Paper AS17

The Role of Quantitative Bias Analysis (QBA) in Ensuring the Reliability of Real-World Data (RWD) in Regulatory Submissions

Tasha Nagamine MS, Droice Labs, New York, USA
Brian Gillette PhD, Droice Labs, New York, USA

ABSTRACT

Real-world data (RWD) has tremendous potential to increase clinical trial efficiency by reducing required sample size. However so far there has been limited use of RWD for primary evidence in regulatory submissions due to RWE reliability issues stemming from rampant missingness and misclassification in RWD. To make RWE reliable for regulatory decision-making, FDA's recently finalized guidance recommends quantitative bias analysis (QBA) to adjust for biases that occur due to RWD reliability issues. However, measuring the bias parameters needed for QBA in RWD is extremely challenging. Droice Labs developed a unique approach to measuring bias parameters in order to rely on RWD inferences in regulatory decision-making and recently discussed it with the FDA. Here we present our approach where patient-level data from multiple completed clinical trials from several major pharma companies were integrated with patient-level RWD to demonstrate potential sample size reductions in each trial while making it reliable for regulatory submissions by using QBA.

INTRODUCTION: RWD FOR FASTER, MORE EFFICIENT TRIALS

Randomized controlled trials (RCTs) are the gold standard for regulatory evidence generation but are expensive and often take years to complete. Real-world data (RWD) from sources like electronic health records (EHR) contains rich information about patient disease progression and treatment response with huge potential for advancing disease understanding [1-2]. If RWD can be used reliably in clinical trials for regulatory approvals, the cost and time to complete RCTs can be brought down significantly by reducing the number of patients required to enroll by using external controls (e.g., external control arms, hybrid control arms [3]) (Figure 1).

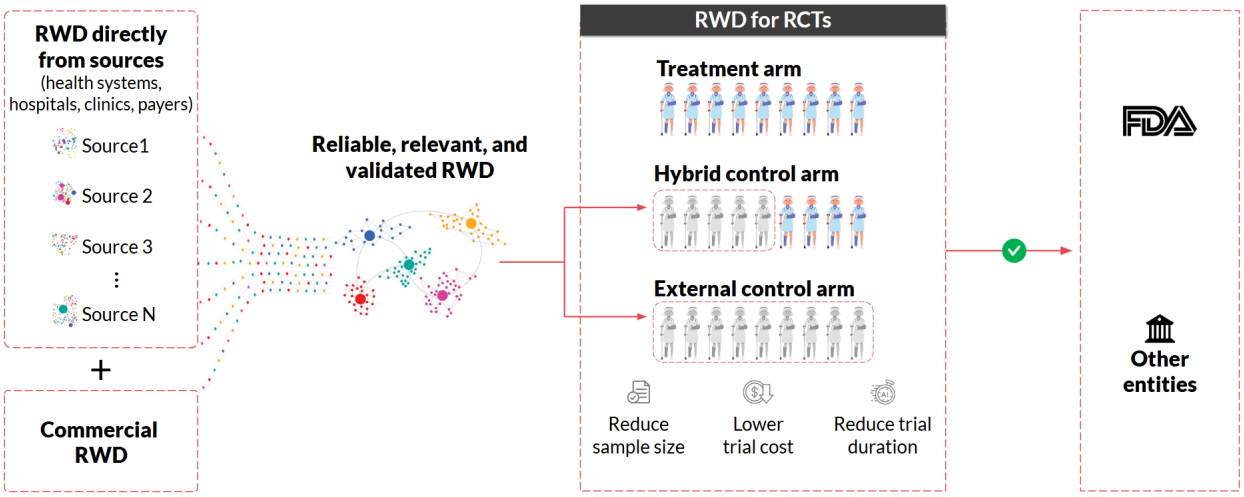

Figure 1. Reliable, relevant, and validated RWD can be harnessed to reduce required enrollment in trials for regulatory decisions

