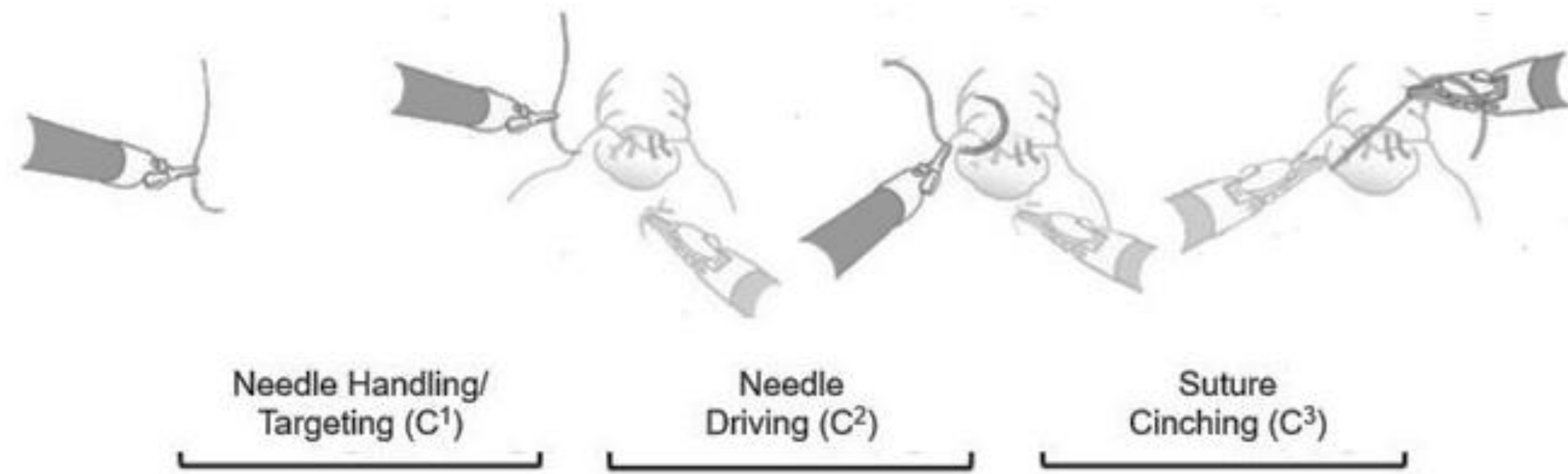


Introductions and objectives

- Automated performance metrics (APMs) are a validated objective measure of surgeon performance
 - APMs include instrument motion tracking, Endowrist manipulation, and system events data
- Machine Learning (ML) has shown that APMs, especially during the vesico-urethral anastomosis (VUA) of the RARP, are predictive of outcomes such as continence recovery time.
- We now divide the VUA into individual stitches and sub-stitches and use ML to analyze APMs to distinguish surgeon experience.

