BMJ Global Health

Organising knowledge to prevent global health crises: a comparative analysis of pandemic preparedness indicators

Alexander Kentikelenis [1], Leonard Seabrooke [2,3]

ABSTRACT

Preparing for the possibility of a global pandemic presents a transnational organisational challenge: to assemble and coordinate knowledge over institutionally diverse countries with high fidelity. The COVID-19 pandemic has laid these problems bare. This article examines the construction of the three main cross-national indicators of pandemic preparedness: a database with self-reported data by governments, external evaluations organised by the WHO and a global ranking known as the Global Health Security Index. Each of these presents a different model of collecting evidence and organising knowledge: the collation of self-reports by national authorities; the coordination of evaluation by an epistemic community authorised by an intergovernmental organisation and on the basis of a strict template; and the cobbling together of different sources into a common indicator by a transnational multi-stakeholder initiative. We posit that these models represent different ways of creating knowledge to inform policy choices, and each has different forms of potential bias. In turn, this shapes how policymakers understand what is 'best practice' and appropriate policy in pandemic preparedness.

Summary box

► Three cross-national indicators—self-reporting by governments, WHO-run expert evaluations and the privately run Global Health Security Index—have tried to quantify the degree of health emergency preparedness, and they reflect different systems of organising knowledge.
► These indicators shape what is visible and legible to policy audiences, and influence a range of other actors, including donors, media and the global health community.
► There is a potential bias in the construction of indicators arising from political gaming, groupthink in expert groups and blind spots in data selection.
► Our framework emphasises trade-offs between the development of pandemic preparedness indicators by governmental authorities or independent expert networks.
► Global health policy debates in the aftermath of the COVID-19 pandemic need to consider the limitations of different measurement approaches that seek to capture health emergency preparedness.

[1]Department of Social and Political Sciences, Bocconi University, Milano, Italy
[2]Department of Organization, Copenhagen Business School, Frederiksberg, Denmark
[3]Norwegian Institute of International Affairs, Oslo, Norway

Correspondence to
Dr Alexander Kentikelenis;
alexander.kentikelenis@unibocconi.it

INTRODUCTION

The COVID-19 pandemic took the world by surprise and led countries to scramble to contain the spread of the virus and mitigate its health and socioeconomic consequences. The pandemic also revealed large discrepancies in countries' preparedness infrastructures. Some of the world's richest nations were slow to respond and evidently lacked appropriate public health guidelines, health system capacity, and reserves of necessary medical and personal protective equipment. In contrast, other countries—including many at lower levels of economic development and purportedly limited state capacity—were initially successful at limiting the scale of the outbreak.

For readers of the Global Health Security Index (GHSI), these patterns were likely surprising. Released in October 2019, this Index is the first multi-stakeholder attempt to evaluate and benchmark the world's preparedness vis-à-vis health security threats, including transborder infectious disease outbreaks. To do this, it was co-developed by the Nuclear Threat Initiative (NTI) think-tank, the Johns Hopkins Center for Health Security and the Economist Intelligence Unit (EIU), and was funded by major donors like the Bill and Melinda Gates Foundation. In its inaugural report, the three highest ranked countries in terms of pandemic preparedness were the USA, the UK and the Netherlands. But such supposed preparedness did not correlate with adequate responses to the pandemic.[1] Indeed, scholars have more broadly questioned the kinds of expertise that inform what is deemed measurable and becomes measured in global health.[2–8]

These debates point to inherent challenges of generating reliable cross-national knowledge on important policy issues: how can we collect evidence from different countries and

BMJ

organise them into uniform and comparable formats? These considerations have direct real-world governance implications, as such knowledge can influence country approaches towards preparing for or responding to health emergencies. For instance, GHSI findings were used by then US President Donald Trump early on in the outbreak, in February 2020, as evidence of his country's readiness to tackle the approaching crisis. Pointing to a chart plotting GHSI data, he explained in a public briefing, 'The Johns Hopkins, I guess—is a highly respected, great place—they did a study, comprehensive: "The Countries Best and Worst Prepared for an Epidemic". And the United States is now—we're rated number one. We're rated number one for being prepared'.[9]

In this article, we open the black box of the construction of cross-national indicators of pandemic preparedness. We focus on the three main, broad-coverage sources of such information: the Electronic State Parties Self-Assessment Annual Reporting Tool (e-SPAR), relying on country-reported information and hosted on a WHO online platform; the Joint External Evaluation (JEE) mission reports, conducted by experts chosen by the WHO; and the GHSI.

Drawing on scholarship in sociology and international relations, we argue that each of these indicators corresponds to a different model of organising knowledge about the world. The first is a system of *collation* from national self-reporting to an authoritative body that hosts this information. The second form of organising is a system of *coordination* by a transnational epistemic community. The third form is a system of *cobbling* in which a benchmark is produced based on the aggregation of data collected from multiple sources.

Subsequently, we unpack how these indicators for pandemic preparedness were created: where did the underlying data come from, who analysed it and how was it aggregated? To answer these questions, from August 2020 onwards we collected all available reports and data on these metrics, examined correlations in their findings, analysed the composition of work teams producing the metrics and interviewed key participants in the development of the indicators, including from headquarters and in the field.

Our findings reveal that the three pandemic preparedness indicators represent fundamentally different approaches to collecting, systematising and communicating knowledge. First, self-reporting by governments yields information that can contain inaccuracies, given that this system has no safeguards for verifying or validating data provided. These data then get treated as official data beyond public reproach, and is widely used to inform policy design in global health governance. Unsurprisingly, then, the policy recommendations are only as good as the data that underpin them, which can undermine collective governance responses.

Second, assessment via the WHO is intended to convey expert judgement: evaluators are selected from a prescreened roster of experts who assess countries based on a WHO-developed multi-sectoral assessment template. In practice, this process is dominated by experts from high-income countries and major intergovernmental organisations (IGOs), who are the ones able to provide financial backing. This funding set-up limits diversity among experts, which can contribute to a degree of groupthink compared with a model of transparent deliberation among a less tightly controlled pool of experts. Further, the use of a standardised and detailed template can create a degree of cognitive closure, whereby only some policy issues become visible and inform the scoring and recommendations. In the words of one informant, JEE team members operate within a system of 'bounded expertise'—contingent on WHO templates, JEE team composition and negotiations with governments of evaluated countries.

Finally, GHSI's approach to measuring pandemic preparedness represents an effort to rank countries from best to worst prepared, and thus pressure them into changing their policies. This process relies on wide-ranging sources of data to feed into a uniform classification system. But this approach can have blind spots as it relies on a dispersed data search exercise through a very specific lens: a predetermined questionnaire relying on a health security framing that needs to be completed using publicly available evidence.