PROBLEM: RWD ≠ RCT DATA

Patient variables in RCT data are manually collected and often completely verified to ensure high reliability (near perfect accuracy and completeness). In contrast, the scale of RWD often makes manual collection and complete verification of patient variables practically unfeasible. Patient variables in RWD are thus typically derived algorithmically, and as a result are less reliable (less accurate and complete) due to missing data in the source and/or misclassification errors/omissions when transforming source RWD into study variables. These errors arise because of both the mismatch between clinical care-focused EHR schemas and clinical research analysis data models, and the immense practical challenges involved in comprehensive identification, extraction, and transformation of relevant data elements from complex EHR sources (Figure 2) [4-10]. As a result, misclassification in RWD can be much more extensive than RCT data and lead to significant biases that are not encountered or handled in typical RCT analyses.

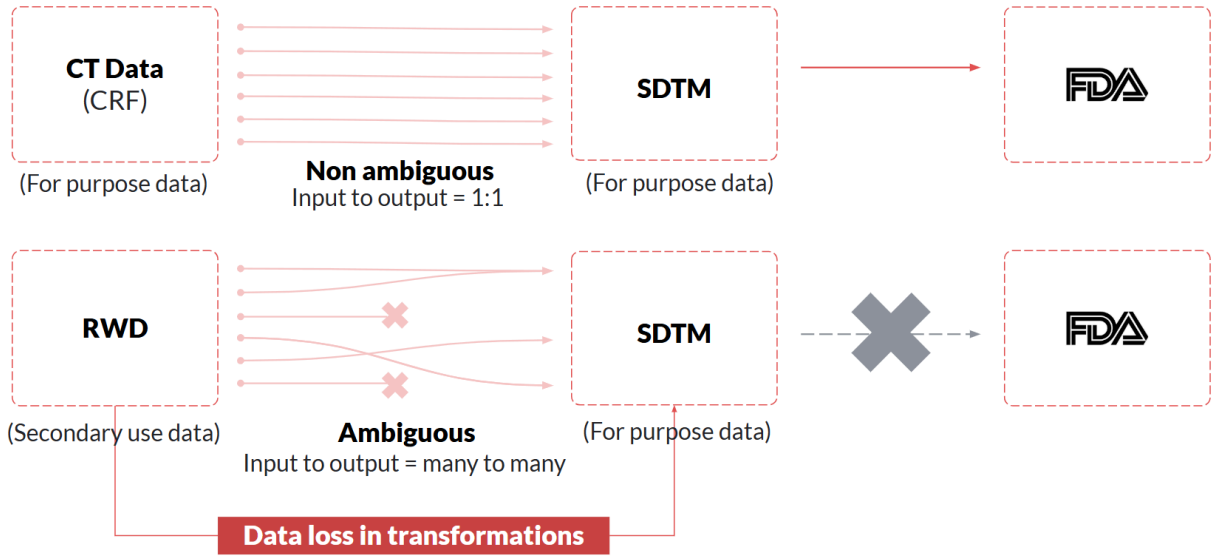

Figure 2. RWD ≠ RCT data. Generating evidence using RWD requires transforming the data from complex and messy sources into organized and accurate variables. However, such transformations are error-prone due to the complex nature of typical RWD schemas that are optimized for clinical care rather than research, and such errors render the data unreliable for rigorous clinical science and regulatory decision-making.

This extensive misclassification across different types of RWD study variables leads to different types of biases that render inferences unreliable:

- Misclassified inclusion/exclusion criteria change the study cohort, resulting in selection bias
- Misclassified exposures change arm assignment (exposed or unexposed), biasing the exposure-outcome relationship
- Misclassified prognostic covariates distort the real profiles of confounding factors, making confounding control ineffective or inaccurate
- Misclassified outcomes can change the potential efficacy and safety profile across study groups, biasing the exposure-outcome relationship

Impact: If misclassification bias cannot be addressed, analyses using RWD will not be robust enough for regulatory decision-making and will face rejection.

