Received: 7 February 2022 | Revised: 17 June 2022 | Accepted: 28 June 2022

DOI: 10.1002/pds.5507

ORIGINAL ARTICLE

WILEY

HARmonized Protocol Template to Enhance Reproducibility of hypothesis evaluating real-world evidence studies on treatment effects: A good practices report of a joint ISPE/ISPOR task force

Shirley V. Wang[1] | Anton Pottegård[2] | William Crown[3] | Peter Arlett[4] | Darren M. Ashcroft[5] | Eric I. Benchimol[6,7,8] | Marc L. Berger[9] | Gracy Crane[10] | Wim Goettsch[11,12] | Wei Hua[13] | Shaum Kabadi[14] | David M. Kern[15] | Xavier Kurz[4] | Sinead Langan[16] | Takahiro Nonaka[17] | Lucinda Orsini[18] | Susana Perez-Gutthann[19] | Simone Pinheiro[13] | Nicole Pratt[20] | Sebastian Schneeweiss[1] | Massoud Toussi[21] | Rebecca J. Williams[22]

[1]Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, USA

[2]University of Southern Denmark, Odense, Denmark

[3]Brandeis University, Waltham, Massachusetts, USA

[4]European Medicines Agency, London, UK

[5]School of Health Sciences, University of Manchester, Manchester, UK

[6]Child Health Evaluative Sciences, SickKids Research Institute, Toronto, Ontario, Canada

[7]ICES, Toronto, Ontario, Canada

[8]Department of Paediatrics and Institute of Health Policy, Management and Evaluation, The Hospital for Sick Children, University of Toronto, Toronto, Ontario, Canada

[9]Independent Consultant, New York, USA

[10]Roche Pharmaceuticals, Camberley, UK

[11]The National Health Care Institute, Diemen, The Netherlands

[12]Utrecht University, Utrecht, The Netherlands

[13]US Food and Drug Administration, Silver Springs, Maryland, USA

[14]Sanofi-Aventis US LLC, North Potomac, Maryland, USA

[15]Janssen Research & Development, LLC, Philadelphia, Pennsylvania, USA

[16]London School of Tropical Hygiene, London, UK

[17]Pharmaceutical and Medical Devices Agency, Japan

[18]COMPASS Pathways Ltd, Skillman, UK

[19]RTI Health Solutions, Barcelona, Spain

[20]Quality Use of Medicines and Pharmacy Research Centre, Clinical and Health Sciences, University of South Australia, South Australia, Australia

[21]IQVIA, France

[22]Bethesda, Maryland, USA

This paper has been shared with ISPE and ISPOR membership for review by society members as part of the endorsement process from both research professional societies. It has been presented in seminars for CNODES, Duke-Margolis, and ISPOR. ISPE presentations will take place in August 2022.

Correspondence
Shirley V. Wang, Brigham and Women's Hospital, Harvard Medical School, 1620 Tremont St, Suite 303, Boston, MA 02021, USA.
Email: swang1@bwh.harvard.edu

Funding information
Wellcome Trust Senior Research Fellowship in Clinical Science, Grant/Award Number: 205039/Z/16/Z; NHMRC, Grant/Award Numbers: GNT1196900, GNT1157506; NIA, Grant/Award Number: R01AG053302; NHLBI, Grant/Award Number: RO1HL141505; National Institutes of Health

Abstract

Problem: Ambiguity in communication of key study parameters limits the utility of real-world evidence (RWE) studies in healthcare decision-making. Clear communication about data provenance, design, analysis, and implementation is needed. This would facilitate reproducibility, replication in independent data, and assessment of potential sources of bias.

What We Did: The International Society for Pharmacoepidemiology (ISPE) and ISPOR–The Professional Society for Health Economics and Outcomes Research (ISPOR) convened a joint task force, including representation from key international stakeholders, to create a harmonized protocol template for RWE studies that evaluate a treatment effect and are intended to inform decision-making. The template builds on existing efforts to improve transparency and incorporates recent insights regarding the level of detail needed to enable RWE study reproducibility. The overarching principle was to reach for sufficient clarity regarding data, design, analysis, and implementation to achieve 3 main goals. One, to help investigators thoroughly consider, then document their choices and rationale for key study parameters that define the causal question (e.g., target estimand), two, to facilitate decision-making by enabling reviewers to readily assess potential for biases related to these choices, and three, to facilitate reproducibility.

Strategies to Disseminate and Facilitate Use: Recognizing that the impact of this harmonized template relies on uptake, we have outlined a plan to introduce and pilot the template with key international stakeholders over the next 2 years.

Conclusion: The HARmonized Protocol Template to Enhance Reproducibility (HARPER) helps to create a shared understanding of intended scientific decisions through a common text, tabular and visual structure. The template provides a set of core recommendations for clear and reproducible RWE study protocols and is intended to be used as a backbone throughout the research process from developing a valid study protocol, to registration, through implementation and reporting on those implementation decisions.

KEYWORDS
protocol, real world evidence, reproducibility, transparency

Key Points

- Ambiguity in communication of key study parameters limits the utility of real-world evidence studies in healthcare decision-making.
- The International Society for Pharmacoepidemiology and ISPOR–The Professional Society for Health Economics and Outcomes Research convened a joint task force, including representation from key international stakeholders, to create a harmonized protocol template for RWE studies that evaluate a treatment effect and are intended to inform decision-making.
- The HARmonized Protocol Template to Enhance Reproducibility helps to create a shared understanding of intended scientific decisions through a common text, tabular and visual structure.

Plain language summary

Ambiguity in communication of key study parameters limits the utility of real-world evidence (RWE) studies in healthcare decision-making. The International Society for

Pharmacoepidemiology and ISPOR–The Professional Society for Health Economics and Outcomes Research convened a joint task force to create a harmonized protocol template for RWE studies. The template builds on existing efforts to improve transparency and incorporates recent insights regarding the level of detail needed to enable study reproducibility. The overarching principle was to reach for sufficient clarity to achieve three main goals. One, to help investigators thoroughly consider, then document their choices and rationale for key study parameters that define the causal question, two, to facilitate decision-making by enabling reviewers to readily assess potential for biases related to these choices, and three, to facilitate reproducibility. The HARmonized Protocol Template to Enhance Reproducibility helps to create a shared understanding of intended scientific decisions through a common text, tabular and visual structure. The template provides a set of core recommendations for clear and reproducible RWE study protocols and is intended to be used as a backbone throughout the research process from developing a valid study protocol, to registration, through implementation and reporting on those implementation decisions.

