

Data Supplement S1. Supplemental material

Table S1: Participant Demographics

Variable	Medical Student n=20	Post-graduate Trainee n=40	Faculty n=23
Country, no. in Canada (%)	20 (100)	23 (57.5)	19 (82.6)
Academic center, no. (%)	20 (100)	40 (100)	22 (95.7)
Post-graduate year in training, no. (%)			
PGY 1		10 (25.0)	
PGY 2		7 (17.5)	
PGY 3	N/A	8 (20.0)	N/A
PGY 4		4 (10.0)	
PGY 5		9 (22.5)	
PGY 6		3 (7.5)	
Number of years in practice, no. (%)			
0 – 5			11 (47.8)
6 – 10			6 (26.1)
11 – 15	N/A	N/A	1 (4.3)
16 – 20			3 (13.0)
> 20			2 (8.7)
Most recent post-graduate training, no. (%)			
Pediatrics		16 (40.0)	2 (8.7)
Emergency Medicine	N/A	12 (30.0)	6 (26.1)
Pediatric Emergency Medicine		12 (30.0)	15 (65.2)
Radiology elective greater than 1 month in duration completed, no. (%)	0	15 (37.5)	13 (59.1)

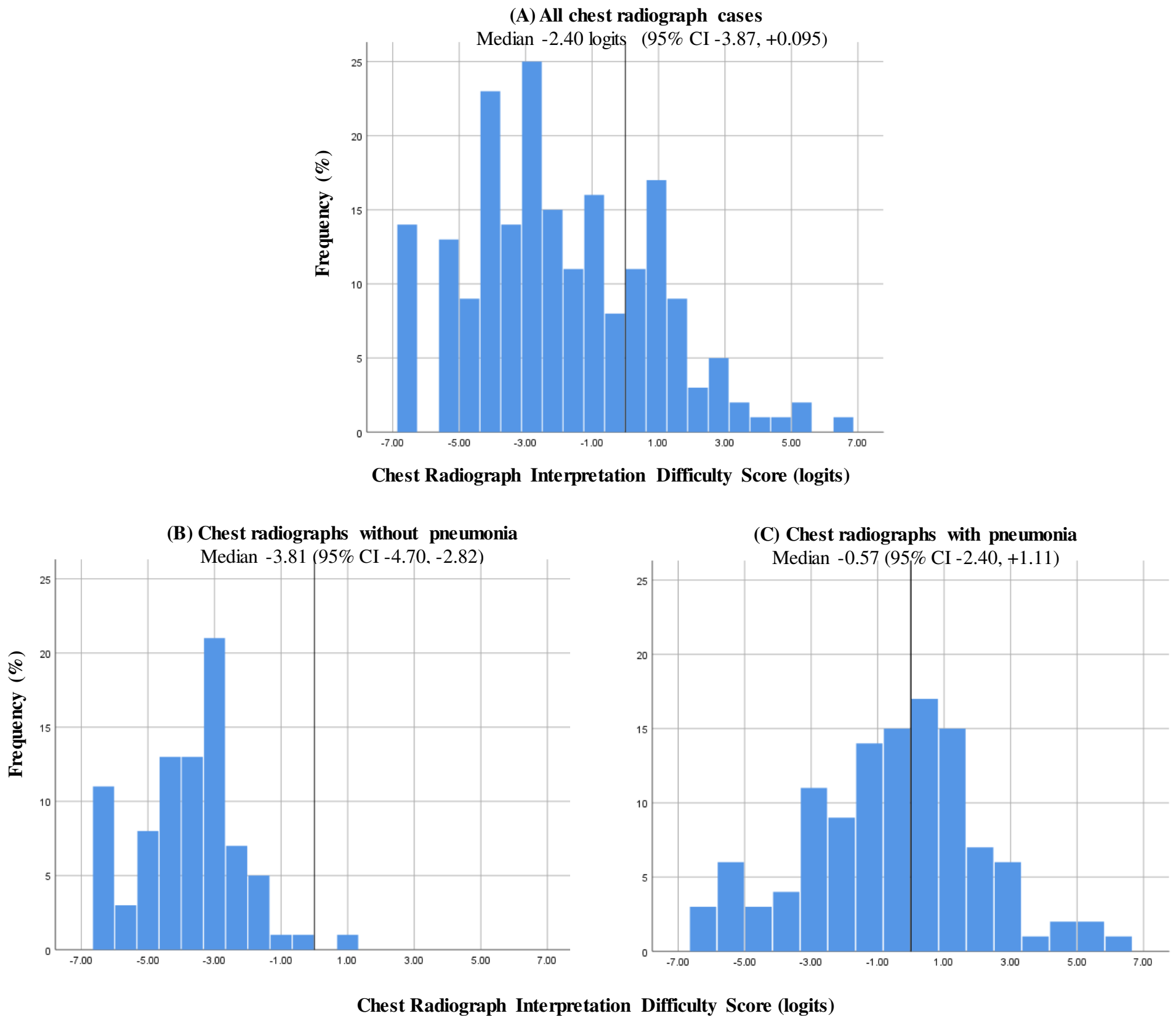
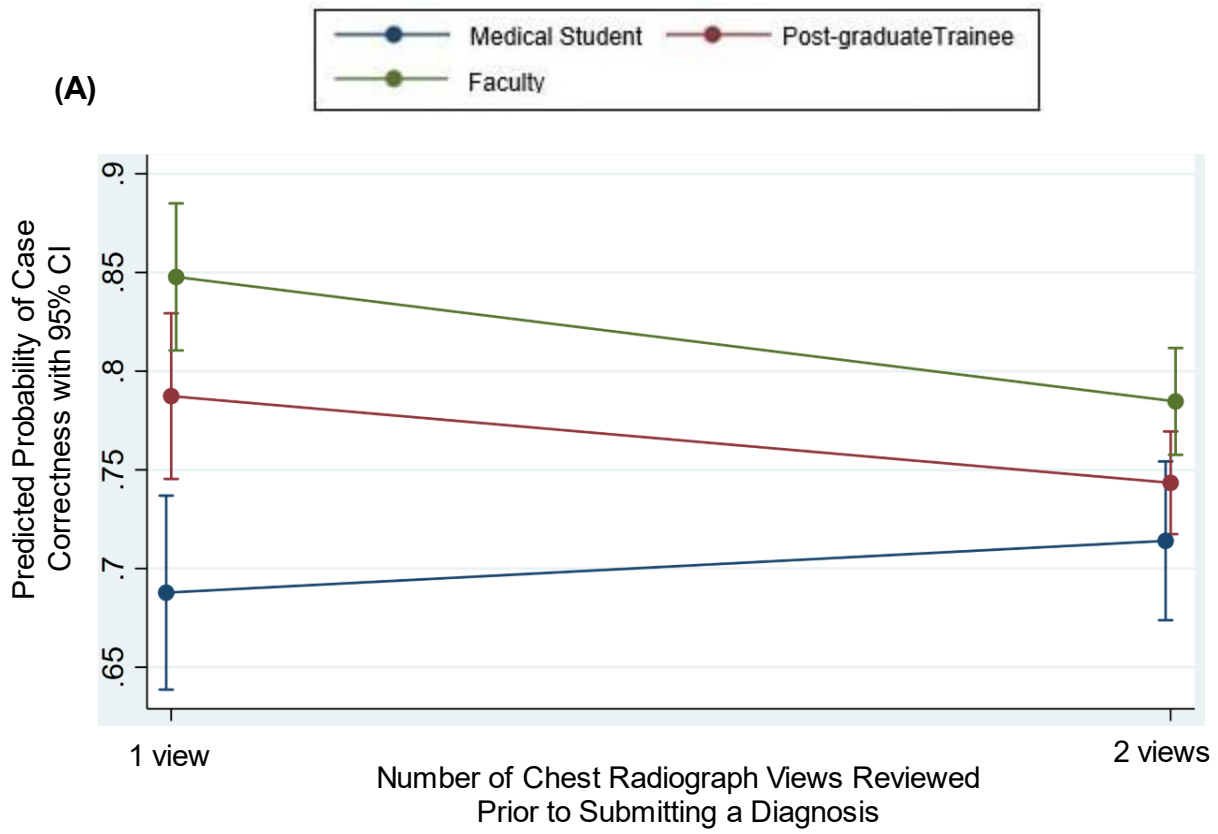
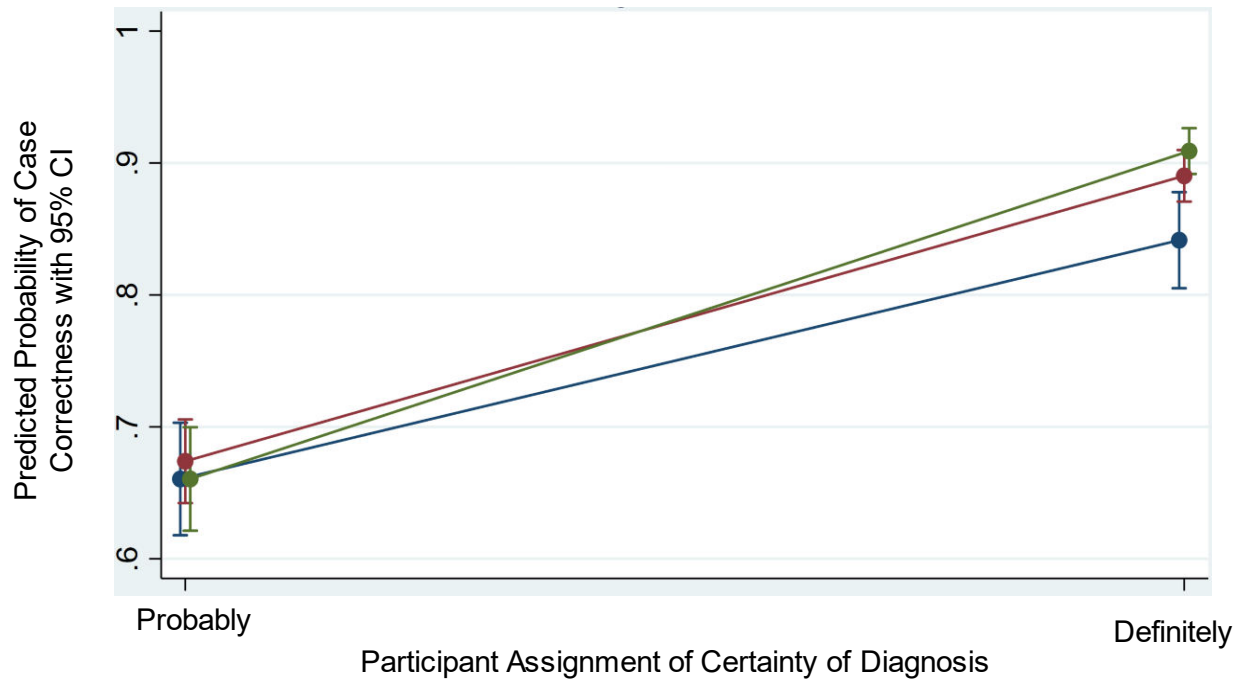