Methods

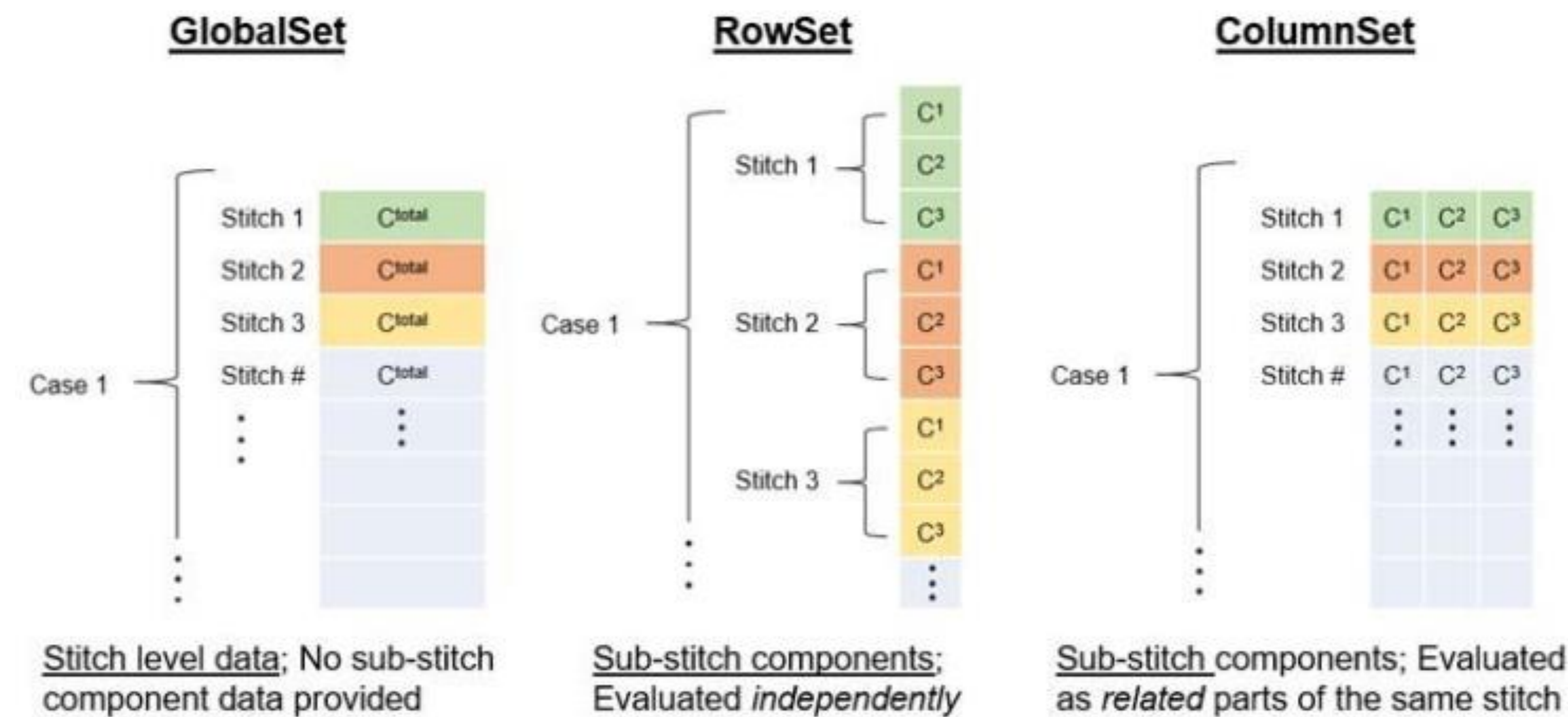
- 68 VUA's were recorded by a da Vinci systems data recorder.
- Each stitch was broken down into 3 sub components.



- APMs were reported for 1,570 stitches and 4,708 sub-stitches
- The 8 faculty surgeons and 9 residents were classified as experts (≥ 100 cases) and novices (< 100 cases).
 - Experts were then further divided into ordinary experts (OE) and super-experts (SE) (≥ 2000 cases).

Methods

- Reported APMs were organized into three datasets: GlobalSet (whole stitch), RowSet (independent sub-stitches), and ColumnSet (associated sub-stitches).



- We applied three ML models (AdaBoost, Gradient Boosting, and Random Forest) to each of the three datasets to solve two classification tests: experts vs novices and ordinary-experts vs super-experts.
- 207 APMs per sub-stitch were evaluated for the stability of their importance to each classification task through a Jaccard index.
- The final feature importance after 20 trials provided the ability to rank variables in order of importance to the model

Results

- In both classification tasks, experts vs novice and OE vs SE, ColumnSet produced the highest accuracy in a single model.

Expert vs Novice Accuracy	ColumnSet	RowSet	GlobalSet
AdaBoost	0.724 +/- 0.016	0.712 +/- 0.009	0.699 +/- 0.018
Random Forest	0.732 +/- 0.005	0.716 +/- 0.003	0.728 +/- 0.009
Gradient Boosting	0.727 +/- 0.010	0.721 +/- 0.006	0.672 +/- 0.001

Ordinary Expert vs Super Expert Accuracy	ColumnSet	RowSet	GlobalSet
AdaBoost	0.801 +/- 0.014	0.772 +/- 0.009	0.774 +/- 0.010
Random Forest	0.761 +/- 0.007	0.761 +/- 0.004	0.769 +/- 0.009
Gradient Boosting	0.770 +/- 0.006	0.784 +/- 0.006	0.759 +/- 0.002

- Stable feature importance scores highlighted Endowrist® articulation in differentiating experts and novices while needle handling/targeting (C¹) was more important when comparing OE and SE

Conclusions

- Surgeon performance measured by APMs on a granular sub-stitch level more accurately distinguishes expertise when compared to summary APMs over whole stitches
- ML accurately classifies expertise when analyzing APMs of the VUA of a RARP