1 | BACKGROUND

Regulatory agencies, health technology assessors, and payers are increasingly interested in studies that make use of real-world data (RWD) to inform regulatory and other policy or clinical decision-making.[1–5] While real-world evidence (RWE) studies using rigorous methods applied to fit-for-purpose RWD can provide critical, timely insights into the safety and effectiveness[6–8] of drugs, devices, and vaccines; high-profile cases of studies conducted with biased methods[9–12] or inadequate reporting on unsuitable data[13–15] have raised concerns over the credibility of RWE studies. These concerns have led to increasing calls from the research community and decision-makers for more transparency on the design and conduct of studies using RWD.[16–18]

Some initiatives are already in place. As an example, the European Medicines Agency (EMA) has, for over a decade, required or recommended registration of a study protocol using a template for observational post-authorization safety studies (PASS) conducted by marketing authorization holders.[19,20] However, a large scale evaluation of the reproducibility of 150 studies highlighted that there remains a great deal of variability in transparency about critical details of RWE study implementation,[21] and recently, the EMA endorsed a strategy for moving toward greater standardization and structure in protocols.[22]

Clear communication within multi-disciplinary study teams and between investigators, decision-makers and other stakeholders is necessary to increase confidence in RWE study design, conduct, and results. The rapid development of fragmented recommendations[23] has highlighted the need for an internationally agreed upon set of core expectations regarding best practices for developing and communicating about study design, analysis, and implementation via transparent, comprehensive, and rigorous RWE study protocols. A joint task force between the International Society for Pharmacoepidemiology (ISPE) and the ISPOR–The Professional Society for Health Economics and Outcomes Research (ISPOR) was convened to meet this need by developing a harmonized protocol template for RWE studies that make secondary use of RWD, evaluate a hypothesis and are intended to inform healthcare decision-making. The task force was comprised of core committee members from both professional societies, and included international stakeholder groups including regulatory agencies, health technology assessment (HTA) organizations, industry, and academia.

The task force was primarily focused on protocols for post-marketing studies that deal with questions of causal inference using RWD because of their importance to decision making and the complexity of design and analysis when addressing causal questions. Examples of such studies include comparative effectiveness or safety studies associated with clinical interventions, studies of the effect of policy interventions such as benefit designs or healthcare delivery models, health care expenditures or value associated with different treatments, and so forth. While it is also important to develop protocols for non-causal inference studies using RWD, that was not the focus of the protocol harmonization effort.

The task force met monthly from July 2021 to January 2022 to develop the harmonized template. The process of developing the harmonized template included both evaluation of external validity (through comparison of existing protocol templates or guidance developed by international multi-stakeholder groups to ensure compatibility with agreed upon scientific principles) and internal validity (through testing and development of example use cases with different designs and data sources by five sub-teams). The final deliverable was a standard template with embedded instruction which harmonized across existing guidance and templates and example protocols for a variety of use cases to illustrate how to use the template.

1.1 | Identification and comparison of protocol templates

Existing protocol templates for RWE studies were identified based on templates known to the core committee of the joint task force,

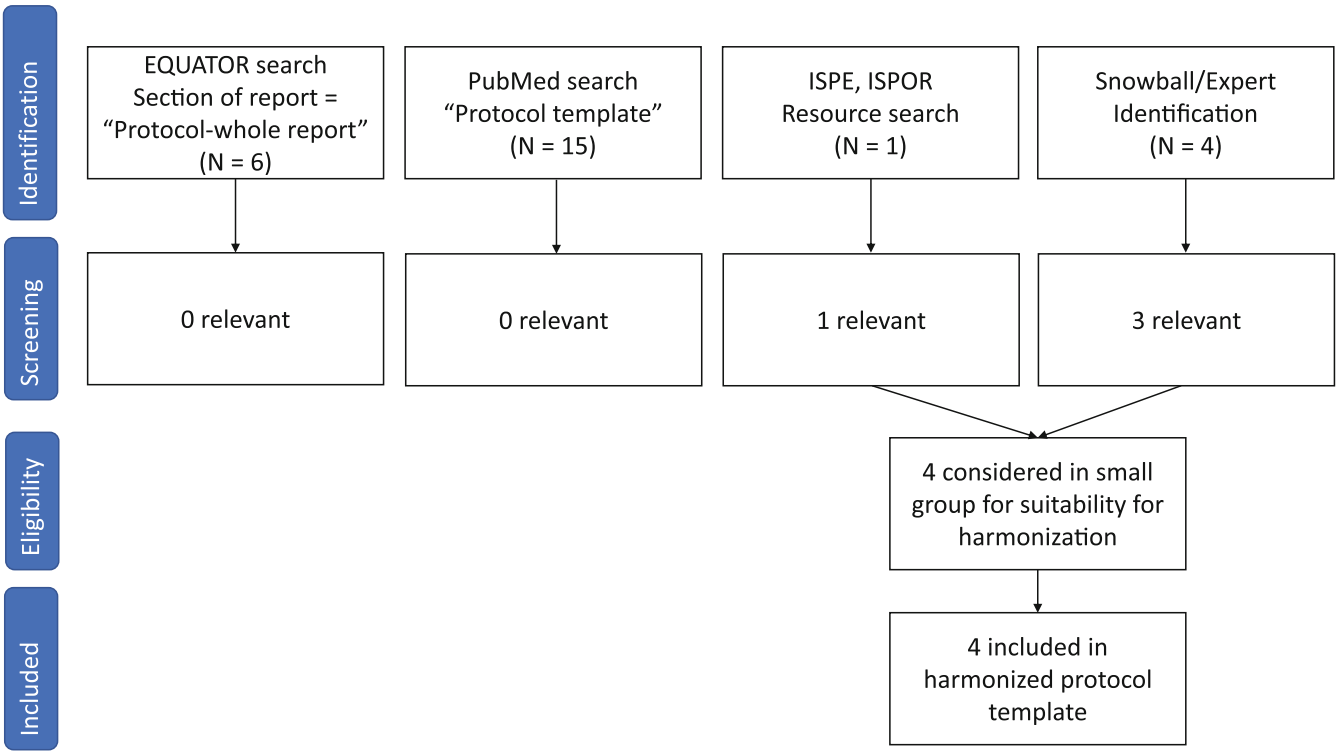

FIGURE 1 PRISMA diagram

coupled with a search for relevant protocol templates in PubMed and the EQUATOR network (Enhancing the QUAlity and Transparency Of health Research) (Figure 1, Appendix S1). Additionally, an extended reviewer group composed of volunteers from ISPE and ISPOR were asked to review the list of identified protocol templates and to supplement the list with other templates that they were aware of. Protocol templates that were not relevant for RWE studies that make secondary use of healthcare data or were not developed by international multi-stakeholder groups were excluded. This resulted in four eligible protocol templates; the European Medicines Agency (EMA) Heads of Medicines Agency Guideline on Good Pharmacovigilance Practices (GVP) Module VIII - post-authorizations safety studies (PASS) template,[20] ISPE's guidelines for good pharmacoepidemiology practice (GPP) section on protocol development,[24] National Evaluation System for health Technology (NEST) protocol guidance,[25] and the Structured Template and Reporting Tool for Real World Evidence (STaRT-RWE).[26]

Section headings of the identified protocol templates were compared and mapped to each other, using the oldest guideline (EMA-GVP Module VIII-PASS) as the starting point (Table 1). The committee observed that at a conceptual level, the major elements of study design and analysis were already largely agreed upon and included in each of the templates. However, the templates differed on the depth and detail of guidance within each section as well as the sequencing of elements within the template. Three of the protocol templates offered a few sentences or paragraphs of guidance on what sort of information to include within each section (EMA-GVP Module VIII-PASS, ISPE-GPP, and NEST), allowing the user flexibility in free-text entries under the section header. The most recently published template (STaRT-RWE) used structured tables to guide the user on where, what and how to specify study implementation details. The STaRT-RWE template tables and figure also had a strong focus on clearly delineating time zero for entry into the study population and orienting baseline and follow up windows around that primary temporal anchor. A high-level summary of other differences in format and depth of detail requested by each template is provided in supplemental appendices (Appendix S2).