HANDLING BIAS IN RWD THROUGH QUANTITATIVE BIAS ANALYSIS (QBA)

Quantitative bias analysis (QBA) is a set of methods to quantify and model the impact of systematic errors on inferences [11, 12], yielding a quantitative assessment of how robust (or not) study inferences are to error in the underlying data. FDA's recently finalized RWE guidance recommends performing QBA to evaluate the impact of misclassification of all types of RWD-derived study variables (inclusion/exclusion criteria, exposures, covariates, and outcomes) on the study findings [8]:

FDA

Real-World Data: Assessing Electronic Health Records and Medical Claims Data to Support Regulatory Decision-Making for Drug and Biological Products

Guidance for Industry

July 2024

> ***FDA also recommends using quantitative approaches, such as quantitative bias analyses... to demonstrate whether and how misclassification, if present, might impact study findings***. The protocol should pre-specify the indices (e.g., sensitivity, specificity) that will be used for quantifying bias and describe how the selected indices will be measured during validation.

QBA generates bias-corrected estimates of treatment effects by utilizing “bias parameters” that quantify the misclassification type and extent in each study variable (e.g., sensitivity and specificity) to adjust the inference results to account for the estimated error (Figure 3).

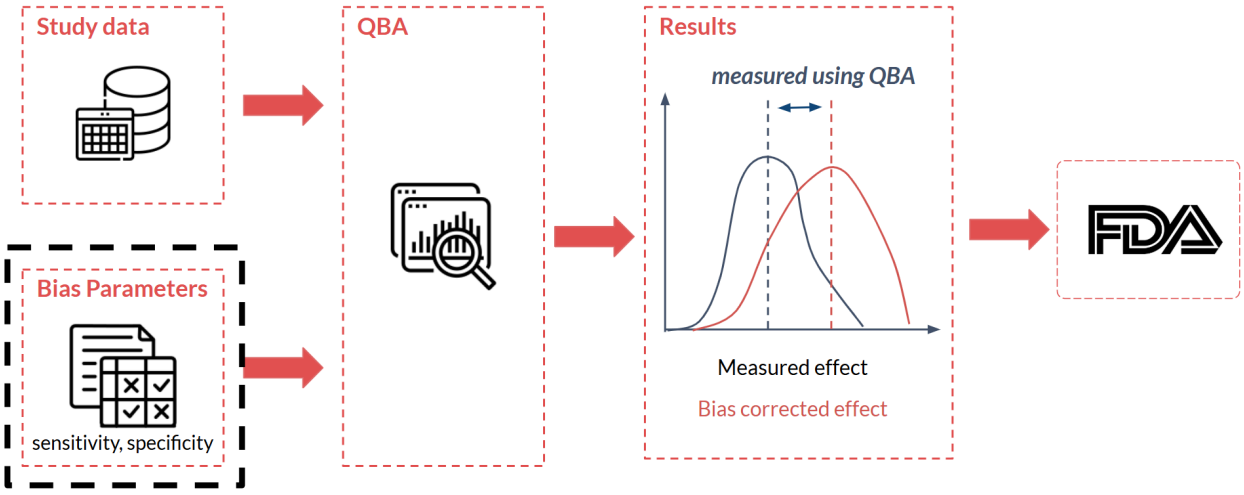

Figure 3. Illustration of the process to use QBA to evaluate robustness of inferences when using RWD in regulatory submissions to FDA.

