

Building An Experience Engine To Make Cancer Treatment Decisions Using Machine Learning

N Ramarajan¹, S Gupta², P Perry¹, G Srivastava¹, A Kumbla³, J Miller³, N Feldman³, NS Nair², RA Badwe²

1. Navya Network, Cambridge (MA); 2. Tata Memorial Centre, Mumbai (India); 3. University of California Los Angeles - Olive View Medical Center, Los Angeles (CA)

Background

Experts provide solutions to complex cases not addressed by high quality evidence
 They intuitively retrieve patterns from years of experience to make treatment decisions
 Short of personal consults, there is no way to access this vast experiential knowledge
 Navya and Tata Memorial Centre designed a machine learning solution

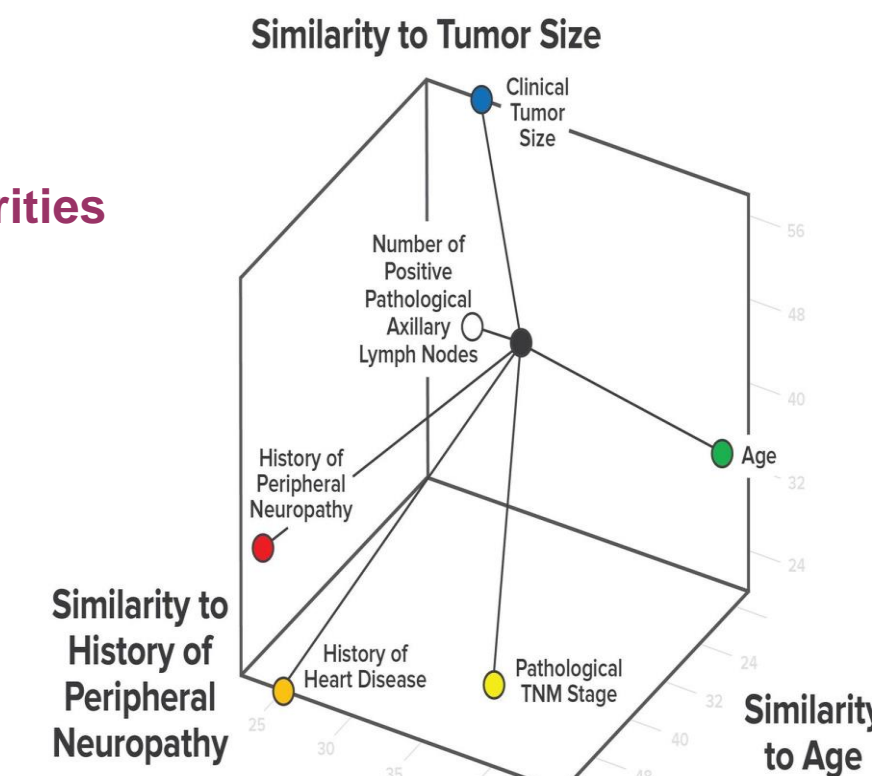
Objective

To build the Experience Engine (XE), a machine learning solution to:

- structure experiential knowledge relevant for decision making
- derive a similarity metric for patients who have received similar treatments
- predict treatment decisions that experts are likely to recommend

Methodology

Training Set: 743 breast cancer tumor board decisions at 2 tertiary care centers
Test Set: 596 decisions for an online expert opinion service
Target Of Prediction: Treatment class (e.g. Anthra) and the specific regimen(e.g. AC)
Primary Endpoint: Accuracy (AUC) of XE's predicted and expert's actual Rx decision
Navya Ontology: 690 structured features designed specifically for breast cancer decisions
Data Grouping for nonlinear similarities: 13 groups by treatment decision point
 e.g. primary Sx, 1st line CT, etc.



Machine Learning to uncover nonlinear similarities
 (e.g., similar treatments for younger patients with multiple comorbidities and elderly patients)
Multiple Similarity Distance Metrics evaluated
 Bhattacharya, Eskin, Goodall, etc.
Multiclass Classification Algorithms evaluated
 XGB, CART, SVM, kNN, c5.0 etc.

Weighted Random Guessing was used as a baseline for prediction
 to compare improvements in accuracy with machine learning
Winning algorithms maximized accuracy (AUC) for each treatment decision point

