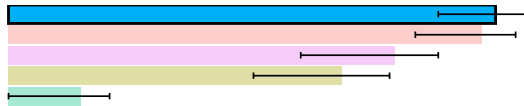


Neonatal Causes of Death, <28 Days (n=477)

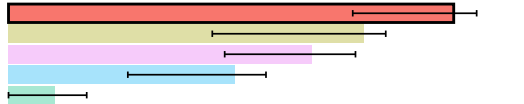
Infections
(n=99, 21%)
[Range=0.73]

InSilicoVA: 0.86 (0.76,0.91)
GPT-5: 0.84 (0.72,0.89)
InterVA-5: 0.68 (0.52,0.76)
GPT-4: 0.59 (0.43,0.67)
GPT-3.5: 0.13 (0,0.18)



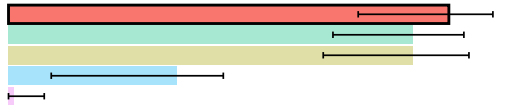
Birth asphyxia and
birth trauma
(n=103, 22%)
[Range=0.70]

GPT-5: 0.78 (0.61,0.83)
GPT-4: 0.63 (0.36,0.67)
InterVA-5: 0.54 (0.38,0.61)
InSilicoVA: 0.40 (0.21,0.45)
GPT-3.5: 0.08 (0,0.14)



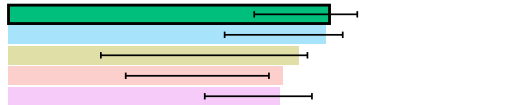
Prematurity and
low birthweight
(n=73, 15%)
[Range=0.77]

GPT-5: 0.78 (0.62,0.85)
GPT-3.5: 0.71 (0.57,0.8)
GPT-4: 0.71 (0.55,0.81)
InSilicoVA: 0.30 (0.08,0.38)
InterVA-5: 0.01 (0,0.06)



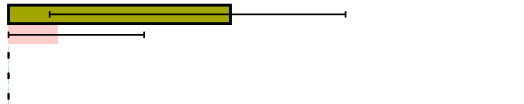
Stillbirth
(n=172, 36%)
[Range=0.09]

GPT-3.5: 0.57 (0.43,0.61)
InSilicoVA: 0.56 (0.38,0.59)
GPT-4: 0.51 (0.16,0.53)
GPT-5: 0.48 (0.21,0.46)
InterVA-5: 0.48 (0.35,0.53)



Ill-defined
(n=23, 5%)
[Range=0.39]

GPT-4: 0.39 (0.07,0.59)
GPT-5: 0.09 (0,0.24)
InterVA-5: 0.00 (0,0)
InSilicoVA: 0.00 (0,0)
GPT-3.5: 0.00 (0,0)



0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0

Partial Chance Corrected Concordance (PCCC)

■ GPT-5 ■ GPT-4 ■ GPT-3.5 ■ InSilicoVA ■ InterVA-5 — 95% Confidence Interval (Bootstrap, 1000 Resamples)