Supplementary Material

Adverse Events Reported With Therapies Targeting the CGRP Pathway During the First Six Months Post-Launch: A Retrospective Analysis Using the FDA Adverse Events Reporting System

Stephen D. Silberstein (ORCID: 0000-0001-9467-5567),[1] Shoshana Reshef,[2] Joshua M. Cohen,[3] Sanjay Gandhi,[2] Michael Seminerio,[4] Verena Ramirez Campos,[3] Yoel Kessler,[5] Stephen F. Thompson,[2] Andrew Blumenfeld[6]

[1]Jefferson Headache Center, Thomas Jefferson University, Philadelphia, PA, USA; [2]Epidemiology and Global Health and Economics Research, Teva Pharmaceutical Industries, West Chester, PA, USA; [3]Global Medical Affairs, Teva Pharmaceutical Industries, West Chester, PA, USA; [4]North American Medical Affairs, Teva Pharmaceutical Industries, West Chester, PA, USA; [5]Pharmacovigilance and Patient Safety, Teva Pharmaceutical Industries, West Chester, PA, USA; [6]Headache Center of Southern California, The Neurology Center, Carlsbad, CA, USA.

Corresponding author:

Stephen D. Silberstein

Email: Stephen.Silberstein@jefferson.edu

DATA CLEANING PROCESS

Advera's Evidex platform uses the following methodology to create a clean, deduplicated version of FAERS data to make analysis accurate and insightful.

FAERS Data Normalization

Data Extraction

Data validation steps are run to verify that all case reports had key identification fields filled in, including the: primaryid/ISR, caseid, drug sequence identification, and MedDRA AE terms. Case reports missing or containing malformed key identification fields are discarded. AE information is coded according to MedDRA® version 21.1.

Drug Mapping Steps

Both automated and manual steps are combined in mapping assignment of the raw drug data collected from each quarterly extract. The automation process matches drug details from each case report based on five data fields: FAERS raw drug name (drugname), route of administration (route), dose form (dose_form), and application number (nda_num) to a manually curated master reference set created and maintained by Evidex.

Entries in the reference set are composed of the same five fields linked to an Evidex drug identifier (drug_id). Cases within a quarterly extract that had drug details that exactly match those in the reference set are automatically mapped to Evidex drug identifiers.

Drug details in a case report within a quarterly extract that are not exact matches are sent to an analyst team for manual review and classification (manual curation). The analysts compare the raw *drugname, route, dose_form, and nda_num* fields in question against available prescribing information and link them to an existing or new drug_id. Completed, manually assigned drug mapping pairs are then added back to the "reference set" to improve future automation.

De-Duplication Process

In the FAERS database, the caseid field is the main identifier for individual patients with one or more reports. Each caseid can have multiple primaryid (or isr in LAERS) to represent different case versions. The first case report for a patient is typically tagged as the initial report (field: *i_f_code* = "I" or "Initial"), with subsequent reports tagged as follow-ups (field: i_f_code = "F" or "Follow-up"). In instances where there is more than one primaryid for the same *caseid*, each *caseid* is only counted once, with only the latest report version retained. A second de-duplication step is implemented to account for scenarios where duplicated case versions are not linked by the FDA processing logic to the same caseid. This is accomplished by implementing AEOLUS (Adverse Event Open Learning through Universal Standardization), an externally validated methodology of de-duplicating the FAERS database. This process captures duplicate cases that may have been listed with different caseid by evaluating four 'key' demographic fields (event date, age, gender, reporter country), a concatenated alphabetic ordered list of raw drugnames, and a concatenated alphabetic list of raw AEs (preferred terms).

These additional duplicate cases are counted only once in all subsequent analyses. At the end of the de-duplication process, the final standardized data set will be consolidated into a clean format to allow for analysis.