OBTAINING BIAS PARAMETERS FROM RWD IS CHALLENGING

Accurate assessment of bias parameters for misclassification in RWD transformations (e.g., sensitivity and specificity) through validation studies is critical for conducting QBA with adequate rigor (Figure 4) [8,11]. Validation studies on RWD are typically performed using two main approaches:

1. **Validation using raw RWD.** Validation studies attempting to use raw RWD as a source of truth face significant practical challenges due to same data complexity issues that cause information loss in data transformations. Traceability of output variables to source data is typically non-existent or limited to field level (e.g., table, column) or value level (e.g., source values in OMOP [13]) rather than individually traceable elements, making it difficult to discern true and false positives, while the lack of lineage for patient source data not used in analysis prevents the determination of true and false negatives. Hence, in this case it is nearly impossible to generate an accurate or complete confusion matrix by validating variables directly on raw RWD.
2. **Validation by manual chart review.** Validation studies utilizing manual chart review via source data interfaces (e.g., EHR user interfaces) are resource intensive, and the diversity, complexity, and limited queryability of such user interfaces makes comprehensive review practically impossible.

Furthermore, in both approaches, each individual RWD source has a unique site-specific transformation error profile due to variations in schemas and clinical/operational workflows, and this variability across sources can induce biases in the validation process that impact the validity of the validation study itself. This lack of standard approaches to validation studies is a hindrance to effective QBA in studies utilizing RWD.

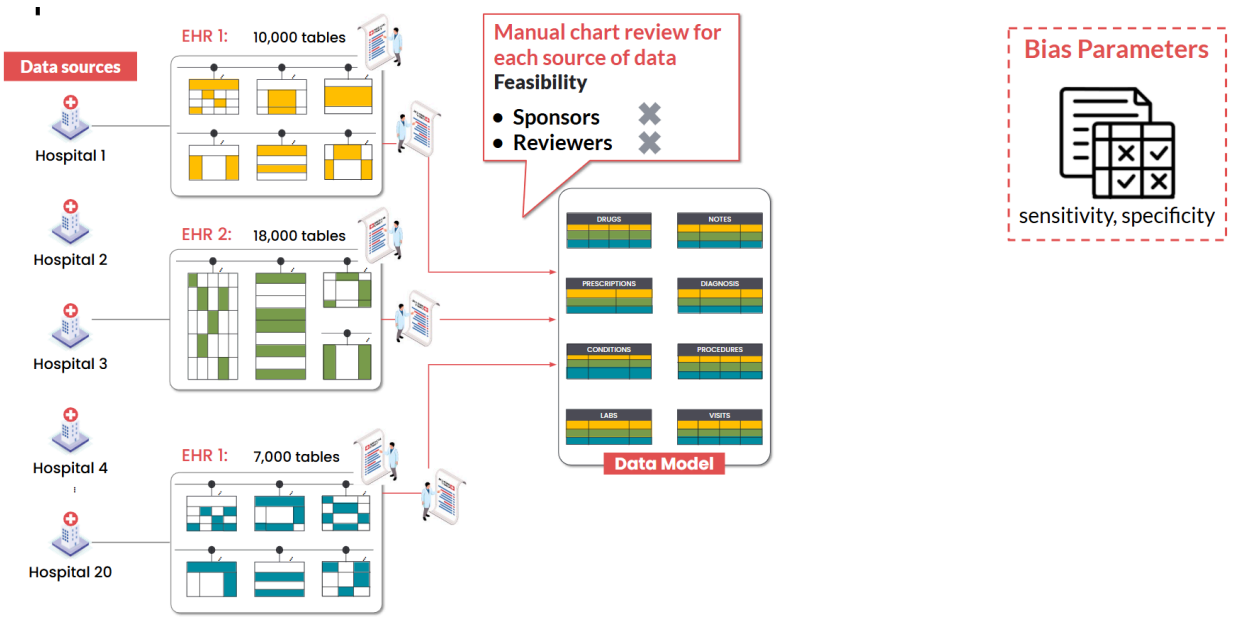

Figure 4. Manual validation on highly variable and complex RWD sources like raw electronic health records (EHR) makes a standardized and reproducible validation process nearly impossible.

