Defining Medication Exposure When Analyzing Real-World Data

Alecia Clary[1], Nancy Lin D[2], Tamar Lasky[3], matthew reynolds[4], Anand Chokkalingam[5], and Carla Rodriguez-Watson[1]

[1]Reagan-Udall Foundation for the FDA
[2]IQVIA Real World Solutions
[3]United States Food and Drug Administration Office of Data Analytics & Research Office of the Commissioner
[4]IQVIA Real-World Solutions
[5]Gilead Sciences

November 29, 2022

Authors:

1. **Alecia Clary** , Ph.D., MSW (AC) **(Corresponding author)**Affiliation: The Reagan-Udall Foundation for the FDA, Washington, DC, USA Address: 1333 New Hampshire Ave, NW Suite 420, Washington DC 20036, USA Email: aclary@reaganudall.org Phone: 202-849-2075 ORCID ID: 0000-0002-7774-9808
2. **Nancy D Lin** , (NDL) Affiliation: IQVIA Real World Solutions, Bridgewater, NJ, USA Address: 77 Corporate Drive, Bridgewater, NJ 08807 Email: Nancy.lin@iqvia.com
3. **Tamar Lasky** , (TL) Affiliation: The United States Food and Drug Administration, Office of Data, Analytics, & Research, Office of the Commissioner, Washington DC, USA Address: 10903 New Hampshire Avenue, Silver Spring, MD, USA. Email: Tamar.lasky@fda.hhs.gov
4. **Matthew W Reynolds** , (MR) Affiliation: IQVIA Real-World Solutions, Washington, DC, USA Address: 201 Broadway, Cambridge, MA, USA. Email: Matthew.Reynolds@IQVIA.com
5. **Anand Chokkalingam** , (ACH) Affiliation: Gilead Sciences, Foster City, California, USA Address: 333 Lakeside Drive Email: Anand.Chokkalingam@gilead.com
6. **Carla Rodriguez-Watson, (CRW)**Affiliation: The Reagan-Udall Foundation for the FDA, Washington, DC, USA Address: 1333 New Hampshire Ave, NW Suite 420, Washington DC 20036, USA Email: crodriguezwatson@reaganudall.org

This paper has not been previously printed, currently has not been submitted for publication in any other journal and is not pending acceptance in any other journal. We have not had any prior correspondence with the journal about the manuscript.

Key points:

- The conceptual definition of medication use includes the context under which the medication is being studied, the research concept of interest, the route of administration of interest, the dose of interest, and specifics regarding elements of timing and duration of medication use.
- The operational definition of medication use includes a description of the data source, the approach used to identify the medication, how timing elements related to medication use were identified and

calculated, the features of medication use, and the algorithm used to identify medication use.
- Adoption and reporting on the use of the proposed considerations may enhance the quality of pharmacoepidemiology research and our ability to contextualize and interpret findings.

Introduction

Medication use is an essential element of studies of drug safety or effectiveness[1]. There have been several successful efforts to define medication use, however, these efforts have focused on specific applications of using medications within particular analyses[2–5]. Although real-world data (RWD) is being increasingly used to generate real-world evidence (RWE) to guide clinical, policy, and regulatory decision making, there are limitations associated with conducting studies using these data. Medication data are not captured for research purposes and often need to be transformed from unstructured to structured data when incorporated into studies. The lack of formal conceptual and operational medication use definitions may result in failing to capture the varied dimensions of patients' medication use. Therefore, it is important to establish conceptual and operational definitions of medication use to aid investigators and the scientific community in designing, conducting, reporting on, and comparing across studies.

To our knowledge, there is no uniform structure to help researchers conceptualize or operationalize medication use in their studies. To address this gap, we developed a structure that includes considerations to address when developing medication use definitions. The structure can be used by researchers and policy makers to understand and contextualize findings from pharmacoepidemiology studies. We categorized our considerations in two major categories: considerations for a conceptual definition, or how medication use would *ideally* be defined; and considerations for an operational definition, or how the medication use *was* defined.

Considerations for the Conceptual Definition of Medication Use

As illustrated in Figure 1, we suggest five key considerations for developing the conceptual definition of the medication use.

Consideration 1: *The context under which the medication is being studied* (or what about medication use the researcher trying to capture*).* Consider whether the medication use is the central exposure variable or a covariate in a study assessing effectiveness or safety of a treatment, or an outcome in a study assessing prescribing practices.

Consideration 2: *The research concept of interest.* That is, whether the research focuses on a specific ingredient that may be found in more than one medication, on a specific medication, or in a class of medications.

Consideration 3: *The routes of administration that may be of interest.* Many medications have different indications and uses in their different formulations and have correspondingly different routes of administration. For example, corticosteroids may be administered via many routes and for myriad indications. The conceptual definition should also note whether any routes of administration are excluded.

Consideration 4: *Medication dose if of interest.* If of interest, the conceptual definition should include a description of the dose for each administration or daily dose, and an estimated cumulative dose. It may also be important to describe how different dosage forms (e.g., parenteral versus oral) will figure into the dose calculation if multiple forms are available. Researchers should also consider that specific dosages of some drugs are highly difficult to identify. They might want to incorporate definitions that consider dose changes over time.

Consideration 5: *The ideal timing and duration of medication use.* Considerations about whether the date the patient initiated the medication, the timing of subsequent uses of that medication, relative timing of

medication use and health events, whether continuous or cumulative exposure (i.e., duration), and whether discontinuation or changes to a different medication (as a proxy for standard of clinical care, lack of drug tolerance, or success/failure of a treatment) are of interest will depend on the study question. It is also useful to consider that the drug effect may persist after treatment discontinuation or that there may be gaps in treatment, which may have implications for the timing of administration in the analysis. This will usually require an understanding of the pharmacological characteristics of the medication.

Considerations for the Operational Definition of Medication Use

As illustrated in Figure 2, we suggest five key considerations for developing the operational definition of the medication use.

Consideration 1: *The underlying pattern of health seeking behavior and its documentation within the health system that give rise to the observable data.* There are various types of encounters patients may have as they move within the health care ecosystem that may influence the operational definition. A strategy to screen for the records or occurrences of the medication use, should be informed by the format in which the medications data are captured and stored. For example, during hospitalization, a patient's medication use may be collected in different ways; hospital clinical staff use patients' charts information to enter patient medication use into electronic health record (EHR) (e.g., clinician-generated prescription orders, medication administrations, pharmacy dispensing); hospital clinical staff ask patients about past and current medication use; or prescription orders recorded as part of hospital discharge instructions. This information may be transformed to a common data model, used for billing and payment purposes, and potentially included in medical claims (though detail related to medication use in claims may be limited).

Consideration 2: *The characteristics of data source that were used (e.g., location,) and approach used to identify the medication used.* As can be seen in Figure 3, identification of a medication requires finding of the likely location (e.g., specific data tables/views and fields) where relevant medications-related information is stored, which is heterogenous. Therefore, we suggest that decisions regarding appropriate sources of medication information first be guided by knowledge of clinical practice and how the data are produced and sorted (e.g., clinician decision to order a medication versus administration of the medication to the patient). The format of the sorted data should also be considered. For example, whereas administrative and claims data are highly structured and standardized in format by virtue of the specific requirements to which providers must adhere to receive payment, other data sources such as EHRs are often more heterogeneous. Medication information may be stored as coded data within structured fields, text within semi-structured or discrete fields, or within free-text clinical notes. Coded medications data may take the form of organization-specific or local codes, or may include standard terminology or classifiers (e.g., RxNorm, NDC, Multum, FirstDataBank). Whether medication information is stored as codes or as text, it is useful to submit supplementary materials include the structured code lists or list of text strings used to identify occurrences of the medication use of interest. Similarly, it is essential to report operational definitions and methods when using information from semi-structured or unstructured data. Regardless, researchers might consider engaging individuals with expertise of the specific data source being used.

