

THE PERFORMANCE OF THE BV SCORE FOR DIFFERENTIATING BETWEEN BACTERIAL AND VIRAL INFECTION IS ROBUST TO METHODOLOGICAL ALTERNATIVES IN BUILDING THE REFERENCE STANDARD

Cihan Papan*, Alberto Argentiero*, Susanna Esposito#, Tobias Tenenbaum#

*contributed equally

#contributed equally

Background/Objectives: A host-protein score (BV score) based on TNF-related apoptosis-induced ligand (TRAIL), interferon gamma-induced protein-10 (IP-10), and C-reactive protein (CRP) was shown to accurately differentiate bacterial from viral infections. We assessed the impact of reference standard methodology on performance results.

Methods: Sub-analysis of children selected randomly from AutoPilot-Dx study. Five reference standards were generated based on two separate adjudication processes, where independent clinical experts were provided with comprehensive patient data (except BV score). Adjudicators were either 'blinded' or 'unblinded' to CRP and procalcitonin. The reference standards were built as follows: (1) Majority unblinded – $\geq 2/3$ experts assigning same classification (bacterial or viral) with confidence $>90\%$ or all with confidence $\geq 70\%$; (2) Unanimous unblinded – all experts assigning same classification with confidence $>90\%$ cases; (3) Majority blinded – (1) with blinded experts; (4) Unanimous blinded – (2) with blinded experts; (5) All inclusive – indeterminate cases from majority blinded cohort were subjected to additional adjudications until majority attained. BV score performance was compared to each reference standard.

Results: There were no changes of viral to bacterial reference standard outcomes or vice versa comparing blinded vs. non-blinded. Performance was comparable across the 5 reference standards (Table).

Conclusions: BV score performance is robust to alternative reference standards.

Table: Performance across different cohorts

	All-inclusive blinded	Majority blinded	Unanimous blinded	Majority unblinded	Unanimous unblinded
n	270	224	125	231	137
Sensitivity %	90.0 (78.6-100.0)	93.3 (79.0-100.0)	100.0 (100.0-100.0)	100.0 (100.0-100.0)	100.0 (100.0-100.0)
Specificity %	82.2 (77.1-87.4)	85.8 (80.8-90.8)	91.9 (86.7-97.0)	88.0 (83.4-92.7)	94.9 (90.9-98.9)
Positive predictive value %	41.5 (29.2-53.8)	34.1 (19.0-49.3)	40.0 (11.9-68.1)	46.5 (31.0-62.0)	57.1 (27.5-86.8)
Negative predictive value %	98.3 (91.6-100.0)	99.4 (93.2-100.0)	100.0 (94.1-100.0)	100.0 (94.4-100.0)	100.0 (95.6-100.0)
Equivocal %	9.6	8.5	6.4	8.2	8.0