We also document that the development of pandemic preparedness indicators varies in how much autonomy is given to evaluated governments. This variation includes whether their inputs to the process are received passively, whether they are negotiated within the formal and expert governance architecture of the adjudicating organisation, or whether they are subject to third-party external assessments to which they have not provided input but may be compelled to respond. This spectrum of government autonomy then speaks to coordination problems and power asymmetries in the global system: Who should be empowered to best judge what pandemic preparedness information is valid? Can powerful countries legitimately lead the process? What is the role of non-governmental and private organisations in the global health security agenda?

Drawing on our findings, we argue that the choice of model of knowledge organisation shapes what becomes visible and legible: biases like inaccuracies, cognitive closure, groupthink or blind spots are not bugs in such indicators but features of their design. In turn, this has direct implications for policy. Doubts around the validity of different metrics have led to the consideration of how to include socio-political and governance metrics in assessments. This wider belief that collecting more non-health regulatory data could improve pandemic preparedness may have knock-on effects in creating further biases. For example, there are active discussions about the inclusion of pandemic preparedness indicators in developing financing decisions or donor allocations. Our analyses generate a set of considerations for experts and policymakers seeking to revamp pandemic

preparedness measurement for the post-COVID era—an activity both the WHO and the GHSI are currently engaged in.

ORGANISING GLOBAL KNOWLEDGE ON PUBLIC HEALTH

The organisation of global knowledge on public health requires the establishment of comparable metrics as well as consensus on how they should be interpreted. Scholarship in sociology and international relations stresses three difficulties in achieving these aims. The first is that the process of providing common metrics introduces bias, with those interpreting the metrics increasingly relying on assumptions made within them to guide their analyses.[10–12] The second issue is that national governments have clear incentives to manipulate the data that is compared across countries, viewing these metrics as judgements on their international status.[13] The third concern is that while metrics are often produced by epistemic communities that may operate at some distance from political agendas,[14] the international organisations that oversee what is emphasised and reported are also political institutions[15–17] and the same is true for transnational non-governmental organisations (NGOs).[18] Applied to global public health, we can readily identify difficulties with the establishment of common metrics,[19] the independence of epistemic communities[20–22] and the politics of intergovernmental and transnational organisations.[23–26]

Given this context, we posit that there are three main options for the organisation of global knowledge on health systems and their preparedness for emergencies, like pandemics. The first is a process of *collation*, where an authority passively receives information from national reporting bodies. The obvious danger here is that governments have an incentive to exaggerate their capacities to signal their strength to other countries, or the opposite to attract funding by alerting donors to dire circumstances. Through this lens, national positions on global health policy also become foreign policy and subject to diplomatic and political gaming.[27] The collation method underpins the e-SPAR system that is hosted by the WHO and provides information from national self-reports on how the country complies with International Health Regulations (IHR). While the SPAR exercise provides valuable information on governments' preparedness, we can expect that those overseeing the system are aware that it is subject to political gaming and deliberate inaccuracies.

The second method to organise global knowledge is through *coordination* among experts. This requires a clear mandate given to an IGO or transnational body, that in turn will select and deploy these experts to collect and analyse data, consult with policymakers and report findings. Such organisation is also a political process in accommodating the interests of key contributing governments with the relevant best practices from technical experts.[28] These processes involve the political determination of what purpose the knowledge will serve—be it underpinning sustainable development,[29] supporting the observance of labour rights,[30] or other issues. On global public health, the WHO is the leading IGO providing a coordination role, and internal struggles between political interests and expert staff are well documented.[15] In the context of pandemic preparedness, the WHO has successfully established a coordination mechanism through the JEE assessments conducted by a transnational network of experts drawn from a roster. The key task here is to not only assess the capacity to comply with IHRs but to have a conversation about effective responses to possible health emergencies. Given the reliance on a transnational expert network, we can expect network structure—that is, which experts and organisations are prominent in work teams—to reveal some key characteristics of this method of organising global knowledge.

The third system can be referred to as *cobbling*—often called 'commensuration' in sociology.[10] Here, a third party is typically involved in collecting a range of indicators to produce a common metric, with the legitimacy of the metric affirmed by a combination of expertise and organisational prestige. The cobbling of such benchmarks has been on the rise as it is an effective way for organisations to team-up to pursue a particular agenda.[31] In global public health, the GHSI reflects this way of organising knowledge for health emergency preparedness, fusing an assessment of pandemic preparedness with a health security agenda on response capacity. Formed through a collaboration between an NGO, a university and a private forecaster, the GHSI is the first comprehensive benchmark for health security across 195 countries and 'which solely monitors and documents the presence or absence of critical health security capacities'.[32] This index draws on a range of publicly available information to compile a benchmark based on a questionnaire developed with an international panel of experts drawn from different scholarly fields related to pandemic preparedness. To trace the characteristics of this form of organising global knowledge it is important to unpack two issues: who is compiling the indicators and what sources they use, as this signals what becomes visible to coders seeking to standardise information that can be used in cross-country comparisons.

THREE MODELS OF MEASURING PANDEMIC PREPAREDNESS

How can we understand the processes underpinning the construction of emergency preparedness metrics? As noted earlier, we elaborate on the construction of the three main indicators that have tried to systematise such knowledge: e-SPAR, JEE reports and the GHSI. Initially, we downloaded all the preparedness metrics from the three indicators (all also had a host of other subindicators—eg, on antimicrobial resistance, immunisation or epidemiology workforce; such data were beyond our focus and not collected). First and most straightforwardly, we downloaded e-SPAR data from the WHO website. For

Table 1 Correlations between three indicators of emergency preparedness

	e-SPAR*	JEE†	GHSI‡
e-SPAR	1.00		
JEE	0.75	1.00	
GHSI	0.45	0.50	1.00

*Score for capacity 8.1: planning for emergency preparedness and response mechanism.
†Score for indicator R.1.1.: national multi-hazard public health emergency preparedness and response plan is developed and implemented.
‡Score for indicator 3.1: emergency preparedness and response planning.
e-SPAR, Electronic State Parties Self-Assessment Annual Reporting Tool; GHSI, Global Health Security Index; JEE, Joint External Evaluation.

our purposes, we collected the scores on the 'national health emergency framework' capacity. This measures whether countries are 'prepared and operationally ready for response to any public health event, including emergencies' on the basis of three indicators: having an emergency preparedness and response mechanism, having management plans for the emergency response operations and being able to mobilise the necessary resources.[33]

Second, we extracted the JEE scores for their two indicators on preparedness: having developed a multi-hazard national public health emergency preparedness and response plan, and having mapped priority public health risks and resources. As the purpose of JEE missions is to provide in-depth information on a multi-sectoral basis rather than comparative rankings, the scores are only available in the corresponding sections of each report and not in an accessible dataset format, as is the case of e-SPAR and GHSI. Finally, we collected the detailed GHSI 2019 data (the first and—at the time of writing—only available year), and extracted the scores and underlying information on the policy area 'emergency preparedness and response planning'.

