADE's teaching modules, but also through UpToDate, a resource for authors and health practitioners "taught" to students, in the hope that they keep using it after completing their education. Clinical modes of reasoning have become the norm in medicine. According to a practitioner, most medical students get familiarized with the GRADE framework since their early undergraduate epidemiology classes.[42] Although GRADE, as well as the EBM movement more broadly, initially emerged as a critique, they have now "become mainstream." As lucidly argued by a member of GRADE:

> Before we were the angry young hippies, now we are at the top of the hierarchy, we have evidence as the excuse. We moved from a movement which was grassroots, we were trying to force people in power to justify their decisions, and now we are in power. It happens to all revolutionaries.[43]

Clinical modes of reasoning have also extended to other governance domains such as environment and food and nutrition research, or development economics. Recently, a special series in *Environment International* was entirely devoted to reflection on the use of GRADE in environmental exposures, with the aim of providing

[38] Interview with a former WHO staff member, May 24, 2022.
[39] Interview with a GRADE methodologist, December 3, 2020.
[40] Interview with an expert clinician, October 10, 2020.
[41] Interview with a public health scholar, June 1, 2022
[42] Interview with an expert clinician, October 10, 2020
[43] Interview with a Cochrane staff member, November 24, 2020.

guidance for the application of the framework for environmental and occupational decision-making" (Morgan et al. 2019). Similarly, GRADE has contributed to clinicalize existing academic research in food and nutrition. In 2018, a consortium of nutritionists and researchers called the "Nutritional Recommendations" (NutriRECS) was launched to promote "trustworthy" systematic reviews based on GRADE (NutriRECS 2018). The consortium has been actively disseminating its outputs via open-access publications, training experts in nutritional guideline development worldwide, and advancing the methodology of systematic reviews in the field of nutrition more broadly (NutriRECS 2018).

Clinical reasoning and its experimental models have significantly shifted the politics of knowledge production in a number of domains (Donovan 2018). Despite being contested by critical health scholars who propose to shift research toward the development of interventions that recognize both the complexity of localized realities and the value of iterations between "implementers" and "researchers" in processes of knowledge generation (Hawe 2015), the model is in fact spreading out to other domains. In the realm of international development, those interventions that have been tested following the clinical model are also portrayed as superior. Banerjee and Duflo, winners of the 2020 Nobel Prize in Economics, argue that "the cleanest way to answer such questions (about poverty) is to mimic the randomized trials that are used in medicine to evaluate the effectiveness of new drugs" (Banerjee and Duflo 2011, 26). Not only has the clinical model largely reshaped what constitutes policy-relevant knowledge but it has also reconfigured the standards of how research is carried out. Questions that cannot be addressed through the dominant paradigm come to be seen as "uncomfortable knowledge" (Rayner 2012).

Conclusion

Despite evidence-based policy being fashioned as a rationalization of global governance, specific epistemological assumptions inform the selection of the evidence that is taken into account, its hierarchization, and the way it is synthetized and eventually translated into recommendations. Regardless, such decisions are always justified through claims of resorting to "the best available evidence," without ever making explicit what choices inform the prioritization of different forms of evidence. Thus, we set ourselves to reflect on how the evidence hierarchy upon which GRADE and its inscriptions are based shapes what is considered as policy-relevant knowledge in global health and beyond.

Through an exploration of the role of GRADE as a "methods regime," we have shed light on the processes that intervene in the making of what serves as the evidence base of global health recommendations. Through the epistemological assumptions concealed behind a seemingly procedural and mechanical functioning, which looks void of substantial content, the work of a dense network of epistemic operators, and the effects of its inscriptions and devices, which sustain the regime, GRADE shapes and polices the making of global health knowledge. The fact that influential academic clusters, such as MacMaster, knowledge centers, prestigious journals, and Cochrane, all supported GRADE made it possible for this specific sociomaterial arrangement, the GRADE "methods regime," to become powerful and effective. The dominance of GRADE has worked to the effect of empowering a new cast of methodologists, seen as more objective and portable across domains, sidelining certain forms of evidence that do not conform with its own methodological criteria of scientificity, and "clinicalizing" research in medicine and beyond. The "procedural" aura of GRADE makes this displacement in epistemic authority, as well as these knowledge effects, appear natural, thus hiding their political implications in the governing of global problems.

