

Supplementary Table 2. Information of the manual verification workflow.

Manual Verification	Label Issue Module (n=3767)	Outlier Module (n=894)
κ (95% CI) at Stage 1	0.34 (0.31-0.36)	0.46 (0.40-0.52)
Consensus at Stage 1 (%)	2262 (60.05%)	673 (75.28%)
Samples referred to Reviewer C at Stage 2 (%)	1505 (39.95%)	221 (24.72%)
Majority consensus at Stage 3	1375	221
Three-way disagreement samples at Stage 3	130	0

Note: Stage 1 κ was calculated from independent assessments by Reviewer A and Reviewer B before adjudication. For Label Issue verification, Reviewer C independently assessed the 1505 cases showing initial discordance between Reviewers A and B. Majority consensus at Stage 3 refers to the majority opinion among the three reviewers. In samples where all three reviewers disagreed, the opinion of Reviewer C was adopted. For Outlier verification, Reviewer C independently assessed the 221 discordant cases. The Outlier module used a binary verification protocol, precluding three-way disagreement. Abbreviations: κ , Cohen's Kappa; CI, confidence interval.