1.2 | Creation of HARmonized Protocol Template to Enhance Reproducibility

In order to create a harmonized template, the core committee of the joint task force discussed each section header in the mapped table of protocol templates. Again, starting with the EMA-GVP Module VIII-PASS template, the committee evaluated the different sections, guidance and/or structure of more recently developed protocol templates under the same section header, jointly deciding how to incorporate these updates into the harmonized protocol template. The committee then categorized the sections as core elements required for any RWE study protocol and non-core elements that may provide important context, administrative and other information (Table 1). Core elements of the protocol were defined as sections that were either considered key for the purposes of reproducibility and validity assessment or were common elements that were found in multiple protocol

TABLE 1 Comparison of four protocol templates for real-world evidence studies developed by multi-stakeholder, international organizations.

High level summary	Largely free text, with guidance on what to include under section headers - details in the *Guidance for the format and content of the protocol of non-interventional post-authorisation safety studies* (www.ema.europa.eu)	Largely free text, with guidance on what to include under section headers - details in report https://www.pharmacoepi.org/resources/policies/guidelines-08027/	Largely free text, with guidance on what to include under section headers - details in NESTcc report	Structured tables that lay out operational parameters to be specified - details in Wang et al, BMJ	Combination of free text with structured tables under section headers
Section Header					
Title page with administrative information (e.g. title, registry ID, drug/device ID, sponsor)	1	A, C, M	2, 10	Table 1	1, table
Table of contents	2			Table of Contents	Table
Abbreviations	3			Table 9	
Glossary of terminology				Table 8	
Responsible parties	4	B		Table 1	1, table
Abstract	5	D			2, free text
Amendments and updates	6	L		Table 2	3, table
Milestones/timeline	7	E			4, table
Rationale and background	8	G	1		5, structured free text, table
Research question and objectives	9	F	3	Table 1	6, table
Study design	9.1	H1	7	Figure 1, Table 3	7.1, 7.2, free text, figure
Setting	9.2	H2	4	Table 3A, 3B, 3C, 3D, Table 6	7.3, free text, table
Variables	9.3	H4	5, 6	Table 3B, 3E, 3F, 3G, 3H	7.4, free text, table
Device description			2		
Data sources	9.4	H3		Table 3A, Appendices	7.6, free text, table
Study size	9.5	H5	9	Table 7	7.9, free text, table
Data management	9.6	H6, H7		Table 3A, Appendices	7.7, free text, table
Data analysis	9.7	H8	12	Table 4, Table 5	7.5, free text, table
Quality control	9.8	H9	11		7.8 free text, table
Limitations of the methods	9.9	H10			
Other aspects	9.1				
Protection of human subjects	10	I	8	Table 1	9, free text
Management and reporting of adverse events	11				10, free text
Plans for disseminating and communicating study results	12	J			
References	13	K			References
Appendices	Annex			Appendices	Appendices
ENCePP Checklist for study protocols	Annex				Appendix

Note: Shaded gray area within bold black lines reflects core protocol components.

templates and were important to consider core for administrative or other reasons.

After populating an initial mock-up template, the core committee discussed and concluded that a combination of free-text and structured tables would increase the rigor and clarity in communication about study implementation decisions. Therefore, the structure of the harmonized protocol template largely follows the headers of the EMA GVP Module VIII-PASS template, with free-text and structured free-text prompts (in the form of helper text) aligned with the ISPE and NEST protocol guidance. These free text sections are where context and rationale for scientific decisions are entered. In sections about study methods, free-text is accompanied by structured tables from the STaRT-RWE template. The tables are where details of operational study implementation are specified. The free-text and structured tables are supplemented by detailed clinical code lists, algorithms, and descriptions of data linkage or data transformation in appendices.

1.3 | Piloting the usability of HARPER

To pilot the usability of the draft harmonized template, the core committee formed five subgroups. These subgroups had the task of populating the

TABLE 2 Example use cases

Use case	Question	What is unique
Empagliflozin versus DPP4i on 3P-MACE	Effectiveness	Existing protocol was reconstructed with more details added to match expectations in harmonized template (cohort study)
New cancer therapy compared to standard of care	Effectiveness	Harmonized template used for new protocol in development (cohort study)
Pioglitazone and risk of bladder cancer	Safety	Nested case–control design
Topiramate and oral clefts	Safety	Pregnancy cohort study with more complexity in design parameters
Palivizumab and RSV	Effectiveness	Self-controlled design

draft harmonized protocol template for a variety of use cases that involved different study designs, data sources, and types of data elements. Four of these use cases were based on published effectiveness and safety studies and one was for a study that was in the planning/design phase (Table 2). The members of each subgroup worked together to populate the initial version and relayed any issues to the core committee at large for discussion. The harmonized protocol template was revised to improve usability following this group exercise. These revisions included expanding the set of sections that were considered core, re-labeling of some structured prompts, and the addition of more helper text to guide investigators in use of the template. The abbreviated protocols for each use case was transferred onto the final version of the template to provide guidance and examples for future users (Appendix S3).

1.4 | Core sections of HARPER

Following the title page, abstract, and a table for amendments and updates, there are nine sections for the harmonized template, each of which includes structured free text, a structured table, or a figure (Table 3, Appendix S4). The majority of the protocol is focused on the critically important research methods (Section 1.4.7), where there are numerous subsections organized in the same way. A free-text section to lay out context and rationale for scientific choices is coupled with a table or figure to provide details on operational definitions.

1.4.1 | Title page

The title page includes a table for administrative details, such as the title of the protocol, brief objectives, a protocol version date, names of investigators and sponsor, study registration, and potential conflicts of interest.