Our in-depth case study is highly relevant beyond global health. In finance, food safety, or development economics, methods experts participate in a number of

expert panels that also "govern" evidence-making processes. They act as quality assurance officers, auditors, or risk assessors, who advise organizations on the design and performance of their evidence gathering or evaluation procedures. The rise of RCTs, systematic reviews, cost–benefit types of analysis, or meta-analyses that claim to condense all existing research in a given domain has become new dogma that prevails in an increasing number of domains (Sending 2015; Kelly and McGoey 2018). In education or economics, researchers are measuring the *large n* effects of *specific interventions*, such as giving micro-credits to people living under a certain poverty threshold or changing the size of school classrooms to improve educational objectives. However, the way such epistemological regimes operate and come to shape and police the production, assessment, and selection of what counts as valid evidence in global governance *fora* had so far been largely overlooked. Given that IOs and other sites of global governance make it their daily business to produce guidelines, standards, and recommendations, which they claim are based on the "best" available evidence, it is necessary to understand what is actually meant by "best." In shedding light on the political work that determines the making of such evidence, we get a grasp of how global standards and forms of guidance are being produced. We also shed light on the way IOs themselves are embedded into broader sociomaterial arrangements, those networks, knowledge techniques, and objects, which they embody but also sustain.

These findings are highly relevant to existing research on the politics of knowledge and expertise in IR. We build upon a research agenda that has moved beyond examining the impact of expert evidence on policy-makers and points instead to the enmeshment between the scientific and the political in processes of knowledge production (Sending 2015; Bueger 2018; Littoz-Monnet 2022). By shedding light on the sociopolitical dimensions of methods regimes and their effects on knowledge-making and the governing of problems, we shed light on a novel aspect of the politics of expertise. Methods are central to evaluating certain forms of knowledge as "expert," and dismissing others. They are also part of the processes through which certain forms of expertise become authoritative (Sending 2015), circular, and exclusive (Littoz-Monnet 2022). Shedding light on the role of methods regimes, their networks, highly codified and mobile procedures, and material devices also reintroduces the role of power, resources, and hierarchies into the study of knowledge-making in global governance, which is key to understanding how certain forms of arrangements emerge, and become stable, while others do not.

Acknowledgments

We would like to thank the three anonymous referees for their thorough and insightful comments on our article. We are also thankful to Jan Aart Scholte, Christian Bueger, and Jens Steffek, as well as the other participants in the workshop "Standardization of Good Governance" at the University of Duisburg-Essen in June 2021.

References

Adams, Vincanne. 2002. "Randomized Controlled Crime: Postcolonial Sciences in Alternative Medicine Research." *Social Studies of Science* 32 (5–6): 659–90.

Alonso-Coello, Pablo, Holger J. Schünemann, Jenny Moberg, Romina Brignardello-Petersen, Elie A. Akl, Marina Davoli, Shaun Treweek, Reem A. Mustafa, Gabriel Rada, and Sarah Rosenbaum. 2016. "GRADE Evidence to Decision (EtD) Frameworks: A Systematic and Transparent Approach to Making Well Informed Healthcare Choices. 1: Introduction." *BMJ* 353: i2016.

Aradau, Claudia, and Jef Huysmans. 2014. "Critical Methods in International Relations: The Politics of Techniques, Devices and Acts." *European Journal of International Relations* 20 (3): 596–619.

Aue, Luis. 2021. "How Do Metrics Shape Polities? From Analogue to Digital Measurement Regimes in International Health Politics." *International Political Sociology* 15 (1): 83–101.

Balshem, Howard, Mark Helfand, Holger J. Schünemann, Andrew D. Oxman, Regina Kunz, Jan Brozek, Gunn E. Vist, Yngve Falck-Ytter, Joerg Meerpohl, and Susan Norris. 2011. "GRADE Guidelines: 3. Rating the Quality of Evidence." *Journal of Clinical Epidemiology* 64 (4): 401–406.

