

# Abstract

Network proximity is widely used to prioritize compound–disease associations, yet proximity Z-scores are sensitive to target set size, network density, and network construction parameters. While often interpreted as functional indicators, these metrics are susceptible to inferential instability, where ranking significance is driven by target-count-dependent null distributions rather than physical reachability.

Here, we systematically evaluate the robustness of proximity-based and influence-based metrics using a liver-expressed human protein–protein interaction network. Using two constituents of *Hypericum perforatum* as a controlled comparison, we show that proximity-based rankings are unstable and threshold-dependent. Although Quercetin (62 targets) achieves higher proximity Z-scores than Hyperforin (10 targets) in high-confidence networks, this significance is an artifact of the law of large numbers acting on the null distribution. In physical distance space, Hyperforin is closer to drug-induced liver injury (DILI) genes, a ranking that is stably captured by random walk–based influence propagation.

We quantify this disparity by calculating the average influence per target, revealing that Hyperforin achieves 3.7-fold greater efficiency in DILI-module perturbation than Quercetin. This efficiency advantage remains robust across topology-only and expression-weighted analyses, as well as bootstrap resampling and chemical similarity controls. Our results demonstrate that proximity Z-scores may lead to misleading prioritizations when target set sizes differ, and that influence-based propagation provides a more robust and theoretically consistent framework for network toxicology. More broadly, this work provides a methodological template for identifying and correcting metric artifacts in polypharmacological risk assessment.

**Keywords:** network propagation, proximity metrics, metric robustness, drug-induced liver injury, polypharmacology, Z-score confounding, perturbation efficiency.