

The DIVERSE Tool: A Standardized Tool to Describe Data Source Diversity

Real-World Data (RWD) are generated through a variety of complex and heterogeneous processes that reflect the diversity of healthcare delivery systems, data capture methods, and patient populations. Understanding and accurately describing these processes is critical for ensuring the reproducibility and reliability of real-world evidence (RWE) generated from such data sources.

To address this need, the International Society for Pharmacoepidemiology (ISPE) Databases Special Interest Group (SIG) launched the **DIVERSE Initiative**. This initiative aims to provide a standardized, practical framework and tool to describe data source diversity systematically. By doing so, it supports researchers, regulators, and other stakeholders in better assessing the quality, applicability, and limitations of RWD sources.

The DIVERSE Framework and Tool

The DIVERSE framework was developed based on a scoping literature review and piloted by the ISPE Databases SIG to capture key attributes and characteristics of data sources. This structured approach facilitates transparent documentation of how RWD are created, processed, and maintained, enabling improved reproducibility of studies and comparability across different data sources.

The practical outcome of this initiative is an interactive tool designed to guide users through the characterization of their data sources based on the DIVERSE framework. This tool helps standardize descriptions and promotes consistent terminology across studies.

Access the DIVERSE Tool

The DIVERSE tool is available online and can be accessed here:

<https://docs.google.com/spreadsheets/d/1yRCmjb71smATl5bk1yJTNcy-o4HyKZfD/edit?gid=2146261533#gid=2146261533>

This spreadsheet provides a template for the standardized description of data sources, allowing for consistent and transparent reporting critical for advancing the use of RWD in healthcare research. Two example entries for a variety of data sources are provided as guidance for users. More examples will be provided in the final publication.

Contributors of the DIVERSE initiative: Romin Pajouheshnia, Rosa Gini (co-lead), Shahab Abtahi, Dimitri Bennett, Mehmet Burcu, Carlos Duran Salinas, Heidi Elkjær Sonne, Helga Gardarsdottir, Claudia Gulea, Gillian Hall, Sigal Kaplan, Tomas Lasys, Lin Li, Carolina Moriello, Anton Pottegård, Robert Platt, Giuseppe Roberto, Soko Setoguchi, Daniel Tsai, Xuerong Wen, Angelika Wientzek-Fleischmann, Lei Zhu