Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Original Research

Making science computable: Developing code systems for statistics, study design, and risk of bias

Brian S. Alper a,*, Joanne Dehnbostel a, Muhammad Afzal b, Vignesh Subbian c, Andrey Soares d, Ilkka Kunnamo e, Khalid Shahin a, Robert C. McClure f, For the COVID-19 Knowledge Accelerator (COKA) Initiative

a Computable Publishing LLC, Ipswich, MA, United States
b Sejong University, Seoul, South Korea
c The University of Arizona, Tucson, AZ, United States
d University of Colorado, Aurora, CO, United States
e Duodecim Medical Publications Ltd, Helsinki, Finland
f MD Partners, Inc., Lafayette, CO, United States

ARTICLE INFO

Keywords:
- Ontology
- Code system
- Terminology
- Evidence-based medicine
- Science communication
- Research literature

ABSTRACT

The COVID-19 crisis led a group of scientific and informatics experts to accelerate development of an infrastructure for electronic data exchange for the identification, processing, and reporting of scientific findings. The Fast Healthcare Interoperability Resources (FHIR®) standard which is overcoming the interoperability problems in health information exchange was extended to evidence-based medicine (EBM) knowledge with the EBMonFHIR project. A 13-step Code System Development Protocol was created in September 2020 to support global development of terminologies for exchange of scientific evidence. For Step 1, we assembled expert working groups with 55 people from 26 countries by October 2020. For Step 2, we identified 23 commonly used tools and systems for which the first version of code systems will be developed. For Step 3, a total of 368 non-redundant concepts were drafted to become display terms for four code systems (Statistic Type, Statistic Model, Study Design, Risk of Bias). Steps 4 through 13 will guide ongoing development and maintenance of these terminologies for scientific exchange. When completed, the code systems will facilitate identifying, processing, and reporting research results and the reliability of those results. More efficient and detailed scientific communication will reduce cost and burden and improve health outcomes, quality of life, and patient, caregiver, and healthcare professional satisfaction. We hope the achievements reached thus far will outlive COVID-19 and provide an infrastructure to make science computable for future generations. Anyone may join the effort at https://www.gps.health/covid19_knowledge_accelerator.html.

1. Introduction

Crisis leads to innovations. The COVID-19 crisis stimulated collaborative efforts resulting in a breakthrough in the communication of evidence in scientific literature. Today the evidence is not reported in a form that computers can understand. Evidence is not yet expressed in precise, unambiguous format (i.e., computable formats). The near-infinite variations in how evidence can be expressed using natural language means that it requires substantial expertise and contextual awareness for people to determine if the evidence matters, to interpret what the evidence means, and to determine the certainty of these interpretations. To make scientific evidence shareable, interoperable, and computable, it is essential to use standardized concepts from controlled terminologies and vocabularies. This article introduces early efforts to develop an infrastructure for electronic data exchange for the identification, processing, and reporting of scientific findings, and presents a 13-step Code System Development Protocol created to support global development of terminologies for exchange of scientific evidence.
2. Background

2.1. Introduction to Fast Healthcare Interoperability Resources

Fast Healthcare Interoperability Resources (FHIR®) is rapidly overcoming the seemingly intractable interoperability problem in the sharing and exchange of health information [1]. FHIR solves the interoperability problems by breaking down key units of data exchange into resources. Each FHIR resource instance describes a distinct identifiable entity, and each FHIR resource has a FHIR StructureDefinition Resource instance that describes the set of data element definitions and their rules of use that define the FHIR specification itself. Rather than forcing all health-related knowledge to fit one organizational pattern for a common structural model, FHIR enables resource-specific structure definitions to enable the most efficient and flexible approach. Health Level 7 International (HL7®), the standards developing organization that created and maintains FHIR, addresses the human problem in universal agreement to a technical standard by supporting open, transparent, logical processes and systems for people from all perspectives to participate [2].

2.2. Extension of FHIR to evidence-based medicine

There is currently no widely implemented standard that overcomes the seemingly intractable interoperability problem of sharing and exchange of computable representations of scientific knowledge. Facing such challenges with the communication of scientific knowledge to inform healthcare decision making, communities within and across researchers, systematic reviewers, guideline developers, and healthcare professionals have advanced human-interpretable expectations for trustworthy interpretation and application of scientific knowledge [3]. This area is often labeled evidence-based medicine (EBM), evidence-based practice, or evidence-based healthcare [4,5].

HL7 approved a project in 2018 to develop FHIR Resources for Evidence-Based Medicine Knowledge Assets (EBMonFHIR) [6]. In the following 18 months via weekly web meetings and five Connectathons, the EBMonFHIR project created FHIR StructureDefinition Resources for Evidence, EvidenceVariable, Statistic, and OrderedDistribution FHIR resources.

- The EvidenceVariable Resource is used to describe a variable used in statistical expressions, with one or more of defining characteristics expressed using standardized concept codes (i.e., codable concepts [7]).
- The Statistic Resource supports the expression of a statistic, including the numerical values, the related attributes which are also statistics, and the type of statistic as a codable concept [8].
- The OrderedDistribution Resource supports expression of a statistical array [9].
- The Evidence Resource supports expression of the statistics for a distinct combination of variables and the certainty of the interpretation of the statistics [10].

2.3. Extension of EBMonFHIR to COVID-19 Knowledge Accelerator

Multiple groups in the EBM community sought to use EBMonFHIR resources to support efforts related to global collaboration, cooperation and coordination for identification, evaluation, and reporting of COVID-19 evidence. Participating and related efforts include Agency for Healthcare Research and Quality (AHRQ) evidence-based Care Transformation Support initiative (ACTS) [11], ACTS COVID-19 Guidance-to-Action Collaborative [12], Australian National Clinical Evidence Taskforce [13], Centers for Disease Control and Prevention’s Adapting Clinical Guidelines for the Digital Age [14], COVID-19 Advanced Literature Classifier (CALC) [15], COVID-19 DistillerSR Access [16], COVID-19 Evidence Alerts from McMaster Plus [17], COVID-19 Evidence Network to support Decision making (COVID-END) [18], COVID-19 Open Research Dataset (CORD-19) [19], HL7 Biomedical and Regulation Research Group (BRR) Work Group [20], HL7 Clinical Decision Support Work (CDS) Group [21], Librarian Reserve Corps [22], the LIVING Project [23], Mobilizing Computable Biomedical Knowledge (MCBK) [24], and Systematic Review Data Repository (SRDR) [25].

On March 30, 2020, we started the COVID-19 Knowledge Accelerator (COKA) and by July had more than 150 working meetings with more than 40 active participants from more than 25 organizations from academia, industry, government, and nonprofits in 7 countries [26]. The COKA developed 10 active working groups meeting virtually 12 times per week. COKA efforts revised the FHIR Statistic Resource to include expressions of the statistical model. COKA efforts also created two more FHIR StructureDefinition resources: Citation Resource to support exchange of about 100 elements used to identify articles referenced for scientific reporting [27], and EvidenceReport Resource to support compositions of all the other resources in many combinations [28].