TABLE 3 HARPER table of contents

1. Title Page
2. Abstract
3. Amendments and updates
4. Timeline
    - Table 1 Milestones and Timeline
5. Rationale and background
6. Research question and objectives
    - Table 2 Primary and secondary research questions and objective
7. Research methods
    - 7.1. Study design
    - 7.2. Study design diagram
    - 7.3. Setting
        - 7.3.1 Context and rationale for definition of time 0 (and other primary time anchors) for entry to the study population
            - Table 3 Operational Definition of Time 0 (index date) and other primary time anchors
        - 7.3.2 Context and rationale for study inclusion criteria:
            - Table 4. Operational Definitions of Inclusion Criteria
        - 7.3.3 Context and rationale for study exclusion criteria
            - Table 5. Operational Definitions of Exclusion Criteria
    - 7.4. Variables
        - 7.4.1 Context and rationale for exposure(s) of interest
            - Table 6. Operational Definitions of Exposure
        - 7.4.2 Context and rationale for outcome(s) of interest
            - Table 7. Operational Definitions of Outcome
        - 7.4.3 Context and rationale for follow up
            - Table 8. Operational Definitions of Follow Up
        - 7.4.4 Context and rationale for covariates (confounding variables and effect modifiers, e.g. risk factors, comorbidities, comedications)
            - Table 9. Operational Definitions of Covariates
    - 7.5. Data analysis
        - 7.5.1 Context and rationale for analysis plan
            - Table 10. Primary, secondary, and subgroup analysis specification
            - Table 11. Sensitivity analyses – rationale, strengths and limitations
    - 7.6. Data sources
        - 7.6.1 Context and rationale for data sources
            - Table 12. Metadata about data sources and software
    - 7.7. Data management
    - 7.8. Quality control
    - 7.9. Study size and feasibility
        - Table 13. Power and sample size
8. Limitation of the methods
9. Protection of human subjects
10. Reporting of adverse events
11. References
12. Appendices

1.4.2 | Abstract

The abstract is a free text section that includes a description of the background, research question and objectives, study design, and data sources.

1.4.3 | Amendments and updates

The protocol is intended to be a living document over the lifecycle of the study. Therefore, it is important to keep track of what changed and why the changes were made. This table documents what is changed, when it is changed, and why. For example, over the process of developing and implementing a protocol, investigators could start with an initial version of inclusion–exclusion criteria for doing an initial set of feasibility counts (looking at outcome counts that are not stratified by exposure), in version 2 using revised algorithms to generate a second set of feasibility counts to evaluate whether there is enough power and assess diagnostics such as propensity score overlap and balance, and in version 3 using finalized algorithms to create the analytic cohort.

1.4.4 | Milestones

This section includes a table to outline the anticipated timeline for study milestones.

1.4.5 | Rationale and background

This section includes structured free-text prompts to encourage inclusion of important key contextual information. For example, a paragraph about what is known about the condition and the exposures being investigated, knowledge gaps, and the expected contribution from the study described in the protocol.

1.4.6 | Research question and objectives

The prompts ask the user to summarize PICOT information – that is the population, intervention/exposure, comparator, outcome, and time horizon for the study (when follow up begins and ends) - as well as the main measure of effect. The text prompts closely align with the information needed to compare a RWE study design to a theoretical trial designed to address the same question (e.g., a target trial[27]). This section includes structured free-text prompts to specify the primary and secondary objectives, as well as the hypotheses being tested for each.

1.4.7 | Research methods

Study design and study design diagram

This section includes structured free-text prompts that ask the user to name the design and rationale for the choice of design. The structured free-text prompts are coupled with a design diagram, which the joint task force members agreed would be a critically important part of the harmonized protocol template because this figure provides a concise visual abstract to summarize the design of the study. We recommend a recently developed graphical framework for depicting study design for studies conducted with RWD,[28] but other visualizations can be used as appropriate. Conceptual models or directed acyclic graphs may be included as well.

Setting and variables

These sections include a free-text component to discuss rationale and context for choices relating to setting (selection of time zero [1], inclusion [2], exclusion criteria [3]) and variables (exposure [a], outcome [b], follow up [c], and covariates [d]). Each free-text component is followed by a structured table which prompts users to specify what is measured, the timing of measurement, the care setting (e.g., inpatient, outpatient, emergency department), type of codes that are used to define the measure (for example, drug, diagnosis, procedure or lab codes), as well as the sources for any algorithms used to derive study measures, for example, defining exposures, outcomes or covariates (whether that be from a publication or clinician review). For algorithms based on diagnosis codes, there is a section to define whether codes are required to be in the primary position (suggesting that the diagnosis is the main reason for the encounter). The clinical codes used to define each measure are specified later, in structured, machine-readable appendices as part of Section 1.9. Each table also includes fields to indicate whether the study parameter was pre-specified, whether it was varied for sensitivity, and the source of the algorithm to define that measure. Examples of how to populate the template for algorithms that are not based on clinical codes are available in the examples provided in supplemental appendices (Appendix S3).

In the outcome table (b), there is an additional field to specify the performance or validation of outcome algorithms, as well as a field to indicate which are the primary and which are the secondary outcomes. For the covariate table (d), there are also fields to specify things like how the variables are modeled (for example, as continuous, or categorical variables).

The structure of the follow up table (c) is different. The table uses structured fields to define when follow up begins relative to cohort entry and how it ends. The prompts help the investigator to consider each option and also makes it clear for the reviewer what is and is not used to end follow up. The table has fields to specify a variety of conditions that could end follow up such as death, disenrollment, a fixed calendar date, or end of exposure, with prompts to provide details. For example, on how duration of therapy is defined, it can be helpful to specify decisions regarding how to handle early refills or conversely short gaps in between dispensations.

Data analysis

This section is where the primary, secondary and subgroup analyses are specified. The context and rationale are discussed in the free-text component. The structured table includes fields for the hypothesis being tested, software packages, the specific models that are fit, the type of confounding adjustment, with prompts for specification of key parameters such as matching ratio and caliper for matched analyses, formulas for weights, trimming and truncation rules. Also reported in this table are fields to specify how missing data are handled in the analysis and subgroup analyses. For example, the investigator might choose to exclude patients with missing or unknown age and use multiple imputation for missing laboratory values.

There is also a structured table for detailing what sensitivity analyses are conducted and the rationale for conducting them (in other words, stating what might be learned from the sensitivity analysis). This rationale is especially important to help end users make sense of and interpret the results, particularly in a discipline where it is easy to run many sensitivity analyses.

Data sources

There is a free text component followed by a structured table for specifying data sources. The free text includes structured prompts to state the reasons for selecting the data, strengths and limitations of the data source(s) and information about data source provenance/curation. As shown in the examples, users may refer to detailed materials developed by data providers. This section can also include a detailed evaluation of the fitness-for-purpose of data source options, as outlined in the SPIFD[29] framework for identifying fit-for-purpose data or the EUnetHTA REQUEST[30] tool. The structured table outlines details such as the data source name, the data version, extraction date, sampling criteria (if relevant), data linkage,[31,32] or conversion to a common data model[33–35] are summarized in this table with more detail and data dictionaries in appendices as needed.

Data management

This section includes only a free-text component where the investigator can specify procedures for securely receiving, quality checking, storing, backing up and preparing data.

Quality control

This section includes only a free-text component where the investigator describes steps for quality assurance or quality check procedures, such as double programming or assessment of the reliability of the data (e.g., missing or miscoded data).

Study size and feasibility

This is a free text section where the appropriate precision, power and study size calculations are delineated to address the research questions, with description of the assumptions being made and sources that were used to make the assumptions. A table may be used to provide the selected parameters used in the power/sample size calculation if relevant.

