Researcher Difficulties Using Secondary Data Sources to Generate Real-World Evidence: Results from an Online Survey

RWD45

David Thompson, PhD

Founder & Principal Consultant, Rubidoux Research LLC

BACKGROUND & RATIONALE

- The past four decades have witnessed widespread proliferation of real-world data (RWD) sources, technological innovations to enable data linkages, refinements of analytic methods to limit bias and confounding, and articulation of real-world evidence (RWE) use cases by payers, regulators, and other health system stakeholders
- With all these developments, it is an open question to what extent generation of RWE from secondary RWD sources has become an easier endeavor—or more difficult
- The objective of this study was to gain insights into the degree of difficulty researchers have identifying, evaluating, and analyzing secondary RWD sources, such as claims and electronic health records, to generate RWE

SURVEY DESIGN & IMPLEMENTATION

- The anonymous online survey was fielded during the period February-May 2024
- The survey asked researchers to draw on their experiences conducting RWD analyses over the last five years to indicate how difficult they found each of the following elements, on average, using a 7-point Likert Scale (from 1 = "very difficult" through 7 = "very easy"):
  - ✓ Identifying fit-for-purpose RWD sources
  - ✓ Assessing the quality & completeness of the RWD sources
  - ✓ Identifying a rigorous study design
  - ✓ Identifying appropriate codes (eg, ICD, CPT, NDC, etc) and developing algorithms to select study patients
  - ✓ Identifying appropriate codes (eg, ICD, CPT, NDC, etc) and developing algorithms to select interventions of interest and assign patients to treatment groups
  - ✓ Identifying appropriate codes (eg, ICD, CPT, NDC, etc) and developing algorithms to specify patient covariates of interest (eg, comorbidities, concomitant medications)
  - ✓ Identifying appropriate codes (eg, ICD, CPT, NDC, etc) and developing algorithms to specify outcomes of care
  - ✓ Selecting the statistical methods
- A final item asked respondents to rank-order these items directly from most to least difficult

SURVEY RESPONDENT CHARACTERISTICS

- A total of 53 researchers completed the survey
- Most survey respondents worked for HEOR consultancies or life sciences companies (Figure 1) & the vast majority had performed at least six analyses of RWD in the past five years (Figure 2)

Figure 1. Distribution of Survey Respondents by Professional Affiliation

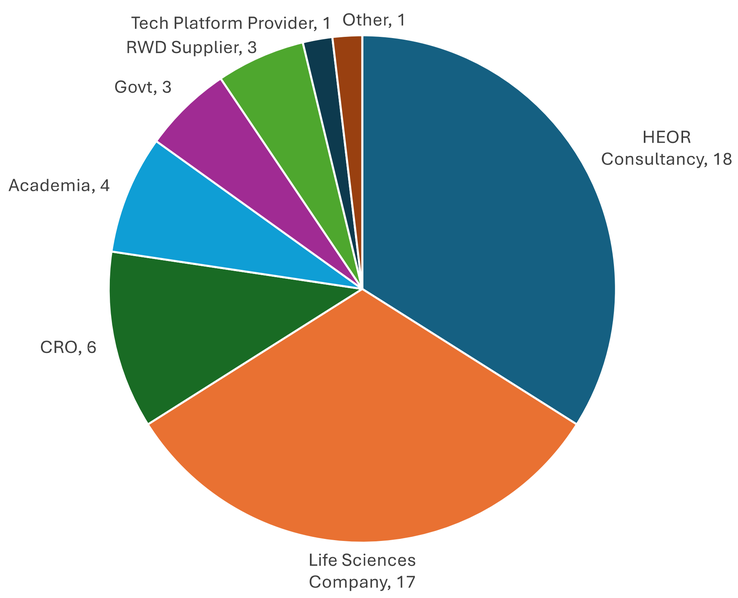
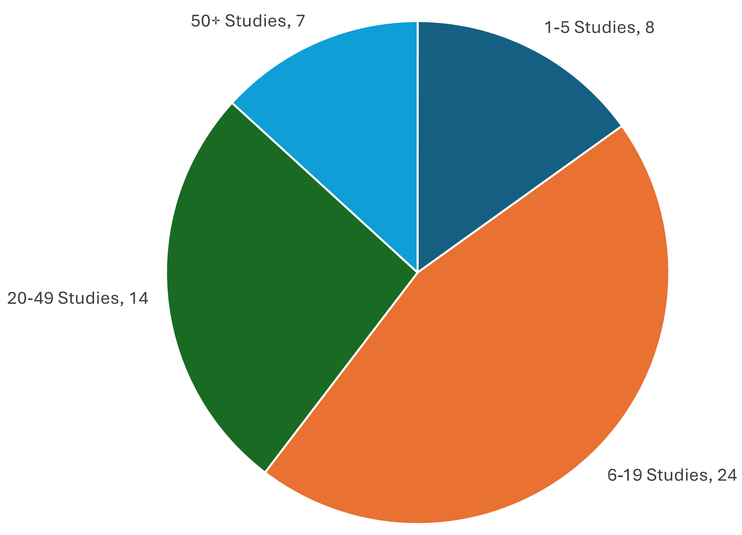