As presented in table 1, we ran simple correlations between the comparable components of these metrics that capture policies and procedures in place to respond to health emergencies. We find a high correlation between the e-SPAR and JEE scores, while much lower correlations between GHSI and the other two indicators. This suggests that these metrics capture somewhat different aspects of the underlying emergency preparedness realities of different countries, and—thereby—points to the need for better understanding how they are constructed.

To elaborate on indicator construction, we collected extensive data from all three sources. As both the data and the methods used to analyse them are specific to each indicator, we discuss these issues within the upcoming relevant subsections. Additionally, we conducted eight semi-structured 'elite' interviews with individuals heavily involved in the compilation of the three indicators. After identifying the key actors involved, we contacted them via email or LinkedIn to request an interview. All responded although there was a linguistic constraint in establishing an interview with one person. Our aim with the interviews was to complement publicly available data sources with insights of individuals with hands-on experience, as well as ensuring the integrity of our findings by checking in with the relevant communities. We conducted interviews until we reached a saturation point, with multiple interviews confirming a narrative on the production of the different metrics. Interviews generally lasted 60–75 min, were granted on the condition of anonymity, and informed consent was secured at the beginning of each interview.

Collation: states' self-reporting to the WHO

The e-SPAR dataset is based on self-reporting by states, in line with their obligations according to the IHR which form binding commitments. All states have a responsibility to submit information, however, the 2019 e-SPAR data contain information on only 171 countries and 24 did not report scores: these are mostly small island states, but also a few larger countries (Italy, New Zealand, Iran and Malaysia). The information collected pertains to 24 indicators, which themselves are organised into 13 core capacities. For example, a core capacity refers to the ability to handle chemical events, and the associated indicator captures whether a country has appropriate resources to detect and respond to such an event.

Initially, binary indicators of compliance were used for each indicator, but the limits of this approach—attempting to compress complex national-level infrastructures into answers to simple yes/no questions—became clear (Interview #2). So, since 2018, countries score themselves on each indicator on a scale ranging from having no policies or strategies in place to implement IHR guidance (score 1) to have such policies and strategies at the national, intermediate and local levels in all relevant sectors and subject to regular revision and updates (score 5). These scores are then converted into a 1–100 scale and averaged per capacity.

The main responsibility for compiling this indicator lies with teams within health authorities of reporting countries, known as 'national IHR focal points', which coordinate an intensive process through which they are supposed to arrive at the scores. Underpinning this process is the principle of 'multi-sectoral collaboration' which is intended to bring together multiple country agencies—including ones that do not have a direct role on health, like civil protection or crisis management offices—that then must discuss and decide on appropriate scores. Although this process is driven by countries, the WHO can still have a supporting role, especially in resource-poor settings with established WHO country teams present, as it can facilitate these multi-sectoral workshops and encourage the collection of valuable data (Interviews #2 and #6).

However, even though the ultimate scores are hosted on a WHO online platform, the WHO has no role in

the scoring process. The organisation treats submitted scores as official data and therefore beyond contestation. As one informant explained, 'we need to use the information as an official report by the government that has been cleared, whatever clearance procedure they have at the national level, and then we use that information to develop reports for the Health Assembly' (Interview #3). In other words, the WHO has no mandate to second-guess or verify the scores. To be sure, WHO officials are aware that many such scores are 'guesstimates', that political and cultural context are important for determining what scores countries give themselves, and that data from some countries are implausibly similar across time (Interviews #4 and #5). However, given no mandate to validate scores, the data are used in official WHO reports, notwithstanding staff scepticism over their veracity. Doubts were also raised about the purpose of deliberate exaggerations of over-capacity or under-capacity in reporting. In most cases, those over-reporting are engaged in status games, normally with regional peers. Cases of under-reporting are associated with attempts at drawing attention from donors to fund public health action plans (Interview #5).

Coordination: the WHO's management of JEEs

We collected all JEE reports that were publicly available as of August 2020. In total, these were 96 reports, primarily on low-income and middle-income countries, with some notable exceptions (eg, the USA and Canada). No reports were published for Latin American countries and very few for European Union countries (Belgium, Finland, Latvia, Lithuania and Slovenia), and no country in our sample had more than one JEE report prepared for it. All reports followed a standardised template, and included scores given by the evaluation team across a range of technical areas, like 'antimicrobial resistance', 'zoonotic disease' and 'biosafety and biosecurity'.

The JEE scoring process is composed of two steps. First, in advance of the visit of the external experts, countries need to plan a series of internal consultations that ultimately yield the documentation needed to underpin JEE scores, as well as suggestions for what these scores should be. This is often an intensive process, involving multiple public authorities in different policy areas who feed into this work (Interviews #1, #2 and #3). In principle, country authorities have the sole responsibility to manage this process, however in resource-poor settings, the WHO may send staff or consultants who aid the countries in preparing for the evaluation.

Subsequently, the JEEs are written over the period of 1 week, when a team of experts commissioned by the WHO visits the country-under-evaluation to conduct their assessment. No team members hail from the evaluated country to avoid conflicts of interest. For instance, the US JEE report was co-led by a Finnish and a Ugandan expert, and included members from France, Kenya, Indonesia, the World Bank and the Food and Agriculture Organization. These team members have diverse disciplinary backgrounds, and each has the prime responsibility for drafting the segment of the report that is closest to their expertise, following the template provided by the WHO. In addition, WHO staff and contracted copy-writers are also formally part of these teams, although their input in the actual evaluations is limited. This is done by design, as an informant from the WHO explained: 'we try to have few WHO staff on a team in order to enable peer-to-peer knowledge transfer' (Interview #3). The technical writers have duties limited to compiling and copy-editing the technical area reports, without contributions on substance (Interview #2). In the network analysis presented later, we exclude all individuals identified as 'technical writer', 'communications consultant', 'editor' or similar in their JEE work description, as including them would give them undue influence in the network, which we were assured by our informants that did not actually take place (Interviews #2 and #5).

Even though labour-intensive for both evaluated countries and the external experts (Interview #8), this process entails very limited costs for the WHO: experts do not get paid for their participation (excluding flight, hotel and per diem costs), and their home-institutions cover any other costs of their involvement (eg, time devoted to preparation and the country visit). This funding structure means that participants overwhelmingly hail from high-income countries or large IGOs that have budgets to support their staff—an issue we return to below.

During their visit to the country, JEE experts meet with a host of country authorities, and discuss precise scoring and recommendations. Overall, the experts we interviewed reported a high degree of faith in this process. The expert team has ultimate authority to decide on scoring, but governments can exert pressure or challenge these scores (as well as the broader analysis) during consultations. External assessors working on the JEE reports work within a system of, as one interviewee described it, 'bounded expertise' where they stick to their knowledge base and allocated tasks (Interview #4). As such the JEE missions combine a desire of the WHO to have multi-sectoral analyses of pandemic preparedness, while external assessors are generally expected to report on their own issue area rather than intrude into findings and analysis of others (Interview #6).

Informants reported several reasons why governments might wish to influence report content. On the one hand, some policymakers in resource-poor settings use these reports to attract additional financing from donors, and this means there is an incentive to score low. As one interviewee who participated in several missions recalled, 'countries can use these reports as motivators to go to donors and say 'this is what the external assessment by the WHO says, so this is where you should place your funding" (Interview #5). On the other hand, some countries treat these reports as global status competitions, thereby wanting consistently high scores and complaining to the WHO when the expert team assigned low scores. Informants tended to attribute pressures from governments on cultural factors (Interview #5) or the desire

to project strength to neighbours (Interview #2). In any case, pressures by the government in either direction are taken into consideration but need not be yielded to.

To understand the dynamics within the expert teams that drafted the evaluations and scores, we collected data for all team members conducting JEE evaluations: each report listed the names and affiliations of the experts involved. In total, the team members of the 96 reports were 1172, but the unique individuals included in our dataset are 636, as several individuals participated in more than one JEE missions. To get analytical leverage on these data we employed social network analysis techniques, which allow us to see both the structure of the JEE work team network as well as likely forms of interaction. The networks are composed of nodes and edges, the former denoting the actor and the latter the tie or relationship. Social network analysis is especially appropriate for identifying actors who are brokers within and between work teams—those with centrality.

We assessed the eigenvector centrality of JEE experts in their network. This measure establishes who is not only the most central in the network but the key actors with the most 'friends of friends'. It is these actors who are best able to spread information to others through the network structure.[34] This approach is consistent with the JEE aim to also facilitate knowledge transfer between evaluators, as noted earlier. In our context the capacity to spread information matters because shared knowledge on pandemic preparedness can not only more easily move within the network but will also be judged as appropriate given the known influence of cliques in affirming particular bodies of knowledge (applied to public health see [35 36]). Furthermore, actors in similar organisations are likely to facilitate the spread of knowledge, and multi-organisational work teams will also permit the diffusion of knowledge across different platforms. Differences in geographic settings and the extent of cross-relationship collaborative work will also affect how knowledge on pandemic preparedness is shared and diffused.

To gain insight into the JEE network, we assessed the types of organisation and individuals involved. Table 2 shows the frequency of JEE evaluators' organisational backgrounds. Health-related public agencies were most prominent in the network, amounting for 45% of all JEE team members. Importantly, approximately two-thirds of individuals with this professional affiliation hailed from high income countries—most notably, the USA, Sweden, Finland and Norway. This dominance of the Global North in the roster of experts should come as no surprise: on the one hand, high-income countries have ample resources to support their staff in taking on JEE responsibilities, which are not remunerated and take staff time away from other organisational duties. On the other hand, there is an element of path dependency at play. The JEE originated from a post-Ebola joint policy push by the USA and Finland—the former instigated the Global Health Security Agenda, a forum established in 2014 and endorsed by the G7; while the latter became the lead country in its steering group. This work yielded the first pilots of what became the JEE and was incorporated into the WHO's toolkit for monitoring health emergency preparedness, including the development of its assessment template (Interviews #5 and #8), so—predictably—the countries most involved in its creation are highly represented in the expert roster. An informant told us that the Gates Foundation put up funding for the administrative costs of the JEE in its start-up phase (Interview #8).

After public authorities, the WHO and then other IGOs (primarily the Food and Agriculture Organization and the World Organisation for Animal Health) are the largest contributors of experts. The WHO's prevalence is expected given they coordinate the JEE missions, with WHO regional offices selecting experts through a roster system. The WHO is also involved in training as well as editing and publishing the JEE reports. Finally, professionals from NGOs and private firms, including consultants, are also present in the network, then followed

Table 2	Organisational background of individuals in JEE teams		
Organisation type		**Observations**	**Frequency**
Public agency		486	43% of total team members
Public agency based in a high-income country		319	66% of public agency officials
… in particular, the USA		*101*	
… in particular, a Nordic country		*64*	
Public agency based in a low-income or middle-income country		167	34% of public agency officials
WHO		386	35% of total team members
International organisation (non-WHO)		164	15% of total team members
NGO/private		54	5% of total team members
Independent		28	3% of total team members

We present these data based on individuals' participation in JEE teams (1118 observations, excluding copywriters as noted above), and not based on unique individuals in our dataset, as we are interested in the mix of expert backgrounds present in JEE visits. Country income classifications are based on World Bank data from 2016, the first year of JEE roll-out.
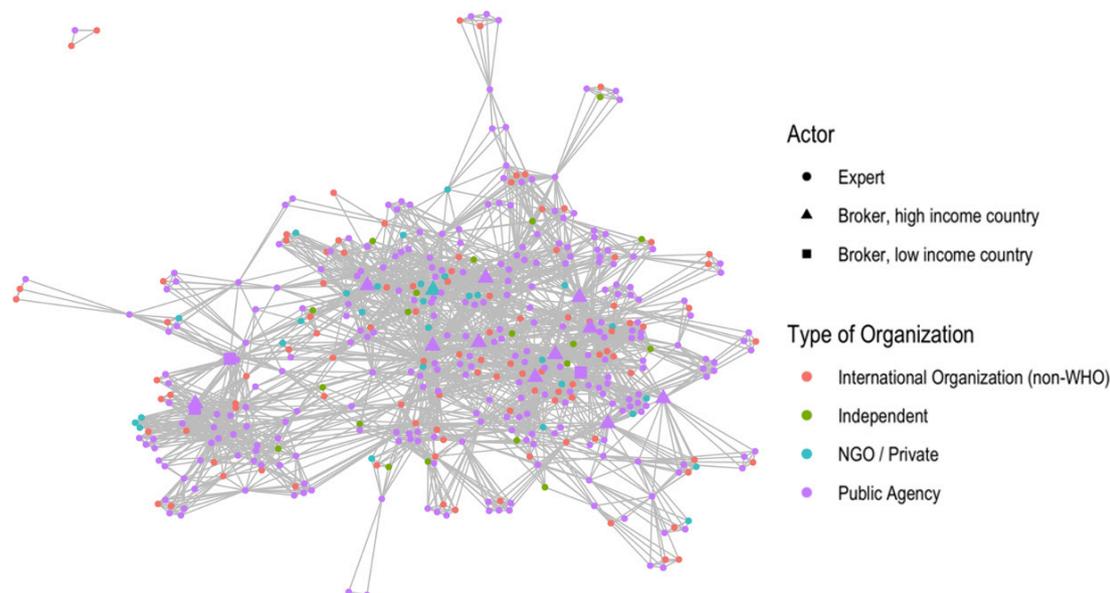JEE, Joint External Evaluation; NGO, non-governmental organisation.

