Standards and guidelines for observational studies: quality is in the eye of the beholder

Sally C. Morton\textsuperscript{a,}\*, Monica R. Costlow\textsuperscript{a}, Jennifer S. Graff\textsuperscript{b}, Robert W. Dubois\textsuperscript{b}

\textsuperscript{a}Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh, 130 DeSoto Street, Pittsburgh, PA 15261, USA
\textsuperscript{b}National Pharmaceutical Council, 1717 Pennsylvania Avenue, NW, Suite 800, Washington, DC 20006, USA

Accepted 26 October 2015; Published online 5 November 2015

Abstract

**Objectives:** Patient care decisions demand high-quality research. To assist those decisions, numerous observational studies are being performed. Are the standards and guidelines to assess observational studies consistent and actionable? What policy considerations should be considered to ensure decision makers can determine if an observational study is of high-quality and valid to inform treatment decisions?

**Study Design and Setting:** Based on a literature review and input from six experts, we compared and contrasted nine standards/guidelines using 23 methodological elements involved in observational studies (e.g., study protocol, data analysis, and so forth).

**Results:** Fourteen elements (61\%) were addressed by at least seven standards/guidelines; 12 of these elements disagreed in the approach. Nine elements (39\%) were addressed by six or fewer standards/guidelines. Ten elements (43\%) were not actionable in at least one standard/guideline that addressed the element.

**Conclusion:** The lack of observational study standard/guideline agreement may contribute to variation in study conduct; disparities in what is considered credible research; and ultimately, what evidence is adopted. A common set of agreed on standards/guidelines for conducting observational studies will benefit funders, researchers, journal editors, and decision makers.

Keywords: Comparative effectiveness research; Observational studies; Standards

1. Introduction

Patient care and policy decisions demand high-quality evidence. Some have hypothesized that the volume, velocity, and veracity of observational data from electronic health records, administrative claims, and investment in data networks can be positioned to meet this demand. Despite the increased availability of data, lack of acceptance of observational studies by clinical and nonclinical decision makers may be due to unfamiliarity with the methods used or variability in study quality [1–4]. Deficient studies may lead to misuse of resources and result in poor health care outcomes for patients.

To improve the quality of observational studies conducted and reported, standards and guidelines by scientific, regulatory, and government organizations have proliferated over the past 5 to 10 years. Some documents serve as prerequisites for accessing specific data sets or obtaining funding [5]. A recent international initiative, STRengthening Analytical Thinking for Observational Studies, has been launched to provide guidance in the design and analysis of observational studies [4]. Among the various standards/guidelines, differences may exist based on purpose (e.g., conducting vs. reporting vs. evaluating research), philosophy (e.g., best practices vs. minimum standards), or format (checklist vs. explanatory primer).

The use of real-world evidence is critical to transforming health care, and policies need to address this issue. Although the availability of multiple standards/guidelines has the potential to improve the quality and credibility of observational studies, there may be unintended consequences. If those guidelines differ from one another, decisions based on one guideline may subsequently be found deficient if measured against a different one. Not having a “gold standard” becomes problematic for: (1) funders evaluating the merits of proposed research, (2) researchers seeking to use the most appropriate methods, (3) journal editors and reviewers assessing the quality of research conducted, and (4) decision makers who may adopt evidence from observational studies.

Funding: This research was funded by the National Pharmaceutical Council and the University of Pittsburgh (grant number: 709233).

\* Corresponding author. Tel.: 412-624-9939; fax: 412-624-2183.
E-mail address: scmorton@pitt.edu (S.C. Morton).
What is new?

Key findings
- Nine major standards/guidelines for observational studies have areas of disagreement based on a comparison of 23 common elements.

What this adds to what was known?
- Comparison of the standards/guidelines is presented, along with an assessment of how actionable each is, and a policy discussion of next steps is provided.

What is the implication and what should change now?
- Lack of standard/guideline agreement may contribute to variation in study conduct.
- Common standards/guidelines for conducting observational studies will benefit funders, researchers, journal editors, and decision makers.
- A consensus process to determine a common set of standards/guidelines needs to be established.

from academia, government, and industry to provide insight and advice throughout each stage of the project.

