Supplemental Digital Appendix 1
Category-specific agreement, all comments (n = 25,714)

(Sub)competency	Agreement (%)	Most frequent co-occurring category
PC1	48.7	PC2
PC2	31.3	PC1
PC3	48.3	PC4
PC4	38.7	N
PC5	71.9	PC4
PC6	47.6	PC3
PC7	36.6	PC4
PC8	43.7	D
PC9	15.6	PC4
PC10	72.0	D
MK1	62.9	N
MK2	30.5	PC4
P1	26.6	N
P2	21.2	P1
P3	18.4	N
ICS1	68.1	ICS2
ICS2	50.9	SBP3
ICS3	23.5	SBP2
PBLI1	31.7	MK1
PBLI2	35.0	N
SBP1	28.3	PC4
SBP2	16.4	PC2
SBP3	9.9	ICS2
D	63.6	N
N	59.1	P1

Abbreviations: PC, patient care; MK, medical knowledge; P, professionalism; ICS, interpersonal and communication skills; PBLI, problem-based learning and improvement; SBP, systems-based practice; D, demographic; N, not useful.

Supplemental Digital Appendix 2
Internal Validation Metrics

(Sub-)competency	Positive predictive value	Sensitivity	F1	AUC
PC1	0.72 (0.49, 0.93)	0.57 (0.33, 0.77)	0.61 (0.47, 0.71)	0.78 (0.66, 0.87)
PC2	0.58 (0.38, 0.81)	0.33 (0.08, 0.61)	0.38 (0.13, 0.54)	0.66 (0.54, 0.79)
PC3	0.72 (0.48, 0.93)	0.53 (0.29, 0.73)	0.60 (0.45, 0.71)	0.76 (0.65, 0.86)
PC4	0.58 (0.39, 0.82)	0.44 (0.11, 0.71)	0.46 (0.18, 0.57)	0.71 (0.55, 0.83)
PC5	0.88 (0.76, 0.96)	0.77 (0.56, 0.89)	0.81 (0.69, 0.86)	0.88 (0.78, 0.94)
PC7	0.44 (0.21, 0.88)	0.26 (0.05, 0.54)	0.28 (0.08, 0.42)	0.63 (0.52, 0.76)
PC8	0.59 (0.26, 1.00)	0.35 (0.11, 0.61)	0.41 (0.18, 0.62)	0.67 (0.55, 0.80)
PC10	0.88 (0.72, 0.97)	0.77 (0.54, 0.91)	0.81 (0.69, 0.87)	0.88 (0.77, 0.94)
MK1	0.83 (0.64, 0.95)	0.70 (0.51, 0.84)	0.75 (0.65, 0.81)	0.84 (0.75, 0.91)
MK2	0.51 (0.27, 0.86)	0.33 (0.14, 0.53)	0.38 (0.21, 0.51)	0.66 (0.57, 0.76)
P	0.59 (0.38, 0.82)	0.63 (0.44, 0.79)	0.59 (0.50, 0.65)	0.79 (0.71, 0.84)
ICS	0.76 (0.52, 0.94)	0.60 (0.34, 0.79)	0.65 (0.50, 0.73)	0.79 (0.67, 0.87)
PBLI	0.59 (0.35, 0.86)	0.52 (0.31, 0.71)	0.53 (0.42, 0.62)	0.75 (0.65, 0.84)
SBP	0.31 (0.11, 0.78)	0.19 (0.01, 0.40)	0.19 (0.03, 0.31)	0.59 (0.51, 0.69)
D	0.78 (0.64, 0.94)	0.77 (0.55, 0.91)	0.76 (0.69, 0.82)	0.87 (0.77, 0.92)
N	0.68 (0.53, 0.84)	0.89 (0.77, 0.97)	0.76 (0.68, 0.82)	0.87 (0.85, 0.90)