Consideration 3: *Actual timing elements related to medication use.* The ability to define various aspects of timing of medication use is an integral component of pharmacoepidemiologic research questions. The researcher should describe how they identified the date of patient initiation of a medication and the timing of subsequent uses of that medication, date of medication discontinuation, and relative timing of medication use and health events of interest, as they were defined in the conceptual definition. They should also specify how they calculated the duration of medication use if it is of interest, including any assumptions of medication use tied to the definition (e.g., identifying a prescription within an EHR and assuming the medication was filled and used for the full days supplied).

Consideration 4: *Other features of medication use.* The features of medication use that are of interest should be defined. These features might include whether prevalent or incident use is of interest and how it

was defined; whether patients can have multiple qualifying episodes of medication use or only one; and how dose was calculated, if of interest. Researchers should also describe how they addressed missing data and discrepancies if they are using more than one data source.

Consideration 5: *Algorithm validation.* If an algorithm was used to identify the medication, researchers should specify whether the algorithm for the medication measure was validated.

Discussion

As RWD is increasingly used to generate evidence to make safety and efficacy decisions, the challenges and lack of clarity associated with incorporating medication use into analyses should be addressed. We offered points to consider when developing medication use definitions, with a focus on electronic health care data sources. These considerations may help researchers understand the strengths and limitations of different conceptual and operational approaches (e.g., selection of data sources, method of identifying medication) and contextualize their findings. These points are intended to complement considerations from existing rubrics, such as STaRT-RWE[6], the ISPE-ISPOR Consensus Document[7], RECORD-PE[8], and the ENCEPP Guide on Methodological Standards[9], which aid investigators in designing, conducting, reporting on, and comparing across studies.

Our proposed points included considerations for investigators who incorporate medication use into their study. We have shown the importance of the considering context under which the medication is being used in the analyses, the volume of medications formulas, dosages, routes of administration, and the often-longitudinal nature of medication use may influence a definition of medication use.

In summary, adopting and reporting on the use of the proposed considerations may enhance the quality of pharmacoepidemiology research and our ability to contextualize and interpret findings.

Funding statement: This activity is a project supported by the Food and Drug Administration (FDA) of the U.S. Department of Health and Human Services (HHS) as part of an award of $570,206 in federal funds (100% of the project). The contents are those of the author(s) and do not necessarily represent the official views of, nor an endorsement, by FDA, HHS, or the U.S. Government. For more information, please visit FDA.gov.

Conflict of interest disclosure: Anand P. Chokkalingam is an employee of and own stock in Gilead Sciences. The rest of the authors have no competing interests to declare that are relevant to the content of this article.

Ethics approval statement: Because this wasn't a research activity it is exempt from review and the requirement for informed consent.

Patient consent statement: Not applicable

Permission to reproduce material from other sources: Not applicable

References:

1. Richesson RL, Smith SB, Malloy J, Krischer JP. Achieving standardized medication data in clinical research studies: two approaches and applications for implementing RxNorm. Journal of medical systems. Springer; 2010;34:651–7.

2. What is the Phenotype KnowledgeBase? | PheKB [Internet]. [cited 2022 Apr 26]. Available from: https://phekb.org/

3. SEER*Rx - Cancer Registrar's Interactive Antineoplastic Drugs Database - SEER Tools [Internet]. SEER. [cited 2022 Apr 26]. Available from: https://seer.cancer.gov/tools/seerrx/index.html

4. Warner MD MS. HemOnc CC BY subset [Internet]. Harvard Dataverse; 2022 [cited 2022 Apr 26]. Available from: https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/9CY9C6