Figure 2. Distribution of Survey Respondents by RWE Experience*

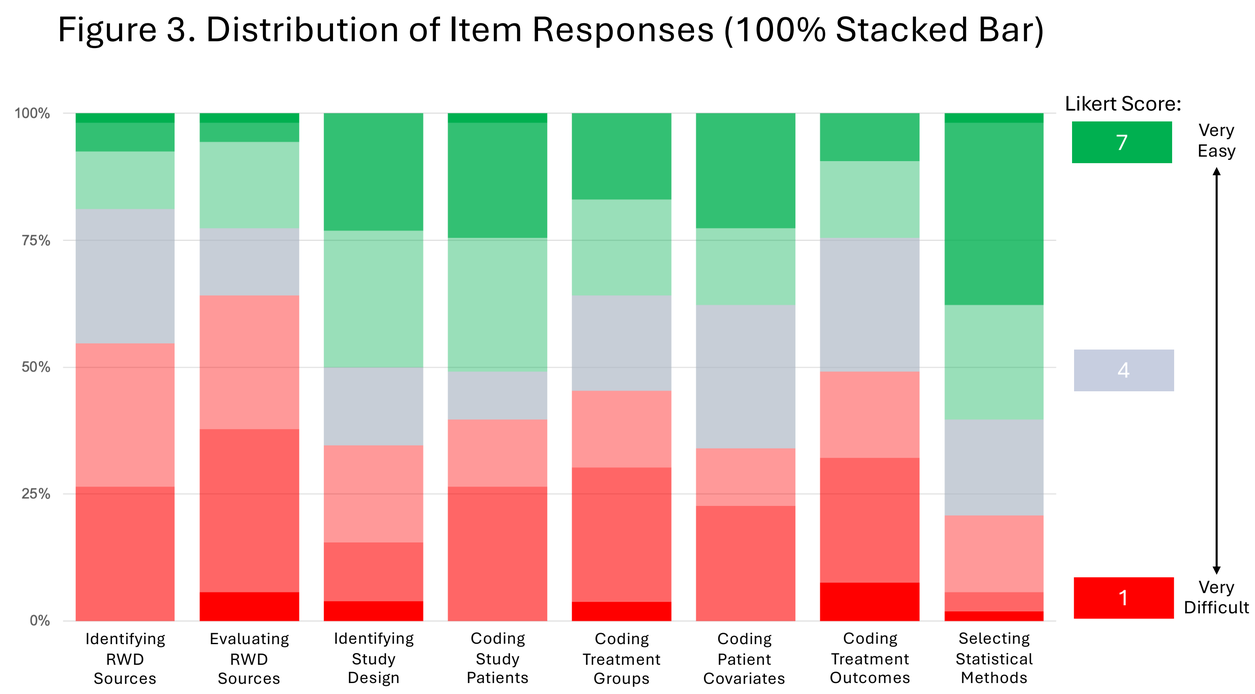


* Respondents' RWE experience defined as number of studies performed in past five years using RWD to generate RWE

OVERALL SURVEY FINDINGS

- Respondents directly ranked Evaluating RWD Sources as the most difficult task, followed (in descending order) by Identifying RWD Sources, Coding Treatment Outcomes, Coding Treatment Groups, Coding Study Patients, Identifying Study Design, Coding Patient Covariates, and Selecting Statistical Methods
- These rankings are consistent with the distribution of item responses summarized in the 100% stacked bar chart depicted in Figure 3

Figure 3. Distribution of Item Responses (100% Stacked Bar)



SUBGROUP ANALYSES

- Subgroup analyses RWD experience showed that experience matters, as there was a clear trend across all items towards higher Likert scores (signifying less difficulty) over the progressive experience categories
- Differences in Likert scores also were observed by reported professional affiliation, with those working in HEOR consultancies or CROs reporting it more difficult to identify and evaluate RWD sources, and less difficult to code the analytic files and select the statistical methods; the opposite was the case for those working in life sciences companies

CONCLUSIONS & IMPLICATIONS

- Researchers find identifying and evaluating secondary data sources to be the most vexing aspects of using RWD to generate RWE
- Developing code-based algorithms to create the analytic data files are somewhat less difficult, with selection of study design and statistical methods relatively straightforward in comparison
- These findings underscore the continuing need for tools and guidance to alleviate difficulties in using secondary data sources to generate RWE, particularly for inexperienced researchers

FULL REPORT

- This study was published in the September/October issue of *Value & Outcomes Spotlight*, available here:

AUTHOR CONTACT DETAILS

- Email: dave@rubidouxresearch.com
- LinkedIn: