EVALUATION OF A READING PANEL'S PERFORMANCE USING WORLD HEALTH ORGANIZATION CHEST RADIOGRAPH INTERPRETATION METHODOLOGY AMONG CHILD PNEUMONIA CASES IN A PNEUMOCOCCAL CONJUGATE VACCINE EFFECTIVENESS STUDY FROM SYLHET, BANGLADESH

Syed Jafar Raza Rizvi, Nazma Begum, Salahuddin Ahmed, Arun Dutta Roy, Nabidul Haque Chowdhury, Ahad Mahmud Khan, Lawrence H. Moulton, William Checkley, Abdullah H Baqui, Eric D. McCollum

for the Projahnmo Study Group: a partnership of Johns Hopkins University with the Bangladesh Ministry of Health and Family Welfare; Child Health Research Foundation, Dhaka, Bangladesh; icddr,b, Dhaka, Bangladesh; and Shimantik, Bangladesh

Introduction

We evaluated the performance of a panel to read chest radiographs (CXR) using World Health Organization (WHO) interpretation methodology in the context of a pneumococcal conjugate vaccine (PCV) effectiveness study conducted in rural Sylhet, Bangladesh.
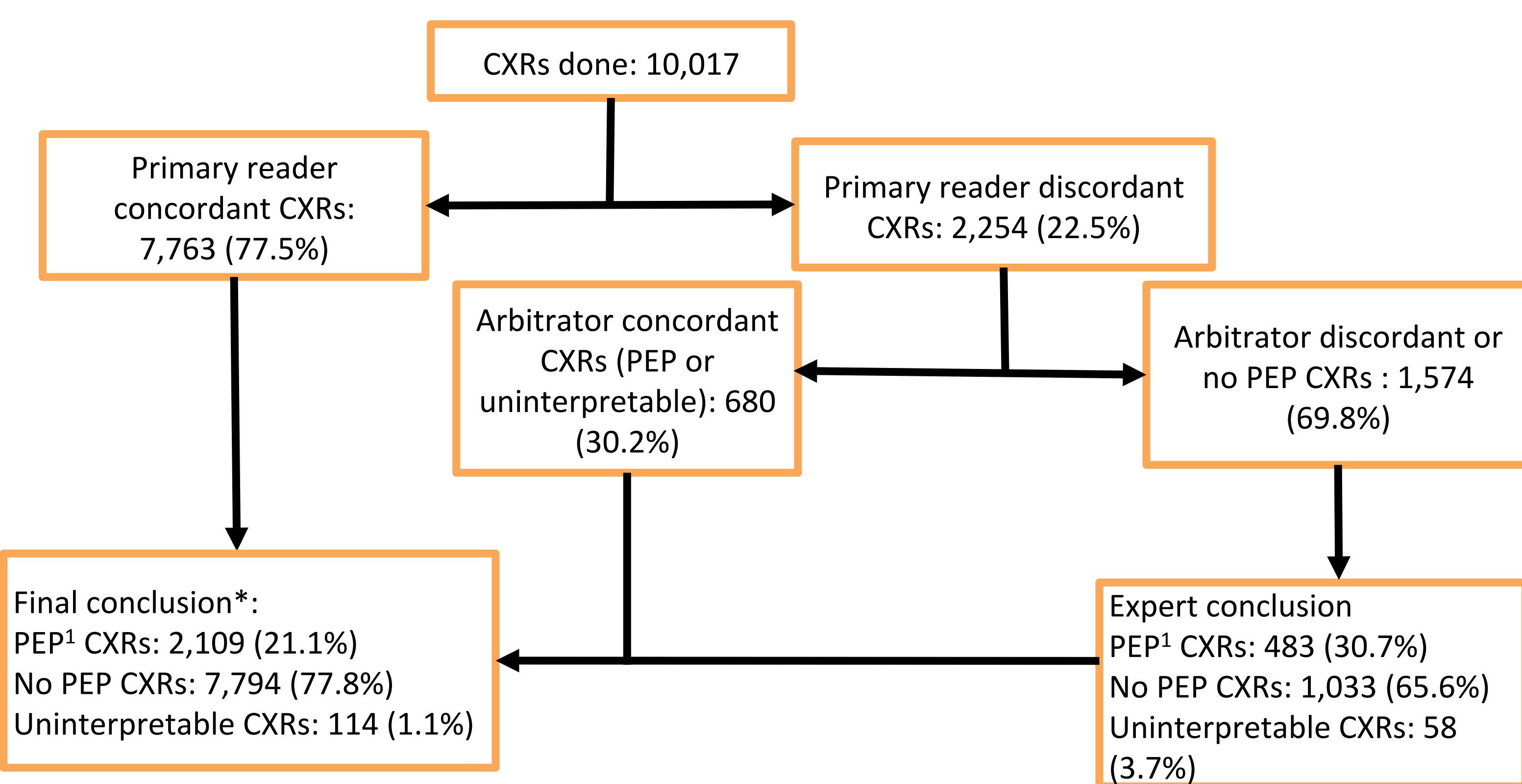
Methods