First, we conducted a detailed literature search to identify applicable standards/guidelines for observational studies from regulatory bodies, government agencies, professional organizations, initiatives, or collaborations. With the assistance of a medical librarian, we searched PubMed from January 1, 1947, to August 6, 2015, and also received gray literature citations from our experts, retrieving 3,812 citations (see Preferred Reporting Items for Systematic Reviews and Meta-Analyses [PRISMA] diagram [6], Fig. 1). A single reviewer examined titles and abstracts to identify articles that referred to standards, guidelines, consensus development conferences, or outstanding issues in the design, conduct, or reporting of observational studies. Dual data abstraction was conducted on the resulting 211 articles. One hundred forty-nine articles were excluded because they were either duplicative or did not address observational study research.

Next, based on input from our experts, we developed a set of inclusion criteria to focus on standards/guidelines for doing observational research. We thus excluded articles that were focused on methods for evaluating a specific condition (e.g., rheumatology), did not directly address observational study designs (e.g., were related to standards for systematic reviews or meta-analyses), or were standards/guidelines for reporting studies in journals (Fig. 1). After applying these criteria, 11 articles remained. These articles corresponded to nine standards/guidelines because three articles are content related and created by the same organization [7–9]. The nine standards/guidelines are as follows: Agency for Healthcare Research Quality (AHRQ): Developing a

2. Methods

We conducted a detailed literature search, identified salient standards/guidelines for observational studies, and compared and contrasted the standards/guidelines. We engaged a team of six international methodology experts conducted, and (4) patients, providers, and health plans as they assess the merits of the research to inform their decisions.

We sought to understand the commonalities and differences among standards/guidelines for observational studies and to determine how great a challenge exists. Based on this review, we outline the policy implication of standards/guidelines that are not uniform and focus on areas for future methods development and consensus and describe the benefits and potential harms associated with a common or harmonized set of standards/guidelines and optimal approach. To accomplish these objectives, we first conducted a detailed literature search to identify standards and guidelines. We then compared and contrasted the existence or lack thereof of specific methodological elements (and the actionability of those elements) associated with the research plan development, data collection and study implementation, analysis, discussion, and conclusions. Finally, we examine whether there are benefits and harms associated with a common or harmonized set of standards/guidelines and the optimal approach for harmonization.

![Fig. 1. PRISMA flow diagram.](image-url)
2.1. Data collection in a primary content table

Two team members independently reviewed all nine standards/guidelines. We then created a framework to compare them (Supplemental Material/Appendix at www.jclinepi.com). The framework was modeled after the comprehensive GRACE framework which outlines various steps associated with conducting a study and are typically found in a study publication (e.g., research plan, study methods, data analysis, discussion, interpretation, and conclusion). After testing this approach with our experts, we added unique elements from the other selected standards/guidelines. In this framework, each standard/guideline represented a column and each row corresponded to 23 observational study elements. The table cells contain the related standard/guideline content. We duplicated a standard/guideline’s content across elements, if appropriate, because we wanted to understand if specific standards/guidelines addressed a topic and if there was variation in the elements addressed. For example, potential bias was addressed by some standards/guidelines throughout the standard/guideline (e.g., when selecting the treatment comparators, in the study protocol of the research plan, in the analysis phase via conducting sensitivity analysis, and in the interpretation and bias section of discussion and conclusion), whereas others addressed bias in only a single element.

Some standards/guidelines did not offer a single recommendation for an element, but instead offered a choice in recommended actions. For example, GRACE 2.0 notes that “if one or more comparison groups were used, were they concurrent comparators? If not, did the authors justify the use of historical comparison group(s)?” allowing for choice. We recorded such choices in content within table cells.

2.2. Element summarization

We summarized each element (row) by taking into consideration the number of standards/guidelines (columns) addressing the issue, as well as similarities, differences, and choices. We categorized each element based on the number of standards/guidelines that addressed the element:

- 7–9 standards/guidelines addressed the element and content agreed across the standards/guidelines (green);
- 7–9 standards/guidelines addressed the element, but content did not agree across the standards/guidelines (blue);
- 4–6 standards/guidelines addressed the element (yellow);
- 1–3 standards/guidelines addressed the element (red).

Table 1 lists the results of this element summarization. The colors in Table 1 correspond to the four categories shown previously.

2.3. Actionability assessment

For each standard/guideline, we analyzed each element to determine if the standard/guideline advice (cell) was actionable. We defined actionable as, “A researcher has a two-part evaluation process for each component outlined in the nine standards/guidelines when doing an observational study. First, the researcher should clearly know that an element should be addressed. Second, if the element is included, then the researcher should understand the steps of that element and the level of detail required for fulfillment. If both these criteria are met, then the element is actionable.” For example, the FDA Guidance for Industry Good Pharmacovigilance Practices and Pharmacoepidemiologic Assessment notes that a hypothesis should be prespecified, but the guidance does not meet the second requirement describing how this should be done. However, for a separate element, the elements recommended by the FDA Guidance for industry are laid out clearly. If the element itself is not mentioned, we note that it is not addressed. For those standards/guidelines that addressed each element, we calculated the percent which provided actionable information. Table 2 lists the results of this actionability assessment.