FDA guidance on the conduct of RWD validation studies notes that a "standard and reproducible process is critical for minimizing intra- and inter-rater variability, especially for multi-site studies in which medical records… cannot be shared across systems and a centralized medical record review is not possible." [8] However, as is illustrated in Figure 4, this is extremely difficult to achieve in practice on highly messy, noisy, source RWD.

STANDARDIZED AND SCALABLE RWD VALIDATION

Droice Labs developed SuperLineage (Superset of Positive and Negative Lineage) to solve the challenges in RWD validation (Figure 5). SuperLineage contains lossless and comprehensive lineage for each source patient data element that specifies either its location(s) in the output analysis dataset (Positive Lineage) or that the element was not used in the output analysis dataset (Negative Lineage). Positive Lineage allows for assessment of true and false positives, while Negative Lineage allows for assessment of true and false negatives, which are all required to adequately assess misclassification of study variables and calculate bias parameters (sensitivity and specificity) for QBA. By converting diverse source RWD losslessly into a standard and queryable format, SuperLineage enables purpose-built approaches to error quantification in data transformations to scale and reduce bias in validation studies, providing a standard solution for RWD reliability. Droice discussed this approach with FDA as a solution to providing the accurate bias parameters needed for reliable RWD without the drawbacks of validating raw RWD databases or manual chart review [14].

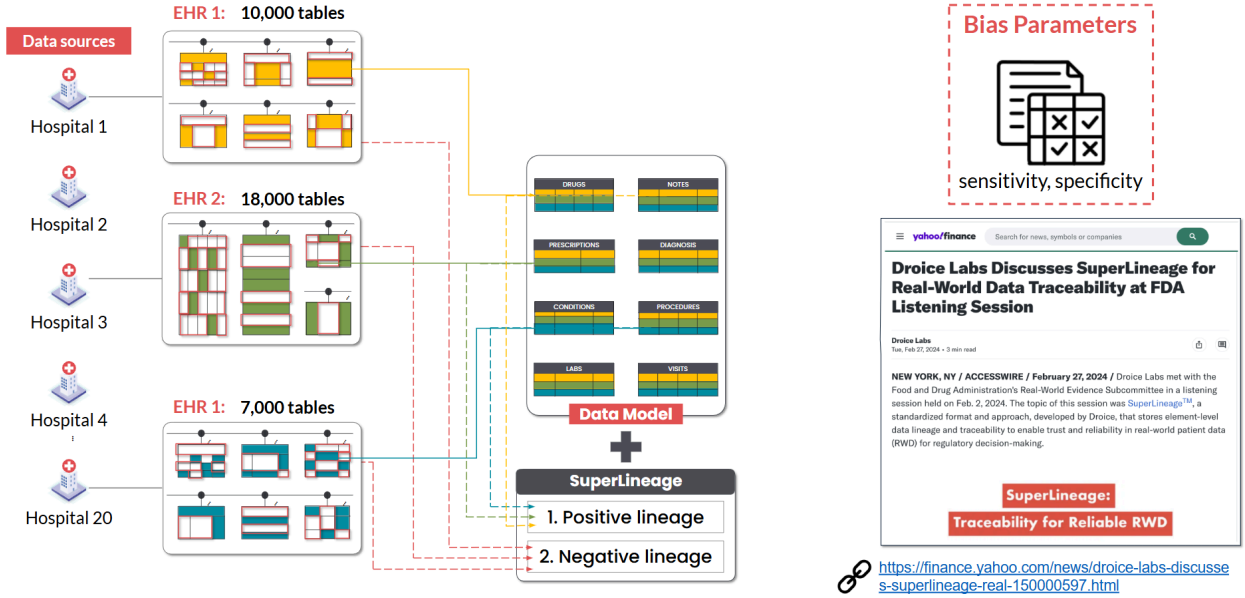

Figure 5. Droice SuperLineage, a standardized and scalable approach to RWD validation.