Figure 1 Non-WHO broker network. NGO, non-governmental organisation.

by independent organisations such as hospitals and think-tanks.

To obtain more information on dynamics in the network we looked for the presence of brokers who (i) are not from the WHO and (ii) are not working in their home-region. Theoretically, it is these people who are able to collect and spread shared information through the network. Consequently, we made two methodological decisions. First, we removed from the analysis all 386 WHO staff included in our population. In line with interview findings, including WHO staff in the network analysis would have inflated their role, which was primarily to provide technical support, rather than to author the substantive, expertise-based analyses, recommendations and scores. Second, we denoted as brokers those who worked on four or more reports with at least three outside their home region. These brokers, engaged in evaluations of several countries around the world, hold and diffuse the type of expertise favoured by the WHO.

Figure 1 presents the results of our network analysis. One notable feature of the network is how densely connected it is. It is easy to imagine the JEE as a series of dispersed work teams with little connection to each other (such as the isolated triad at the top left of the figure, or the individual work teams sprouting to the bottom-right). Instead, we show that the network is tightly connected. Another notable feature is that other than the presence of one NGO/consultant broker in the centre of the network (the turquoise triangle), all other brokers in the network are from public agencies. The most prominent brokers are from high-income countries (12 out of the 15) that have been heavily involved in policy planning on pandemic preparedness and share similar goals to the WHO in spreading the JEE.

In sum, our analysis reveals that the JEE network forms a bounded epistemic community. Experts are preselected based on a skillset valorised by the WHO, that then

provides them with a template circumscribing the scope of their review and assessment. This template fosters cognitive closure around IHR compliance and is criticised by outsiders for excluding broader assessments of health systems (Interview #8). Even though the number of individuals who have been involved in these processes is high, there is a small subset of brokers who help ensure that knowledge travels within the network. In this context, the funding constraints for JEE mean that high-income countries—especially the US and Nordic governments—oversupply experts, suggesting limited diversity within JEE teams. When combined with an evaluation template that already encourages cognitive closure, homogeneity among the brokers can lead to groupthink and the legitimation of some conceptions of health emergency preparedness at the expense of others.[2]

Cobbling: the construction of GHSI indicators

The GHSI was spurred by the Obama Administration's interest in developing pandemic preparedness metrics following the West Africa Ebola crisis of 2014. Framed under an umbrella of 'health security', initiatives were taken to provide more information on threats and risks, including from zoonotic disease. With the entry of the Trump administration in 2016 the vanguard of pandemic preparedness indicators left government (Interview #6), and then joined a multistakeholder initiative already underway (Interview #8). Key policy entrepreneurs honed the GHSI as a collective effort between the NTI, the Johns Hopkins Centre for Health Security and the EIU. This extended a previous relationship between the NTI and the EIU in the production of the Nuclear Security Index from 2008 (Interview #7). As with the index on nuclear security, the ambition of the GHSI was address the 'enthusiasm gap' from countries in providing public knowledge on their pandemic preparedness capacities, and creating a ranking which could have a motivating

effect for policymakers (Interviews #7 and #8). Furthermore, the index was developed to overcome a perceived lack of objectivity in JEEs, given its reliance on a relatively small expert network (Interview #7). Instead, the GHSI would rely on public information, both providing a more objective metric of preparedness that was independently coded, and also spurring those not making such information publicly available to do so. As such, the GHSI is an 'advocacy tool' developed with a particular audience in mind: politicians, the media and countries being scored (Interview #8).

We collected detailed GHSI 2019 data and extracted the scores and underlying information on the policy area 'emergency preparedness and response planning' (like the aforementioned indicators, GHSI also scores countries on a host of other indicators such as antimicrobial resistance, immunisation or epidemiology workforce—such data were beyond the scope of our analysis). The preparedness subindex contains data on 195 countries, of which 88 scored zero on this indicator and two (the USA and Finland) scored 100. We extracted detailed information on two questions that fed into this subindex: 'Does the country have a national public health emergency response plan in place which addresses planning for multiple communicable diseases with epidemic or pandemic potential?' (question 3.1.1a) and 'Does the country have a specific mechanism(s) for engaging with the private sector to assist with outbreak emergency preparedness and response?' (question 3.1.2a). The answers to each of these questions contained a set of references, commonly three or more. We extracted these references as well. Finally, we collected the names, affiliations and educational background of the individuals involved in the GHSI: this included an 11-person core team and a 21-person international panel of experts.

As our objective is to understand the process of compiling the GHSI, we focus our attention on two issues: the team responsible for data collection and the types of information that fed into the indicator. The GHSI rests on quantitative and qualitative data, which is gathered through 140 questions (of which 89 are qualitative, requiring the coding of legislation or guidelines).

Data collection—summarized in figure 2—is primarily handled by the EIU, using its considerable international network to activate approximately 110 staff, with various linguistic skills to assess the data against a 'scoring guidance template'. The majority of this staff were economists, but occasionally other specialties, like health economists or public health experts, were also present (Interview #7). These data are then aggregated at regional and international levels by the EIU, and then transferred to the core GHSI team to conduct 'quality assurance/quality control checks' and subsequently assign scores on the basis of a scoring sheet.

The international expert panel is composed of people with senior roles in policymaking, think-tanks and academia, and commonly holding advanced degrees in relevant areas (PhD or MD). These individuals were selected based on their diverse disciplinary and regional backgrounds and being leaders in their field, which would grant high degrees of legitimacy to the Index. As one informant explained, 'we wanted people seen as trusted, credible experts in the field that could be critical in helping us build a framework that was rigorous and had credibility behind it' (Interview #7). The panel's involvement was particularly pronounced in the early stages of building the assessment framework, where they were asked to provide extensive critiques and feedback. Subsequently, panel members have been 'great advocates of the Index and they have brought it into their home institutions' (Interviews #7 and #8). Table 3 details those involved.

Second, the core team of the GHSI was co-led by Elizabeth Cameron and Jessica Bell, senior staffers at the NTI think-tank, and Jennifer Nuzzo, associate professor at the Johns Hopkins School of Public Health. The background of all three co-leaders has been in the field of health security, commonly having career trajectories at the intersection of policy roles and research (whether in academia or think-tanks). For example, Cameron—prior to joining the NTI in 2017—was Senior Director for Global Health Security and Biodefense on the US National Security Council, where she chaired the International Review Council of the Global Health Security Agenda (as noted
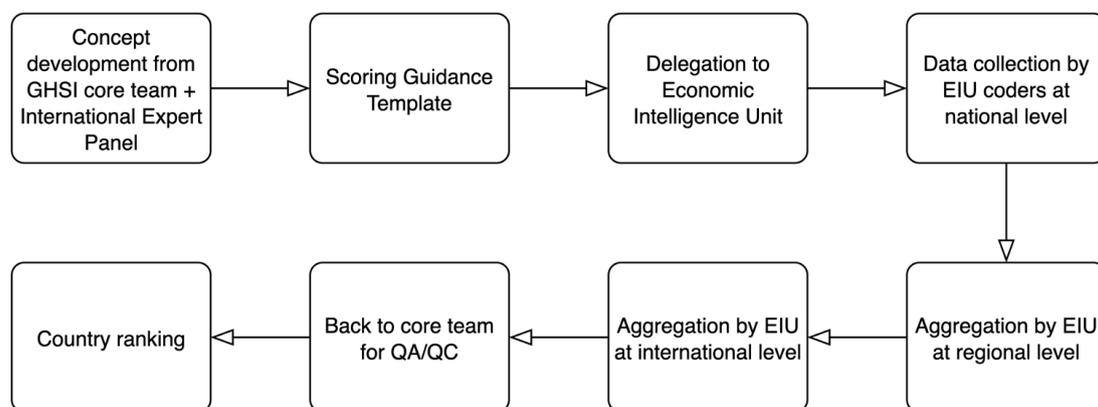

Figure 2 GHSI creation workflow. EIU, Economist Intelligence Unit; GHSI, Global Health Security Index; QA/QC, quality assurance/quality control.

Table 3 Global Health Security Index's international panel of experts

Name	Qualification	Role	Employer	Country
Wu Fan	MD, PhD	Deputy Director General	Shanghai Municipal Health Commission	China
Simo Nikkari	PhD Medicine	Chief Physician	Centres for Military Medicine and Biothreat Preparedness	Finland
Indira Nath	MD, DSc	Former Senior Professor	All India Institute of Medical Sciences	India
Pretty Multihartina	PhD	Director	Center for Health Determinants, Ministry of Health	Indonesia
Mukesh Chawla	PhD Economics	Chief Adviser	World Bank	International organisation
Lee Myers	Doctor of Veterinary Medicine	Manager	Food and Agriculture Organization (on behalf of USDA)	International organisation
Tomoya Saito	MD, MPH, PhD Medicine	Chief Senior Researcher	Department of Health Crisis Management	Japan
Tolbert G Nyenswah	MPH	Director General	National Public Health Institute	Liberia
Oyewale Tomori	PhD Virology	Professor of Virology	Nigerian Academy of Science	Nigeria
Malik Muhammad Safi	PhD	Director	System Strengthening and Information Analysis Unit	Pakistan
Ernesto Gozzer	MD, MSc	Associate Professor	Universidad Peruana Cayetano Heredia	Peru
Chaeshin Chu	PhD	Managing Director	Korea Centers for Disease Control and Prevention	South Korea
David Barash	MD	General Executive	General Electric Foundation	USA
David Blazes	MD	Physician Epidemiologist	Bill & Melinda Gates Foundation	USA
Scott F Dowell	MPH	Deputy Director	Bill & Melinda Gates Foundation	USA
Dylan George	PhD Biology	Vice President	In-Q-Tel	USA
Lawrence O Gostin	JD, LLD	Director	O'Neill Institute for National and Global Health Law, Georgetown University	USA
Wilmot James	PhD Sociology and African History	Senior Research Scholar	Columbia University	USA
Robert Powell	MA History	Editorial Director	The Economist Intelligence Unit	USA
Julius J Lutwama	PhD Entomology	Senior Principal Research Officer	Uganda Virus Research Institute	Uganda
Issa Makumbi	MD	Director	Public Health Emergency Operations Centre	Uganda

earlier, this Agenda led to the eventual establishment of the JEE). At the time of writing, Cameron resigned from the think-tank to return to her previous role at the National Security Council.[37] Below the leadership team are analysts and interns who contribute to the GHSI 'hive mind' along with the expert panel (Interview #8).

To achieve more nuance on the input into GHSI scores, we analysed the citations underpinning the scoring of countries. For the emergency preparedness indicators analysed here, we extracted the individual references, and classified them into four types of origin sources: a national-level public health agency (PUB), the WHO (including JEEs or other reports), other IGOs or other independent sources (IND) like hospitals, think-tanks or journal articles. Our desire with this data collection exercise was to trace possible discrepancies in the sources of knowledge employed for the scoring exercise—for

example, it could be that for low-resource settings the main data sources would be WHO reports, given the important role of the organisation in these contexts, while high-income country scores would draw more on official government reports. Table 4 presents descriptive statistics on our findings. In line with what our informants told us, we see that 82% of all references originated from evaluated countries' public authorities. The scores of only nine countries relied primarily on non-public authority documents (ie, such documents accounted for less than 50% of references cited): Oman, Cuba, Bangladesh, Saudi Arabia, Laos, Trinidad and Tobago, Somalia, Democratic Republic of Congo, and Antigua and Barbuda.

Given the emphasis on the production of an advocacy tool, the GHSI's form follows its function. To avoid criticism from the WHO that the GHSI would undermine their toolkits, JEE metrics were deliberately included

Table 4 Descriptive statistics on reference sources in Global Health Security Index

Source of reference	Mean per country	Median per country	SD	Minimum	Maximum	Reference count
Domestic public authority	7.3	7	4.2	0	41	1425
WHO	1.0	1	1.2	0	6	194
Intergovernmental organisation	0.5	0	1.1	0	7	88
Independent source	0.2	0	0.6	0	6	34
Total	*8.9*	*8*	*4.3*	*0*	*41*	*1741*

in the GHSI (Interview #8). And the explicit objective to influence media and policymakers was aided by the publication of explicit rankings that have a greater potential to capture attention (Interview #7). The GHSI's addition scoring then extends the range of factors included in the ranking, with weighting of these factors provided by the expert panel in deliberation with the core team.