1.4.8 | Limitation of the methods

This section is free-text and provides space for the investigators to summarize the anticipated limitations of the methods and data described in Section 1.4.4.

1.4.9 | Protection of human subjects

This free-text section is intended for the investigators to describe patient privacy protections and the plan to maintain data confidentiality or prevent re-identification. For example, investigators may report how the data were anonymized or pseudo-anonymized, whether small cell sizes were suppressed (if the data holder requires), and/or whether the study protocol underwent ethics review. For many studies using RWD, the latter may not be applicable. If the study is considered exempt by the relevant ethics board this should be stated with the reason it is considered exempt.

1.4.10 | Reporting of adverse events

This free-text section is for investigators to state the plan to report adverse events. This reporting is mandated for certain types of post-authorization studies.[36] If it is not applicable, that can be stated here.

1.4.11 | References

This section is for providing a bibliography for cited work.

1.4.12 | Appendices

The structured, human readable tables in the harmonized template are intended to be accompanied by appendices that list out the clinical code algorithms in a way that can be directly read in by programming code to facilitate creation of study variables. An example is provided in supplemental Appendix S3 Example 1. Appendices to detail other things, like decisions made when converting source data to a common data model or doing data linkage may also be relevant, depending on the study. Some appendices (e.g., specifying clinical code algorithms used for covariates), may not be developed until later versions of the protocol as the study progresses. Likewise, over the course of the conduct of the study, algorithms included in the appendices may be amended, with the changes documented in the amendments table. Some investigators may use code algorithms that they consider proprietary. If that is the case, this should be so noted in the protocol, thus allowing the reviewer to weigh the potential impact of not having this information on their ability to evaluate the validity or relevance of the study results.

2 | DISCUSSION

A joint task force between ISPE and ISPOR, including representation from key international stakeholders was formed to create a harmonized protocol template for RWE studies that evaluate a treatment effect and are intended to inform decision-making. HARmonized Protocol Template to Enhance Reproducibility (HARPER) builds on existing efforts, providing clarity, structure, and a common denominator regarding the level of operational detail, context, and rationale necessary in a protocol to produce a transparent, reproducible study and to support assessment of fitness-for-purpose. The overarching principle

was to reach for sufficient clarity in the protocol regarding data, design, analysis, and implementation over the lifecycle of a study to achieve three main goals. One, to help investigators thoroughly consider, then document their decisions and rationale regarding key study parameters that define the causal question (e.g., target estimand[37]). In this way, the template could help investigators to think more carefully about their choices and be used to help train a future generation on best practices. The second goal was to facilitate decision-making by enabling reviewers to readily assess potential for biases related to the clearly communicated investigator choices and rationale. The third goal was to facilitate reproducibility of results.

While the primary focus was on hypothesis evaluating RWE studies, HARPER can also be used as the basis of protocols for descriptive, utilization, predictive or other types of RWE studies. However, there may be some variation regarding which sections are considered core versus optional for different stakeholders (e.g., regulatory, HTA, academic, etc.).

2.1 | Parallel workstreams, relationships to checklists/bias assessment tools for RWE

In addition to issues of transparency, many professional associations, regulatory bodies, and health technology assessment agencies have issued best practice guidelines and checklists for the analysis of RWD. ISPOR,[38–42] ISPE,[24,43,44] the FDA,[1,45,46] the EMA,[17,47] the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP®),[5,48–50] and the European Network for Health Technology Assessment (EUnetHTA)[51] and the Japanese Pharmaceuticals and Medical Device Agency (PMDA)[52] have all published guidance documents on good practice. Widely used checklists for the reporting of observational studies include RECORD-PE,[53] STROBE[54] and CHEERS.[55] Several bias assessment tools[56–58] have been developed as well. However, these rely on provision of sufficient details on study methods to enable effective assessment of what was planned, what was done, and how these methods relate to what was found (e.g., validity). The harmonized protocol template aims to help investigators communicate clearly and effectively with reviewers and is consistent with current recommendations and strategies from key stakeholders.

In addition to improving transparency about hypothesis evaluating studies conducted with RWD, it is our hope that use of a harmonized protocol template will guide investigators in thinking about issues of study design, epidemiological and statistical methods, thereby reducing avoidable mistakes. Indeed, issues with study design may be even more important than confounding due to lack of randomization in explaining inconsistencies between RWE and RCT results.[59,60] Relatedly, assessing whether the data used for the study are fit-for-purpose is critical for considering issues of bias introduced by measurement error and inadequate control for confounders. Data that are fit-for-purpose for some study questions and designs may not be fit for others (e.g., an outpatient claims based data source is not fit for evaluating a study on the effectiveness of alternative inpatient therapies). The structured, harmonized protocol template outlines necessary details on the study question, design, and data to enable assessment of their collective fitness-for-purpose.

2.2 | Limitations

There are several limitations to HARPER. First, there is a trade-off between setting common standards for communicating about study design, analysis, and implementation versus full freedom to describe these in whatever fashion the investigator chooses. For example, the structured tables of the template may be challenging for studies that use complex, emerging designs. However, there is always the option of using the free-text sections to provide context. Relatedly, the sequencing of the sections in HARPER cannot align with all existing templates and processes from potential users across different subdisciplines. If users find that an alternative sequence better fits their needs (e.g., data sections before design, tables at the end instead of integrated in each section), they can re-order as needed. Second, our focus was on creating a harmonized protocol template to document and enable clear communication of scientific decisions for studies that make secondary use of real-world data for causal inference. This does not cover every aspect of transparency over the lifecycle of a research study, which may involve sharing of protocol, code, data, as well as results. Third, the guidance documents and templates identified and used in this harmonization effort may not include all relevant guidance that have been developed by different organizations around the world, however, among the major guidance documents that were identified, we have observed a great deal of concordance in the main elements. Fourth, as described earlier, this effort to change the status quo will rely on successful integration of the template into existing processes and guidance by key stakeholders. Thus, we have outlined a plan to pilot use of the template with multiple international regulatory, HTA, and payer stakeholders. Finally, real-world data analytics is a rapidly evolving field and while the template is flexible, it may need iterative revision. Therefore, we plan to have the harmonized template reviewed and updated as needed through a standing review process that is part of ISPE's policy for endorsed papers or products.

2.3 | Strategies to disseminate and facilitate use

In addition to this publication introducing HARPER, presentation at and endorsement from prominent research societies, we are engaging with international, inter-disciplinary stakeholders to lay the foundation toward routine use of the template for development of clearly specified protocols for RWD studies intended to inform decision-making. Once published, HARPER will be freely available for anyone to download and use. However, recognizing that the impact of HARPER relies on uptake, we have outlined a strategy to introduce and pilot the template with numerous key stakeholders over the next 2 years. We will be presenting the template directly to international regulatory agencies and HTA groups and are laying the groundwork to pilot test

the template with ongoing demonstration projects that are evaluating or guiding the use of RWE to support decision-making.

