4.3.4 – Validity Testing Results

Table 17. Preliminary validity testing of data elements: Encounter-level criterion validity of data elements used for measure specification among a VA and UU among a sample with either an initial or discharge pneumonia diagnosis (by ICD-10 code or NLP).

Classification of data elements among hospitalizations with initial or discharge diagnosis of pneumonia (identified by ICD code or NLP).				
**	Performance characteristics for 2 reviewers			
**	VA N=50		UU N=50	
**	Se	PPV	Se	PPV
Hospital discharge ICD-10 diagnosis code for pneumonia (any position)	0.43 0.48	1.0 1.0	0.52 0.59	0.75 0.83
Chest imaging consistent with pneumonia (NLP result vs. physician reviewer)*	0.91 0.89	0.91 0.85	0.85 0.89	0.86 0.68

Note: Two sensitivity and positive predictive values are reported for criterion validity for each data element where two different clinician reviewers reviewed each sampled chart. Individual agreement of the data element obtained electronically is reported for each clinician reviewer.
**Cells intentionally left blank.

Table 18a. Final validity testing: Encounter-level criterion validation of data elements used for final measure specification among random sample of patients meeting criteria for denominator (discharge diagnosis of pneumonia and antimicrobial treatment).

Classification Among hospitalizations meeting denominator criteria (with discharge diagnosis code pneumonia identified by ICD code and receipt of antimicrobials.						
***	Performance characteristics for single reviewer					
	VA N=50		UU N-50		MICHIGAN N=835	
	Se	PPV	Se	PPV	Se	PPV
Hospital discharge ICD-10 diagnosis code for pneumonia (any position)	*	1.00	*	1.00	*	**
Receipt of antimicrobials	*	1.00	*	1.00	*	**
Discharge diagnosis and treatment confirmed by chest imaging	0.982	1.00	0.978	1.00	0.90	0.98

* = Sensitivity not obtained as all cases in the denominator received an ICD code and a denominator.
** = ICD code and receipt antimicrobials were previously validated in the Michigan system and not re-verified.
***Cells intentionally left blank.

Details of the most recent chart review validation of data elements in the enriched sample (50:50 concordant:discordant).

Table 18b. Raw criterion validation of final eCQM score validation (discharge diagnosis and treatment confirmed by chest imaging). Sample of n=25 electronically identified concordant scores per site and n=26 electronically identified discordant scores per site; random sample of VA and University of Utah hospitalizations with ICD code-identified discharge diagnosis of pneumonia (2015-2022).

**	**Chart Review ("gold standard") for concordance between imaging and diagnosis+ treatment**		**		
VA dataset validation					
Measure-based concordance of chest imaging and pneumonia diagnosis+treatment	**Yes**	**No**	**		
Yes (+ chest imaging)	25	0	25	Se: 25/29 = 0.862	Sp: 22/22= 1
No	4	22	26	PPV: 25/25= 1	NPV: 22/26 = 0.846
University of Utah validation					
Yes	25	0	25	Se: 25/30 = 0.833	Sp: 21/21= 1
No	5	21	26	PPV: 25/25= 1	NPV: 21/26= 0.808
University of Michigan validation*					
Yes	723	14	737	Se: 723/799= 0.904	Sp: 17/31= 0.548
No	76	17	98	PPV: 723/737= 0.981	NPV: 17/98= 0.173

*No correction for weighted sampling was necessary for University of Michigan as full chart-reviewed dataset was used to assess chest imaging concordance (N=830; 1 hospitalization missing due to inability to match between EDW and chart-reviewed datasets).
**Cells intentionally left blank.

Table 18c. Correction for weighted sampling of electronically defined discordant eCQM score for criterion validation of eCQM score (discharge diagnosis and treatment without chest imaging consistent with pneumonia). Sample of n=25 electronically identified concordant scores per site and n=26 electronically identified discordant scores per site. Concordant cases 9 times more frequent than the 1:1 concordant:discordant sampling conducted so concordant cases multiplied by nine.

**	**Chart Review ("gold standard") for concordance between imaging and diagnosis+ treatment**		**		
VA dataset validation					
Measure-based concordance of chest imaging and pneumonia diagnosis+treatment	**Yes**	**No**	**		
Yes (+ chest imaging)	225	0	225	Se: 225/229 = 0.983	Sp: 22/22= 1
No	4	22	26	PPV: 225/225= 1	NPV: 22/26 = 0.846
University of Utah validation					
Yes	225	0	225	Se: 225/230 = 0.978	Sp: 21/21= 1
No	5	21	26	PPV: 225/225= 1	NPV: 21/26= 0.808

**Cells intentionally left blank.

Table 19: Document-level NLP results – Sample of Chest Images Performed among All Hospitalizations.