Abbreviations: PC, patient care; MK, medical knowledge; P, professionalism; ICS, interpersonal and communication skills; PBLI, problem-based learning and improvement; SBP, systems-based practice; D, demographic; N, not useful.
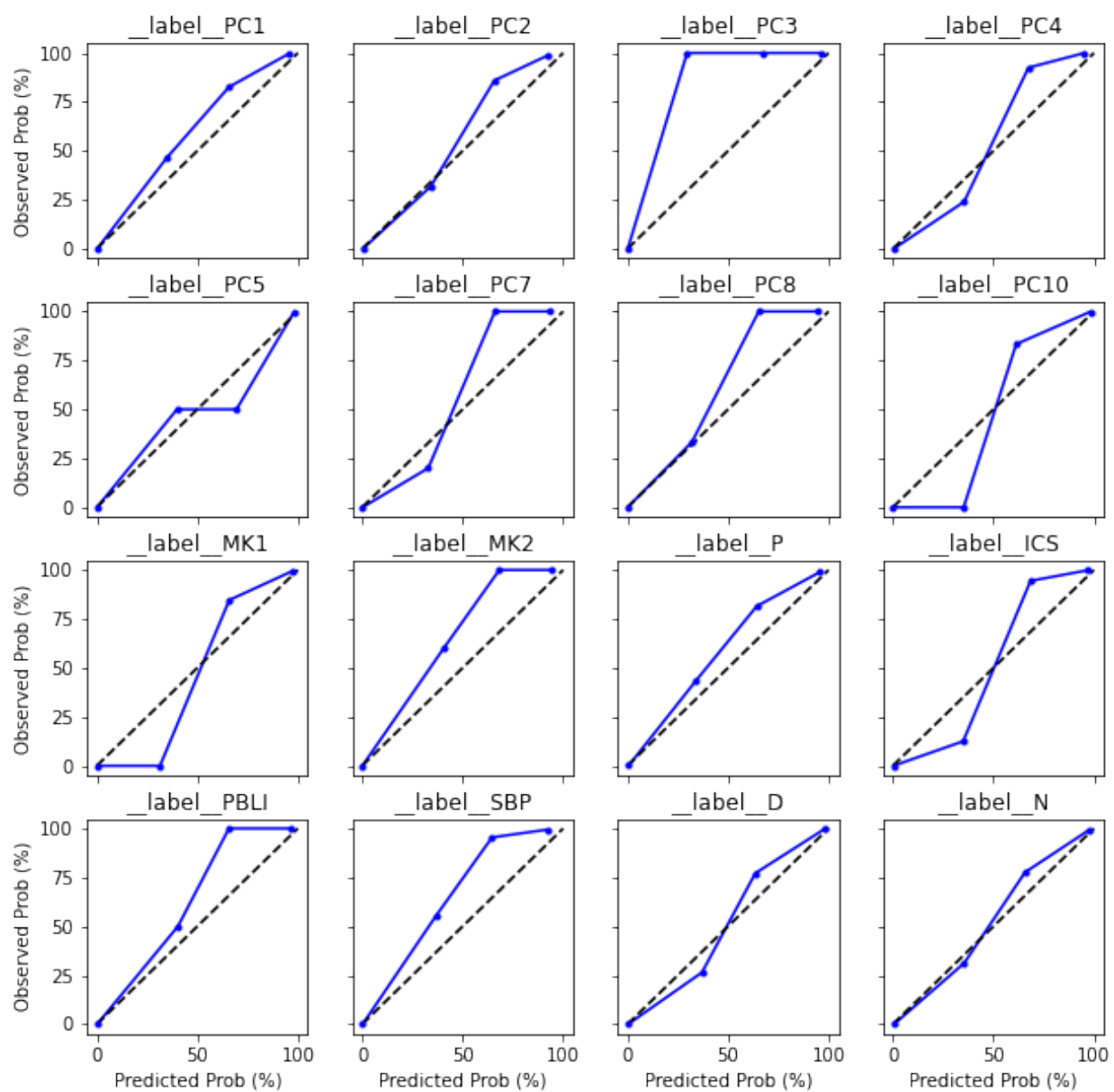
Supplemental Digital Appendix 3
External Validation Metrics