Across the six FHIR resources maintained by the EBMonFHIR/COKA efforts, there were more than 30 elements that would benefit from the use of standardized encoded concepts. Some concepts can be expressed with commonly used code systems such as SNOMED CT® [29], RxNorm [30], and LOINC® [31]. However, we discovered many situations where we could not find a comprehensive code system that was functionally applicable for the concepts commonly communicated.
3. Methods

3.1. Development of code system development protocol

We initially developed code systems [32] with pragmatic approaches by using codable concepts found in other code systems where available (such as the STATistics Ontology [STATO] [33] and National Cancer Institute thesaurus [NCIt] [34]) and developing mnemonic codes for terms commonly used by the EBMonFHIR and COKA participants. Though functional for the growing but small community, the desire for interoperability with many related communities included those represented in the HL7 CDS, Clinical Quality Information, BRR, and Vocabulary Work Groups. This demanded development of methods to support open, multinational, multidisciplinary input; comprehensive attention to existing ontologies; global consensus development; and sustainability planning.

Through multiple open virtual web meetings and shared documents, we developed a Code System Development Protocol (full protocol in Appendix A, related image in Fig. 1) which includes 13 steps [35]:

1) Assemble an expert working group.
2) Identify tools or systems commonly used today to express relevant concepts.
3) Map out a single list of non-redundant concepts to support common uses.
4) Identify existing ontologies that are openly available without restrictions.
5) Map related terms and definitions across the ontologies.
6) Define preferred terms, alternative terms, and definitions for the new code system.
7) Identify code system entries with universal agreement by the expert working group.
8) Deliberate suggested changes and reach universal agreement for code system entries where possible.
9) Deliberate unresolved disagreements and reach at least 80% agreement for code system entries where possible.
10) Determine the relative contribution of ontologies to the code system and seek further collaboration for heavily used ontologies.
11) Publish the initial version of the new code system.
12) Evaluate implementation of the code system and refine the system as needed.
13) Maintain continued support to adjust the code system based on changes in the prior 12 steps.

3.2. Scope setting

We selected four domains for initial application of the Code System Development Protocol and defined them as [35]:

- **The Statistical Type Code System** will be used to precisely classify univariate statistics (such as mean, median, and proportion), comparative statistics (such as relative risk, mean difference, and odds ratio), and statistic attribute estimates (such as confidence interval, p value, and measures of heterogeneity). Consistent reporting across systems will facilitate interoperability for science communication.

- **The Statistical Model Code System** will precisely communicate characteristics that define the model used for a statistic. Science reports often do not convey complete information about statistical models. Model characteristics may include concepts such as fixed-effects analysis, linear regression, and Mantel-Haenszel method for pooling. Consistent reporting of statistical models will facilitate interoperability for science communication.

- **The Study Design Code System** will be used to precisely describe methodology characteristics of scientific observations including exposure introduction (such as interventional or observational), cohort definition (such as parallel, crossover or case-control), and group assignments (such as block randomization, every-other quasi-randomization, or non-randomized). Consistent reporting of research study design across systems will facilitate interoperability for science communication.

- **The Risk of Bias Code System** will be used to precisely describe concerns with methods or reporting of scientific observations including selection bias (such as gaps in randomization or allocation concealment), performance bias (such as gaps in blinding), and analysis bias (such as gaps in intention to treat analysis or selective analysis reporting). Consistent reporting of risk of bias across systems will facilitate interoperability for science communication.

3.3. Step 1: Assemble an expert working group

For Step 1, we developed an Invitation to Join an Expert Working Group for any of the four code systems (Statistic Type, Statistical Model, Study Design, Risk of Bias). Joining the group was open to anyone and group members could self-identify their expertise. Relevant expertise for a code system could include without limitation experience evaluating or expressing the concepts to be included in the code system, either for human interpretation or for machine interpretation.

We shared the invitations through multiple communities (mostly via email distribution lists) including the COKA Initiative, COVID-END, the evidence-based healthcare (EBH) listserve, Grading of Recommendations, Assessment, Development and Evaluation (GRADE) Working Group, the Developing and Evaluation Communication strategies to support Informed Decisions and practice based on Evidence (DECIDE) project participants, the AHRQ evidence-based practice centers (EPCs), the HL7 CDS and BRR work groups, the Society for Clinical Trials, the Society for Participatory Medicine, International Society for Clinical Biostatistics, and Patient-Centered Outcomes Research Institute (PCORI).

3.4. Step 2: Identify commonly used tools and systems

For Step 2, we asked Expert Working Group members to identify sources to signal the scope of (or common need for) a code system, namely tools or systems in common current use for reporting concepts relevant to the code system.

3.5. Step 3: Create lists of non-redundant concepts

For Step 3, we started with one of the common tools or systems, identified a series of non-redundant concepts for expression to support it, and provided a categorical classification. We then mapped the next identified tool or system, matched concepts where possible, added more concepts where needed, and adjusted the categorical classification. The process was shared openly during weekly Steering Group web meetings and summarized for the Expert Working Group by email distribution lists with open links to the Step 3 mapping spreadsheets.

3.6. Time course for initial development

The COVID-19 Knowledge Accelerator consists of 10 active working groups meeting a total of 12 times weekly in open web meetings. Several working groups were developing code systems and the discussions about a common approach started on August 24, 2020. The first draft of a Code System Development Protocol with 11 steps was created on August 28. The protocol was finalized on September 17. Initial efforts were started ahead of wider dissemination of invitations. Invitations to join the expert working groups were sent widely during the week of September 21. All participants were asked to comment by an October 14 cutoff date for communicating the degree of contribution to Step 3 for version 1.0.0 of the code systems.

We report here the results of Steps 1–3 of this effort as of October 14,
2020. These results are not complete in terms of code system development as they do not include definitions or codes and may change through the remaining steps. These remaining steps, and the overall protocol, share and build upon principles and practices in existing ontology development methods [36,37,38]. Key aspects such as reusing existing ontologies, enumerating important terms (ie, concepts) across ontologies, and the overall iterative and agile nature of ontology development are well represented in our code system development protocol. We presented our protocol and preliminary findings in an October 30 Workshop on COVID-19 Ontologies (https://github.com/CIDO-ontology/WCO). In November of 2020, we met with ontology developers of the STATO and Ontology of Biological and Clinical Statistics (OBCS), both of which are Open Biological and Biomedical Ontologies (OBO) Foundry recognized ontologies. The ontology developers found our work valuable for identifying gaps, alignments, new terms, and other improvements for existing ontologies and potentially for creating an application ontology.