Tool	Sensitivity	PPV	F1
VA NLP	96.2	71.4	81.7
UU NLP – off the shelf	100.0 (+3.8)	78.3 (+6.9)	87.9 (-5.9)
UU NLP - customized	100.0 (+3.8)	87.0 (+15.6)	93.1 (+11.1)

Figure 6. Facility-level correlations between measure performance and care processes (VA facilities, 2015-2021; n =89,767).

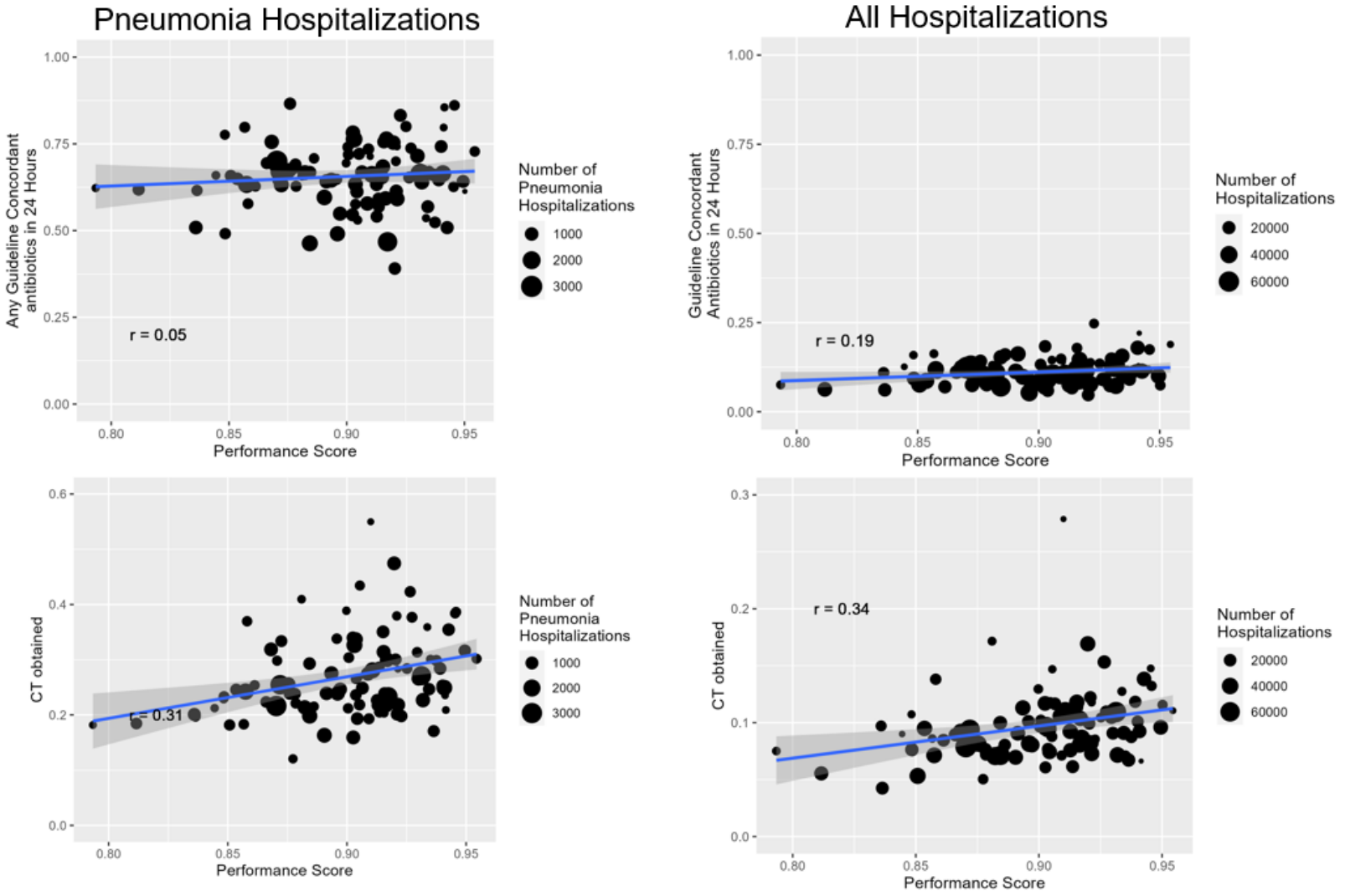

Figure 7. Facility-level correlations between measure performance and outcomes (VA facilities, 2015-2021; n =89,767).