3. Results

Each of the nine standards/guidelines had an intended, self-identified audience: five targeted researchers, three decision makers, one stakeholders, and one industry. The focus of the standards/guidelines varied with five focused on comparative effectiveness research (CER), three on pharmacoepidemiological or pharmacovigilance studies, and one on patient-centered outcomes research (PCOR), or arguably PCOR-CER. The format of the guidelines/stands also varied from journal articles, documents with web resources, a multiple-chapter document,
3.1. Comparison of standards/guidelines

Of 23 methodological elements, 14 (61%) were addressed by seven or more standards/guidelines (green or blue in Table 1). For two of the elements (9%), content agreed among at least seven standards/guidelines (green in Table 1). An example of agreement is the “study objectives and research questions” element. Eight standards/guidelines include reference to research questions and/or objectives of the study being defined before the study being conducted. Two standards/guidelines specifically outline the study objective to be based on the PICOTS/PICO [e.g., Population (P), Intervention (I), Comparator (C), Outcomes (O), Timing (T) and Setting (S)]. For 12 of the elements (52%), the content was addressed by seven to nine standards/guidelines, but those standards/guidelines disagreed on the approach or recommended actions (blue in Table 1). The standards/guidelines that addressed the element thus considered it important but did not agree on how to best address the element. An example is the missing data element. Although seven standards/guidelines mention missing data and six want information on the management of missing data, they differed in how missing data should be handled. For example, one standard/guideline suggested using mean value imputation, whereas others suggest other imputation methods, such as substitution of a predicted value from a regression model, hotdeck imputation, imputation based on a Bayesian approach, or multiple imputation.

Five elements (22%) were addressed by between four to six standards/guidelines (yellow in Table 1). An example is data linkage with five standards/guidelines addressing the element. Data linkage encompasses the study’s plan on how to link data from two or more sources, such as registries and data networks, while taking into consideration quality and accuracy. Across the standards/guidelines, groups noted the challenges associated with merging

---

**Table 1. Summary**

<table>
<thead>
<tr>
<th>Element of Observational Study</th>
<th>Propriety Addressing*</th>
<th>Similarities</th>
<th>Differences</th>
<th>Critical elements**</th>
</tr>
</thead>
<tbody>
<tr>
<td>Research Plan</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Review of Prior Research</td>
<td>6/9</td>
<td>Review of literature required</td>
<td>Specificity, e.g., ROM standards for SR versus &quot;review&quot;</td>
<td>Yes</td>
</tr>
<tr>
<td>Study Objectives and Research Questions</td>
<td>8/9</td>
<td>Research questions and objectives defined before study conducted</td>
<td>Use of PICOTS/PICO</td>
<td>No</td>
</tr>
<tr>
<td>Hypotheses</td>
<td>6/9</td>
<td>Pre-specification of hypotheses</td>
<td>Meaningful hypothesized difference</td>
<td>Yes</td>
</tr>
<tr>
<td>Communication With Stakeholders</td>
<td>9/9</td>
<td>Existence of multiple stakeholders; importance in outcomes/measures; stakeholders should take action</td>
<td>Stakeholder activities</td>
<td>Yes</td>
</tr>
<tr>
<td>Study Protocol</td>
<td>8/9</td>
<td>Have a study protocol and describe modifications</td>
<td>Elements required in protocol and detail</td>
<td>Yes</td>
</tr>
<tr>
<td>Study Design</td>
<td>8/9</td>
<td>Design based on study objectives</td>
<td>Choice of study design</td>
<td>Yes</td>
</tr>
<tr>
<td>Study Limitations and Potential Confounders</td>
<td>8/9</td>
<td>Mention study limitations and confounders</td>
<td>Methodology in dealing with confounders</td>
<td>Yes</td>
</tr>
<tr>
<td>Population Including Inclusion/Exclusion Criteria</td>
<td>9/9</td>
<td>Population mentioned</td>
<td>Subgroup characteristics; comparison between target and study populations</td>
<td>Yes</td>
</tr>
<tr>
<td>Comparators, Comparison Groups and Exposure</td>
<td>9/9</td>
<td>Description of intervention or exposure and comparison group</td>
<td>Choice in inclusion of comparator; historical control group; usual care</td>
<td>Yes</td>
</tr>
<tr>
<td>Sample Size and Statistical Power</td>
<td>8/9</td>
<td>Address sample size and statistical power</td>
<td>Assumptions, parameters, subgroup analyses</td>
<td>Yes</td>
</tr>
<tr>
<td>Measures, Endpoints and Outcomes</td>
<td>8/9</td>
<td>Address measurement characteristics</td>
<td>Statistical measures, accuracy, intermediate outcomes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

