Translating National Survey Data into Usable Pharmacoepidemiologic Evidence: Methodological Considerations from Costa Rica

Yu-Hua Fu[1]; Sofia Segura[2]; Danya M. Qato[3]

[1]Pharmaceutical Health Services Research, University of Maryland, Baltimore; [2]University of Costa Rica; [3]School of Pharmacy & School of Medicine, University of Maryland

Background: Self-reported medication data from national surveys provide invaluable insights into medication utilization; however, transforming these data into analytically tractable pharmacoepidemiologic measures requires deliberate and structured operationalization strategies.

Objectives: To develop a systematic framework to reconcile, standardize, operationalize self-reported medication data from wave 3 of the Spanish-language Costa Rican Longevity and Healthy Aging Study (CRELES) into measures suitable for epidemiologic analysis.

Methods: We developed a four-step framework to convert raw self-reported medication data into standardized analytic variables. First, medications reported under brand names or in Spanish were standardized to generic names by study team members (all pharmacists and one who is a native speaker of Spanish). A binary “first occurrence” indicator was created to identify newly appearing medications; unique codes were assigned to first occurrences and reused for subsequent identical entries to ensure consistency. Second, combination products were flagged and disaggregated into individual active ingredients whenever appropriate to enable ingredient-level analyses. Third, medications were classified by route of administration as systemic (oral or injectable), otic, ophthalmic, inhaled, topical, or unknown (those with multiple or unclear routes). Agents intended exclusively for otic, ophthalmic, cutaneous, or inhaled use were categorized as non-systemic. Finally, all coding underwent independent and iterative cross-validation for more than 10 cycles to ensure accuracy.

Results: Final study sample included 1851 older adults, of whom 85% reported use of medications (total of 7309 records). The raw data contained 417 unique medication names (including one unknown category). After reconciliation, entries were standardized to 333 unique generic ingredients, reducing variability by 20%. Fifty-one combination products were identified; combination products (e.g., levodopa and carbidopa) were retained as unified products. Despite systematic reconciliation and verification against national drug listings, some entries (e.g., certain herbal supplements) lacked sufficient information for further classification and thus we retained original coding. Medications were classified as systemic (98.4%), non-systemic (1.2%), or unknown (0.4%). The finalized dataset enabled derivation of reproducible ingredient-level exposures such as potentially inappropriate medication indicators.

Conclusions: This replicable framework, which will be available for public use, supports population medication safety research, particularly where administrative prescribing data are limited.