Table 1
Demographics of 55 Members of Expert Working Groups.

| Country (total 26) | Australia (1), Bangladesh (2), Brazil (2), Canada (5), Costa Rica (1), Czech Republic (1), Egypt (1), Finland (2), France (1), Ghana (1), Greece (2), India (2), Ireland (1), Italy (2), Japan (1), Lebanon (1), Malaysia (1), Nigeria (4), Peru (1), Romania (2), South Africa (1), South Korea (1), Sri Lanka (1), Switzerland (2), United Kingdom (2), United States (14) |
| Type of expertise*%, n (%) | Researcher 42 (76%), Evaluate scientific concepts 34 (62%), Systematic Reviewer 32 (58%), Statistician 23 (42%), Guideline developer 14 (25%), Developer of reporting systems 12 (22%), Learner 10 (18%), Software engineer/Informatics specialist 10 (18%), Write-in responses Librarian (3), Teacher of medical literature evaluation (2), Clinician/health professional, Terminologist, Standards developer, Qualitative researcher, Book author |
| Age, n (%) | 18–25 years 2 (4%), 26–40 years 16 (29%), 41–55 years 21 (38%), 56–69 years 13 (24%), 70+ years 1 (2%), Not shared 2 (4%) |
| Sex, n (%) | Female 18 (33%), Male 36 (65%), Not shared 2 (2%) |
| Race/ethnicity*, n (%) | Asian 11 (20%), Black 6 (11%), Hispanic/Latino 6 (7%), Indigenous 2 (4%), White 27 (49%), Not stated 7 (13%) |

*More than one selection may apply to each person.

4. Results

4.1. Expert working groups

As of October 10, 2020, a total of 55 people from 26 countries in 6 continents joined an Expert Working Group for up to four code system development efforts (see Table 1 and Appendix B).

4.2. Initial results (Step 2 and Step 3)

Twenty-three commonly used tools and systems were applied across the four code systems, ranging from 2 to 12 per code system (Table 2). There were 368 non-redundant concepts (draft display terms for a code system) identified across the four code systems, ranging from 53 to 170 per code system (Table 2, Appendices C, D, E and F).

Table 2
Step 2 and Step 3 Results to Inform Code System Development.

| Code System | Tools and Systems Considered | # draft codable concepts |
|-------------|------------------------------|-------------------------|
| Statistic Type | - StatisticalType code system defined by the FHIR project [39] | 88 |
| Model | - StatisticalAttributeEstimateType code system defined by the FHIR project [40] | 53 |
| Design | - ObservationsMethodAggregate value set from HL7 V3 ObservationMethod code system [41] | 57 |
| Risk of Bias | - Cochrane Review Manager (RevMan) [42] | 170 |

As of October 10, 2020, a total of 55 people from 26 countries in 6 continents joined an Expert Working Group for up to four code system development efforts (see Table 1 and Appendix B).
5. Discussion

5.1. Progress toward code system development

Coordinating 55 experts from 26 countries to identify 198 concepts for the development of code systems for scientific methodology (statistics and study design) and 170 concepts for the assessment of quality of evidence (risk of bias) is an early step in what is needed to support interoperable data exchange for scientific communication.

Next steps include mapping concepts across ontologies, reaching universal or near-universal agreement for common code systems for data exchange, and continuous adaptation to meet needs discovered in implementation.

The COKA effort will benefit from the newly crafted HL7 Unified Terminology Governance (UTG) process wherein terminology artifacts, such as the code systems and mappings we are creating, are published by HL7 [62]. The UTG approach aligns with our protocol by subjecting the artifacts created to an open comment and review process. The UTG process starts with transforming the code system and concept map terminology content into FHIR code system and concept map artifacts, typically represented in FHIR JSON or XML [63]. Once the content is entered into the UTG environment, it exists as a set of proposed changes to the core HL7 terminology. Those proposals are made available for review and comment within the UTG environment, consistent with steps 12 and 13 of our protocol. Once comments on the proposed artifacts are resolved and voting requirements are met, if approved, the terminology additions are merged into the HL7 terminology environment at terminology.hl7.org, which is updated and made available through a continuous integration process [63]. In this way, updates and improvements for any content can be developed, proposed, reviewed, improved, voted on and released within a documented environment aligned with the American National Standards Institute (ANSI)-sectioned HL7 ballot process, and ultimately published as part of the official HL7 terminology content.

Our protocol (step 6) includes entering data into an ontology web editor which by design would include top-level ontology concepts (classes, hierarchy, attributes) such as those represented in the Basic Formal Ontology [64] to help refine the classes and hierarchy. The consideration of the FHIR CodeSystem Resource StructureDefinition [65] in preparation for the UTG approach helped us realize we can represent these top-level ontology concepts as property elements within the CodeSystem Resource and we are currently considering modifying step 6c of our protocol to use FHIR tooling directly instead of a web ontology editor.

5.2. Strengths and limitations

Strengths of our approach include a substantial spirit of comradery across many diverse people facing a common challenge, multidisciplinary engagement, and coordination with global systems for standards development. In addition, use of FHIR as the underlying standard provides support from a method demonstrated to meet the interoperability needs for a similarly complex global community.

Limitations include the rapid timeline for development, having processed the initial listing of hundreds of concepts in just a month or so. There will undoubtedly be multiple revisions. The current list does not include outcome-specific statistic types (such as mortality for observed proportion or incidence related to death) or application-specific statistic types (such as recall instead of sensitivity for the application to information retrieval). This approach was purposefully taken to maximize simplicity and flexibility. Also, it is not yet established what resources will be needed to complete and maintain the code systems. For the initial effort, the degree of volunteerism and availability was influenced substantially by COVID-19 and we hope the spirit will continue for application across other domains.

5.3. Example for computable evidence

We demonstrate a computable expression of evidence [66] with the results (summary effect estimate) of a meta-analysis of three randomized trials [67,68,69] for the effect of remdesivir on 14-day mortality in patients with COVID-19 pneumonia. This example includes 43 instances of a “coding” element to express codable concepts with a “system” element to denote the code system, a “code” element to denote the specific code, and a “display” element for human-readable interpretation of the code. For example, the JSON includes (see Table 3):

This example of computable evidence uses existing codes in published code systems where available, and these may differ from the code systems in development. Where not available, we use “system”: “not yet published” and “code”: “not yet defined” and this shows the need for creation or extension of code systems. One can search the JSON in this example to find 1 code related to study design (“display”: “randomized trial”), 8 codes related to statistic type (“display” values of “Relative Risk”, “Confidence Interval”, “Z-score”, “P-value”, “I-squared”, “Cochran’s Q statistic”, “degrees of freedom”, and “Tau squared”), 4 codes related to statistic model (“display” values of “Meta-analysis”, “Fixed-effects”, “Random-effects”, and “Dersimonian-Laird method”), and 1 code related to risk of bias (“display”: “Lack of blinding”). In this example, the effect estimate is statistically significant using a fixed-effect model and not statistically significant using a random-effects model for the meta-analysis, a situation for which explicit representation of the statistic model is necessary for proper interpretation.

5.4. Benefits of code system development

When completed, the code systems will make finding knowledge easier. For example, systematic reviewers may specify study design concepts to facilitate identification of articles meeting their inclusion criteria. The code systems will facilitate re-use of scientific results. For example, clinical trial reporters who express their results for regulatory purposes could re-use the data to express their results for publication, and the systematic reviewers could directly re-use these results without the need for manual data extraction. All of these code systems will expedite recognition of the trustability of scientific knowledge whether seeking the data parameters (as expressed with statistic type codes), the methods for data creation (as expressed with study design and statistic model codes), or the assessments of others (as expressed with risk of bias codes). Someday, via explicit encoded study results, data within published papers can integrate with clinical decision support systems, particularly when reporting meta-analysis results.

We hope the processes, systems, and accomplishments we have produced so far in response to the COVID-19 crisis are sufficient to provide an infrastructure that will endure to make scientific communication accessible for a long time.

6. Conclusion

We started with efforts to support each other to accelerate knowledge transfer for COVID-19, and then developed solutions with expansive potential. We identified non-redundant concepts to support computable expression of scientific methods. Mapping these concepts to existing ontologies, selecting preferred terms and definitions by the global community, evaluating the implementation of the code systems,
and supporting continued development of the systems will support an extensive ecosystem for communicating scientific evidence. More efficient scientific communication will reduce cost and burden and improve health outcomes, quality of life, and patient, caregiver and healthcare professional satisfaction. Anyone who is communicating these concepts may join the effort at https://www.gps.health/covid19_knowledge_accelerator.html [70].

CRediT authorship contribution statement

Brian S. Alper: Conceptualization, Methodology, Investigation, Data curation, Writing - original draft, Writing - review & editing, Supervision, Project administration. Joanne Dehnost: Methodology, Investigation, Data curation, Writing - review & editing, Writing - original draft, Project administration. Muhammad Afzal: Methodology, Investigation, Data curation, Writing - review & editing, Writing - original draft, Project administration. Vignesh Subbian: Methodology, Data curation, Writing - review & editing. Andrey Soares: Methodology, Data curation, Writing - review & editing. Ilkka Kunnamo: Methodology, Investigation, Data curation, Writing - review & editing. Khalid Shahin: Methodology, Data curation. Robert C. McClure: Methodology, Writing - review & editing, Visualization.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: All authors are members of the COVID-19 Knowledge Accelerator (COKA) Initiative. The COKA Initiative is a volunteer virtual organization with no funding or contractual relations. The non-software content created by the COKA Initiative (including the data shared in this manuscript) is openly and freely available by Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0) license. BSA is the owner of Computable Publishing LLC which may commercialize software services related to this content. JD and KS are employed by Computable Publishing LLC. MA, VS, AS, IK, and RCM have no conflicts to report.

Acknowledgement

We would like to thank Karen A. Robinson, Harold Lehmann, Zbys Fedorowicz, and Lehana Thabane for substantial contributions to the Code System Development Protocol. We thank Cheow Peng Ooi for assistance with citations. We also acknowledge contributions by Expert Working Group members with the level of contribution provided in Appendix B. We also thank Asiyah Yu Lin, Yongguin “Oliver” He, Jie Zheng, and Philippe Rocca-Sera for providing feedback for the extended value of our work for ontology development.

Funding sources

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Appendix A. Code system development protocol

For the COVID-19 Knowledge Accelerator (COKA) Initiative Protocol as of September 17, 2020

This protocol will be applied to the development of Risk of Bias Code System, Statistic Type Code System, Statistic Model Code System, and Study Design Code System.

Code System Descriptions

The Statistic Type Code System will be used to precisely classify univariate statistics (such as mean, median, and proportion), comparative statistics (such as relative risk, mean difference, and odds ratio), and statistic attribute estimates (such as confidence interval, p value, and measures of heterogeneity). Consistent reporting across systems will facilitate interoperability for science communication.

The Statistic Model Code System will precisely communicate characteristics that define the model used for a statistic. Science reports often do not convey complete information about statistical models. Model characteristics may include concepts such as fixed-effects analysis, linear regression, and Mantel-Haenszel method for pooling. Consistent reporting of statistical models will facilitate interoperability for science communication.

The Study Design Code System will be used to precisely describe methodology characteristics of scientific observations including exposure introduction (such as interventional or observational), cohort definition (such as parallel, crossover or case-control), and group assignments (such as block randomization, every-other quasi-randomization, or non-randomized). Consistent reporting of research study design across systems will facilitate interoperability for science communication.

The Risk of Bias Code System will be used to precisely describe concerns with methods or reporting of scientific observations including selection bias (such as gaps in randomization or allocation concealment), performance bias (such as gaps in blinding), and analysis bias (such as gaps in intention to treat analysis or selective analysis reporting). Consistent reporting of risk of bias across systems will facilitate interoperability for science communication.

Protocol Steps:

1. Assemble an expert working group for each code system.
   a. Expert working group membership will be open to any individual who self-identifies as a relevant expert for the code system. Relevant expertise for a code system may include but is not limited to experience evaluating or expressing the concepts to be included in the code system, either for human interpretation or for machine interpretation.
   b. We will post open invitations as email messages to the distribution lists for the COKA Initiative, COVID-END, EBH listserv, GRADE Working Group, DECIDE project participants, AHRQ EPC listserv, HL7 CDS and BRR work groups, the Society for Clinical Trials, the Society for Participatory Medicine, International Society for Clinical Biostatistics, and PCORI.
   c. With the invitation we will share an introduction to what is a code system, why we are doing this, a link to the protocol, and a link to a data entry form to sign up. Sign up at Code System Development Intake Form.
   d. The data entry form will include optional demographic questions (age, gender, race/ethnicity) for the sole purpose of reporting demographic distribution of the expert working group in submitted publications of the code system.
   e. Set up a code system steering group from the most actively engaged participants, specifically those who join open weekly work group meetings.
2. For each code system, identify sources to signal the scope of (or common need for) a code system, namely tools or systems in common current use for reporting the concepts relevant to the code system. Expert working group members will be asked to identify such sources.
3. Create a list of non-redundant concepts that convey the concepts in commonly used tools and systems.
   a. Categorical classifiers (names of code sets) may be added. (A concept may be a member of a code set.)
   b. A concept may be marked as “also serves as a categorical classifier” in which case the concept may be a “parent” in one or more IS-A relationships with other concepts. (A name of a code set may be a member of another code set.)
c. A concept may be marked as being a “child” in an IS-A relationship with another concept by listing the “parent” concept as a categorical classifier.

d. This list will be reviewed in the open work group meetings.

4. Identify ontologies likely to include concepts on the lists created in step 3. Expert working group members will be asked to identify such ontologies. We will limit the effort to ontologies available for use without restrictions (or limited to Category 0 or 1 Restrictions per UMLS Restriction Levels described at https://uts.nlm.nih.gov/help/license/licensecategoryhelp.html).

5. For each concept, from each ontology, extract the display (or preferred term), synonym list (or alternative terms), and definition(s) that best match the concept, and note closely related variations.

6. For each concept:
   a. Review the displays, synonyms lists, and definitions available from ontologies.
   b. Draft a preferred display, synonym list, and definition, and note matches to the ontologies to measure relative contributions.
   c. Enter the draft preferred display, synonym list, and definition into an ontology web editor (such as WebProtege). If approved, the dataset can be shared with National Cancer Institute (NCI) Enterprise Vocabulary Services (EVS) for entry in the NCI Thesaurus and exported for use with WebProtege.

