

Risk of Bias Assessment (Figure 1)

Patient Selection and Confounding

Summary: Nearly 50% of studies demonstrated high risk of bias in patient selection, with only about 20% showing low risk. This suggests prevalent issues with how participants were recruited and selected across the evidence base.

Tool/Index Test

Summary: Approximately 60% of studies exhibited high risk of bias in the index test domain, with only about 10% showing low risk, indicating widespread concerns about how diagnostic tools were applied and interpreted.

Reference Standard

Summary: The reference standard domain showed the most favorable profile, with about 60% of studies at low risk of bias and about 20% at high risk, suggesting relatively good quality in how the diagnostic “gold standard” was implemented.

Flow and Timing

Summary: Nearly 60% of studies showed high risk of bias in flow and timing, with only about 15% at low risk, indicating significant concerns about the sequence and intervals of test administration and analysis.

Applicability Assessment (Figure 2)

Population

Summary: Only about 10% of studies had no applicability concerns regarding population, with the vast majority having narrow eligibility criteria or including more complex patients than typical of community settings.

Test

Summary: Approximately 50% of studies had no applicability concerns regarding the test, though many others used tests not common in current practice or employed highly selected teams not representative of typical clinical settings.

Reference Standard

Summary: About 50% of studies had no applicability concerns for the reference standard, with remaining studies showing unclear DSM-5 diagnostic criteria or other reference standard issues limiting generalizability.

Outcomes

Summary: About 45% of studies had no applicability concerns regarding outcomes, with many others using surrogate outcomes that may not directly translate to clinical practice.

Setting

Summary: Only about 20% of studies had no applicability concerns regarding setting, with the vast majority conducted in care levels different from community settings, limiting their generalizability to routine practice.