Form of evidence	Examples of quality-assessment tools
Types of evidence for which quality-assessment tools exist	
Data analytics	ROBINS-I (riskofbias.info) for observational studies, such as those that have examined associations between selected factors (including interventions) and selected outcomes, where there is a risk of bias from: confounding (where the observed relationship between a factor and an outcome, differs from the true relationship because of one or more additional factors that are not accounted for) selection of participants into the study classification of intervention(s) deviations from intended intervention missing data measurement of outcomes selection of the reported result
Evaluation	Risk of Bias (RoB) 2 (riskofbias.info) for randomized-controlled trials, where the risk of confounding is less, but where there is a risk of bias from some (albeit fewer) of the same sources as above: randomization process deviations from the intended interventions missing (outcome) data measurement of outcomes selection of the reported result
Behavioural/ implementation research	See other rows for the relevant types of studies or synthesis
Qualitative research	JBI critical appraisal checklist for qualitative research (bit.ly/31Lsib1), where different considerations come into play, such as: congruity between the research methodology and the research question, data-collection methods, data representation and analysis, and results interpretation, as well as between stated philosophical perspective and the methodology reflexiveness on the part of the researcher, such as statements locating the researcher culturally and theoretically, and addressing the researcher's influence on the research and vice versa representation of study participants and their voices flow of conclusions from the analysis and interpretation of the data
Evidence synthesis	See above for the relevant types of studies considered in the evidence synthesis
	A MeaSurement Tool to Assess systematic Reviews (AMSTAR; amstar.ca) for the quality of the evidence synthesis, where the risk of bias can arise from: identification of all potentially relevant studies through a comprehensive search of both published and greater literature and without language restrictions selection of all studies addressing the research question using explicit criteria about study designs and about participants, interventions/factors, comparisons and outcomes, and with at least two reviewers applying the criteria quality appraisal of and data extraction from all included studies synthesis of findings from all included studies Note that there are two versions of AMSTAR: 1) the original version that can be applied across all types of syntheses, albeit with some criteria removed from both the numerator and denominator; 2) a second version of AMSTAR that is more specifically relevant to syntheses of randomized-controlled trials
	Grading of Recommendations, Assessment, Development and Evaluations (GRADE; bit.ly/3C9pMrx) for the certainty of evidence for the outcomes of an intervention, with: certainty rated down because of risk of bias (with evidence from randomized-controlled trials starting at high certainty and evidence from observational studies starting at low quality and then being adjusted based on RoB2 or ROBINS-I), imprecision (e.g., one or two small studies), inconsistency (e.g., two studies showing different findings), indirectness (e.g., surrogate measures used or atypical settings studied), and publication bias (e.g., more common with observational studies because of the lack of studies registries or with industry-funded studies because of the commercial incentive to publish positive studies) certainty rated up for large magnitude of effect, dose-response gradient, and when all residual confounding would decrease the magnitude of effect
	GRADE CERQual (cerqual.org) for the certainty of evidence for the qualitative representation of a phenomenon of interest, with: certainty rated down because of concerns about methodological limitations (because problems in the studies were designed or reported were identified using a critical-appraisal tool like the JBI one above), relevance (because the context in which the primary studies were conducted are substantively different from the context of the synthesis question), coherence (because some of the data contradict the findings or are ambiguous), and adequacy (because the data are not sufficiently rich or only come from a small number of studies or participants) [illegible]

Technology assessment / cost-effectiveness analysis	In ernational Ne ork of Agencies for Heal h Technolog Assessmen (INAHTA) checklis (bit.ly/2YJVMVK) for he q ali of echnolog assessmen s, h o of he 14 q es ions addressing he approach o s nthesi ing he e idence (h promp s similar o AMSTAR) and and her q es ion addressing he her he assessmen as con e ali ed hro gh an accompan ing cos -effec i eness anal sis (h local meaning na ional or s b-na ional cos ing da a), and considera ion of local legal, e hical and social implica ions
	Dr mmond checklis of cos -effec i eness anal ses (bit.ly/3FbnB8R), and for economic e al a ions more generall , h q es ions abo s d design, da a colle ion, and he anal sis and in erpre a ion of res s
	Philips checklis for cos -effec i eness anal ses ha incl de a decision-anal ic modeling componen (bit.ly/3FcWBGc) h q es ions abo he s r c re of he model (e.g., e plici ra ionale, j s ified ass mp ions and appropria e ime hori on), he da a sed (e.g., baseline probabil ies from obser a ional s dies, rea men effec s from randomi ed-con rolled rials, and assessmen s of fo r pes of nce ain , namel he s r c re of he model, he me hodological s eps follo ed, he he erogene in he pop la ion s died, and he parame ers sed), and he consis enc (in ernal and e ernal) here is also he complemen ar TRUST ool o assess nce ain ies in decision-anal ic models (bit.ly/3quFSKp)
Guidelines	AGREE II ool (bit.ly/30qyFAb) for assessing he de elopmen , repor ing and e al a ion (or q ali appraisal) of g idelines, hich ses 23 ems gro ped in o si domains, each of hich is scored independen l : scope and p rpose described s akeholder (ci i en/pa ien and professional) in ol emen rigo r of de elopmen (h e idence s n heses sed as an inp , a rob s recommenda ions-de elopmen process, and recommenda ions linked o he s ppor ing e idence) clari of presen a ion applicabil edi orial independence (in rela ion o f nder and panel members' conflic s of in eres)
	Grading of Recommenda ions, Assessmen , De elopmen and E al a ions (GRADE; bit.ly/3C9pMrx) for assessing he s reng h of recommenda ions, hich ses fo r ke considera ions: balance be een desirable and ndesirable o comes (rade-offs), aking in o acco n bes es ima es of he magni de of effec s on desirable and ndesirable o comes, and he impor ance of hose o comes (es ima ed pical al es and preferences) confidence in he magni de of es ima es of effec s of he in er en ions on impor an o comes (see GRADE in a pre io s ro) confidence in al es and preferences and heir ariabil reso rce se
Types of evidence for which quality-assessment tools don't yet exist	
Modeling	No idel accep ed ool e e is s for mos pes of models, ho e er, here are some general q es ions ha can be asked abo models (m ch like hose lis ed as par of he Philips checklis abo e), s ch as: s r c re of he model (e.g., e plici ra ionale, j s ified ass mp ions, and appropria e ime hori on) da a sed (e.g., baseline probabil ies from obser a ional s dies, in er en ion effec s from a range of so rces*, and assessmen s of fo r pes of nce ain , namel he s r c re of he model, he me hodological s eps follo ed, he he erogene in he pop la ion s died, and he parame ers sed) consis enc (in ernal and e ernal) a ailabil of he so are or ool so ha i can be assessed b o hers *One of he challenges h COVID-19 as ha s d designs pical sed o cap re in er en ion effec s, s ch as randomi ed-con rolled rials, ere e hicall or logis icall diffic l and/or ook ime o comple e, so o her s d designs needed o be sed and e per opinion needed o be so gh (and here are approaches ha enable his o be done in a a ha is s s ema ic and ransparen , s ch as SHELF see bit.ly/30nteC4)
Approaches used with certain types of evidence for which quality-assessment tools don't yet exist	
Artificial intelligence	No idel accep ed ool e e is s