**Methods**

| Data Sources, including Collections, Coding, Capture and Storage | 9/9 | Detail on data sources | Registry requirements, coding, non-United States dataset, dataset choices | Yes |
| Data Linkage | 5/9 | Descriptive plan for linking data | Process, quality, replication of data linkage | Yes |
| Data Quality, Validity and Privacy Assurance | 8/9 | Description of reliability and validity, privacy assurance | Medical record and literature review; inclusion in protocol, data protection requirements | Yes |
| Missing Data | 7/9 | Information on management of missing data | Imputation method; sensitivity analysis | Yes |

**Analysis**

| Descriptive Analysis, Inferential Analysis and/or Modeling | 7/9 | Describe descriptive analysis, inferential analysis and/or modeling | Choice in statistical tests to use | No |
| Confounders and Modifiers | 9/9 | Identify confounders and modifiers; address confounding and explain how | Process; how confounding and modifiers will be addressed | Yes |
| Heterogeneity of Treatment Effect | 8/9 | Mention the need to assess heterogeneity of treatment effect | Limitation to subgroup analysis | Yes |
| Sensitivity Analysis | 7/9 | Description of sensitivity analysis | Result implications | Yes |

**Element of Observational Study**

<table>
<thead>
<tr>
<th>Propriety Addressing*</th>
<th>Similarities</th>
<th>Differences</th>
<th>Choice offered**</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discussions and Conclusions</td>
<td>8/9</td>
<td>Identify and address bias</td>
<td>Type of bias and how in addressed</td>
</tr>
<tr>
<td>Ethics Committee, IRB Approval or Other Ethical Considerations</td>
<td>3/9</td>
<td>Address ethical considerations</td>
<td>Type of ethical considerations</td>
</tr>
<tr>
<td>Interpretation</td>
<td>5/9</td>
<td>Internal validity</td>
<td>External validity</td>
</tr>
<tr>
<td>Dissemination</td>
<td>1/9</td>
<td>Plan and strategy description for dissemination</td>
<td>Authorship guidelines</td>
</tr>
</tbody>
</table>

*Green= 7-9 standards/guidelines addressed the element and the content agreed; Blue= 7-9 standards/guidelines addressed the element and the content did not agree; Yellow= 4-6 standards/guidelines addressed the element; Red= 1-3 standards/guidelines addressed the element

**Choice of actions was offered in at least one or more standards or guidelines
various data sets, privacy issues, data quality, and transferability of data elements or definitions.

Four (17%) of the elements were addressed by three or fewer standards/guidelines (red in Table 1). This outcome does not necessarily indicate disagreement between the standards/guidelines. However, there is a lack of information provided on the element by the majority, which is evident by as many as six empty or silent standards/guidelines. An example is communication with stakeholders, which only two standards/guidelines addressed.

Eighteen of 23 elements (78%) included options or choices for the researcher conducting the research. In terms of all possible elements across all standards/guidelines, 19% (40 cells of the total 207 cells) offered choices. The element of addressing confounders and modifiers was the most frequently associated with choices (six standards/guidelines). For example, rather than offering a single approach for dealing with differences between the comparator and intervention groups, a number of standards/guidelines identified different methods or techniques (e.g., restriction, stratification, interaction, and so forth) to adjust for these differences.

The level of detail and degree to which the standards/guidelines are actionable differed. Within the 23 elements, 13 (57%) had actionable entries in all standards/guidelines that addressed the element, such as study objectives and research questions, sample size and statistical power, and confounders and modifiers. Elements associated with the research plan, such as the hypotheses, study protocol, study design, study limitations and potential confounders, population, data sources, data linkage, sensitivity analysis, bias and interpretation, had at least one standard/guideline entry that was not actionable.

