Objective:
To evaluate stability and clinical relatedness of nearest neighbors identified via vector representations for adverse events and drugs derived from global pharmacovigilance reporting patterns.

Background:
Pharmacovigilance requires identification of adverse event terms related to the same clinical condition, and solely relying on hierarchies in medical terminologies is often not sufficient. Similar challenges exist for the analysis of medicinal products. By utilizing advancements in machine learning UMC has developed vigiVec, yielding vector representations of MedDRA preferred terms and WHODrug active substances.

Methods
We train a Word2Vec Skipgram neural network model with 20 dimensions to predict co-reported drugs and adverse events on each report in VigiBase. The weights from the trained neural network are then used as the vector representations for each drug and adverse event.

Conclusions
The semantic representations of vigiVec are stable and show a high level of clinical relatedness. Data-driven identification of clinically related adverse events and drugs may complement existing medical hierarchies, supporting domain experts in pharmacovigilance.