Furthermore, we intend to engage with study registration sites (EU-PAS, ClinicalTrials.gov, ISPOR-ISPE-Open Science Framework RWE Registry) to address registration of comprehensive protocols for RWE studies that estimate causal effects of clinical or policy interventions. On EU-PAS, the longest established registration site for observational studies, 57% of studies were registered without a protocol.[5] The work stream of this task force was specifically aimed at setting expectations regarding what needs to be in a study protocol to ensure more reproducible and reliable results. However, addressing key issues in protocol registration[61] and having policies that support it as an expectation for the field has potential to increase the ability of end-users to evaluate RWE study quality and therefore, their utility for decision-making.[1,38,43,62,63]

3 | CONCLUSION

Ambiguity in communication about the design and conduct of RWE studies that make secondary use of RWD limits their utility in healthcare decision-making. Clear communication about data provenance, design, analysis and implementation is needed. This would facilitate reproducibility, replication in independent data, and assessment of potential sources of bias.

HARPER was designed to reduce ambiguity by helping research teams be clear about the scientific decisions made in the design and conduct of an RWE study and to allow other investigators or reviewers to have a shared understanding of those decisions. It achieves this by creating a common text, tabular and visual structure so that multidisciplinary research teams and reviewers of their work will know what to look for and where to find it. The template provides a set of core expectations for clear, reproducible RWE study protocols and is intended to be used as a backbone throughout the research process from developing a study protocol, to registration, through implementation and reporting on those implementation decisions.

ACKNOWLEDGMENTS

Shirley V. Wang and Sebastian Schneeweiss were supported by NIH: NHLBI RO1HL141505 and NIA R01AG053302. Nicole Pratt was supported by NHMRC GNT1157506 and GNT1196900. Sinead Langan was supported by a Wellcome Trust Senior Research Fellowship in Clinical Science (205039/Z/16/Z). This research was funded in whole or in part by the Wellcome Trust (205039/Z/16/Z). For the purpose of Open Access, the authors have applied a CC BY public copyright license to any Author Accepted Manuscript (AAM) version arising from this submission. We would like to thank ISPE and ISPOR for their support and endorsement of this work.

CONFLICT OF INTEREST

Drs. Wang, Pottegård, Crown, Arlett, Ashcroft, Berger, Goettsch, Hua, Kurz, Orsini, Pratt, have no conflicts of interest to declare. Dr. Schneeweiss is principal investigator of the FDA Sentinel Innovation Center funded by the FDA, co-principal investigator of an investigator-initiated grant to the Brigham and Women's Hospital from Boehringer Ingelheim unrelated to the topic of this study. He is a consultant to Aetion Inc., a software manufacturer of which he owns equity. His interests were declared, reviewed, and approved by the Brigham and Women's Hospital and Partners HealthCare System in accordance with their institutional compliance policies. Dr. Benchimol has acted as a legal consultant for Hoffman La-Roche Limited and Peabody & Arnold LLP and acted as a consultant to McKesson Canada for matters unrelated to this manuscript. Dr. G Crane is an employee of F. Hoffmann-La Roche Ltd and holds stocks in F. Hoffmann-La Roche Ltd. Dr. Kabadi is an employee and shareholder of Sanofi. Dr. Kern is an employee of Janssen R&D, and stockholder of Johnson & Johnson. Dr. Perez-Gutthann is an employee of RTI Health Solutions, a division of the independent nonprofit Research Triangle Institute, who conducts research under contract for pharmaceutical companies, private, public organizations. Dr. Pinheiro worked on this research and article while employed by the US FDA, but is now employed by Abbvie. Dr Toussi works for IQVIA, a Human data science company who receives funds from pharmaceutical industry, governments and non-profit organizations to conduct research. Dr. Williams worked on this research and article while employed with the National Library of Medicine, NIH, but is no longer employed with NIH. The views expressed in this article are the personal views of the authors and may not be understood or quoted as being made on behalf of or reflecting the position of the European Medicines Agency, the Food and Drug Administration, the Pharmaceutical and Medical Devices Agency or the National Institutes of Health.

ORCID

Shirley V. Wang https://orcid.org/0000-0001-7761-7090
Anton Pottegård https://orcid.org/0000-0001-9314-5679
Eric I. Benchimol https://orcid.org/0000-0001-8855-3598
Xavier Kurz https://orcid.org/0000-0002-9838-7754
Susana Perez-Gutthann https://orcid.org/0000-0001-5798-3691
Nicole Pratt https://orcid.org/0000-0001-8730-8910
Sebastian Schneeweiss https://orcid.org/0000-0003-2575-467X

REFERENCES

1. US Food and Drug Administration. Framework for FDA's Real World Evidence Program 2018 2018. Accessed January 31, 2019. https://www.fda.gov/downloads/ScienceResearch/SpecialTopics/RealWorldEvidence/UCM627769.pdf
2. Real world evidence: experience and lessons from China. *BMJ*. 2018; 361:k1580. doi:10.1136/bmj.k1580
3. Elements of real world data/evidence quality throughout the prescription drug product life cycle. Government of Canada 2019. Accessed May 5, 2020, https://www.canada.ca/en/services/health/publications/drugs-health-products/real-world-data-evidence-drug-lifecycle-report.html.
4. Ball R, Robb M, Anderson SA, Dal Pan G. The FDA's sentinel initiative—a comprehensive approach to medical product surveillance. *Clin Pharmacol Therap*. 2016;99(3):265-268. doi:10.1002/cpt.320
5. Kurz X, Perez-Gutthann S, The ESG. Strengthening standards, transparency, and collaboration to support medicine evaluation: ten years of the European network of Centres for Pharmacoepidemiology and