ACHIEVING ROBUST, FASTER, MORE EFFICIENT TRIALS WITH RWD AND QBA

Droice Labs is developing approaches to use RWD to reduce sample size in RCTs without sacrificing statistical power and inference robustness through comprehensive validation and QBA (Figure 6). In this project, participant-level data from >60 completed RCTs from multiple pharma are being combined with deidentified patient-level RWD from a diverse set of >100 hospitals across the US and Europe to evaluate the potential for RWD to reduce required enrollment in RCTs while meeting FDA's requirements for RWD reliability and inference robustness using QBA.

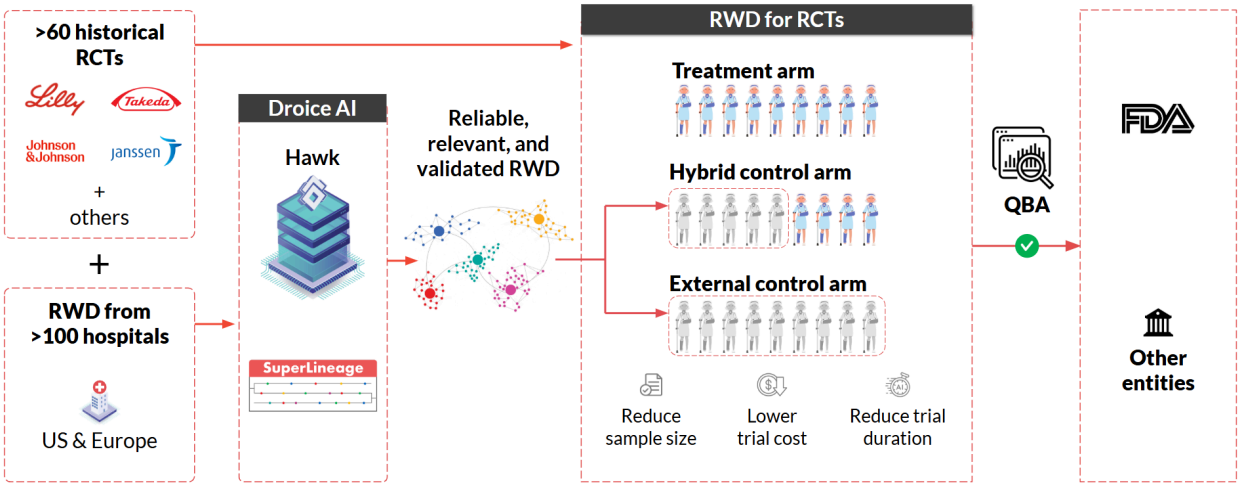

Figure 6. Droice Labs is developing approaches to leverage its technologies for scalable and reliable RWD to accelerate clinical trials using real world data (RWD) from electronic health records (EHR).

CONCLUSION

Using RWD in clinical trials for regulatory decision-making requires quantitative assessment of misclassification in RWD-derived study variables and its downstream impacts on study inferences to ensure that misclassification bias does not invalidate study findings. By modeling how misclassification influences effect estimates, QBA provides a framework for assessing whether inferences drawn using RWD are robust. Validation studies are essential for measuring the bias parameters that characterize misclassification across study variables so that the impacts of error on inferences can be accurately accounted for through QBA. While a standardized and reproducible validation process for diverse, messy, and voluminous RWD sources is very difficult to achieve using traditional manual approaches, the standardized and scalable approach to RWD validation enabled by SuperLineage allows for efficient bias parameter derivations to inform QBA with more extensive and accurate error profiles to achieve robust evidence when using RWD to accelerate clinical trials.