### 3.2. Policy discussion

We used 23 criteria to compare nine prominent standards/guidelines for conducting observational studies and found both similarities and important differences among them. Over half (61%) of the 23 methodological elements were included in at least seven of the standards/guidelines. However, those standards/guidelines often disagreed based on how those elements should be addressed or acted on. Depending on which standard/guideline was followed, research may miss important elements when viewed by others or address an element differently than expected.

For some elements, these inconsistencies may have little impact on variability in research quality or interpretation (e.g., the use of PICOT to outline the research question). However, for other elements (e.g., how to handle missing data, the need for a systematic review before conducting the research, or adequacy of subgroup analysis to assess heterogeneity of treatment effects), these inconsistencies may be problematic. Inconsistencies across standards/guidelines may contribute to methodological variation; disparities in what may be considered credible research by journal editors and peer reviewers; and ultimately, what evidence is adopted by clinicians.
With the increased availability of data for conducting observational research, a wide range of researchers have the ability to conduct observational studies and increasing numbers of studies will become available. A lack of consensus regarding how observational studies should be conducted may contribute to variability in interpretation by clinicians, policy makers, and patients and whether that research can or should influence decisions.

Different disciplines (e.g., epidemiology, pharmacovigilance, and health economics) each have similar, but different specific standards/guidelines. Bridging standards/guidelines across the disciplines may increase awareness across disciplines. To bridge these differences, we outline considerations below on how and who should drive alignment and consensus for observational study methodologies. In addition, we outline key considerations to ensure ongoing support and adherence to established standards/guidelines. This may serve to improve research conducted as well as raise confidence in those trying to interpret study results and to decide whether to use them.

3.2.1. Gaining alignment and consensus

Divergent views from different disciplines (e.g., biostatistics, econometrics, pharmacovigilance) and types of stakeholders (e.g., patients, funders, researchers, and decision makers) need to be considered in a consensus process. Given the differences in audience and philosophy (best practice vs. minimum standards), it is unlikely that any one set of standards/guidelines will be acknowledged by all stakeholders.

The nine standards/guidelines in our review were created from six different organizational perspectives representing public, private, and professional societies in the United States and internationally. With a limited number of organizations generating and updating standards/guidelines, creating a set of commonly agreed on methods is feasible. However, with the exception of PCORI, few organizations are likely to have the authority, resources, or staff to convene public and private thought leaders and a variety of stakeholders repeatedly. Furthermore, there is diversity in focus for each organization (e.g., PCOR-CER vs. pharmacovigilance). Therefore, bringing all relevant organizations together in a voluntary manner may be best to gain consensus. Audiences and research focuses may differ, but the elements of good practice in conducting observational studies should not be different. For example, with respect to missing data, CER and epidemiology may vary, but both would agree there is a need to consider this element.

Consensus-based approaches have previously been used to gain alignment across diverse stakeholders and perspectives. These approaches have been used to develop quality of care indicators, define appropriate use of “high cost/big ticket” procedures, create clinical practice guidelines and reporting guidelines such as PRISMA, and prioritize research questions [18–22]. Similar processes to gain alignment on appropriate methods for the conduct, reporting, and evaluation of observational studies are feasible and could bridge gaps among stakeholders.

A voluntary, consensus-based organization could foster this effort. There is a long history of this approach in health care, specifically in United States Pharmacopeia, such as the National Council of Prescription Drug Programs, Health Level 7, the National Quality Forum, and the Council for Affordable Quality Healthcare Committee on Operating Rules for Information Exchange, among others. These organizations are typically open to all parties, allow for debate across various stakeholders, include a broad representation of different stakeholders, are facile and iterative, and allow for open dialogue and discussion.

3.2.2. Identifying the level of consensus

Once a process is selected, it is important to determine what level of consensus should be sought. Should consensus standards/guidelines reflect best practices? Or should they reflect a minimum bar? Observational studies are used to inform different types of decisions, in which more or less rigor may be required. Although some stakeholders may advocate a set of aspirational best practices, a consensus process is unlikely to identify a single set that all stakeholders will agree on. For this reason, many of the existing standards/guidelines (e.g., AHRQ, PCORI, and so forth) have been outlined as a minimum set of standards/guidelines. Some researchers have noted that a minimum set of standards/guidelines may encourage rather than discourage methods innovation.

We believe a set of minimum standards/guidelines rather than best practices may be best. Beyond the minimum set, groups could choose to add additional elements and steps for their specific purposes. Over time, there is an opportunity for consensus to develop on newer or more advanced standards/guidelines gradually raising the bar. A set of minimum standards/guidelines could be developed to encourage uniformity, accessibility, and understandability across both experts and nonexperts, thereby improving the overall quality of research.