Pharmacovigilance (ENCePP). *Pharmacoepidemiol Drug Saf*. 2018; 27(3):245-252. doi:10.1002/pds.4381
6. Patorno E, Schneeweiss S, Gopalakrishnan C, Martin D, Franklin JM. Using real-world data to predict findings of an ongoing phase IV cardiovascular outcome trial: cardiovascular safety of linagliptin versus glimepiride. *Diabetes Care*. 2019;42(12):2204-2210. doi:10.2337/dc19-0069
7. Franklin JM, Patorno E, Desai RJ, et al. Emulating randomized clinical trials with nonrandomized real-world evidence studies: first results from the RCT DUPLICATE initiative. *Circulation*. 2020;143:1002-1013. doi:10.1161/circulationaha.120.051718
8. Pottegard A, Kristensen KB, Ernst MT, Johansen NB, Quartarolo P, Hallas J. Use of N-nitrosodimethylamine (NDMA) contaminated valsartan products and risk of cancer: Danish nationwide cohort study. *BMJ*. 2018;362:k3851. doi:10.1136/bmj.k3851
9. Suissa S. Immortal time bias in observational studies of drug effects. *Pharmacoepidemiol Drug Saf*. 2007;16(3):241-249. doi:10.1002/pds.1357
10. Suissa S. Reduced mortality with sodium-glucose cotransporter-2 inhibitors in observational studies: avoiding immortal time bias. *Circulation*. 2018;137(14):1432-1434. doi:10.1161/CIRCULATIONAHA.117.032799
11. Kosiborod M, Cavender MA, Fu AZ, et al. Lower risk of heart failure and death in patients initiated on SGLT-2 inhibitors versus other glucose-lowering drugs: the CVD-REAL study. *Circulation*. 2017;136:249-259. doi:10.1161/circulationaha.117.029190
12. Udell JA, Yuan Z, Rush T, Sicignano NM, Galitz M, Rosenthal N. Cardiovascular outcomes and risks after initiation of a sodium glucose co-transporter 2 inhibitor: results from the EASEL population-based cohort study. *Circulation*. 2017;137:1450-1459. doi:10.1161/circulationaha.117.031227
13. Mehra MR, Desai SS, Kuy S, Henry TD, Patel AN. Retraction: cardiovascular disease, drug therapy, and mortality in Covid-19. *N Engl J Med*. 2020;382(26):2582. doi:10.1056/NEJMoa2007621
14. Mehra MR, Desai SS, Ruschitzka F, Patel AN. RETRACTED: hydroxychloroquine or chloroquine with or without a macrolide for treatment of COVID-19: a multinational registry analysis. *Lancet*. 2020. doi:10.1016/S0140-6736(20)31180-6
15. Benchimol EI, Moher D, Ehrenstein V, Langan SM. Retraction of COVID-19 pharmacoepidemiology research could have been avoided by effective use of reporting guidelines. *Clin Epidemiol*. 2020;12:1403-1420. doi:10.2147/CLEP.S288677
16. International Council on Harmonization. ICH Reflection Paper: Strategic Approach to International Harmonization of Technical Scientific Requirements for Pharmacoepidemiological Studies Submitted to Regulatory Agencies to Advance More Effective Utilization of Real-World Data, 2019. Accessed April 15, 2020, https://admin.ich.org/sites/default/files/2019-08/ICH_ReflectionPaper_Pharmacoepidemiology_2019_0605.pdf
17. Arlett P, Brun, N, Steering committee. HMA-EMA joint big data steering group: mandate. European Medicines Agency 2020. Accessed January 21, 2021. https://www.hma.eu/fileadmin/dateien/HMA_joint/00-_About_HMA/03-Working_Groups/Big_Data/2020_09_HMA-EMA_Big_Data_SG_Mandate.pdf.
18. COVID-19 Evidence Accelerator. Reagan-Udall Foundation for the Food and Drug Administration and friends of cancer research. Accessed January 29, 2021, https://evidenceaccelerator.org/
19. Guideline on good pharmacovigilance practices (GVP) Module VIII – post-authorisation safety studies (rev 3) section VIII.B.2. Study registration. European Medicines Agency; 2017.
20. European Medicines Agency. Guidance for the format and content of the protocol of non-interventional post-authorisation safety studies. Accessed January 4, 2022, https://www.ema.europa.eu/en/documents/other/guidance-format-content-protocol-non-interventional-post-authorisation-safety-studies_en.pdf
21. Wang SV. Reproducible evidence: practices to enhance and achieve transparency (repeat). 2020. https://www.repeatinitiative.org/.
22. European Medicines Regulatory Network Data Standardisation Strategy. European medicines agency. Accessed February 7, 2022. https://www.ema.europa.eu/en/documents/other/european-medicines-regulatory-network-data-standardisation-strategy_en.pdf.
23. Jaksa A, Wu J, Jónsson P, Eichler H-G, Vititoe S, Gatto NM. Organized structure of real-world evidence best practices: moving from fragmented recommendations to comprehensive guidance. *J Compar Effectiv Res*. 2021;10(9):711-731. doi:10.2217/cer-2020-0228
24. The International Society of Pharmacoepidemiology. Guidelines for good pharmacoepidemiology practices (GPP). Guideline. *Pharmacoepidemiol Drug Saf*. 2008;17(2):200-208. doi:10.1002/pds.1471
25. National Evaluation System for health Technology Coordinating Center (NESTcc) Methods Framework A Report of the Methods Subcommittee of the NEST Coordinating Center – An initiative of MDIC. 2020. https://nestcc.org/data-quality-and-methods/
26. Wang SV, Pinheiro S, Hua W, et al. STaRT-RWE: structured template for planning and reporting on the implementation of real world evidence studies. *BMJ*. 2021;372:m4856. doi:10.1136/bmj.m4856
27. Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. *Am J Epidemiol*. 2016;183(8):758-764. doi:10.1093/aje/kwv254
28. Schneeweiss S, Rassen JA, Brown JS, et al. Graphical depiction of longitudinal study designs in health care databases graphical depiction of study designs. *Ann Intern Med*. 2019;170(6):398-406. doi:10.7326/m18-3079
29. Gatto NM, Campbell UB, Rubinstein E, et al. The structured process to identify fit-for-purpose data: a data feasibility assessment framework. *Clin Pharmacol Ther*. 2021;111:122-134. doi:10.1002/cpt.2466
30. REQueST tool and its vision paper. EUnteHTA Accessed January 31, 2022. https://www.eunethta.eu/request-tool-and-its-vision-paper/
31. Pratt NL, Mack CD, Meyer AM, et al. Data linkage in pharmacoepidemiology: a call for rigorous evaluation and reporting. *Pharmacoepidemiol Drug Saf*. 2020;29(1):9-17.
32. Rivera DR, Gokhale MN, Reynolds MW, et al. Linking electronic health data in pharmacoepidemiology: appropriateness and feasibility. *Pharmacoepidemiol Drug Saf*. 2020;29(1):18-29. 10.1002/pds.4918
33. PCORnet Common Data Model. http://www.pcornet.org/pcornet-common-data-model/
34. Brown JB, Curtis N, Raebel L, Haynes MA, Rosofsky K, Sentinel R Common data model. Accessed June 14, 2018, https://www.sentinelinitiative.org/sentinel/data/distributed-database-common-data-model/sentinel-common-data-model
35. OMOP common data model. Observational medical outcomes partnership. Accessed March 3, 2016, http://omop.org/CDM
36. European Medicines Agency. Guideline on good pharmacovigilance practices (GVP) module VI – Collection, management and submission of reports of suspected adverse reactions to medicinal products (Rev 2). 2017. Accessed November 18, 2021. https://www.ema.europa.eu/en/human-regulatory/post-authorisation/pharmacovigilance/good-pharmacovigilance-practices.
37. Addendum on estimands and sensitivity analysis in clinical trials to the guideline om statistical principles for clinical trials E9(R). International Council on Harmonization. 2019, Accessed March 29, 2022, https://database.ich.org/sites/default/files/E9-R1_Step4_Guideline_2019_1203.pdf
38. Berger ML, Sox H, Willke RJ, et al. Good practices for real-world data studies of treatment and/or comparative effectiveness: recommendations from the joint ISPOR-ISPE special task force on real-world evidence in health care decision making. *Pharmacoepidemiol Drug Saf*. 2017;26(9):1033-1039.
39. Berger ML, Mamdani M, Atkins D, Johnson ML. Good research practices for comparative effectiveness research: defining, reporting and