Eight physicians (CXR panel) were standardized to WHO CXR interpretation methodology and read CXRs of children 3-35 months old from May 2015 to October 2017. Each CXR was randomized to two primary readers masked to clinical data. If the readings of two primary readers were discordant for CXR interpretability or the presence or absence of primary endpoint pneumonia (PEP), then the image was sent to another randomly selected reader to adjudicate (arbitrator). If the arbitrator's interpretation disagreed with both primary readers, or concluded no PEP, then a masked expert reader established the final conclusion. The expert reader also conducted blinded quality control (QC) on 20% of CXRs. We evaluated primary reader agreement and expert QC agreement by percentage, unadjusted kappa, and a kappa adjusted for prevalence and bias.

Projahnmo study site
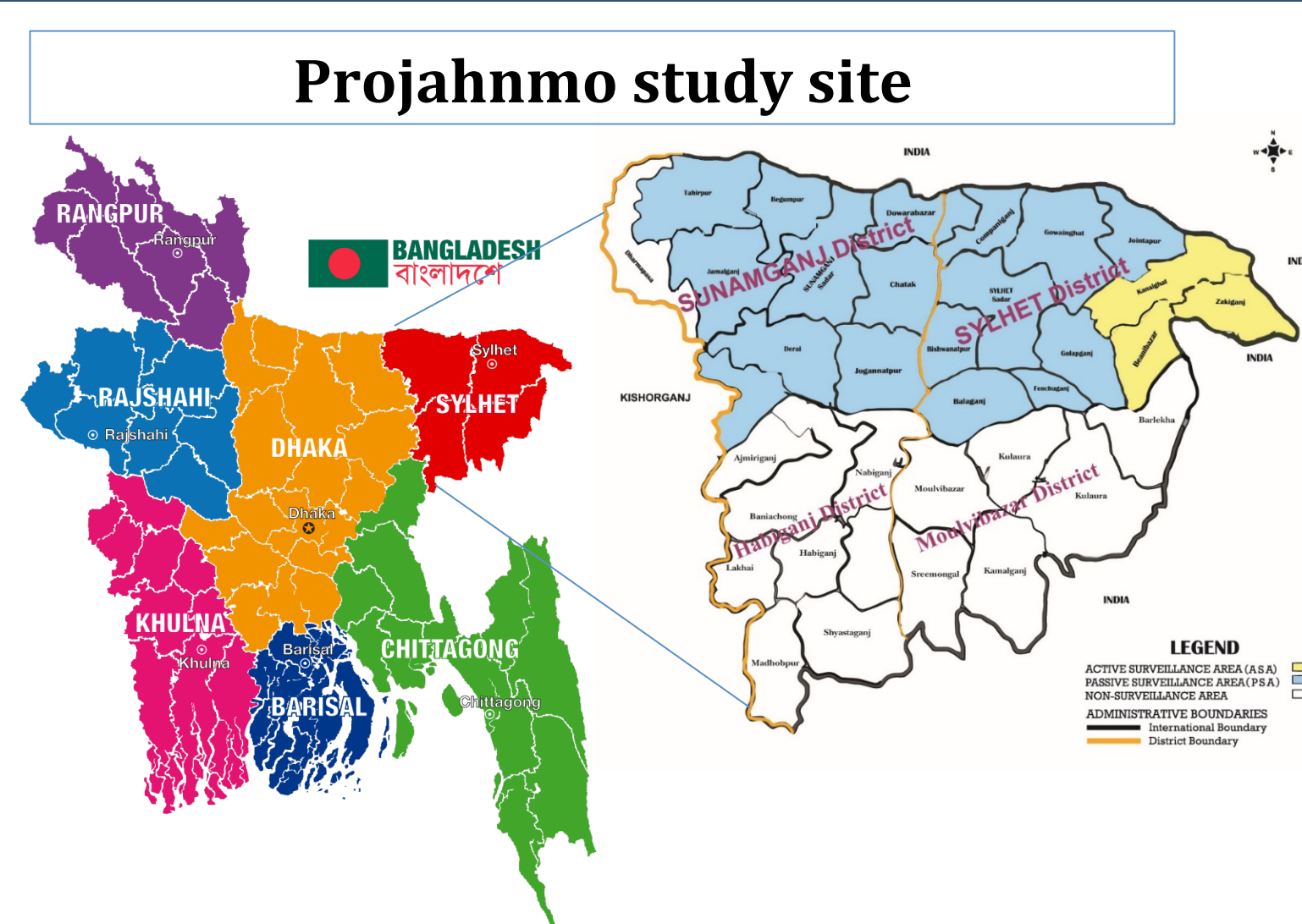
Use of X-ray machine in Hospital

Figure 1: Outline of the interpretation process for CXRs of children aged 3-35 months

CXRs done: 10,017

Primary reader concordant CXRs: 7,763 (77.5%)

Primary reader discordant CXRs: 2,254 (22.5%)

Arbitrator concordant CXRs (PEP or uninterpretable): 680 (30.2%)

Arbitrator discordant or no PEP CXRs : 1,574 (69.8%)

Final conclusion*:
PEP[1] CXRs: 2,109 (21.1%)
No PEP CXRs: 7,794 (77.8%)
Uninterpretable CXRs: 114 (1.1%)

Expert conclusion
PEP[1] CXRs: 483 (30.7%)
No PEP CXRs: 1,033 (65.6%)
Uninterpretable CXRs: 58 (3.7%)
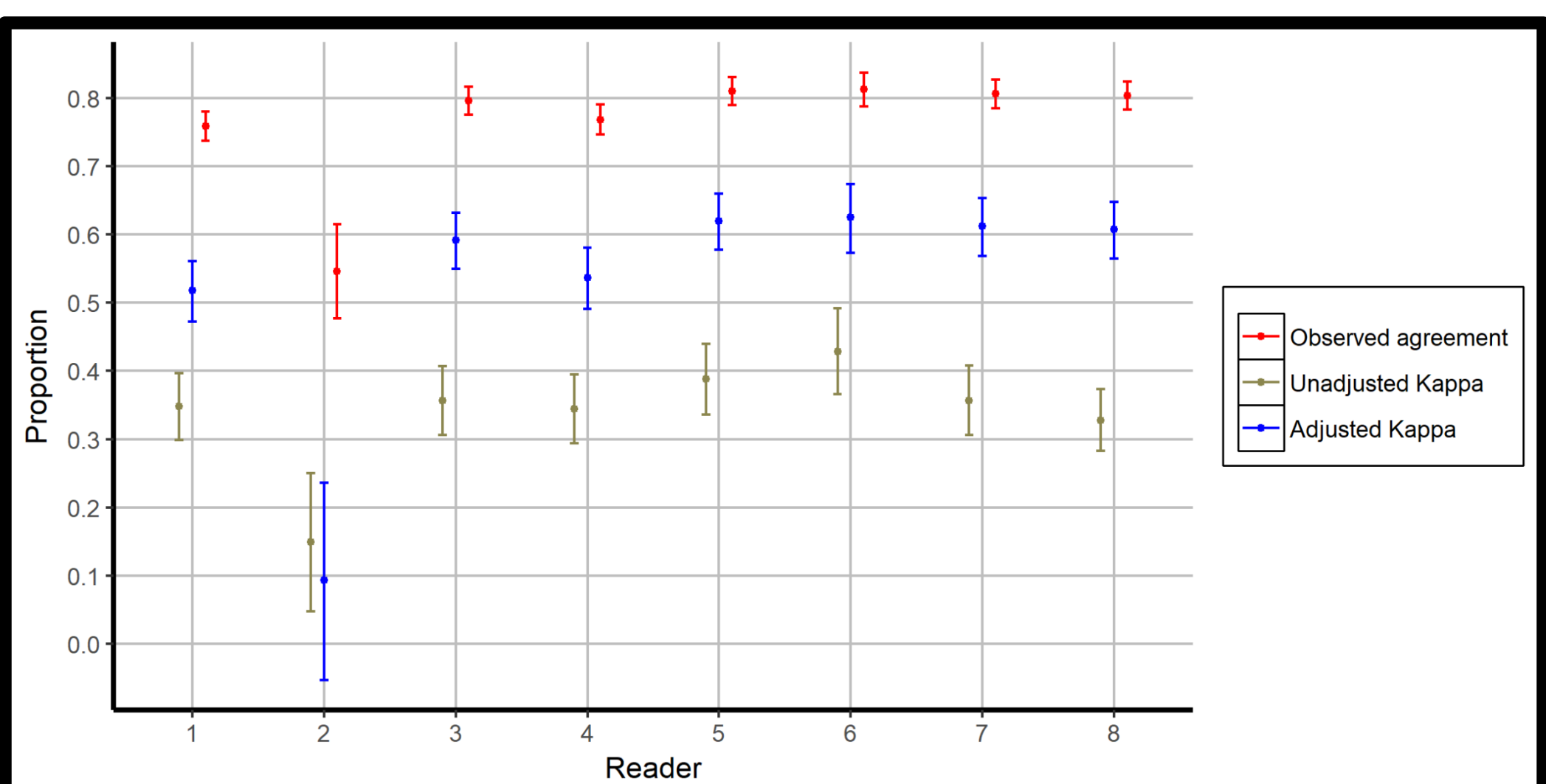
Note: PEP[1]: WHO primary endpoint pneumonia.

