

Study-level QUADAS-2 risk-of-bias and applicability judgments

Study	Risk of bias				Applicability concerns			Overall risk of bias
	Patient selection	Index test	Reference standard	Flow and timing	Patient selection	Index test	Reference standard	
Huang et al. 2024 [14]	H	L	U	U	L	L	L	H
Ye et al. 2025 [15]	U	L	H	L	L	L	L	H
Lu et al. 2021 [16]	H	H	H	H	L	H	L	H
Lu et al. 2023 [29]	H	H	H	H	L	H	L	H
Seth et al. 2023 [24]	U	L	L	L	L	L	L	U
Jiang et al. 2022 [17]	H	H	L	H	L	L	L	H
Shu et al. 2024 [4]	H	H	H	H	L	H	L	H
Kim et al. 2020 [12]	H	L	L	H	L	L	L	H
Fei et al. 2021 [13]	H	L	H	H	L	L	L	H
Lin et al. 2024 [25]	H	H	U	H	L	H	L	H
Park et al. 2025 [26]	H	L	U	U	L	L	L	H
Zeng et al. 2022 [18]	H	H	H	H	L	L	L	H
Zhao et al. 2020 [23]	H	L	U	H	L	L	L	H
Gupta et al. 2020 [27]	H	H	H	H	H	L	H	H
Sharma et al. 2022 [17]	L	L	L	L	H	L	H	L
Zhu et al. 2021 [19]	H	H	H	H	L	H	L	H
Hilmi et al. 2023 [21]	H	L	H	U	L	L	L	H
Huang et al. 2020 [28]	H	L	H	H	H	L	H	H
Brehm et al. 2023 [22]	H	L	H	H	H	L	H	H

*Judgments: L = low; U = unclear; H = high. Overall risk of bias was rated high when any risk-of-bias domain was high, unclear when no domain was high but at least one was unclear, and low when all four risk-of-bias domains were low. Applicability judgments do not contribute to the overall risk-of-bias classification.

QUADAS-2 domains were assessed independently by two reviewers; disagreements were resolved by consensus and, when required, consultation with the research supervisor.