



4.12 Weak evidence in a health-evidence

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 selected forms of evidence, many of the insights also applied 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 springing 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 redactable to roughly 600 'best' evidence syntheses in the COVID-END in error (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 clinical 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 hydrocholoric acid and seven addressing ocularly) - 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 in error, where 'living' status was a criterion used to identify 'best' evidence syntheses) and more than 10 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 three or more, while the mean and median time between searches for syntheses was 49 and 31 days, respectively - Many COVID-19 evidence syntheses addressed single drug treatments, so the COVID-END in error transitioned to relying primarily on COVID-NMA and others looking across all drug treatments (and including only syntheses of prognostic studies that include all available prognostic factors)	For 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-e ising s dies and s nheses made i diffic i o nders and ha o make of reports abo blood clds being e periened b selecd accine recipien s • A COVID-END eam cond d ed a s s ematic re ie o comple e a ca saly assessmen of thrombotic thromboc openia ha is temporall relaed o accine adminis ration
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	• Man reports doc men ed he lack of sharing of indi id al participan da a (e.g., one re ie of 140 s dies earl in he pandemic fo nd ha da a ere shared from onl one s d see bit.ly/31WQUxM) • The COVID-19 Kno ledge Accelerat or ad anced he methods needed o share comp able e pressions of e idence and g idance across pla forms, and Vi lie ended s pla form o enable he sharing of COVID-19 rials da a
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	• Man reports described ho cli ens ere less in ol ed in COVID-19 research han he had been in o her pes of research before he pandemic, as ell as abo plain-lang age s mmaries of e idence s nheses no being a ailable earl in he pandemic (e.g., bit.ly/3kwCHhr) • The National COVID-19 Clinical E idence Task Force in ol ed man heal h professionals (and heir associa ons) and pa ien s in heir li ing g idelines, and he orked in partnership h e idence comm nies main aining li ing ne rk me a-anal ses • Man gro ps engaged in modeling o help choose among a ailable op ions (e.g., lockdo ns) based on a ailable e idence and e per opinion, and in some cases he con e pro ided b decision-makers • Man gro ps prepared con e ali ed rapid s nheses a he req es of decision-makers (h cli en partners in he case of man COVID-END rapid s nheses)
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 han 18,000 s dies had been ploaded o j s one preprin ser er (medR i) b J l 2021, dramai call shorening he ime o p blica ion (hile ha ing ncer ain harms d e o he lack of peer re ie) • Man se cases for machine-learning approaches in COVID-19 responses ere iden ified in a medi m-q all scoping re ie of 183 reports (bit.ly/3D7bTeV), b ere no idel sed earl in he pandemic • L*VE (Li ing O er ie of E idence) sed machine learning o main ain a reposi or of primar s dies and e idence s nheses, and he EPPI-Cen re sed machine learning o main ain a li ing e idence 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 res s of man primar s dies ha e been made a ailable hro gh media releases instead of hro gh f ll research reports ha can be cricaly appraised • Man reports no ed ha primar s dies ere fo nd o ha e an intermedia e o high risk of bias (e.g., 81% of he 713 articles incl ding original pa ien da a from a pool of 10,516 COVID-19 articles see bi l /3HiI90X) and o ha e been re aded beca se of scien ific miscon d • COVID-END prepared reports abo e idence s nheses lack of c renc (91% and 61% in he f ll da abase and in en or of bes e idence s nheses, respec i el , ere based on searches comple ed more han 180 da s earler), medi m or lo q all (75% and 55%, respec i el), and lack of an e idence profile (81% and 42%, respec i el), as ell as ho rapid s nheses ere more likel o be lo q all han f ll s nheses (43% compared o 13%) • RECOVERY (reco er rial ne) and WHO COVID Solidar Therape ics Trial pro ided pla forms for ra-rapid, high-q all , m i-co n r rials of COVID-19 dr g reat men s • COVID-19 E idence Ale s profiled q all -ra ed primar s dies