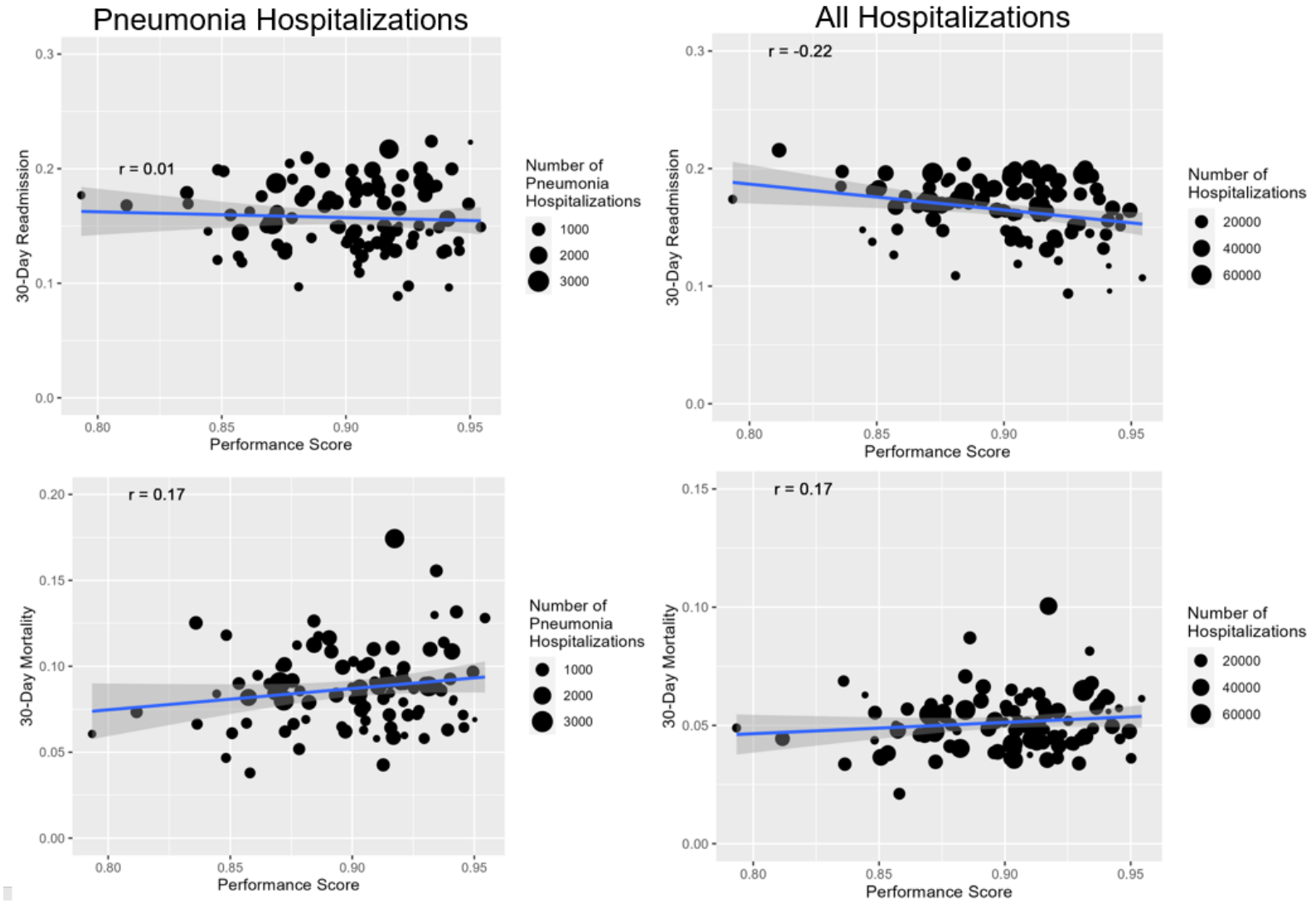

Table 20. Patient and accountable entity construct validity (association with outcomes) within VA system (N=89,767; 2015-2021).

Associations with Processes and Outcomes				
Patient-level analysis				
**	**Observed Proportion**	**Expected Proportion***	**Individual Risk Difference**	**p value**
Receipt of any antimicrobial within first 24 hours of hospitalization	93.2%	49.6%	43.6%	<0.001
30-day mortality	4.8%	5.3%	0.4%	0.08
30-day readmission (among patients surviving to discharge)*	14.3%	16.9%	-2.6%	<0.001
Facility-level analysis				
**	**Correlation**		**Spearman's correlation coefficient (R)**	**95% Confidence Intervals**
Receipt of guideline concordant antibiotics – pneumonia	Non-significant		0.05	(-0.14, .23)
Receipt of any antimicrobial- all hospitalizations	Weakly Positive		0.19	(0.02, 0.34)
CT obtained - pneumonia	Positive		0.31	(0.16, 0.45)
CT obtained - all hospitalizations,	Positive		0.34	(0.19, 0.50)
30-day mortality – pneumonia	Non-significant		0.18	(-0.04, 0.34)
30-day mortality	Non-significant		0.17	(-0.02, 0.33)
30-day readmission - pneumonia	Non-significant		0.01	(-0.19, 0.19)
30-day readmission (all hospitalizations, all diagnoses)	Negative		-0.22	(-0.40, -0.02)

* = proportion of patients experiencing the expected event if they had not experienced discordance between discharge diagnosis and chest image. Patient-level readmission risk should be interpreted with caution since it is a function of surviving to discharge.
**Cells intentionally left blank.