Results

Figure 2a: Inter-reader CXR agreement for uninterpretable vs interpretable CXRs among the 8 individual primary CXR readers
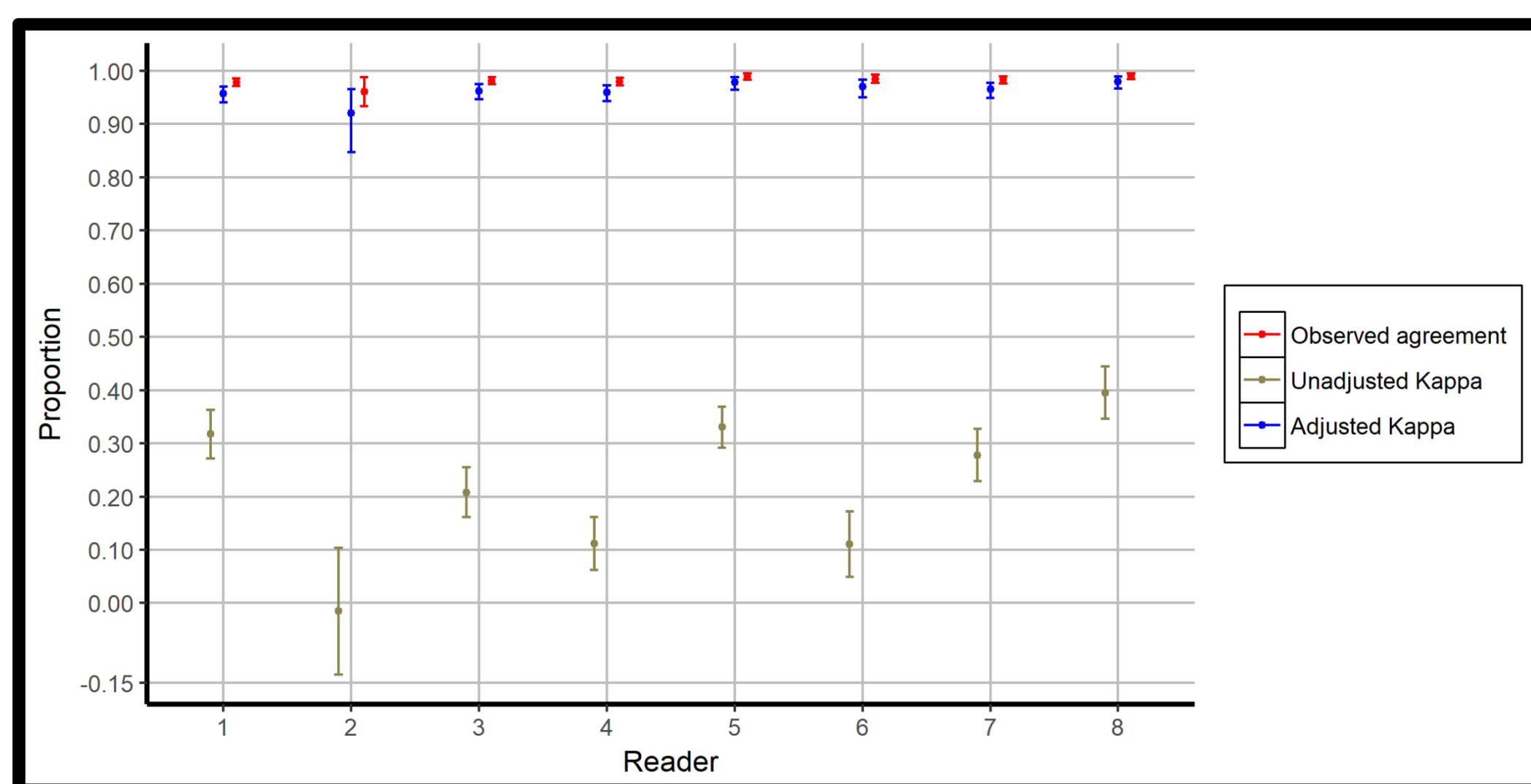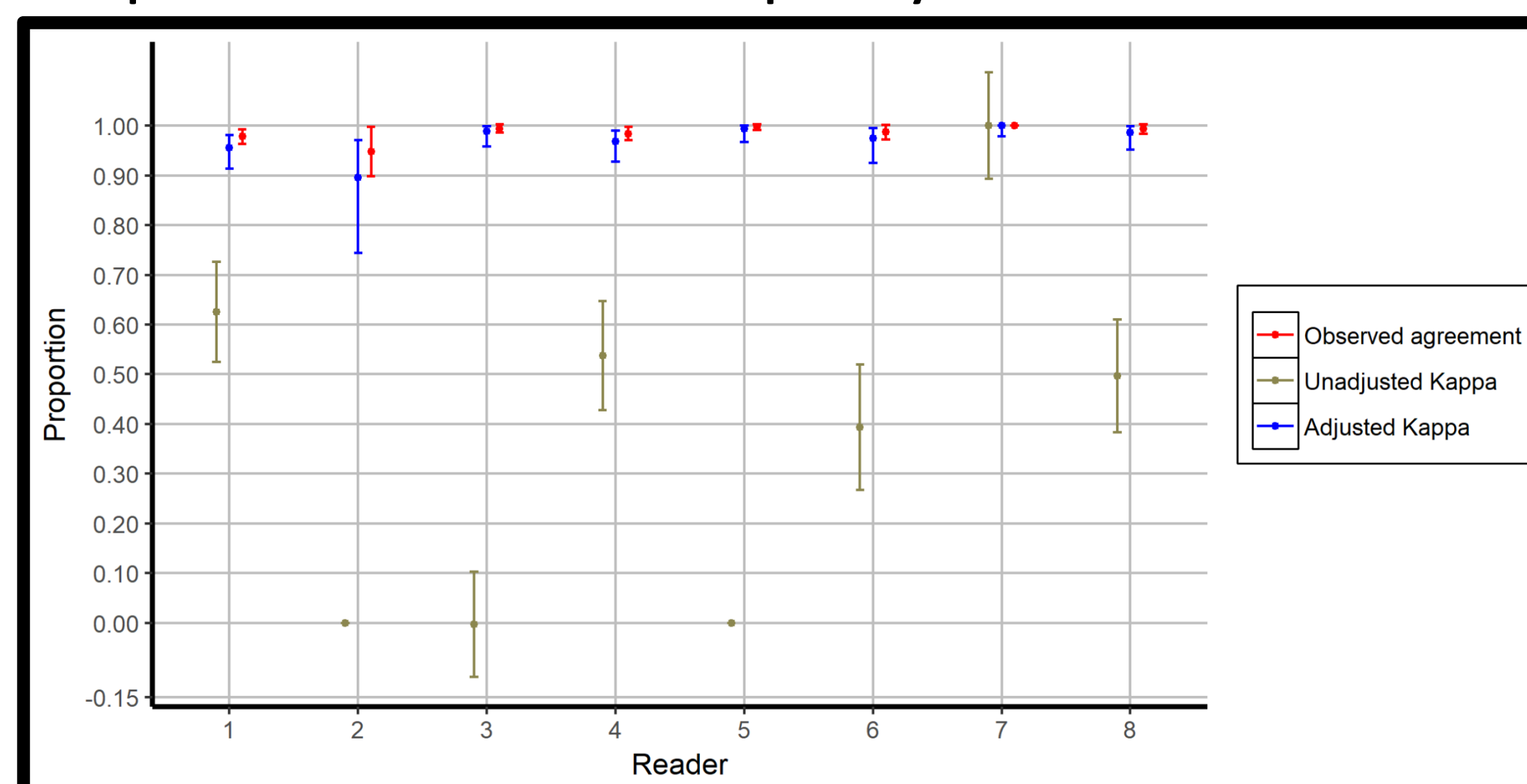
Observed Agreement : 96%-98%
Kappa: 0.11-0.39
Adjusted Kappa: 0.92-0.97

