4.12 Weaknesses in a health-research system

Prior to the start of the COVID-19 pandemic, a group of researchers documented the weaknesses in the health-research system. They called for a reorganization of the system, including the structures (e.g., global collaborations like Cochrane) and incentives (e.g., from universities, funders and journals) that underpin it, in order to better meet the needs of decision-makers.(15-17) They were primarily concerned with three of the forms of evidence that decision-makers typically encounter, namely primary research (and specifically evaluation, especially randomized-controlled trials), evidence syntheses, and guidelines (and to a lesser extent technology assessments).

While some of the weaknesses became more apparent through the COVID-19 evidence response, the pandemic response also generated notable examples of efforts to address many of the weaknesses. Although the researchers were originally focused on health challenges and on select forms of evidence, many of the insights also apply to other societal challenges and to other forms of evidence. That said, a similar exercise will need to be undertaken for societal challenges and forms of evidence that are quite different from those described here. For example, the Intergovernmental Panel on Climate Change (IPCC) has helped a great deal with global coordination in their area of focus, and with spurring new approaches to modeling over long time horizons. However, the IPCC may also benefit from complementing these approaches with post-hoc evaluations of climate-change response options.

Pre-COVID weaknesses in the health-research system	Examples of weaknesses that became more apparent through the COVID-19 evidence response	Examples of efforts to address weaknesses through the COVID-19 evidence response
Lack of **global coordination** of evidence communities, with each ideally addressing a globally prioritized challenge using systematic and transparent methods and a full array of data sources *(e.g., study registries, regulatory agencies, and administrative databases)*	• Many topics prioritized by COVID-END's global horizon-scanning panel were never addressed by one or more 'best' evidence syntheses • Low signal-to-noise ratio: nearly 11,000 evidence syntheses about COVID-19 were reduceable to roughly 600 'best' evidence syntheses in the COVID-END inventory (as of 7 November 2021) based on four criteria: addressing a unique decision-relevant question, recency of the search for evidence, quality of the synthesis, and availability of a GRADE evidence profile	• COVID-END engaged 55 leading evidence-synthesis, guideline-development and technology-assessment groups, as well as citizen partners and evidence intermediaries, in efforts to reduce duplication and enhance coordination • PROSPERO encouraged those registering a protocol for a COVID-19 evidence synthesis to search for already registered protocols and to pick a new topic if duplication was likely (although 138 teams proceeded with a topic already registered by one of 57 other teams, including 14 addressing hydroxychloroquine and seven addressing tocilizumab) • GloPID-R (Global Research Collaboration for Infectious Disease Preparedness) engaged leading research-funding organizations in coordinating their rapid funding of primary research about COVID-19

Lack of focus of evidence communities on maintaining **living evidence syntheses** that examine all interventions addressing a prioritized challenge *(e.g., a network meta-analysis rather than pairwise comparisons only)*	• Only 13% of COVID-19 evidence syntheses self-identified as a living evidence synthesis (versus 52% in the COVID-END inventory where 'living' status was a criterion used to identify 'best' evidence syntheses) and more than two thirds addressed clinical management (rather than public-health measures, health-system arrangements, and economic and social responses) • Only 21% of living COVID-19 evidence syntheses had one update (after the first publication), 8% had two, and 13% had two or more, while the mean and median time between searches for syntheses with updates was 49 and 31 days, respectively • Many COVID-19 evidence syntheses addressed single drug treatments, so the COVID-END inventory transitioned to relying primarily on COVID-NMA and others looking across all drug treatments (and to including only syntheses of prognostic studies that include all available prognostic factors)	• Four evidence communities maintained high-quality living meta-analyses of all drug treatments, with one (COVID-NMA) supporting weekly updates of risk-of-bias assessments and GRADE certainty assessments
Lack of focus of evidence communities on identifying **harms** arising from interventions as well as benefits *(and more generally including a broader array of study designs and types of data)*	• Then-existing studies and syntheses made it difficult to understand what to make of reports about blood clots being experienced by select vaccine recipients	• A COVID-END team conducted a systematic review to complete a causality assessment of thrombotic thrombocytopenia that is temporally related to vaccine administration
Lack of sharing of **individual participant data** and its use to examine how findings vary by type of participant, setting or other factors, and hence how interventions can be better personalized or contextualized	• Many reports documented the lack of sharing of individual participant data (e.g., one review of 140 studies early in the pandemic found that data were shared from only one study – see bit.ly/31WQUxM)	• The COVID-19 Knowledge Accelerator advanced the methods needed to share computable expressions of evidence and guidance across platforms, and Vivli extended its platform to enable the sharing of COVID-19 trials data
Lack of **inclusion** in evidence communities of representatives from all relevant evidence groups *(e.g., researchers conducting primary studies like trials, evidence synthesizers and guideline developers)*, all relevant types of decision-makers, and all relevant types of evidence intermediaries	• Many reports described how citizens were less involved in COVID-19 research than they had been in other types of research before the pandemic, as well as about plain-language summaries of evidence syntheses not being available early in the pandemic (e.g., bit.ly/3kwCHhr)	• The National COVID-19 Clinical Evidence Task Force involved many health professionals (and their associations) and patients in their living guidelines, and they worked in partnership with evidence communities maintaining living network meta-analyses • Many groups engaged in modeling to help choose among available options (e.g., lockdowns) based on available evidence and expert opinion, and in some cases the context provided by decision-makers • Many groups prepared contextualized rapid syntheses at the request of decision-makers (with citizen partners in the case of many COVID-END rapid syntheses)
Lack of use by evidence communities of a range of **new approaches** to become more efficient and timely in their work *(e.g., machine learning and crowd-sourcing contributions to their work)*	• More than 18,000 studies had been uploaded to just one preprint server (medRxiv) by July 2021, dramatically shortening the time to publication (while having uncertain harms due to the lack of peer review) • Many use cases for machine-learning approaches in COVID-19 responses were identified in a medium-quality scoping review of 183 reports (bit.ly/3D7bTeV), but were not widely used early in the pandemic	• L*VE (Living Overview of Evidence) used machine learning to maintain a repository of primary studies and evidence syntheses, and the EPPI-Centre used machine learning to maintain a living evidence map

Lack of **reporting** about the gaps in and quality and transparency of primary studies (including conflicts of interest) as part of a feedback loop meant to support learning and improvement – for more details, see box 1 in this paper: (17)	• The results of many primary studies have been made available through media releases instead of through full research reports that can be critically appraised • Many reports noted that primary studies were found to have an intermediate to high risk of bias (e.g., 81% of the 713 articles including original patient data from a pool of 10,516 COVID-19 articles – see bit.ly/3HiI90X) and to have been retracted because of scientific misconduct • COVID-END prepared reports about evidence syntheses' lack of currency (91% and 61% in the full database and inventory of 'best' evidence syntheses, respectively, were based on searches completed more than 180 days earlier), medium or low quality (75% and 55%, respectively), and lack of an evidence profile (81% and 42%, respectively), as well as how rapid syntheses were more likely to be low quality than full syntheses (43% compared to 13%)	• RECOVERY (recoverytrial.net) and WHO COVID Solidarity Therapeutics Trial provided platforms for ultra-rapid, high-quality, multi-country trials of COVID-19 drug treatments • COVID-19 Evidence Alerts profiled quality-rated primary studies

evidencecommission@mcmaster.ca
www.evidencecommission.org
@evidencecomm