Predictive models for clopidogrel outcomes using prescription records and diagnosis codes

In Gu Lee[1], Samuel Kim[1], Mijeong Ban[1], Min Kim[1], and Jane Chiang[1]
[1] Cipherome Inc., San Jose, CA, U.S.A.

Background.

Clopidogrel, a P2Y12 inhibitor, is one of the most widely prescribed drugs in the U.S. Clopidogrel and aspirin, referred to as dual antiplatelet therapy (DAPT), are prescribed for coronary artery disease requiring stent placement, and is the standard of care after percutaneous coronary intervention (PCI). Among various risk factors impacting clopidogrel response, FDA's black box label warns that *CYP2C19* poor metabolizers may be at increased risk for treatment failure (TF)[1]. However, realistic testing options have been limited, making genetic testing difficult to implement[2]. This study aims to explore the medical history, especially prescription data and prior diagnosis, associated with clopidogrel outcomes. These factors will then be used to quickly evaluate and improve drug response.

Methods.

We used UK Biobank's hospital inpatient and primary care data, which includes prescription records. We included subjects who had prescription records of at least 180 clopidogrel tablets within one year of initial PCI. TF was defined as myocardial infarct (MI), cardiovascular death, transient ischemic attack (TIA), ischemic stroke, stent thrombosis, PCI, or coronary artery bypass grafting (CABG) within one year of the initial PCI and was identified using ICD-9/10 and OPCS-4 codes. After excluding subjects with missing data, inconsistent prescription records, and ambiguous clinical outcomes, we analyzed 8,744 subjects (1,826 with TF). We used multi-hot encoded features to represent the existence of clinical factors and the Random Forest algorithm to learn patterns in the features. We set aside 670 subjects (137 with TF) as test data. The test data were not included in the training procedure and used only for evaluation.

Results.

In TF prediction tasks with the prescription and diagnosis records prior to clopidogrel exposure, the area under receiver operating characteristics (AUROC) curves were 66.9% with ICD-9/10 data, 64.6% with OPCS-4 data, and 66.9% with prescription data. We limited data availability to one year of medical history data before clopidogrel exposure. In this setting, the AUROCs were 61.1% with ICD-9/10 data, 64.7% with OPCS-4 data, and 67.2% with prescription data. The results show that the performance in terms of AUROC does not degrade in experiments with OPCS-4 and prescription data even with only one year of data availability. This implies that the most recent medical history of a subject, which is less challenging than accessing extensive historical records, is sufficient in predicting outcomes.

As prescription data and diagnostic data are complementary to describing patients' overall health, we examined a fusion algorithm to concatenate all the features from ICD-9/1, OPCS-4, and prescription data. With the concatenated features, the AUROCs were 68.9% with the data prior to clopidogrel exposure and 67.5% with the most recent year of medical records.

Conclusion.

Drug response is triggered by an undetermined balance of genetic, clinical and external factors. For clopidogrel, genetics is a critical factor associated with treatment failure, as shown in other previous studies[3]. However, when genotyping results are unavailable, leveraging patients' medical history may result in a rapid assessment of treatment failure risk. The experimental results show that prescription records and diagnostic codes are an accurate reflection of patients' clinical risk factors which also contribute significantly to outcomes.

Future analyses will include incorporating the time element of the medical history data, fusion with genomic data, and categorizing the factors that contribute most to drug response to generate a risk assessment tool which can be implemented into the current standard of care.

Study limitations include self-selection for UKB study participants, the inherent inaccuracies associated with registry data, including inconsistent ICD reporting.

---

[1] Pan Y, Chen W, Xu Y, *et al*. Genetic polymorphisms and clopidogrel efficacy for acute ischemic stroke or transient ischemic attack. A systematic review and meta-analysis. *Circulation* 2017;135:21–33.

[2] Gower MN, Ratner LR, *et al*. Clinical Utility of CYP2C19 Genotype-Guided Antiplatelet Therapy in Patients at Risk of Adverse Cardiovascular and Cerebrovascular Events: A Review of Emerging Evidence. *Pharmgenomics Pers Med*. 2020;13:239-252.

[3] Galli M, Benenati S, Capodanno D, *et al*. Guided versus standard antiplatelet therapy in patients undergoing percutaneous coronary intervention: a systematic review and meta-analysis. *Lancet*. 2021 Apr 17;397(10283):1470-1483.