Figure 2b: Inter-reader CXR agreement for WHO PEP vs no PEP CXRs for 8 individual primary CXR readers
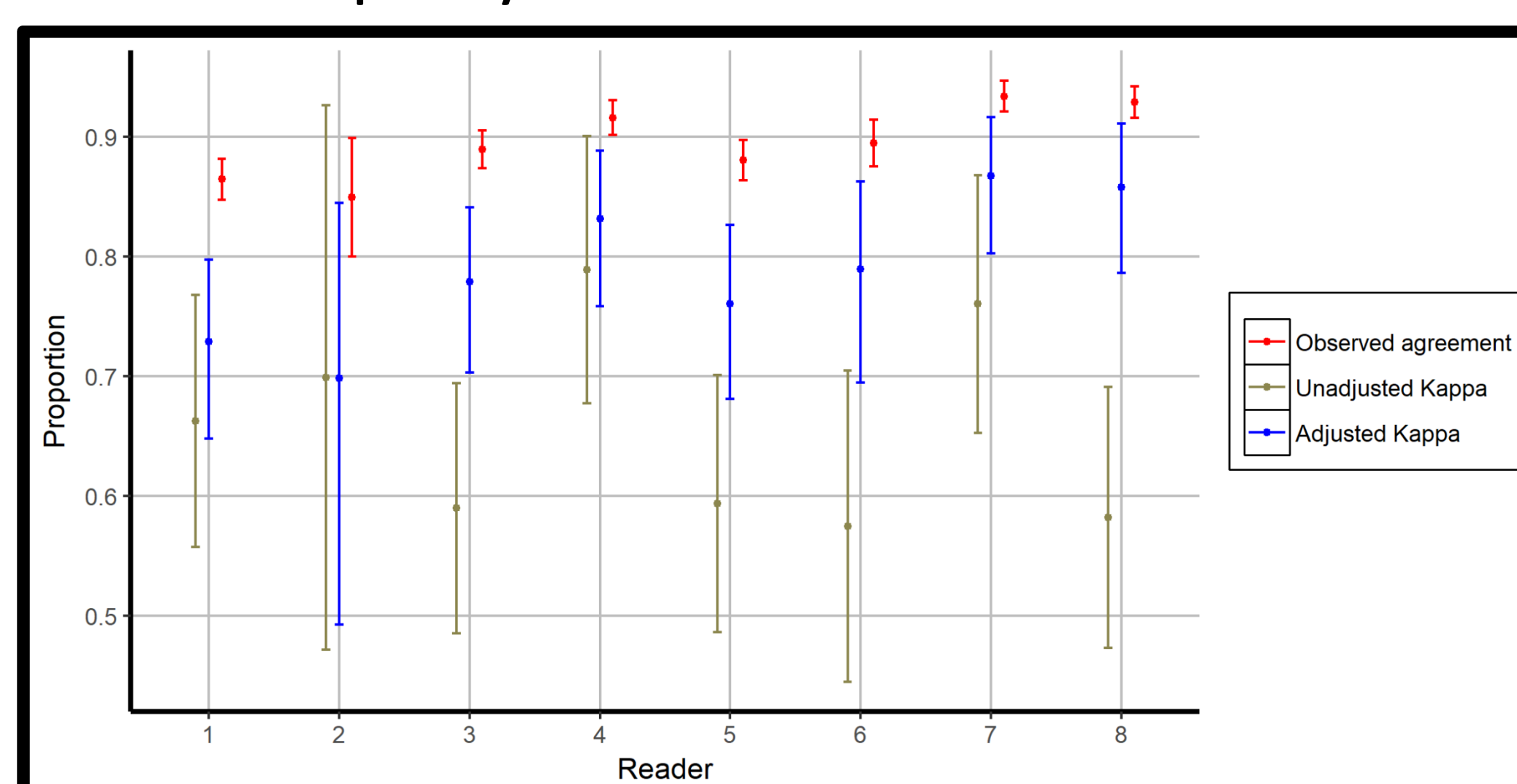
Observed Agreement : 54%-81%
Kappa: 0.14-0.43
Adjusted Kappa: 0.10-0.62