Figure 8. Facility-level correlations between measure performance and care processes or outcomes (VA facilities, 2015-2021; n =89,767).

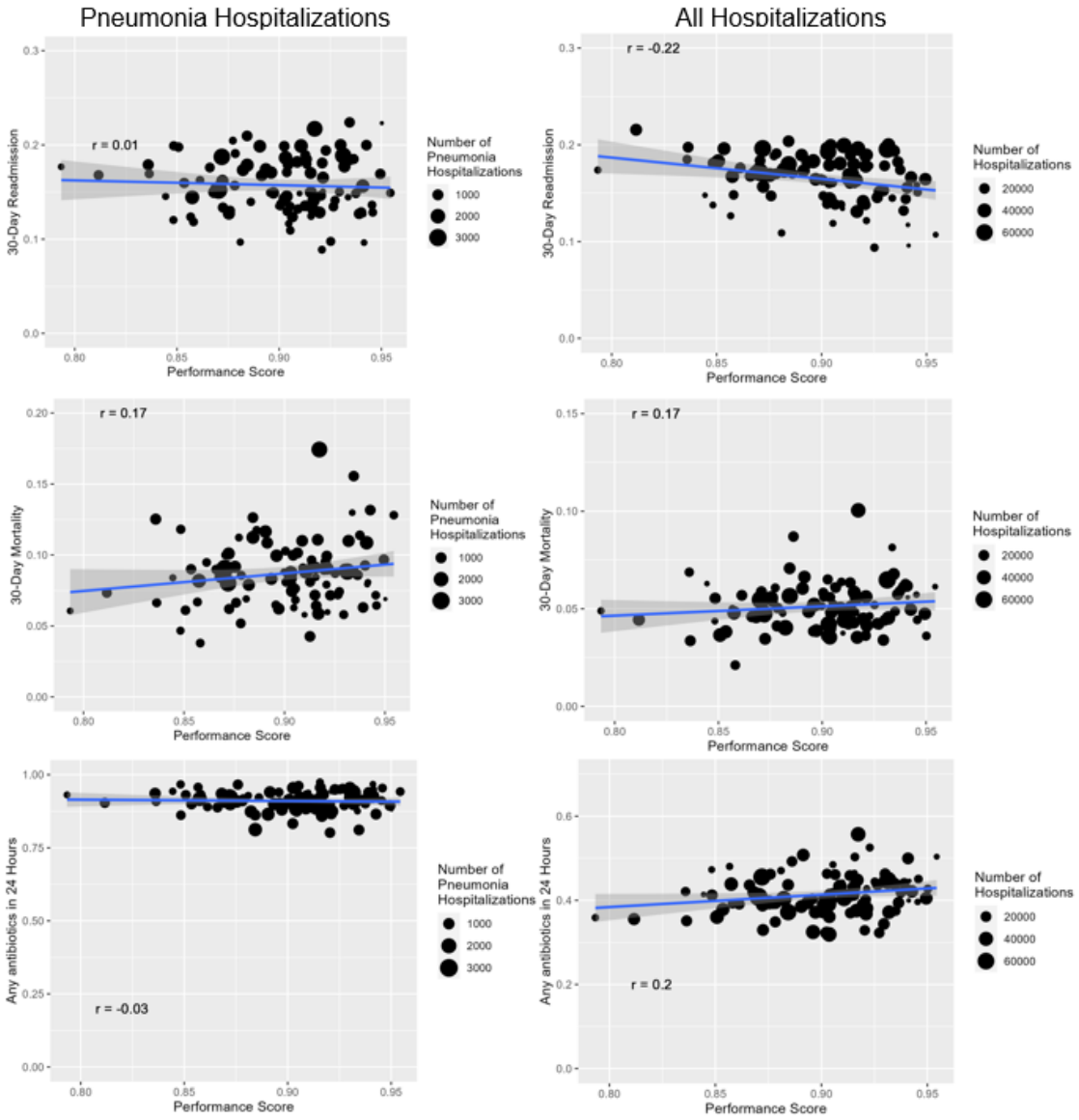

Figure 8 (continued)