Banerjee, Abhijit, and Esther Duflo. 2011. *Poor Economics*, vol. 619. New York: Public Affairs.

Barnett, Michael N., and Martha Finnemore. 2004. *Rules for the World: International Organizations in Global Politics*. Ithaca, NY: Cornell University Press.

Bédécarrats, Florent, Isabelle Guérin, and François Roubaud. 2020. *Randomized Control Trials in the Field of Development: A Critical Perspective*. Oxford University Press.

Beisheim, Marianne, and Andrea Liese. 2014. *Transnational Partnerships: Effectively Providing for Sustainable Development?* New York: Springer.

Bero, Lisa A., Susan L. Norris, and Mark A. Lawrence. 2019. "Making Nutrition Guidelines Fit for Purpose." *BMJ* 365: l1579.

Best, Jacqueline. 2014. *Governing Failure: Provisional Expertise and the Transformation of Global Development*. Cambridge: Cambridge University Press.

BMJ. 2018. "BMJ Guidance for Authors."

BMJ OPEN. 2022. "BMJ Open Editorial Board." Accessed October 14, 2022. https://bmjopen.bmj.com/pages/editorial-board/.

Broome, André, and Leonard Seabrooke. 2012. "Seeing Like an International Organisation." *New Political Economy* 17 (1): 1–16.

Bueger, Christian. 2015. "Making Things Known: Epistemic Practices, the United Nations, and the Translation of Piracy." *International Political Sociology* 9 (1): 1–18.

———. 2018. "Territory, Authority, Expertise: Global Governance and the Counter-Piracy Assemblage." *European Journal of International Relations* 24 (3): 614–37.

Cochrane. 2022. "Use of Cochrane Reviews to Inform WHO Guidelines." Accessed October 14, 2022. https://www.cochrane.org/news/use-cochrane-reviews-inform-who-guidelines.

Cooley, Alexander. 2015. "The Emerging Politics of International Rankings and Ratings: A Framework for Analysis." *Ranking the World: Grading States as a Tool of Global Governance*, edited by Alexander Cooley and Jack Snyder, 1–38. Cambridge: Cambridge University Press.

Daly, Jeanne. 2005. *Evidence-Based Medicine and the Search for a Science of Clinical Care*. Berkeley, CA: University of California Press.

Davies, Huw To, and Sandra M. Nutley. 2000. *What Works?: Evidence-Based Policy and Practice in Public Services*. Bristol Policy Press.

Donovan, Kevin P. 2018. "The Rise of the Randomistas: On the Experimental Turn in International Aid." *Economy and Society* 47 (1): 27–58.

Doshi, Rush, Judith G. Kelley, Beth A., and Simmons. 2019. "The Power of Ranking: The Ease of Doing Business Indicator and Global Regulatory Behavior." *International Organization* 73 (3): 611–43.

Elbe, Stefan. 2014. "The Pharmaceuticalisation of Security: Molecular Biomedicine, Antiviral Stockpiles, and Global Health Security." *Review of International Studies* 40 (5): 919–38.

GIN. 2021. "INGUIDE." Accessed October 14, 2022. https://g-i-n.net/what-is-inguide/.

GRADEPRO. 2021. "Everything You Need from GRADEing Your Systematic Review, to Developing Your Guideline and Beyond." Accessed October 14, 2022. https://www.gradepro.org/product#about.

Grannis, Frederic W. 2009. "Methodologists Are Not Qualified to Write Clinical Guidelines." *The Lancet* 374 (9698): 1325–26.

Grossman, Jason, and Fiona J. Mackenzie. 2005. "The Randomized Controlled Trial: Gold Standard, or Merely Standard?" *Perspectives in Biology and Medicine* 48 (4): 516–34.

Guyatt, Gordon, John Cairns, David Churchill, Deborah Cook, Brian Haynes, Jack Hirsh, and Jan Irvine et al. 1992. "Evidence-Based Medicine: A New Approach to Teaching the Practice of Medicine." *JAMA* 268 (17): 2420–25.

Guyatt, Gordon H., Andrew D. Oxman, Gunn E. Vist, Regina Kunz, Yngve Falck-Ytter, Pablo Alonso-Coello, and Holger J. Schünemann 2008. "Grade: An Emerging Consensus on Rating Quality of Evidence and Strength of Recommendations." *British Medical Journal* 336 (7650): 924–26.

Haas, Peter M. 1992. "Introduction: Epistemic Communities and International Policy Coordination." *International Organization* 46 (1): 1–35.

Haas, Peter M., and Casey Stevens. 2011. "Organized Science, Usable Knowledge, and Multilateral Environmental Governance." In *Governing the Air: The Dynamics of Science, Policy, and Citizen Interaction*, edited by Rolf Lidskog and Goran Sundqvist, 125–61. Cambridge, MA: MIT Press.

HANEMAAYER, ARIANE. 2016. "Evidence-Based Medicine: A Genealogy of the Dominant Science of Medical Education." *Journal of Medical Humanities* 37 (4): 449–73.

HANRIEDER, TINE. 2019. "How Do Professions Globalize? Lessons from the Global South in US Medical Education." *International Political Sociology* 13 (3): 296–314.

HANSEN, HANS KRAUSE, AND TONY PORTER. 2012. "What Do Numbers Do in Transnational Governance?" *International Political Sociology* 6 (4): 409–26.

HAWE, PENELOPE. 2015. "Lessons from Complex Interventions to Improve Health." *Annual Review of Public Health* 36: 307–23.

HEARSON, MARTIN. 2018. "Transnational Expertise and the Expansion of the International Tax Regime: Imposing 'Acceptable' Standards." *Review of International Political Economy* 25 (5): 647–71.

JATTEAU, ARTHUR. 2013. *Les Expérimentations Aléatoires En Économie*. Paris: La découverte.

JCE. 2022. "GRADE Series." Accessed October 14, 2022. https://www.jclinepi.com/content/jce-GRADE-Series.

KELLY, ANN H., AND LINSEY MCGOEY. 2018. "Facts, Power and Global Evidence: A New Empire of Truth." *Economy and Society* 47 (1): 1–26.

KNORR CETINA, KARIN. 1999. *Epistemic Cultures: How the Sciences Make Knowledge*. Cambridge, MA: Harvard University Press.

———. 2007. "Culture in Global Knowledge Societies: Knowledge Cultures and Epistemic Cultures." *Interdisciplinary Science Reviews* 32 (4): 361–75.

KRAUSS, ALEXANDER. 2018. "Why All Randomised Controlled Trials Produce Biased Results." *Annals of Medicine* 50 (4): 312–22.

KUHN, THOMAS S. 1962. *The Structure of Scientific Revolutions*. Chicago, IL: University of Chicago Press.

LAKATOS, IMRE. 1978. "The Methodology of Scientific Research Programmes: Philosophical Papers," edited by John Worrall and Gregory Currie. 102–38. Cambridge: Cambridge University Press.

LAKOFF, ANDREW, AND STEPHEN J. COLLIER. 2008. *Biosecurity Interventions: Global Health and Security in Question*. New York: Columbia University Press.

LANDZELIUS, KYRA. 2006. "Introduction: Patient Organization Movements and New Metamorphoses in Patienthood." *Social Science & Medicine* 62 (3): 529–37.

LATOUR, BRUNO. 2005. *Reassembling the Social. An Introduction to Actor-Network-Theory*. New York: Oxford University Press.

LATOUR, BRUNO, AND WOOLGAR STEVE. 1986. *Laboratory Life: The Construction of Scientific Facts*. Princeton, NJ: Princeton University Press.

LAW, JOHN. 2004. *After Method: Mess in Social Science Research*. London Psychology Press.

———. 2008. "Actor Network Theory and Material Semiotics. In *The New Blackwell Companion to Social Theory*, 3rd ed, edited by Bryan S. and Turner, 141–58. Oxford: Blackwell.