Figure S1: A Histogram of Interpretation Difficulty of Chest Radiographs. The y-axis represents the frequency of items with a particular item difficulty score. The x-axis represents the item difficulty score calculated using one-parameter item response theory, from easiest (negative numbers) to hardest (positive numbers). Panel (A) is all CXR items, while panel (B) represent the CXR without pneumonia and panel (C) represents the items with pneumonia. For purposes of comparison, the vertical black line on each panel delineates where difficulty = 0 logits.

(A)



(B)



(c)

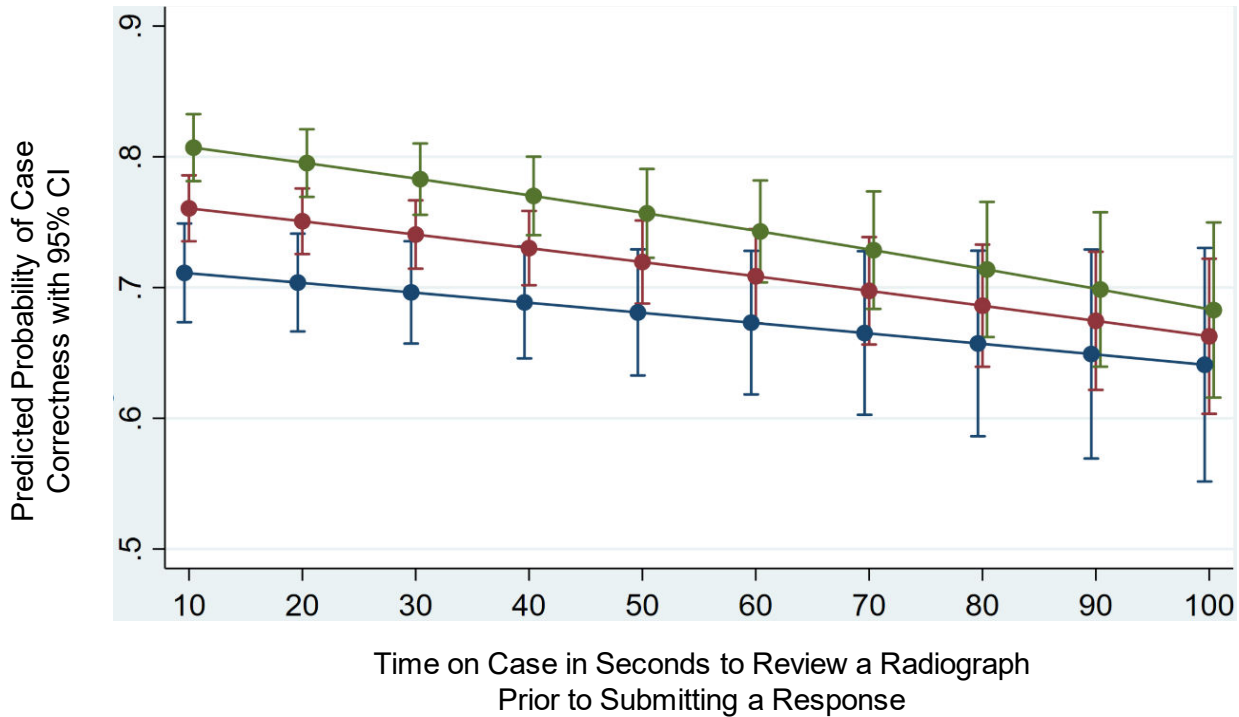


Figure S2: These data represent the effect of the interaction of career level (medical student, post-graduate trainee or faculty) with image review process variables on the dependent variable of case correctness. In Figure 4A, case correctness increased with looking at both views for medical students but decreased for post-graduate trainees and faculty. In Figure 4B, case correctness was not different between the groups when there was participant uncertainty (“probably”) but faculty and post-graduate trainee case correctness significantly increased relative to medical students when they were certain (“definitely”) of their response. In 4C, case correctness decreased similarly for all the groups the longer cases were reviewed.