Reference sets in pharmacovigilance should be designed carefully, as restrictive control choices might cause discrepant effects between methodologies in terms of both direction and order of magnitude, hindering fair comparative evaluation. There is a need for establishment of open and sizable benchmarks that include diverse controls to ensure transparency and limit the amount of bias added to the performance evaluation.