Dissection Assessment for Robotic Technique (DART) to Evaluate Nerve-Spare of Robot-Assisted Radical Prostatectomy

Runzhuo Ma, Alvin Hui, Jiashu Xu, Aditya Desai, Michael Tzeng, Emily Cheng, Loc Trinh, Jessica H. Nguyen, Anima Anandkumar, Jim C. Hu, Andrew J. Hung

Introduction and Objective

High quality nerve-spare (NS) is essential for the preservation of erectile function (EF) after robot-assisted radical prostatectomy (RARP). In a previous study, we developed an assessment tool for tissue dissection, *Dissection Assessment for Robotic Technique* (DART). Herein, we further apply DART scores to the NS step and evaluate whether DART can predict 1-year EF recovery after RARP.

Methods

RARP cases from 2016-2019 with preoperative EF and ≥1 year postoperative follow up were included. Non-nerve-sparing procedures were excluded.

Phase 1: After standardized training, 5 independent assessors used DART to evaluate de-identified NS videos. Inter-rater reliability (IRR) was evaluated by prevalence-adjusted and bias-adjusted Kappa (PABAK). DART scores of surgeons with different experience levels were compared by Kruskal-Wallis test.

Phase 2: DART scores were used to predict 1 year EF recovery after RARP. EF was defined as erections sufficient for intercourse greater than half the time, which corresponds to an answer of 4 or 5 to the third question of the International Index of Erectile Function. 12 clinical features (eg, age, comorbidities) and 6 domains of DART scores were used to construct a machine learning model (XGBoost) to predict EF recovery. We leveraged 4-fold cross-validation to train and evaluate the model, and reported the mean and standard deviation (SD) of area under the curve (AUC) across the 4 folds on a held-out test set.

Results

96 NS videos from 17 surgeons were included. The IRR among 5 assessors was moderately high (PABAK = 0.62, 95% CI [0.50-0.74]), and agreement rate was 75%. Phase 1: the majority of DART sub-domains were able to distinguish the experience levels of surgeons performing the NS, and the total DART scores were significantly higher in the more experienced groups (p=0.003) (**Table 1**). Phase 2: AUC of the model predicting 1-year EF recovery achieved 0.72 (SD±0.16) and 0.78 (SD±0.07) using clinical features and DART scores alone, respectively. Combining both elements, AUC of the prediction model achieved 0.84 (SD±0.08).

Conclusions

DART scores can distinguish surgeons with different experience levels during the NS step of RARP and can accurately predict 1-year EF recovery.

Table 1. DART scores distinguish nerve-spare performance. Estimates were given as median (quartile 1, quartile 3).

Domains	**Trainees** *(Prior Caseload 0-300)* N = 26	**Experts** *(Prior Caseload 301-1999)* N = 16	**Super-experts** *(Prior Caseload ≥2000)* N = 54	**p value***
Gesture selection and efficacy	2.8 (2.6, 3.0)	2.8 (2.7, 3.0)	2.8 (3.0, 3.0)	0.729
Instrument visualization and awareness	3.0 (3.0, 3.0)	3.0 (3.0, 3.0)	3.0 (3.0, 3.0)	0.016
Respect of tissue planes	2.8 (2.8, 3.0)	3.0 (2.8, 3.0)	3.0 (2.8, 3.0)	0.088
Tissue handling	2.8 (2.6, 3.0)	2.8 (2.5, 2.9)	2.8 (2.8, 3.0)	0.361
Tissue retraction	2.8 (2.6, 2.8)	2.8 (2.6, 3.0)	3.0 (2.8, 3.0)	0.024
Efficiency	2.7 (2.6, 3.0)	2.8 (2.5, 2.9)	3.0 (2.8, 3.0)	0.001
Total	16.6 (16.4, 17.3)	17.1 (16.2, 17.4)	17.3 (17.0, 17.6)	0.003

* Three groups comparison.

Characters - 1,980 + 1 table (225), limitation 2,280