

Enabling data-driven clinical quality assurance: predicting Adverse Event reporting in clinical trials using machine learning

Drug Safety

Timothé Ménard¹ • Yves Barmaz¹ • Björn Koneswarakantha¹ • Rich Bowling¹ • Leszek Popko¹

¹ F.Hoffmann- La Roche AG

Correspondence: timothe.menard@roche.com

Electronic Supplementary Material #2 - List of features used for the final model

Feature identification	Description
VISIT_COUNT_CUMULATIVE	Number of study visits to date
MD_COUNT_CUMULATIVE	Cumulative count of concomitant medications prior a study visit
DISEASE_TYPE_malignancy	Disease type = malignancy
DISEASE_TYPE_lung disease other	Pulmonary disease other than malignancy
ADMIN_ORAL	Study drug administered orally
HX_COUNT_CUMULATIVE	Number of concomitant disease
AGEIC_POW	Age
WT_3V_SLOPE	Slope of the weight since last 3 visits
BP_LEVEL_VISIT_MOST_RECENT_UNKNOWN	Blood Pressure at the most recent visit was not reported
WT_VISIT_MOST_RECENT_LOG	Log of the weight reported at the most recent visit
BLINDING_Open Label	Study is open label (not blinded)
BPD_VISIT_MOST_RECENT	Blood Pressure Diastolic value (most recent visit)
PHASE_II	Study phase is II
TEMP_VISIT_MOST_RECENT	Body temperature (most recent visit)
BPS_VISIT_MOST_RECENT	Blood Pressure Systolic value (most recent visit)
ADMIN_SC	Drug administered subcutaneously
PHASE_I	Study phase is I

CONC_AGENTS	Study drug administered with concomitant agents/drugs
RANDOMIZED	Study is randomized
MOLECULE_CLASS_monoclonal antibody	Study drug - monoclonal antibody
PHASE_III	Study phase is III
SEX_Female	Female subject
ADMIN_IV	Drug administered intravenously
MOLECULE_CLASS_kinase inhibitor	Study drug - kinase inhibitor
MOLECULE_CLASS_sodium channel blocker	Study drug - sodium channel blocker
ADMIN_IM	Drug administered intramuscularly
ETHNIC_ORIGIN_WHITE	Subject ethnic origin = white
BP_LEVEL_VISIT_MOST_RECENT_Normal	Blood pressure reported at most recent visit was within normal range
BLINDING_Single Blind	Study blinding of the study = single blinded
ETHNIC_ORIGIN_OTHER	Subject ethnic origin = other
MOLECULE_CLASS_interleukin	Study drug - interleukin
DISEASE_TYPE_respiratory disease	Disease type = respiratory disease
BLINDING_Double Blind	Study blinding of the study = double blinded
DISEASE_TYPE_healthy participants	Study conducted on healthy subject (e.g. Phase I)
ETHNIC_ORIGIN_UNKNOWN	Subject ethnic origin = unknown (not reported)
MOLECULE_CLASS_nucleosid analog	Study drug - nucleoside analog
BP_LEVEL_VISIT_MOST_RECENT_Stage 1 Hypertension	Blood pressure reported at most recent visit was within the range of stage 1 hypertension
ETHNIC_ORIGIN_BLACK OR AFRICAN AMERICAN	Subject ethnic origin = black of african american
DISEASE_TYPE_infectious disease	Disease type = infectious disease
ETHNIC_ORIGIN_ASIAN	Subject ethnic origin = asian
DISEASE_TYPE_other	Disease type other than healthy participants, malignancies, autoimmune diseases, neurodegenerative diseases, respiratory diseases, skin disorders, lung diseases, infectious diseases
BP_LEVEL_VISIT_MOST_RECENT_Stage 2 Hypertension	Blood pressure reported at most recent visit was within the range of stage 2 hypertension
PHASE_IV	Study phase is IV

BLINDING_Observational	Observational study
ETHNIC_ORIGIN_MIXED ORIGIN	Subject ethnic origin = mixed
MOLECULE_CLASS_PPAR modulator	Study drug - Peroxisome Proliferator-Activated Receptor (PPAR) modulator
ADMIN_INTRAVITREAL	Study drug administered intravitreally
BP_LEVEL_VISIT_MOST_RECENT_Prehypertension	Blood pressure reported at most recent visit was within the range of pre-hypertension
DISEASE_TYPE_neurodegenerative disease	Disease type = neurodegenerative disease
MOLECULE_CLASS_glycine reuptake inhibitor	Study drug - glycine reuptake inhibitor
MOLECULE_CLASS_thiol protease inhibitor	Study drug - thiol protease inhibitor
ETHNIC_ORIGIN_AMERICAN INDIAN OR ALASKA NATIVE	Subject ethnic origin = american indian or alaska native
MOLECULE_CLASS_SMO antagonist	Study drug - SMO antagonist
MOLECULE_CLASS_anti-inflammatory	Study drug - anti-inflammatory