(Sub-)competency	Positive predictive value	Sensitivity	F1	AUC
PC1	0.57 (0.47, 0.67)	0.51 (0.40, 0.62)	0.54 (0.45, 0.63)	0.75 (0.69, 0.81)
PC2	0.47 (0.41, 0.53)	0.50 (0.42, 0.58)	0.49 (0.42, 0.55)	0.73 (0.69, 0.77)
PC3	0.54 (0.41, 0.68)	0.42 (0.29, 0.55)	0.47 (0.35, 0.59)	0.71 (0.65, 0.78)
PC4	0.43 (0.36, 0.49)	0.38 (0.31, 0.45)	0.40 (0.34, 0.46)	0.66 (0.63, 0.70)
PC5	0.68 (0.60, 0.77)	0.76 (0.67, 0.85)	0.72 (0.65, 0.79)	0.87 (0.83, 0.92)
PC7	0.43 (0.32, 0.56)	0.28 (0.19, 0.37)	0.34 (0.24, 0.43)	0.64 (0.59, 0.69)
PC8	0.44 (0.21, 0.67)	0.23 (0.10, 0.39)	0.30 (0.13, 0.46)	0.61 (0.55, 0.69)
PC10	0.62 (0.54, 0.72)	0.73 (0.63, 0.82)	0.67 (0.60, 0.74)	0.86 (0.81, 0.90)
MK1	0.64 (0.59, 0.71)	0.73 (0.65, 0.80)	0.68 (0.63, 0.73)	0.85 (0.82, 0.89)
MK2	0.35 (0.22, 0.47)	0.23 (0.14, 0.33)	0.27 (0.17, 0.38)	0.62 (0.57, 0.68)
P	0.51 (0.45, 0.59)	0.29 (0.24, 0.34)	0.37 (0.31, 0.43)	0.62 (0.60, 0.65)
ICS	0.46 (0.40, 0.53)	0.58 (0.50, 0.67)	0.51 (0.45, 0.58)	0.78 (0.74, 0.83)
PBLI	0.53 (0.45, 0.60)	0.45 (0.36, 0.52)	0.48 (0.41, 0.55)	0.70 (0.66, 0.74)
SBP	0.30 (0.19, 0.42)	0.18 (0.10, 0.25)	0.22 (0.14, 0.30)	0.58 (0.54, 0.62)
D	0.65 (0.60, 0.70)	0.68 (0.61, 0.73)	0.66 (0.61, 0.71)	0.82 (0.79, 0.86)
N	0.56 (0.53, 0.59)	0.86 (0.82, 0.89)	0.68 (0.65, 0.70)	0.85 (0.83, 0.88)

Abbreviations: PC, patient care; MK, medical knowledge; P, professionalism; ICS, interpersonal and communication skills; PBLI, problem-based learning and improvement; SBP, systems-based practice; D, demographic; N, not useful.

Supplemental Digital Appendix 4
Calibration curves for training data

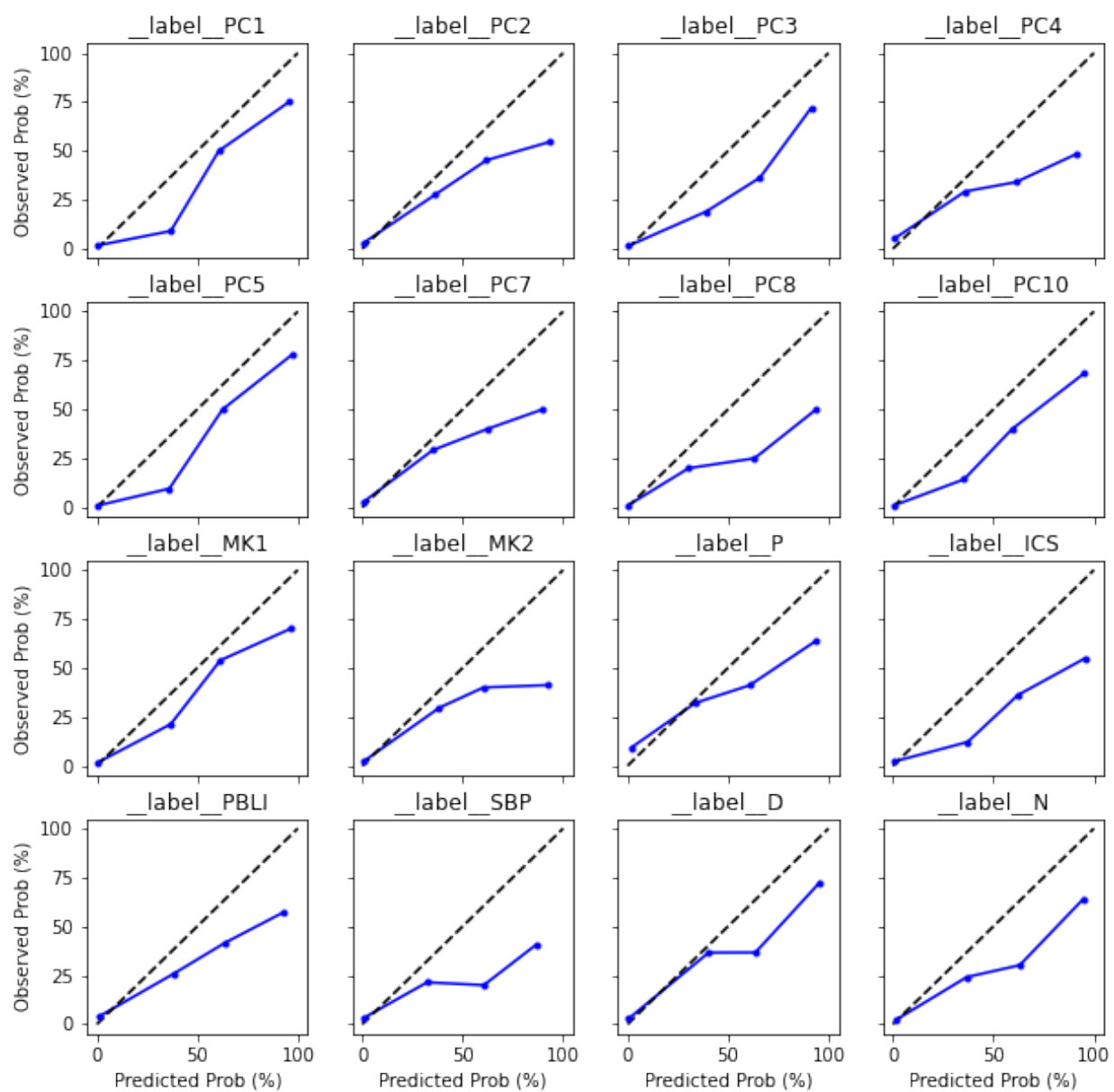## Supplemental Digital Appendix 5
External Validation and Sensitivity Analysis

(Sub)competency	Complete cohort AUC (95% CI), n = 2,255	Sensitivity analysis, complete agreement only, AUC (95% CI), n = 1,108
PC1	0.75 (0.69, 0.81)	0.83 (0.76, 0.89)
PC2	0.73 (0.69, 0.77)	0.75 (0.69, 0.82)
PC3	0.71 (0.65, 0.78)	0.76 (0.65, 0.87)
PC4	0.66 (0.63, 0.70)	0.75 (0.69, 0.81)
PC5	0.87 (0.83, 0.92)	0.90 (0.86, 0.95)
PC7	0.64 (0.59, 0.69)	0.72 (0.62, 0.81)
PC8	0.61 (0.55, 0.69)	0.63 (0.53, 0.73)
PC10	0.86 (0.81, 0.90)	0.87 (0.80, 0.93)
MK1	0.85 (0.82, 0.89)	0.92 (0.89, 0.95)
MK2	0.62 (0.57, 0.68)	0.65 (0.56, 0.78)
P	0.62 (0.60, 0.65)	0.73 (0.67, 0.78)
ICS	0.78 (0.74, 0.83)	0.87 (0.81, 0.92)
PBLI	0.70 (0.66, 0.74)	0.80 (0.74, 0.87)
SBP	0.58 (0.54, 0.62)	0.62 (0.50, 0.73)
D	0.82 (0.79, 0.86)	0.88 (0.85, 0.91)
N	0.85 (0.83, 0.88)	0.91 (0.89, 0.93)

Abbreviations: PC, patient care; MK, medical knowledge; P, professionalism; ICS, interpersonal and communication skills; PBLI, problem-based learning and improvement; SBP, systems-based practice; D, demographic; N, not useful.
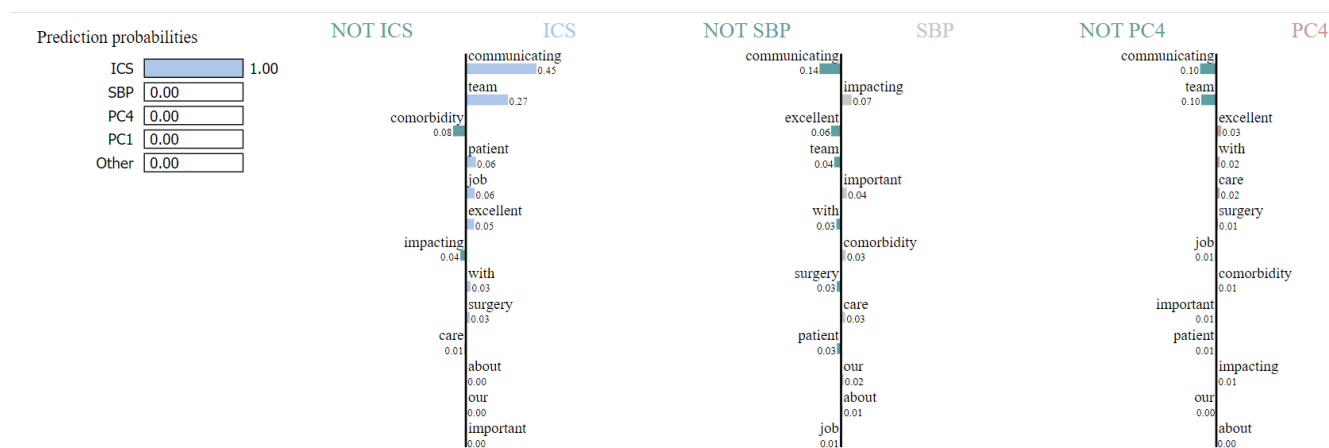
Supplemental Digital Appendix 6
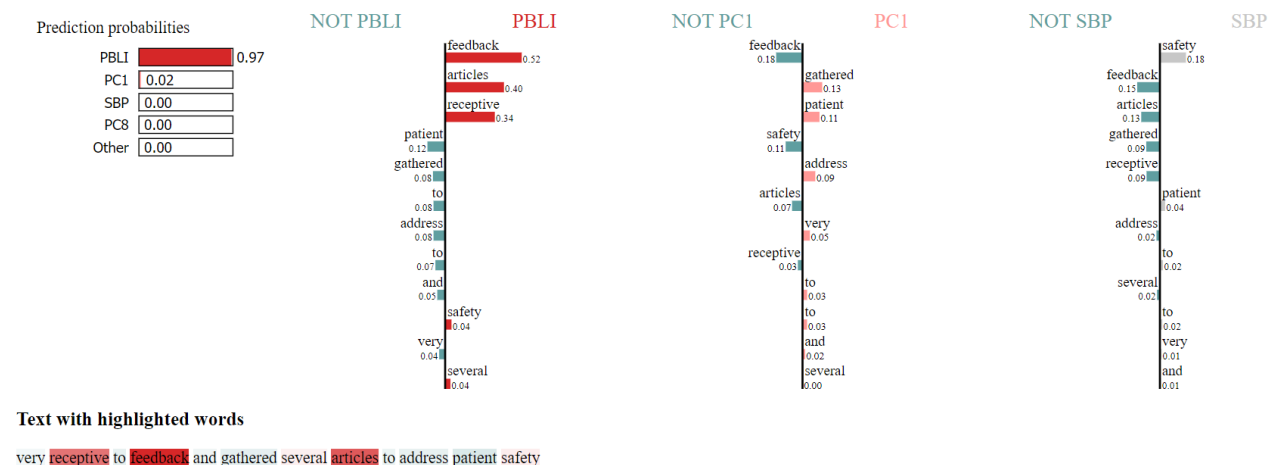## Calibration curves for validation data

Supplemental Digital Appendix 7

Example of model predictions for evaluation: "Excellent job communicating with surgery team about important patient comorbidity impacting our care."
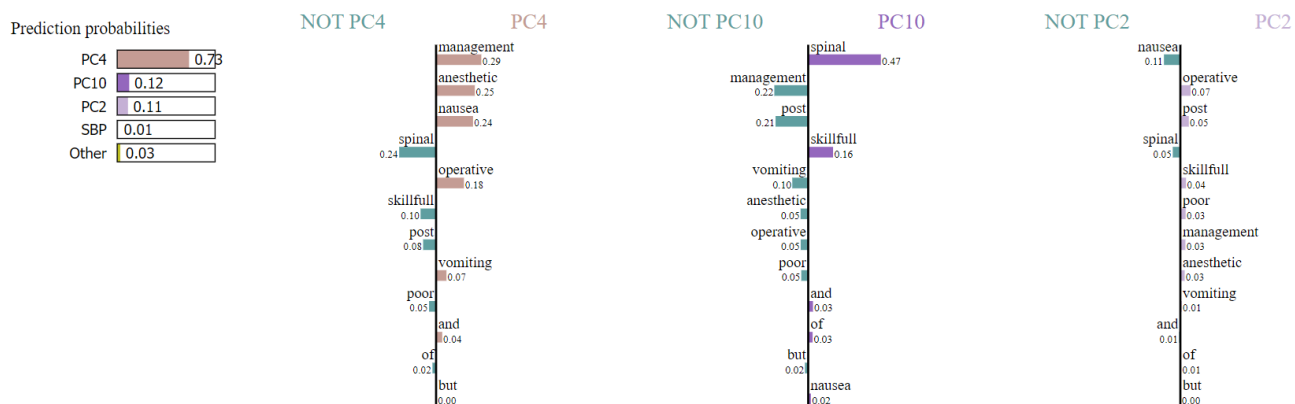

Text with highlighted words

excellent job communicating with surgery team about important patient comorbidity impacting our care

Supplemental Digital Appendix 8

Example of model predictions for evaluation: "Very receptive to feedback and gathered several articles to address patient safety."

Text with highlighted words

very receptive to feedback and gathered several articles to address patient safety

Supplemental Digital Appendix 9
Example of model predictions for evaluation: "Skillful spinal anesthetic but poor management of post operative nausea and vomiting."

Text with highlighted words

skillfull spinal anesthetic but poor management of post operative nausea and vomiting