7. Each member of the expert working group will, for each concept that will be a code system entry, note agreement (with the draft preferred display, synonym list and definition) or suggest changes.
   a. For concepts that are “parents” in IS-A relationships, agreement will also be sought that the concept is useful functionally without subordinate coding.
   b. For concepts that are “children” in IS-A relationships, agreement will also be sought that if the child concept applies then the parent concept must apply AND the parent concept can apply while the child concept does not apply.
   c. This process will be online and asynchronous.

8. For any concepts without universal agreement we will discuss the suggested changes in open meetings, revise as appropriate, then resend for voting as noted in step #7.

9. If a concept does not achieve universal agreement (cycling through steps 7 and 8 with conflicting suggestions): a. Each person recommending changes will write a rationale.
   b. The rationales will be shared with the expert working group prior to a group meeting.
   c. The group meeting will discuss and prepare the preferred version. The preferred version and meeting discussion will be shared with the group.
   d. Group members will have 48 h to vote for the presented version.
   e. The preferred version will become the included version if it achieves at least 80% agreement with at least 5 people voting.
   f. If unable to achieve at least 80% agreement with at least 5 people voting, options may include extending the voting period, dropping the item, or preparing for another group discussion.

10. For the first complete version of the code system with agreement reached for all entries, we will determine the percent contribution from the different ontologies. If an ontology provides >50% contribution across the series of code systems or >75% contribution to a single code system, we may consider deeper collaboration rather than continued maintenance of a new code system.

11. We will publish the code system at terminology.hl7.org and seek publication of introductory articles to the code system in the biomedical literature.

12. For implementation and initial evaluation of the code system:
   a. Identify tools and systems that could use the code system.
   b. Offer support for implementation. Measure proportion of systems that get engaged.
   c. Evaluate ease of use.
   d. Generate code system change requests as needed.
   e. Track systems that implement the code system and set a regular review interval to inquire about usefulness and change requests.

13. For ongoing maintenance and development of the code system:
   a. Maintain an open invitation for code system users to join the expert working group for continued feedback.
   b. Maintain a method for expert working group members to suggest additional tools or systems with common current use of concepts matching the code system.
   c. Code system changes may be initiated by change requests from the community.
   d. The code system steering group will validate that change requests are appropriate for group deliberation (eg, fits the purpose of the code system, has sufficient rationale, avoids duplication).
   e. Valid change requests will lead to drafting a preferred display, synonym list, and definition.
   f. Each member of the expert working group will, for each valid change request, note agreement (with the draft preferred display, synonym list and definition) or suggest changes. This process will be online and asynchronous. (step #7)
   g. For any concepts without universal agreement we will discuss the suggested changes in open meetings, then resend for voting as noted in steps #7 and #8. If not reaching universal agreement, manage as step #9.
   h. Changes to the code system will be published at terminology.hl7.org and released as needed.

Cite as: Alper BS, Dehnbotstel J, Robinson K, Subbian V, Afzal M, Soares A, Kunnamo I, Shahin K, Lehmann H, Fedorowicz Z, McClure R, Thabane L. For the COVID-19 Knowledge Accelerator (COKA) Initiative. Code System Development Protocol. Created September 1, 2020. Last revised September 17, 2020. Available at: https://tinyurl.com/CodeSystemDevelopmentProtocol

Also available at OSF: https://osf.io/3akjv/?view_only=65c2ab5809c1484895d4fb03a2a9ee84

Appendix B. Participants in the COVID-19 Knowledge Accelerator (COKA) Initiative

See Tables B1 and B2.
Table B1
Expert Working Group Contributors to the Code System Development Concept Lists.

| Name                      | Country          | Code System Development Expert Working Group (number of participants) | Study Design (44) | Statistic Type (34) | Statistic Model (32) | Risk of Bias (31) |
|---------------------------|------------------|------------------------------------------------------------------------|------------------|---------------------|----------------------|-------------------|
| Gaelen P. Adam            | United States    | 1                                                                      | 3                | 3                   | 1                    |                  |
| Muhammad Afzal            | South Korea      | 1                                                                      |                  |                     |                      |                  |
| Tanvir Ahammed            | Bangladesh       | 1                                                                      | 1                | 1                   | 1                    |                  |
| Brian S. Alper            | United States    | 3                                                                      | 3                | 3                   | 3                    |                  |
| Eric H. Au                | Australia        | 1                                                                      | 1                | 1                   | 1                    |                  |
| Phillip O. Awodutire       | Nigeria          | 2                                                                      | 2                |                     |                      |                  |
| Sebastien Bailly          | France           | 1                                                                      | 1                | 1                   | 1                    |                  |
| Vusenth Balakrishna       | South Africa     | 1                                                                      | 1                | 1                   | 1                    |                  |
| Sorana D. Bolboaca        | Romania          | 3                                                                      | 3                | 3                   |                      |                  |
| Marek Brabec              | Czech Republic   | 1                                                                      | 1                |                     |                      |                  |
| Stacy B. Brody            | United States    | 3                                                                      |                  |                     |                      |                  |
| Comes Calin-Adrian        | Romania          | 1                                                                      |                  |                     |                      |                  |
| Rachel Coughan            | Canada           | 2                                                                      |                  |                     |                      |                  |
| Keitty Regina C. de Andrade| Brazil          | 1                                                                      | 1                | 1                   | 1                    | 1                 |
| Joanne Dehnbostel         | United States    | 3                                                                      | 3                | 3                   | 3                    | 3                 |
| Sandra Dimitri            | Egypt            | 1                                                                      |                  |                     |                      |                  |
| Marc L. Duteau            | Canada           | 2                                                                      |                  |                     |                      | 2                 |
| Zbyn Federowicz           | United Kingdom   | 1                                                                      |                  |                     |                      | 1                 |
| Emilia J. Flores          | United States    | 1                                                                      | 1                |                     |                      |                  |
| Isaac Fwemba              | Ghana            | 1                                                                      | 1                | 1                   | 1                    | 1                 |
| Abhay M. Gaidhane         | India            | 1                                                                      |                  |                     |                      |                  |
| Eric M. Harvey            | United States    | 3                                                                      |                  |                     |                      | 2                 |
| Danielle Johnson          | United Kingdom   | 1                                                                      |                  |                     |                      |                  |
| Samer A. Kharroubi        | Lebanon          | 1                                                                      |                  |                     |                      |                  |
| Bhagvan Kommandi          | India            | 3                                                                      | 3                | 3                   | 3                    | 3                 |
| Polychronis Kostoulas     | Greece           | 1                                                                      |                  |                     |                      |                  |
| Evangelos Kritsotakis     | Greece           | 3                                                                      |                  |                     |                      |                  |
| Ilkka Kommaso             | Finland          | 1                                                                      |                  |                     |                      |                  |
| Louis E. Leff             | United States    | 2                                                                      | 2                |                     |                      |                  |
| Harold Lehmann            | United States    | 1                                                                      | 1                | 1                   | 1                    | 1                 |
| Jesus Lopez-Alcalde       | Switzerland      | 1                                                                      |                  |                     |                      | 1                 |
| Robert C. McClure         | United States    | 2                                                                      | 2                | 2                   | 2                    | 2                 |
| Matthew D. Mitchell       | United States    | 1                                                                      |                  |                     |                      | 1                 |
| Tamara Navarro-Ruan       | Canada           | 3                                                                      |                  |                     |                      |                  |
| Pentti Nieminen           | Finland          | 1                                                                      |                  |                     |                      | 1                 |
| Akaninyene Patrick Obot   | Nigeria          | 1                                                                      | 1                | 1                   | 1                    |                  |
| Aloysius Odii             | Nigeria          | 1                                                                      | 1                | 1                   | 1                    |                  |
| Cheow Peng Ooi            | Malaysia         | 1                                                                      |                  |                     |                      |                  |
| Alejandro Piscoya         | Peru             | 1                                                                      | 1                | 1                   | 1                    |                  |
| Vivek Podder              | Bangladesh       | 1                                                                      |                  |                     |                      | 1                 |
| K.M. Safi-UR- Rahman      | Japan            | 1                                                                      |                  |                     |                      | 1                 |
| Karen A. Robinson         | United States    | 1                                                                      | 1                | 1                   | 1                    |                  |
| Paola Rosati              | Italy            | 1                                                                      |                  |                     |                      | 1                 |
| Carolyn M. Rutter         | United States    | 1                                                                      | 1                | 1                   | 1                    |                  |
| Khalid S. Shahin          | United States    | 3                                                                      | 3                | 3                   | 3                    | 3                 |
| Roshini Sooriarachchi     | Sri Lanka        | 1                                                                      |                  |                     |                      | 1                 |
| Vignesh Subbian           | United States    | 3                                                                      | 2                | 1                   | 1                    |                  |
| Lehana Thabane            | Canada           | 3                                                                      | 3                | 3                   | 3                    |                  |
| Mario Tristan             | Costa Rica       | 1                                                                      | 1                | 1                   | 1                    |                  |
| Chidi Ugwu                | Nigeria          | 1                                                                      |                  |                     |                      | 1                 |
| Linhu Zhao                | Canada           | 1                                                                      |                  |                     |                      | 1                 |
| Name Withheld             | Brazil           | 1                                                                      |                  |                     |                      | 1                 |
| Name Withheld             | Ireland          | 1                                                                      |                  |                     |                      | 1                 |
| Name Withheld             | Italy            | 1                                                                      | 1                | 1                   | 1                    | 1                 |
| Name Withheld             | Switzerland      | 1                                                                      | 1                |                     |                      | 1                 |