The extended range in the GHSI has also led to its inclusion in judgements made by other organisations. An informant noted that 'the World Bank is using the Index a lot in Latin America, where there are no JEEs. They're using some of the [GHSI] questions, but also adding to them. Then there is a particular loan that countries can apply for to bolster their preparedness' (Interview #8). Similarly, the Gates Foundation is using the index in considering where to place funding to build pandemic preparedness (Interview #7). As such, as an advocacy tool, the GHSI is already being used in policy decisions and feeding into a search for more data.

In sum, GHSI is a multi-stakeholder advocacy tool to provide health security that relies on the aggregation of public data within the parameters established by the core team and a panel of experts, and collected by coders under the direction of the EIU. The 'scoring guidance' and ranking weighting established by the core team and panel of experts is then filtered by the EIU to reduce ambiguity and augment translation across the languages

needed for 195 countries. Data collection relies on publicly available sources, with countries scored negatively for not providing such information, including high-income countries (Interview #8). The danger with the index is that as the purpose is to bridge the 'enthusiasm gap' on pandemic preparedness in pursuit of a health security agenda, the questions contain some blind spots with regard to broader policy capacities and issues with trust in government and trust in science (as seen in the USA's #1 ranking).

CONCLUSION

Our purpose in this article is to open up the black box of indicators on pandemic preparedness, examining the construction of the e-SPAR, JEE and GHSI metrics. We summarise the main findings of our analysis in table 5. As shown there, while all three indicators nominally focus on the same issue, they pursue diverse analytical strategies and have distinct forms of output. We suggest that these strategies reflect logics of *collating*, *coordinating* and *cobbling*, respectively. Each approach carries benefits and drawbacks in how data and analysis are fed into evaluation scoring, and what these metrics then mean for the creation of global health policy.

More broadly, from our findings we can see three very different processes of organising knowledge on

Table 5 Summary of inputs and outputs from pandemic preparedness metrics

	Input			Output		
	Funding	**Data**	**Validation**	**Form**	**Audience**	**Potential bias**
e-SPAR	n/a	Self-reporting	None (no WHO mandate)	Coloured spreadsheet	WHO officials, global health policymakers	Inaccuracies
JEE	Gates Foundation for start-up costs, then countries sending experts, and WHO	Country self-assessment, and external expert assessment during visit	Experts in the mission team then reporting to the WHO on basis of a template	Lengthy policy reports, including scores (not available in a dataset format)	Donors, domestic health policymakers	Cognitive closure and groupthink
GHSI	Gates Foundation, Open Philanthropy Project, Robertson Foundation	Publicly available information collected by the Economist Intelligence Unit	Coders at Johns Hopkins and Nuclear Threat Initiative, following deliberations and on the basis of scoresheet	Coloured rankings and table	Domestic politicians, donors, media	Blind spots

e-SPAR, Electronic State Parties Self-Assessment Annual Reporting Tool; GHSI, Global Health Security Index; JEE, Joint External Evaluation.

e-SPAR   **JEE**   **GHSI**

◄───►

Governmental Authority   Foreign Expert Opinion

Figure 3 Spectrum of authority in pandemic preparedness indicators. e-SPAR, Electronic State Parties Self-Assessment Annual Reporting Tool; GHSI, Global Health Security Index; JEE, Joint External Evaluation.

pandemic preparedness, which reflect different ways in measuring and articulating national health system baseline capacities. These ways of organising can be located on a spectrum of authority in who handles the input and processing of data being considered, from entirely under the control of a government authority to being assessed by foreign expert opinion which governments may not influence. Figure 3 illustrates this, locating e-SPAR, JEE and GHSI on this spectrum of authority. From the left-hand side, e-SPAR allows maximum autonomy for governments to comply with their binding IHR commitments, but inevitably questions over the veracity of information emerge. Governments can exaggerate their scores in a game of status competition with their peers, or to signal to donors that they need financial and administrative support. The development of e-SPAR from a yes/no format into a scale is viewed by those in the WHO as a positive development, but the gaming of the system is protected by sovereign membership status in the IGO: government-provided information is not questioned and just collated into an online database.

In the middle, we have JEE: an assessment and scoring exercise designed and coordinated by the WHO and following a WHO template, but where knowledge is derived from foreign expert opinion. Those running the JEE have to navigate national governments and their preferences while enabling an expert network that relies heavily on staffing from high-income countries. Given this, homogeneity among experts can lead to groupthink that potentially diminishes open deliberation over country specifics, and adherence to a strict template can create cognitive closure in developing policy recommendations. As such, the JEE has a coherent aim in seeking to boost national capacities for pandemic preparedness, but how JEE experts and the WHO can engage an ongoing conversation with national policymakers—rather than intermittently (no country has been evaluated more than once thus far)—is the key issue of concern.

As seen on the right-hand side of figure 3, GHSI is also an indicator guided by foreign expert opinion, constructed by a hybrid body (neither public, nor intergovernmental) focused on health security considerations. By design, there was no active input from governments (although all governments were asked for comments, only 16 governments responded; Interview #7) and all data collection was guided by a predetermined questionnaire that had to be filled in with publicly available data. The form of bias that can emerge here relates to blind spots, especially given the excessive focus on the

procedural aspects of preparedness that misses out on governance issues that shape policy responses even in settings with underdeveloped health emergency procedures (as recognised by its developers[32]).

We recognise that there are limitations in our analysis. First, we could have included the entire universe of data available to add nuance (eg, other available indicators on zoonotic diseases, infectious disease workforce or antimicrobial resistance), but this is beyond the scope of our project that favoured a narrower focus on health emergency preparedness. Second, and most importantly, there are ongoing deliberations about pandemic preparedness metrics, making this topic a moving target. Recent discussions around the JEE and GHSI have included a desire to factor in more governance indicators (Interviews #1 and #8), but with no clear path on how experts could assess these or—if formalised—how the WHO or other actors could enforce compliance with governance objectives.

Our analysis of the construction of pandemic preparedness indicators contributes to a broader discussion on how information is gathered and how knowledge is shared. These forms of organisation shape what is visible and legible to policy audiences, and influence a range of other actors, including donors, media and the global health community. These indicators provide a valuable service in assessing and fostering data collection on national health systems' capacities. However, they can also contain inaccuracies, cognitive closure and blind spots that then enter global health policy conversations. This includes a tendency to focus on health security to the detriment of a conversation on what capacities are needed to improve national resilience to complex emergencies. While it is undoubtedly true that there is scope for improvement by including additional information, simply adding more data points is not the sole way forward. Ongoing discussions—including proposals by the WHO's Independent Panel for Pandemic Preparedness and Response—need to grapple not only with the issue of access to accurate information on preparedness, but also on how both governmental authority and expert opinion on these issues can be challenged and held to account.