LEANDER, ANNA. 2016. "Strong Objectivity in Security Studies: Ethnographic Contributions to Method Development." *International Studies Perspectives* 17 (4): 462–75.

———. 2020. "Collaging as a Method for IR in the Anthropocene." In *International Relations in the Anthropocene: New Agendas, New Agencies and New Approaches*, edited by David Chandler, Franziska Müller, and Delf Rothe, 353–71. New York: Palgrave Macmillan.

LEANDER, ANNA, AND OLE WÆVER. 2018. *Assembling Exclusive Expertise: Knowledge, Ignorance and Conflict Resolution in the Global South*. London: Routledge.

LIESE, ANDREA, JANA HEROLD, HAUKE FEIL, AND PER-OLOF BUSCH. 2021. "The Heart of Bureaucratic Power: Explaining International Bureaucracies' Expert Authority." *Review of International Studies* 47 (3): 353–76.

LITTOZ-MONNET, ANNABELLE. 2020. *Governing through Expertise: The PoliticsR of Bioethics*. Cambridge: Cambridge University Press.

———. 2022. "Exclusivity and Circularity in the Production of Global Governance Expertise: The Making of Global Mental Health Knowledge." *International Political Sociology* 16: 1–20.

LUKE, JOANNE, EBONY VERBUNT, ANGELA ZHANG, MURIEL BAMBLETT, GABRIELLE JOHNSON, CONNIE SALAMONE, AND DAVID THOMAS et al. 2022. "Questioning the Ethics of Evidence-Based Practice for Indigenous Health and Social Settings in Australia." *BMJ Global Health* 7 (6): e009167.

MAGIC. 2020a. "Our Approach." Accessed October 14, 2022. https://magicevidence.org/approach/.

———. 2020b. "COVID-19: MAGIC Making a Difference and MAGICapp Now Available to Develop Living Guidelines." Accessed October 14, 2022. https://magicevidence.org/.

MCMASTER UNIVERSITY. 2022. "GRADE Online Learning Modules." Accessed October 14, 2022. https://cebgrade.mcmaster.ca/.

MERCER, DAVID. 2008. "Science, Legitimacy, and 'Folk Epistemology' in Medicine and Law: Parallels between Legal Reforms to the Admissibility of Expert Evidence and Evidence-Based Medicine." *Social Epistemology* 22 (4): 405–23.

Morgan, Rebecca L., Beverly Brandy, Ghersi Davina, Holger J. Schünemann, A. Rooney Andrew, Whaley Paul, Zhu Yong-Guan, and A. Thayer Kristina. 2019. "GRADE Guidelines for Environmental and Occupational Health: A New Series of Articles in Environment International." *Environment International* 128: 11–12.

Norris, Susan L., Joerg J. Meerpohl, Elie A. Akl, Holger J. Schünemann, Gerald Gartlehner, Yaolong Chen, and Craig Whittington. 2016. "The Skills and Experience of GRADE Methodologists Can Be Assessed with a Simple Tool." *Journal of Clinical Epidemiology* 79: 150–58.

NUTRIRECS. 2018. "NutriRECS Trusted Nutritional Recommendations." Accessed October 14, 2022. https://www.nutrirecs.com/.

Ortolani, Claudio, and Elide A. Pastorello. 2020. "Hydroxychloroquine and Dexamethasone in COVID-19: Who Won and Who Lost?" *Clinical and Molecular Allergy* 18 (1): 1–7.

Oxman, Andrew D., John N. Lavis, and Atle Fretheim. 2007. "Use of Evidence in WHO Recommendations." *The Lancet* 369 (9576): 1883–89.

Parkhurst, Justin. 2017. *The Politics of Evidence: From Evidence-Based Policy to the Good Governance of Evidence*. Abingdon: Routledge.

Porter, Theodore M. 1995. *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*. Princeton, NJ: Princeton University Press.

Power, Michael. 1999. *The Audit Society: Rituals of Verification*. Oxford: Oxford University Press.

Rayner, Steve. 2012. "Uncomfortable Knowledge: The Social Construction of Ignorance in Science and Environmental Policy Discourses." *Economy and Society* 41 (1): 107–25.