Note: 1 — signed up for continued participation
2 — AND approved Step 3 results
3 — AND actively contributed to Step 3 results
Appendix C. Draft term list for statistic type code system version 1.0.0

This list has 88 non-redundant codable concepts for the Statistic Type Code System (items bolded if they are both a classifier and a codable concept):

1. Univariate (CATEGORY ONLY)
   a. Count
   b. Sum
   c. Maximum Observed Value
   d. Maximum Possible Value
   e. Minimum Observed Value
   f. Minimum Possible Value
   g. Cutoff value

h. Central Tendency (CATEGORY ONLY)
   i. Mean
   ii. Median
   iii. Mode

2. Difference (CATEGORY ONLY)
   a. Count Difference
   b. Mean Difference
   c. Standardized Mean Difference
   d. Median Difference
   e. Risk Difference

3. Ratio (CATEGORY ONLY)
   a. Observed (CATEGORY ONLY)
      i. Observed Proportion
      ii. Incidence
      iii. Cumulative Incidence
iv. Incidence Rate (Incidence Density)
v. Period Prevalence
vi. Point Prevalence
f. Effect (CATEGORY ONLY)
i. Hazard Ratio
ii. Incidence Rate Ratio
iii. Odds Ratio
iv. Prevalence Ratio
v. Risk Ratio
vi. Number Needed to Treat (NNT)
vii. Number Needed to Screen (NNS)
viii. Number Need to Diagnose (NND)
ix. Relative Risk Difference
g. Agreement (CATEGORY ONLY)
i. Diagnostic Accuracy
ii. Diagnostic Odds Ratio
iii. Kappa
1. Bennett’s Kappa
2. Cohen’s Kappa
3. Scott’s Kappa
iv. Misclassification Rate
v. F1-score
h. Conditional Probability (CATEGORY ONLY)
i. Predicted Risk
ii. Sensitivity
iii. Specificity
iv. Positive Predictive Value
v. Negative Predictive Value
vi. Likelihood Ratio Positive
vii. Likelihood Ratio Negative
viii. Positive Clinical Utility Index
ix. Negative Clinical Utility Index
4. Correlation (CATEGORY ONLY)
a. Covariance
b. Pearson Correlation Coefficient
c. Regression Coefficient
d. Spearman Rank-Order Correlation Coefficient
e. Matthews Correlation Coefficient
f. Kendall Correlation Coefficient
g. Calibration (CATEGORY ONLY)
i. Mean calibration
ii. Calibration-in-the-large
iii. Calibration intercept
iv. Calibration slope
5. Dispersion (CATEGORY ONLY)
a. Range
b. Interquartile range
c. Standard deviation
i. Standard deviation for population
ii. Standard deviation for sample
iii. Sampling standard deviation
d. Variance
i. Variance for population
ii. Variance for sample
iii. Sampling variance
e. Gini Index
6. Statistical Distribution Measure (CATEGORY ONLY)
a. Dispersion (CATEGORY ONLY)
i. Standard error
i. Standard error of the mean
ii. Standard error of the median
iii. Standard error of the proportion
iv. Standard error of the difference between means
v. Standard error of the difference between proportions
x. Credible interval
xi. Confidence interval
i. Discrimination (CATEGORY ONLY)
i. Area Under the ROC Curve (AUC)
1. C-statistic
c. Heterogeneity (CATEGORY ONLY)
i. Chi square for homogeneity
ii. Cochran’s Q statistic
iii. I-squared
iv. Tau squared
d. Hypothesis Testing (CATEGORY ONLY)
i. Chi square for independence
ii. Chi square for trend
iii. P-value
iv. Z-score
v. T-score
v. Descriptive

The list above was the current list as of October 15, 2020. The list continues to evolve and the current list can be found at https://confluence.hl7.org/display/CDS/COKA+Code+System+Development+Working+Groups.