Figure 3a: Intra-reader CXR agreement for uninterpretable vs interpretable CXRs for 8 individual primary CXR readers

Observed Agreement : 94%-99%
Kappa: 0.39-0.62
Adjusted Kappa: 0.89-0.99

Figure 3b: Intra-reader CXR agreement for WHO PEP vs no PEP CXRs for 8 individual primary CXR readers

Observed Agreement : 85%-93%
Kappa: 0.57-0.78
Adjusted Kappa: 0.70-0.86

Table 1: Summary primary observer agreement for individual conclusions

Characteristic	Number of Observations	Primary CXR reader interpretations					
		n (%)	Expected agreement	Observed agreement	Unadjusted kappa	95% CI	Adjusted kappa
Uninterpretable	10,017	29 (0.3)	0.98	0.98	0.25	(0.23-0.27)	0.97
Rotated*	21	2 (9.5)	0.86	0.72	0.49	(0.07-0.91)	0.71
Blurry*	21	7 (33.3)	0.76	0.50	0.52	(0.09-0.95)	0.52
Over penetrated*	21	0 (0)	0.95	0.95	0.00	(0.0-0.0)	0.90
Under penetrated*	21	3 (14.3)	0.90	0.69	0.69	(0.26-1.12)	0.81
Clipped image*	21	5 (23.8)	0.95	0.61	0.88	(0.45-1.3)	0.90
Any PEP	9,818	973 (9.9)	0.79	0.67	0.35	(0.33-0.37)	0.58
Air bronchogram*	699	15 (2.1)	0.87	0.85	0.19	(0.11-0.26)	0.75
Silhouette sign*	699	362 (51.8)	0.72	0.55	0.38	(0.31-0.46)	0.44
Size criteria*	699	365 (52.2)	0.65	0.58	0.16	(0.09-0.23)	0.29
Pleural fluid only PEP	973	46 (4.7)	1.00	0.91	1.00	(0.94-1.06)	1.00
3-11 months old	5,171	496 (9.6)	0.79	0.68	0.35	(0.32-0.38)	0.58
12-23 months old	3,078	314 (10.2)	0.78	0.66	0.34	(0.30-0.37)	0.56
24-35 months old	1,569	163 (10.4)	0.79	0.67	0.37	(0.32-0.41)	0.58
Right sided PEP	9,818	669 (6.8)	0.82	0.73	0.32	(0.30-0.34)	0.64
Left sided PEP	9,818	261 (2.7)	0.92	0.87	0.35	(0.33-0.37)	0.84
Bilateral PEP	7,069	59 (0.8)	0.97	0.96	0.37	(0.35-0.4)	0.95

Note(*): Variables were included in the data form from June 2016.

Table 2: Overall CXR panel performance versus expert reference[1]

	Sensitivity	Specificity	Positive predictive value	Negative predictive value
N=1,652	232/300 (77.3%)	1,303/1,352 (96.3%)	232/281 (82.5%)	1,303/1,371 (95.0%)

[1]20% random sample of all chest radiographs.
[2]Uninterpretable CXRs excluded: May/June, 2; September, 1; October, 3; December, 3; May 2016, 1; July 2016, 1; October, 1; November, 1; January 2017, 1; February, 2; March, 1; May, 1; August, 1; September, 2.

Table 3: Inter-rater agreement between CXR panel and expert reference for CXR PEP vs no PEP[1]

	Agreement, % (n/N)	Expected agreement, %	Kappa[2]	Adjusted kappa[2,3]
N=1,652	92.9% (1,535/1,652)	71.0%	0.75	0.85

[1]20% random sample of all chest radiographs.
[2]Inter-rater agreement per the scale: <0, poor; 0.01-0.19, slight; 0.20-0.39, fair; 0.40-0.59, moderate; 0.60-0.79, substantial; 0.80-1.0, perfect.
[3]Prevalence-adjusted, bias-adjusted kappa statistic.

Conclusion

Primary reader performance and expert QC results suggest CXR interpretations used to analyze PCV effectiveness in rural Bangladesh meet WHO standards.