Rehfuess, Eva A., Jan M. Stratil, Inger B. Scheel, Anayda Portela, Susan L. Norris, and Rob Baltussen. 2019. "The WHO-INTEGRATE Evidence to Decision Framework Version 1.0: Integrating WHO Norms and Values and a Complexity Perspective." *BMJ Global Health* 4 (Suppl 1): 1–21.

Ruppert, Evelyn, and Stephan Scheel. 2019. "The Politics of Method: Taming the New, Making Data Official." *International Political Sociology* 13 (3): 233–52.

Sayes, Edwin. 2014. "Actor–Network Theory and Methodology: Just What Does It Mean to Say That Nonhumans Have Agency?" *Social Studies of Science* 44 (1): 134–49.

Sending, Ole Jacob. 2015. *The Politics of Expertise: Competing for Authority in Global Governance*. Ann Arbor, MI: University of Michigan Press.

Shapin, Steven, and Simon Schaffer. 2011. *Leviathan and the Air-Pump: Hobbes, Boyle, and the Experimental Life*, vol. 109. Princeton, NJ: Princeton University Press.

Star, Susan Leigh. 1999. "The Ethnography of Infrastructure." *American Behavioral Scientist* 43 (3): 377–91.

Sultan, Shahnaz, L. Morgan Rebecca, M. Hassan Murad, Yngve Falck-Ytter, Philipp Dahm, Holger J. Schünemann, and Reem A. Mustafa. 2020. "A Theoretical Framework and Competency-Based Approach to Training in Guideline Development." *Journal of General Internal Medicine* 35 (2): 561–67.

Thornton, Judith, Philip Alderson, Toni Tan, Claire Turner, Sue Latchem, Elizabeth Shaw, Francis Ruiz, Stefanie Reken, Moira A. Mugglestone, and Jennifer Hill. 2013. "Introducing GRADE across the NICE Clinical Guideline Program." *Journal of Clinical Epidemiology* 66 (2): 124–31.

Timmermans, Stefan, and Alison Angell. 2001. "Evidence-Based Medicine, Clinical Uncertainty, and Learning to Doctor." *Journal of Health and Social Behavior* 42 (4): 342–59.

Timmermans, Stefan, and Marc Berg. 2010. *The Gold Standard: The Challenge of Evidence-Based Medicine*. Philadelphia, PA: Temple University Press.

Towghi, Fouzieyha, and Kalindi Vora. 2014. "Bodies, Markets, and the Experimental in South Asia." *Ethnos* 79 (1): 1–18.

Verbeek, Jos, Marie Eve Heroux, and Emilie Van Deventer. 2017. "WHO Environmental and Occupational Health Guidelines: 8 Challenges." *Cochrane Database of Systematic Reviews*. 9 (suppl 1). Accessed October 14, 2022. https://abstracts.cochrane.org/2017-global-evidence-summit/who-environmental-and-occupational-health-guidelines-8-challenges.

Wahlberg, Ayo, and Linsey Mcgoey. 2007. "An Elusive Evidence Base: The Construction and Governance of Randomized Controlled Trials." *BioSocieties* 2 (1): 1–10.

Walters, William. 2002. "The Power of Inscription: Beyond Social Construction and Deconstruction in European Integration Studies." *Millennium* 31 (1): 83–108.

WHO. 2014. *WHO Handbook for Guideline Development*. Geneva: World Health Organization. Accessed October 14, 2022. https://apps.who.int/iris/handle/10665/145714.

———. 2020. "WHO Updates Clinical Care Guidance with Corticosteroid Recommendations." Accessed October 14, 2022. https://www.who.int/news-room/feature-stories/detail/who-updates-clinical-care-guidance-with-corticosteroid-recommendations.

———. 2022. "WHO Guidelines Review Committee." Accessed October 14, 2022. https://www.who.int/groups/guidelines-review-committee.

Wolters Kluwer. 2022. "UpToDate." Accessed October 14, 2022. https://www.wolterskluwer.com/en/solutions/uptodate/uptodate.