Appendix D. Draft term list for statistic model code system version 1.0.0

This list has 53 non-redundant codable concepts for the Statistic Model Code System (items bolded if they are both a classifier and a codable concept):

1. Determination of Relationship codes (CATEGORY ONLY)
a. Threshold framing (CATEGORY ONLY)
i. one-tailed test (one threshold)
ii. two-tailed test (two thresholds)
b. Parametric tests (CATEGORY ONLY)
i. Z-test
ii. 1-sample t-test
iii. 2-sample t-test
iv. paired t-test
v. chi-squared test
vi. chi-squared test for trend
vii. Pearson correlation
viii. ANOVA (ANalysis Of VAriance)
1. One-way ANOVA
2. 2-way ANOVA without replication
3. 2-way ANOVA with replication
4. 3-way ANOVA
5. multivariate ANOVA (MANOVA)
c. Nonparametric tests (CATEGORY ONLY)
i. sign test
ii. Wilcoxon signed-rank test
iii. Wilcoxon rank-sum test
iv. Mann-Whitney U test
v. Fisher’s exact test
vi. McNemar’s test
vii. Kruskal Wallis test
viii. Spearman correlation
ix. Kendall correlation
x. Friedman test
xi. Goodman Kruska’s Gamma
d. Regression model (CATEGORY ONLY)
i. Linear Regression
ii. Logistic Regression
iii. Poisson Regression
iv. Negative Binomial Regression
v. GLM (Generalized Linear Model)
2. Adjustment of Variables codes (CATEGORY ONLY)
a. Adjusted analysis
b. Zero-cell adjustment with constant
c. Zero-cell adjustment with continuity correction

3. Pooling codes (CATEGORY ONLY)
   a. **Pooling with Meta-analysis**
      i. Mantel-Haenszel method
      ii. Inverse variance method
      iii. Peto method
      iv. Generalized linear mixed model (GLMM)
   
   b. Effects codes (CATEGORY ONLY)
      i. Fixed-effects
      ii. Random-effects
   
   c. Heterogeneity codes (CATEGORY ONLY)
      i. Chi-squared test for homogeneity
      ii. **Tau estimation**
         1. Dersimonian-Laird method
         2. Paule-Mandel method
         3. Restricted Maximum Likelihood method
         4. Maximum Likelihood method
         5. Empirical Bayes method
         6. Hunter-Schmidt method
         7. Sidik-Jonkman method
         8. Hedges method

The list above was the current list as of October 15, 2020. The list continues to evolve and the current list can be found at https://confluence.hl7.org/display/CDS/COKA+Code+System+Development+Working+Groups.

### Appendix E. Draft term list for study design code system version 1.0.0

This list has 57 non-redundant codable concepts for the Study Design Code System (items bolded if they are both a classifier and a codable concept):

1. Method of exposure introduction (CATEGORY ONLY)
   a. Interventional method of exposure introduction
   b. Observational method of exposure introduction
   c. Indirect method of exposure introduction

2. Assignment (CATEGORY ONLY)
   a. **Randomized Assignment**
      i. Simple randomization for assignment
      ii. Stratified randomization for assignment
      iii. Block randomization for assignment
   
   b. **Quasi-Randomized Assignment**
      i. Minimization method of quasi-randomization for assignment
      ii. Every-other method of quasi-randomization for assignment
   
   c. Non-Randomized Assignment
   d. Cluster Assignment
   e. Matched Assignment
   f. Adaptive Assignment

3. Comparator Design Definition (CATEGORY ONLY)
   a. Parallel cohort definition
   b. Crossover cohort definition
   c. **Time series** (multiple time point comparison)
      i. Before-after comparison
      d. Case-Control design approach
   
   e. **Uncontrolled cohort**
      i. Case Report
      ii. Case Set
   
   f. Twin Study
   g. Ecological/Population-based
   h. Tumor vs. Matched-Normal

4. Context (CATEGORY ONLY)
   a. Clinical Trial
   b. Pragmatic Clinical Trial
   c. Clinical Testing
   d. Clinical Care Records
   e. Healthcare Financing Records
   f. Patient Registry
   g. Multicenter Study
   h. Clinical Conference
      i. Collection
      j. Control Set
   
   k. Mendelian
   l. Metagenomics
   m. Xenograft

5. Data Collection Timing (CATEGORY ONLY)
   a. Cross-sectional data collection
   b. Longitudinal data collection

6. Analysis Approach (CATEGORY ONLY)
   a. Quantitative analysis approach
   b. Qualitative analysis approach
   c. Critique analysis approach
   d. Nonsystematic analysis approach

7. Clinical Research Regulatory Subsets (CATEGORY ONLY)
   a. Expanded Access Studies
   b. Early Phase 1 trial
   c. Phase 1 trial
   d. Phase 1/Phase 2 trial
   e. Phase 2 trial
   f. Phase 2/Phase 3 trial
   g. Phase 3 trial
   h. Post-marketing Study
   i. Post-marketing Surveillance Study

8. Study Goal (CATEGORY ONLY)
   a. Equivalence Trial
   b. Evaluation Study
   c. Validation Study
   d. Scientific Integrity Review

The list above was the current list as of October 15, 2020. The list continues to evolve and the current list can be found at https://confluence.hl7.org/display/CDS/COKA+Code+System+Development+Working+Groups.

### Appendix F. Draft term list for risk of bias code system version 1.0.0

This list has 170 non-redundant codable concepts for the Risk of Bias Code System (items bolded if they are both a classifier and a codable concept):

#### Type Classifiers

1. **Participant Selection Bias** (for overall sample, not for comparator group)
   a. Inappropriate selection criteria
   b. Biased sampling strategy
   c. Non-representative sample
   d. Inadequate participation by eligible persons
   e. Post-intervention factors bias selection
   f. Intervention associated with post-intervention factors that bias selection
   g. Outcome associated with post-intervention factors that bias selection
   h. Inadequate Random sequence generation