Twitter Alexander Kentikelenis @Kentikelenis and Leonard Seabrooke @LenSeabrooke

ORCID iDs
Alexander Kentikelenis http://orcid.org/0000-0002-1543-4595
Leonard Seabrooke http://orcid.org/0000-0001-5581-3293

REFERENCES

1 Abbey EJ, Khalifa BAA, Oduwole MO, *et al*. The global health security index is not predictive of coronavirus pandemic responses among organization for economic cooperation and development countries. *PLoS One* 2020;15:e0239398.
2 Dalglish SL. COVID-19 gives the lie to global health expertise. *Lancet* 2020;395:1189.
3 Ekpenyong A, Pacheco MS. COVID-19: reflecting on the role of the who in knowledge exchange between the global North and South. *Glob Soc Policy* 2020;20:388–92.
4 Harris J. Science and democracy reconsidered. *Engaging STS* 2020;6:102–10.
5 Harris J, White A. The sociology of global health. *Sociol Dev* 2019;5:9–30.
6 Holst J. The world expects effective global health interventions: can global health deliver? *Glob Public Health* 2020;15:1396–403.
7 Holst J. Global Health - emergence, hegemonic trends and biomedical reductionism. *Global Health* 2020;16:42.
8 Richardson ET. Pandemicity, COVID-19 and the limits of public health 'science'. *BMJ Glob Health* 2020;5:e002571.
9 White House. Remarks by President Trump, Vice President Pence, and Members of the Coronavirus Task Force in Press Conference [Internet], 2020. Available: https://china.usembassy-china.org.cn/excerpts-rmarks-by-president-trump-vice-president-pence-and-members-of-the-coronavirus-task-force-in-press-conference/ [Accessed 23 Feb 2021].
10 Espeland WN, Sauder M. Rankings and reactivity: how public measures recreate social worlds. *Am J Sociol* 2007;113:1–40.
11 Merry SE. Measuring the world: indicators, human rights, and global governance: with Ca Comment by John M Conley. *Current Anthropology* 2011;52:S83–95.
12 Fukuda-Parr S, Yamin AE, Greenstein J. The power of numbers: a critical review of millennium development goal targets for human development and human rights. *J Human Dev Capabil* 2014;15:105–17.
13 Kelley JG, Simmons BA, eds. *The power of global performance indicators*. New York: Cambridge University Press, 2020.
14 Haas PM. Do regimes matter? Epistemic communities and Mediterranean pollution control. *Int Organ* 1989;43:377–403.
15 Chorev N. *The world Health organization between North and South*. Ithaca: Cornell University Press, 2012.
16 Kamradt-Scott A. WHO's to blame? The World Health Organization and the 2014 Ebola outbreak in West Africa. *Third World Q* 2016;37:401–18.
17 Kentikelenis AE, Babb S. The making of neoliberal globalization: norm substitution and the politics of Clandestine institutional change. *Am J Sociol* 2019;124:1720–62.
18 Keck EM, Sikkink K. *Activists beyond borders: advocacy networks in international politics*. Ithaca, NY: Cornell University Press, 1998.
19 Mackenzie JS, Drury P, Arthur RR, *et al*. The global outbreak alert and response network. *Glob Public Health* 2014;9:1023–39.
20 Kirmayer LJ. Cultural competence and evidence-based practice in mental health: epistemic communities and the politics of pluralism. *Soc Sci Med* 2012;75:249–56.
21 Löblová O. Epistemic communities and experts in health policy-making. *Eur J Public Health* 2018;28:7–10.
22 Georgalakis J. A disconnected policy network: the UK's response to the Sierra Leone Ebola epidemic. *Soc Sci Med* 2020;250:112851.
23 Barnes N. Paradoxes and asymmetries of transnational networks: a comparative case study of Mexico's community-based AIDS organizations. *Soc Sci Med* 2008;66:933–44.
24 Ottersen OP, Dasgupta J, Blouin C, *et al*. The political origins of health inequity: prospects for change. *Lancet* 2014;383:630–67.
25 Kentikelenis A, Rochford C. Power asymmetries in global governance for health: a conceptual framework for analyzing the political-economic determinants of health inequities. *Global Health* 2019;15:70.
26 Shriwise A, Kentikelenis AE, Stuckler D. Universal social protection: is it just talk? *Sociol Dev* 2020;6:116–44.
27 Feldbaum H, Michaud J. Health diplomacy and the enduring relevance of foreign policy interests. *PLoS Med* 2010;7:e1000226.
28 Kentikelenis AE, Seabrooke L. The politics of World Polity: Script-writing in international organizations. *Am Sociol Rev* 2017;82:1065–92.
29 Fukuda-Parr S, McNeill D. Knowledge and Politics in Setting and Measuring the SDG s: Introduction to Special Issue. *Glob Policy* 2019;10:5–15.
30 Koliev F, Sommerer T, Tallberg J. Compliance without coercion: effects of reporting on international labor rights. *Journal of Peace Research* 2020;0022343320910254.
31 Broome A, Quirk J. Governing the world at a distance: the practice of global benchmarking. *Rev Int Stud* 2015;41:819–41.
32 Ravi SJ, Warmbrod KL, Mullen L, *et al*. The value proposition of the global health security index. *BMJ Glob Health* 2020;5:e003648.
33 WHO. Guidance document for the State Party self-assessment annual reporting tool [Internet]. WHO reference number: WHO/WHE/CPI/2018.17. World Health Organization, 2018. Available: http://www.who.int/ihr/publications/WHO-WHE-CPI-2018.17/en/ [Accessed 13 Dec 2020].
34 Bonacich P. Factoring and weighting approaches to status scores and clique identification. *J Math Sociol* 1972;2:113–20.
35 McAneney H, McCann JF, Prior L, *et al*. Translating evidence into practice: a shared priority in public health? *Soc Sci Med* 2010;70:1492–500.
36 Takagi D, Yokouchi N, Hashimoto H. Smoking behavior prevalence in one's personal social network and peer's popularity: a population-based study of middle-aged adults in Japan. *Soc Sci Med* 2020;260:113207.
37 NTI. Statement from Ernest J. Moniz and Joan Rohlfing on the Appointment of Beth Cameron to Serve in the Biden-Harris National Security Council [Internet]. Nuclear Threat Initiative, 2021. Available: https://www.nti.org/newsroom/news/beth-cameron-to-serve-in-biden-harris-administration/ [Accessed cited 2021 Jun 3].