3.2.3. Accomplishing consensus

Gaining consensus could occur in a variety of forms. For example, gathering all relevant stakeholders and organizations could be feasible at a single point in time. However, as evidenced by the multiple iterations or editions of some standards/guidelines (e.g., AHRQ registry handbook), multiple and periodic meetings may be required to adapt and incorporate new method development.

Quicker consensus across stakeholders may be feasible for elements which are already considered important across the nine standards/guidelines. More effort and time would be required where there is disagreement on how to best address the element or where the element has been addressed by just a few organizations. Finally, there may be methods which were not prioritized, but may rise to the level of prioritization when considered by various stakeholders. Among
the methods we identified, some may allow for quick consensus, whereas others may require a more iterative process and resources for an ongoing consensus-based process.

3.2.4. Encouragement or enforcement

How to facilitate adoption of minimum standards/guidelines is the next logical issue. Should it merely be encouraged as guidance or should stronger measures be applied to ensure adherence? Some stakeholders could require or enforce adherence to a set of methodological standards/guidelines. For example, journal editors could support greater use of a common set of standards/guidelines for research through publication requirements. Funders could require adherence to the minimum standards to obtain research funding. The Centers for Medicare and Medicaid Services or other payers could require research adherence to standards/guidelines as a prerequisite to access claims data. Despite these individual leverage points, to reach the greatest adoption, voluntary approaches will also be needed. PCORI has designated funding for dissemination and implementation. These funds could assist in broader acceptance and use.

3.3. Limitations

Several limitations in our research approach may account for some of the differences observed. First, all standards/guidelines were silent on certain topics. For example, GRACE principles, GRACE 2.0 Checklist, and the ENCePP Checklist do not mention review of prior research to describe the study rationale or inform the study design. The absence of information may represent a lack of agreement that the specific elements are important or lower prioritization by the author(s). Second, the information provided by each standard/guideline may differ due to the audience, research questions addressed, format, and philosophy. Each standard/guideline was written for a specific audience, such as researchers or decision makers who may vary in their expertise. Included information was tailored by scope, content, and arrangement for the specific audience. Standards/guidelines focused on answering PCOR research questions would have a distinct focus on stakeholder engagement methods that would most likely not be included in a pharmacovigilance standard/guideline. Some standards/guidelines focused solely on conducting, while others included reporting, and others considered both activities. The content was presented in a variety of formats, ranging from a multichapter manual to a list of yes/no questions. The format and content influenced the level of detail provided and style of writing. Our approach was to excerpt the text verbatim and assume that the user of the standard/guideline would use that single document to assess the quality of a study. Finally, some standards/guidelines sought to inform readers, others outlined best practices, and others set minimum standards/guidelines.

Variation in each of these elements could contribute to the underlying differences across the standards/guidelines.

Some experts are concerned with standardizing methods in a field which is rapidly evolving due to the availability of new clinically rich data sources, analytic techniques, and approaches to dealing with data. There are concerns that standards/guidelines may impair methods innovation. We found that a number of standards/guidelines allowed the researcher, funder, or reader choice rather than a prescriptive approach. This may allow data sources, analytic techniques, and dissemination activities to continue to evolve. The consideration of choice may note differential uses (e.g., pharmacovigilance vs. PCOR-CER) or describe options for conducting research to inform different audiences.

4. Conclusion

From a policy perspective, many are seeking to understand what standards/guidelines are important to base decisions on and further who should be part of that process. The quality of observational studies conducted should be aligned regardless of the eye of the beholder or the standard/guideline by which an observational study is evaluated. A common set of agreed on minimum standards/guidelines could increase the likelihood of high-quality research and adoption of observational studies results. If successful, this effort would benefit funders of research as they evaluate proposals, researchers as they apply agreed on standards for conduct, and decision makers as they evaluate the quality of the research to inform their decisions.

Acknowledgments

The authors are grateful to the following experts for their guidance and assistance: Douglas G. Altman, DSc, University of Oxford; David Atkins, MD, MPH, United States Department of Veterans Affairs; Jesse A. Berlin, ScD, Johnson & Johnson; Nancy Dreyer, MPH, PhD, Quintiles; Mark Helfand, MD, MS, MPH, Oregon Health and Science University; and Cynthia D. Mulrow, MD, MSc, University of Texas Health Science Center. The experts do not necessarily endorse this article, and the authors are solely responsible for the contents and conclusions.

Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.jclinepi.2015.10.014.

References