2. **Comparator Selection Bias**
2b. Inadequate Allocation concealment
2c. Biased selection of the non-exposed cohort
2d. Case-control design
   2d1. Case-control design without appropriate definition of controls
   2d2. Case-control design without appropriate selection of controls
   2d3. Case-control design without description of selection of controls
   2d4. Case-control design with factor-specific concern for comparability of cases and controls
2e. Potential for Confounding
   2e1. Baseline differences
   2e2. Confounding by follow-up time
   2e3. Prognostic factors influencing intervention
   2e4. Post-intervention confounding
3. Performance Bias
   3a. Performance Bias - Blinding of Participants
   3b. Performance Bias - Blinding of Intervention Deliverers
   3c. Performance Bias - Deviations from intended Intervention
   3d. Performance Bias - Imbalance in Deviations from Intended Interventions
   3e. Performance Adherence Bias
      3e1. Performance Adherence Bias - Blinding of Participants
      3e2. Performance Adherence Bias - Blinding of Intervention Deliverers
      3e3. Performance Adherence Bias - Imbalance in Deviations from intended Intervention
      3e4. Performance Adherence Bias - Nonadherence of Implementation
      3e5. Performance Adherence Bias - Nonadherence of Participants
4. Attrition Bias
   4a. Incomplete Outcome Data
   4b. Influence of Incomplete Outcome Data
   4c. Influence of Outcome on Missingness of Data
   4d. Exclusions due to missing data on intervention
   4e. Exclusions due to missing data on measured variables
   4f. Imbalance in missing data
   4g. Sensitivity to missing data
   4h. Inadequate response rate
   4i. Inadequate understanding of missing data
5. Detection Bias
   5a. Detection Bias for Outcomes
   5b. Detection Bias for Exposures
   5c. Detection Bias for Reference Standard
   5d. Detection Bias for Index Test
   5e. Detection Bias for Classifiers
   5f. Detection Bias for Confounders
   5g. Insufficient study characteristics available for proper results interpretation
   5h. Incomplete collection of relevant study results for synthesis
   5i. Methodologic quality assessment inadequate
   5j. Error in risk of bias assessment not minimized
   5x1. Detection Bias for X - Classification Bias (for x = X, substitute a = Outcomes, b = Exposures, c = Reference Standard, d = Index Test, e = Classifiers, or f = Confounders)
      5x1a. Nonrepresentative definition
      5x1b. Risk of misclassification
      5x1c. Definition not prespecified
      5x1d. Threshold not prespecified
      5x1e. Classification potentially influenced by risk of outcome
      5x1f. Definition unclear
   5x2. Detection Bias for X - Assessment Method (for x = X, substitute a = Outcomes, b = Exposures, c = Reference Standard, d = Index Test, e = Classifiers, or f = Confounders)
      5x2a. Inappropriate Measurement Method
      5x2b. Improper conduct of measurement assessment
      5x2c. Incomplete application of measurement assessment
      5x2d. Inadequate follow up period for outcome of interest
      5x2e. Assessment method unclear
      5x2f. Error in data collection not minimized
   5x3. Detection Bias for X - Imbalance (for x = X, substitute a = Outcomes, b = Exposures, c = Reference Standard, d = Index Test, e = Classifiers, or f = Confounders)
      5x3a. Imbalance in Application of Measurement Method
      5x3b. Differential data availability during tests
      5x3c. Inappropriate delay between index test and reference standard
   5x4. Detection Bias for X - Confounding Influence (for x = X, substitute a = Outcomes, b = Exposures, c = Reference Standard, d = Index Test, e = Classifiers, or f = Confounders)
      5x4a. Incorporation bias (eg non-independence of reference standard and index test)
      5x4b. Lack of blinding (eg blinding of index test result during reference test, blinding of outcome assessors)
      5x4c. Influence of Blinding on Measurement
6. Analysis Bias
   6a. Bias controlling for confounding factors
      6a1. Bias controlling for confounding factors and time-varying confounding
      6a2. Adjustment for selection bias
      6a3. Inadequate Intention-To-Treat Analysis
      6a4. Inadequate Adherence Effect Analysis
      6a5. Predictors included in outcome definition
   6b. Analysis Model Selection Bias - improper statistical model
   6c. Inadequate numbers for analysis
   6d. Bias in Handling of Data
      6d1. Incomplete data analysis
      6d2. No accounting for uninterpretable results
      6d3. Inappropriate handling of missing data
      6d4. Inappropriate handling of variables
      6d5. Inappropriate handling of complexities in the data
      6d6. Differential handling of confounder measurement
      6d7. Handling of confounders unclear
      6d8. Inappropriate handling of missing confounder data
   6e. Analysis Selection Bias
      6e1. Selective analysis reporting (from repeated analyses at multiple times)
      6e2. Selective analysis reporting (from multiple analytic models)
      6e3. Early trial termination
      6e4. Preliminary analysis
      6e5. Subgroup analysis
   6f. Analysis bias in predictive model development
      6f1. Selection of predictors based on univariable analysis
      6f2. Inappropriate evaluation of model performance measures
      6f3. Model overfitting and optimism
      6f4. Final model not corresponding to multivariable analysis
7. Reporting Bias
   7a. Reported Result Not Following Pre-Specified Analysis Plan
   7b. Inadequate reporting to assess analytic strategy
   7c. Selective outcome measure reporting (within outcome domains)
   7d. Selective outcome measure reporting (across outcome domains)
   7e. Pre-final publication form
   7f. Subgroup analysis (reporting bias)
   7g. No explanation of withdrawals
   7h. Interpretation of findings not addressing risk of bias
   7i. Relevance of studies to research question not appropriately considered
7j. Results emphasized based on statistical significance

8. Study Selection Bias
8a. Bias in study eligibility criteria
8a1. Study eligibility criteria not prespecified
8a2. Study eligibility criteria not appropriate for review question
8a3. Study eligibility criteria ambiguous
8a4. Study eligibility criteria limits for study characteristics not appropriate
8a5. Study eligibility criteria limits for information sources not appropriate
8b. Database search sources not appropriate
8c. Nondatabase search sources inadequate
8d. Search strategy not sensitive
8e. Search strategy limits for information sources not appropriate
8f. Study eligibility criteria not adhered to
8g. Error in study selection not minimized

9. Synthesis Bias
9a. Synthesis missing eligible studies
9b. Study parameters not appropriate for synthesis
9c. Heterogeneity not addressed
9d. Sensitivity to factors
9e. Biases in studies influence synthesis

L. Qualitative Research (CATEGORY ONLY)
L1. Inappropriate qualitative approach
L2. Inadequate qualitative data collection methods
L3. Inappropriate qualitative analysis
L4. Unsubstantiated interpretation of results
L5. Incoherence between data, analysis and interpretation

M. Mixed Methods Research (CATEGORY ONLY)
M1. Inadequate rationale for mixed methods design
M2. Ineffective integration of study components
M3. Inappropriate interpretation of integration of qualitative and quantitative findings
M4. Inadequate handling of inconsistency

N. Predictive Model Subset (CATEGORY ONLY)
N1. Bias in Predictive Model Development (used to subset classifiers noted elsewhere to be specific to predictive model development)
N2. Bias in Predictive Model Validation (used to subset classifiers noted elsewhere to be specific to predictive model validation)
N3. Absence of any validation
N4. Absence of any external validation

Rating Classifiers
R. Rating of certainty (CATEGORY ONLY)
R1. Low Risk of False Certainty
R2. Moderate Risk of False Certainty
R3. High Risk of False Certainty
R4. Serious Risk of False Certainty
R5. Critical Risk of False Certainty
R6. Some Risk of False Certainty

S. Rating of factor presence (CATEGORY ONLY)
S1. Factor Present
S2. Factor Likely Present
S3. Factor Likely Absent
S4. Factor Absent
S5. No Information
S6. Factor Presence or Absence Unclear

T. Rating of bias direction (CATEGORY ONLY)
T1. Risk of Bias Favoring Experimental
T2. Risk of Bias Favoring Comparator
T3. Risk of Bias Towards Null
T4. Risk of Bias Away from Null

U. Rating of influence (CATEGORY ONLY)
U1. Factor has potential to impact results
U2. Factor likely has potential to impact results
U3. Factor likely does not have potential to impact results
U4. Factor does not have potential to impact results

The list above was the current list as of October 15, 2020. The list continues to evolve and the current list can be found at https://confluence.hl7.org/display/CDS/OKA+Code+System+Development+WorkingGroups.