interpreting nonrandomized studies of treatment effects using secondary data sources: the ISPOR good research practices for retrospective database analysis task force report–part I. *Value Heal J Int Soc Pharmacoeconom Outcom Res*. 2009;12(8):1044-1052. doi:10.1111/j.1524-4733.2009.00600.x
40. Cox E, Martin BC, Van Staa T, Garbe E, Siebert U, Johnson ML. Good research practices for comparative effectiveness research: approaches to mitigate bias and confounding in the design of nonrandomized studies of treatment effects using secondary data sources: the International Society for Pharmacoeconomics and Outcomes Research Good Research Practices for retrospective database analysis task force report–part II. *Value Heal J Int Soc Pharmacoeconom Outcom Res*. 2009;12(8):1053-1061. doi:10.1111/j.1524-4733.2009.00601.x
41. Johnson ML, Crown W, Martin BC, Dormuth CR, Siebert U. Good research practices for comparative effectiveness research: analytic methods to improve causal inference from nonrandomized studies of treatment effects using secondary data sources: the ISPOR good research practices for retrospective database analysis task force report–part III. *Value Heal J Int Soc Pharmacoeconom Outcom Res*. 2009;12(8):1062-1073. doi:10.1111/j.1524-4733.2009.00602.x
42. Berger ML, Martin BC, Husereau D, et al. A questionnaire to assess the relevance and credibility of observational studies to inform health care decision making: an ISPOR-AMCP-NPC good practice task force report. *Value Heal J Int Soc Pharmacoeconom Outcom Res*. 2014;17(2):143-156. doi:10.1016/j.jval.2013.12.011
43. Wang SV, Schneeweiss S, Berger ML, et al. Reporting to improve reproducibility and facilitate validity assessment for healthcare database studies V1.0. *Pharmacoepidemiol Drug Saf*. 2017;26(9):1018-1032. doi:10.1002/pds.4295
44. Cocoros NM, Arlett P, Dreyer NA, et al. The certainty framework for assessing real-world data in studies of medical product safety and effectiveness. *Clin Pharm Therap*. 2021;109(5):1189-1196. doi:10.1002/cpt.2045
45. Real-World Data: Assessing electronic health records and medical claims data to support regulatory decision making for drug and biological products guidance for industry. Food and Drug Administration https://www.fda.gov/media/152503/download
46. Considerations for the use of real-world data and real world evidence to support regulatory decision-making for drug and biological products. Food and Drug Administration Accessed January 4, 2021, https://www.fda.gov/media/154714/download
47. European Medicines Agency. Guideline on registry-based studies, 2021. https://www.ema.europa.eu/en/human-regulatory/post-authorisation/patient-registries
48. ENCePP. ENCePP Checklist for study protocols (revision 4), 2018. https://www.encepp.eu/standards_and_guidances/checkListProtocols.shtml
49. ENCePP. ENCePP guide on methodological standards in pharmacoepidemiology, 2021. https://www.encepp.eu/standards_and_guidances/methodologicalGuide.shtml.
50. Gini R, Fournie X, Dolk H, et al. The ENCePP code of conduct: a best practise for scientific independence and transparency in noninterventional postauthorisation studies. *Pharmacoepidemiol Drug Saf*. 2019;28(4):422-433. doi:10.1002/pds.4763
51. Internal validity of non-randomised studies (NRS) on interventions. EUnetHTA Accessed January 31, 2022. https://www.eunethta.eu/wp-content/uploads/2018/01/Internal-validity-of-non-randomised-studies-NRS-on-interventions_Guideline_Final-Jul-2015.pdf?x16454.
52. RWD WG. Pharmaceutical Devices and Medicines Agency. Accessed March 29, 2022, https://www.pmda.go.jp/english/rs-sb-std/rs/0023.html
53. Langan SM, Schmidt SA, Wing K, et al. The reporting of studies conducted using observational routinely collected health data statement for pharmacoepidemiology (RECORD-PE). *BMJ*. 2018;363:k3532. doi:10.1136/bmj.k3532%J
54. von Elm E, Altman DG, Egger M, et al. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *PLoS Med*. 2007;4(10):e296. doi:10.1371/journal.pmed.0040296
55. Husereau D, Drummond M, Augustovski F, et al. Consolidated health economic evaluation reporting standards (CHEERS) 2022 explanation and elaboration: a report of the ISPOR CHEERS II good practices task force. *Value Heal J Int Soc Pharmacoeconom Outcom Res*. 2022;25(1):10-31. doi:10.1016/j.jval.2021.10.008
56. Sterne JA, Hernán MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ*. 2016;355:i4919. doi:10.1136/bmj.i4919
57. Dreyer NA, Bryant A, Velentgas P. The GRACE checklist: A validated assessment tool for high quality observational studies of comparative effectiveness. *J Manag Care Spec Pharm*. 2016;22(10):1107-1113. doi:10.18553/jmcp.2016.22.10.1107
58. Wells G SB, O'Connell D, Peterson J, Welch V, Losos M, Tugwell P. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. 2013;
59. Franklin JM, Patorno E, Desai RJ, et al. Emulating randomized clinical trials with nonrandomized real-world evidence studies: first results from the RCT DUPLICATE initiative. *Circulation*. 2021;143(10):1002-1013. doi:10.1161/CIRCULATIONAHA.120.051718
60. Franklin JM, Glynn RJ, Suissa S, Schneeweiss S. Emulation differences vs. biases when calibrating real-world evidence findings against randomized controlled trials. *Clin Pharmacol Ther*. 2020;107(4):735-737. doi:10.1002/cpt.1793
61. Zarin DA, Crown WH, Bierer BE. Issues in the registration of database studies. *J Clin Epidemiol*. 2020;121:29-31. doi:10.1016/j.jclinepi.2020.01.007
62. Orsini LS, Monz B, Mullins CD, et al. Improving transparency to build trust in real-world secondary data studies for hypothesis testing-why, what, and how: recommendations and a road map from the real-world evidence transparency initiative. *Pharmacoepidemiol Drug Saf*. 2020;29(11):1504-1513. doi:10.1002/pds.5079
63. Berger ML, Crown W. How can we make more rapid Progress in the leveraging of real-world evidence by regulatory decision makers? *Value Heal J Int Soc Pharmacoeconom Outcom Res*. 2022;25(2):167-170. doi:10.1016/j.jval.2021.09.002

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Wang SV, Pottegård A, Crown W, et al. HARmonized Protocol Template to Enhance Reproducibility of hypothesis evaluating real-world evidence studies on treatment effects: A good practices report of a joint ISPE/ISPOR task force. *Pharmacoepidemiol Drug Saf*. 2023;32(1):44-55. doi:10.1002/pds.5507