REFERENCES

1. Klonoff DC. The New FDA Real-World Evidence Program to Support Development of Drugs and Biologics. J Diabetes Sci Technol. 2020;14(2):345-349. doi:10.1177/1932296819832661
2. Lamberti MJ, Kubick W, Awatin J, McCormick J, Carroll J, Getz K. The Use of Real-World Evidence and Data in Clinical Research and Postapproval Safety Studies. Ther Innov Regul Sci. 2018;52(6):778-783. doi:10.1177/2168479018764662
3. Burcu M, Dreyer NA, Franklin JM, et al. Real-world evidence to support regulatory decision-making for medicines: Considerations for external control arms. Pharmacoepidemiol Drug Saf. 2020; 29: 1228–1235. https://doi.org/10.1002/pds.4975
4. Schneeweiss S, Brown JS, Bate A, Trifirò G, Bartels DB. Choosing Among Common Data Models for Real-World Data Analyses Fit for Making Decisions About the Effectiveness of Medical Products. Clin Pharmacol Ther. 2020;107(4):827-833. doi:10.1002/cpt.1577
5. Facile R, Muhlbradt EE, Gong M, et al. Use of Clinical Data Interchange Standards Consortium (CDISC) Standards for Real-world Data: Expert Perspectives From a Qualitative Delphi Survey. JMIR Med Inform 2022;10(1):e30363 https://medinform.jmir.org/2022/1/e30363. 2022;10(1):e30363. doi:10.2196/30363
6. Zhou X, Murugesan S, Bhullar H, et al. An Evaluation of the THIN Database in the OMOP Common Data Model for Active Drug Safety Surveillance. Drug Saf. 2013;36(2):119-134. doi:10.1007/s40264-012-0009-3
7. Abeysinghe R, Black A, Kaduk D, et al. Towards quality improvement of vaccine concept mappings in the OMOP vocabulary with a semi-automated method. J Biomed Inform. 2022;134:104162. doi:10.1016/j.jbi.2022.104162
8. Center for Drug Evaluation and Research, Center for Biologics Evaluation and Research, Real-World Data: Assessing Electronic Health Records and Medical Claims Data To Support Regulatory Decision-Making for Drug and Biological Products Guidance for Industry FINAL GUIDANCE.; 2024.

https://www.fda.gov/regulatory-information/search-fda-guidance-documents/real-world-data-assessing-electronic-health-records-and-medical-claims-data-support-regulatory
9. Wei WQ, Teixeira PL, Mo H, Cronin RM, Warner JL, Denny JC. Combining billing codes, clinical notes, and medications from electronic health records provides superior phenotyping performance. Journal of the American Medical Informatics Association. 2016;23(e1):20-27. doi:10.1093/jamia/ocv130
10. Knake LA, Ahuja M, McDonald EL, et al. Quality of EHR data extractions for studies of preterm birth in a tertiary care center: Guidelines for obtaining reliable data. BMC Pediatr. 2016;16(1). doi:10.1186/s12887-016-0592-z
11. Fox MP, MacLehose RF, Lash TL. Applying Quantitative Bias Analysis to Epidemiologic Data. Springer International Publishing; 2021. doi:10.1007/978-3-030-82673-4
12. Lash TL, Fox MP, MacLehose RF, Maldonado G, McCandless LC, Greenland S, Good practices for quantitative bias analysis, International Journal of Epidemiology, Volume 43, Issue 6, December 2014, Pages 1969–1985, https://doi.org/10.1093/ije/dyu149
13. Stang PE, Ryan PB, Racoosin JA, Overhage JM, Hartzema AG, Reich C, Welebob E, Scarnecchia T, Woodcock J. Advancing the science for active surveillance: rationale and design for the Observational Medical Outcomes Partnership. Ann Intern Med. 2010 Nov 2;153(9):600-6. doi: 10.7326/0003-4819-153-9-201011020-00010. PMID: 21041580.
14. Droice Labs Discusses SuperLineage for Real-World Data Traceability at FDA Listening Sessio,n https://finance.yahoo.com/news/droice-labs-discusses-superlineage-real-150000597.html

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:
Author Name: Mayur Saxena PhD
Company: Droice Labs
Email: mayur@droicelabs.com
Website: droicelabs.com