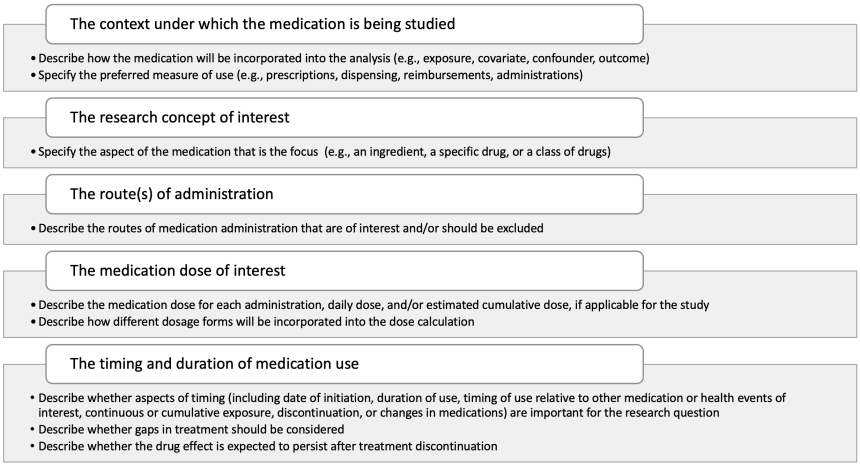
5. Warner MD MS. HemOnc ontology [Internet]. Harvard Dataverse; 2022 [cited 2022 Apr 26]. Available from: https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/FPO4HB

6. Wang SV, Pinheiro S, Hua W, Arlett P, Uyama Y, Berlin JA, et al. STaRT-RWE: structured template for planning and reporting on the implementation of real world evidence studies. bmj. British Medical Journal Publishing Group; 2021;372.

7. Berger ML, Sox H, Willke RJ, Brixner DL, Eichler H-G, Goettsch W, 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. Value in Health. Elsevier; 2017;20:1003–8.

8. Langan SM, Schmidt SA, Wing K, Ehrenstein V, Nicholls SG, Filion KB, et al. The reporting of studies conducted using observational routinely collected health data statement for pharmacoepidemiology (RECORD-PE). bmj. British Medical Journal Publishing Group; 2018;363.

9. Blake KV, Smeraldi C, Kurz X, Arlett P, Blackburn S, Fitt H. The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance: application to diabetes and vascular disease. The British Journal of Diabetes & Vascular Disease. Sage Publications Sage UK: London, England; 2011;11:304–7.



The underlying pattern of health seeking behavior and its documentation within the health system

- Describe the strategy used to screen for medication use, informed by underlying pattern of patient behavior, and formatting documentation within the data source
- Describe the steps taken to transform the data into a common data model (e.g., cleaning, curation, mapping)
- Describe the medication record type(s) (e.g., prescription order, medication administration, dispensing claim, multiple types)

The characteristics of data source that were used and approach used to identify the medication used

- Use knowledge of clinical practice, how the data are produced, and how the data are sorted to guide decisions regarding the appropriate sources of medication information within a given data source
- Use knowledge about the format in which the medications use data are captured and stored (in structured fields, as coded data within structured fields, test within semi-structured or discrete fields, within fee-text clinical notes) to guide the identification of medications used
- Include the structured code list or list of string texts used to identify occurrences of medication use as supplementary material for any study protocols and manuscripts

Timing elements related to medication use

- Describe how elements related to timing (initiation, subsequent uses, discontinuation) were identified
- Describe how the relative timing of medication use (switching to a new product or the health event of interest) was operationalized, based on the conceptual definition
- How the duration of medication use was calculated

Other features of medication use

- If prevalent or incident use is of interest, describe how it was defined. If incident use is of interest, describe whether or how a clean or washout period was applied
- If only one qualifying episode of medication use was allowed (vs. multiple qualifying episodes), describe how this was selected
- If medication dose is of interest, describe whether dose information was available, the measure of interest, and how medication dose was calculated
- Describe how records with missing medication information were handled
- If multiple data sources are used for a given patient, describe how discrepancies and potential duplicates were handled

Algorithm validation

- Provide the relevant references or describe the validation performed.
- Provide the performance metrics from the validation.