Table 1 – Accuracy by Rx Decisions

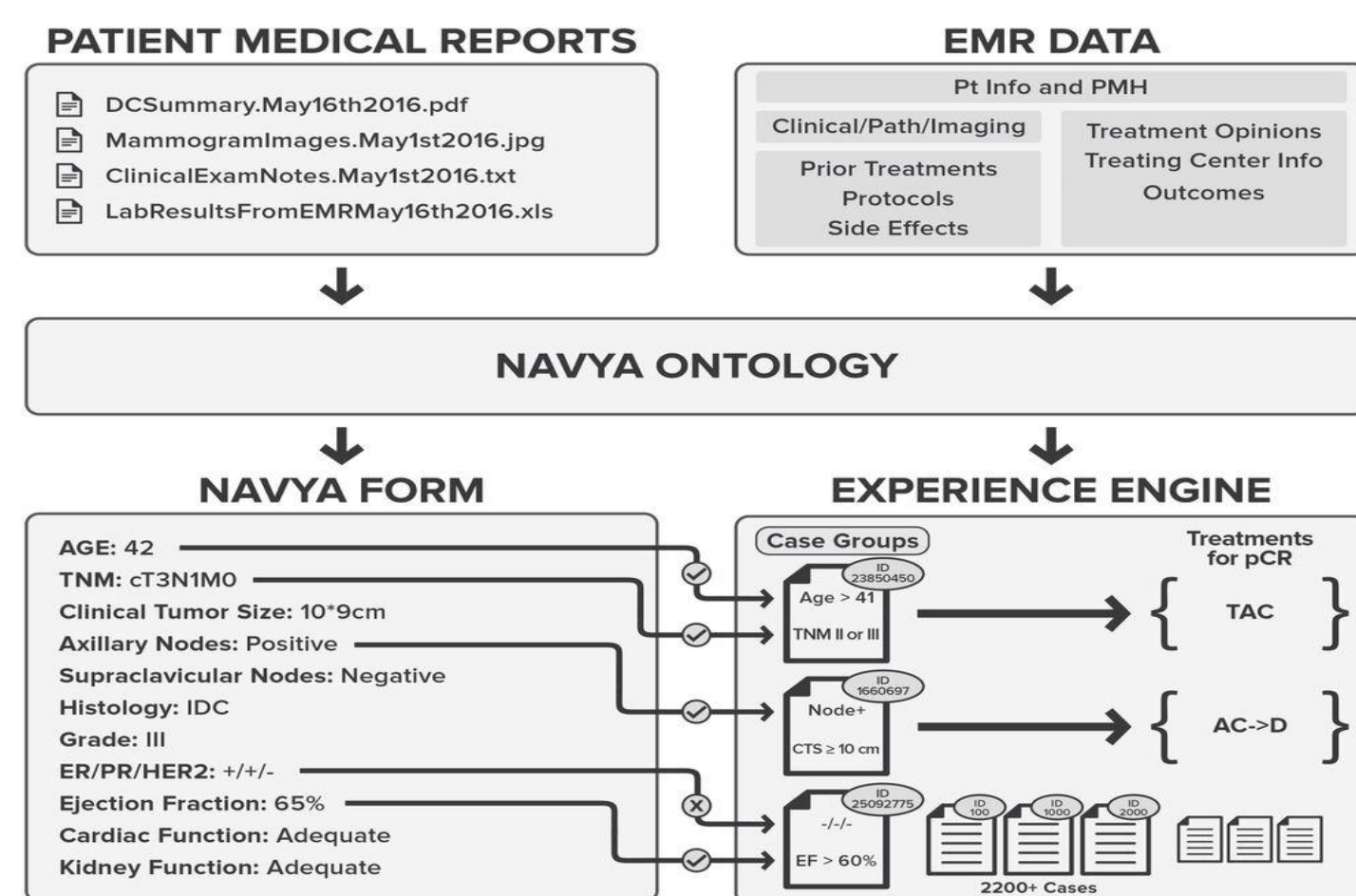
Treatment Decision Point (N)	# of Rx Choices	Baseline Weighted Random Guess	XE Accuracy (AUC) ±CI for All Rx	XE Accuracy (AUC) ±CI for Standard Rx
Primary Breast Surgery (80)	13	27% ±9.7	69% ±10.1 *	69% ±10.1
Primary Lymph Node Surgery (80)	7	65% ±10.4	72% ±9.8	82% ±8.4
Adjuvant HER2- by Chemo Class (88)	12	32% ±9.7	79% ±8.5 *	88% ±6.7
Adjuvant HER2- by Chemo Regimen (88)	38	12% ±6.7	31% ±9.6 *	70% ±9.5 *
Adjuvant HER2+ by Chemo Class (62)	9	18% ±9.5	58% ±12.2 *	70% ±11.4
Adjuvant HER2+ by Chemo Regimen (62)	30	9% ±7.1	50% ±12.4 *	88% ±8 *
Adjuvant Premenopausal Hormone (22)	7	70% ±19.1	75% ±18	99% ±4.1 *
Adjuvant Postmenopausal Hormone (66)	13	53% ±12	99% ±2.3 *	99% ±2.4
Adjuvant Radiation (107)	16	32% ±7.2	82% ±4.8 *	93% ±6.5 *

* No overlap between confidence intervals- Baseline vs XE All and XE All vs Standard

Machine Learning Methodology for the Experience Engine (XE)

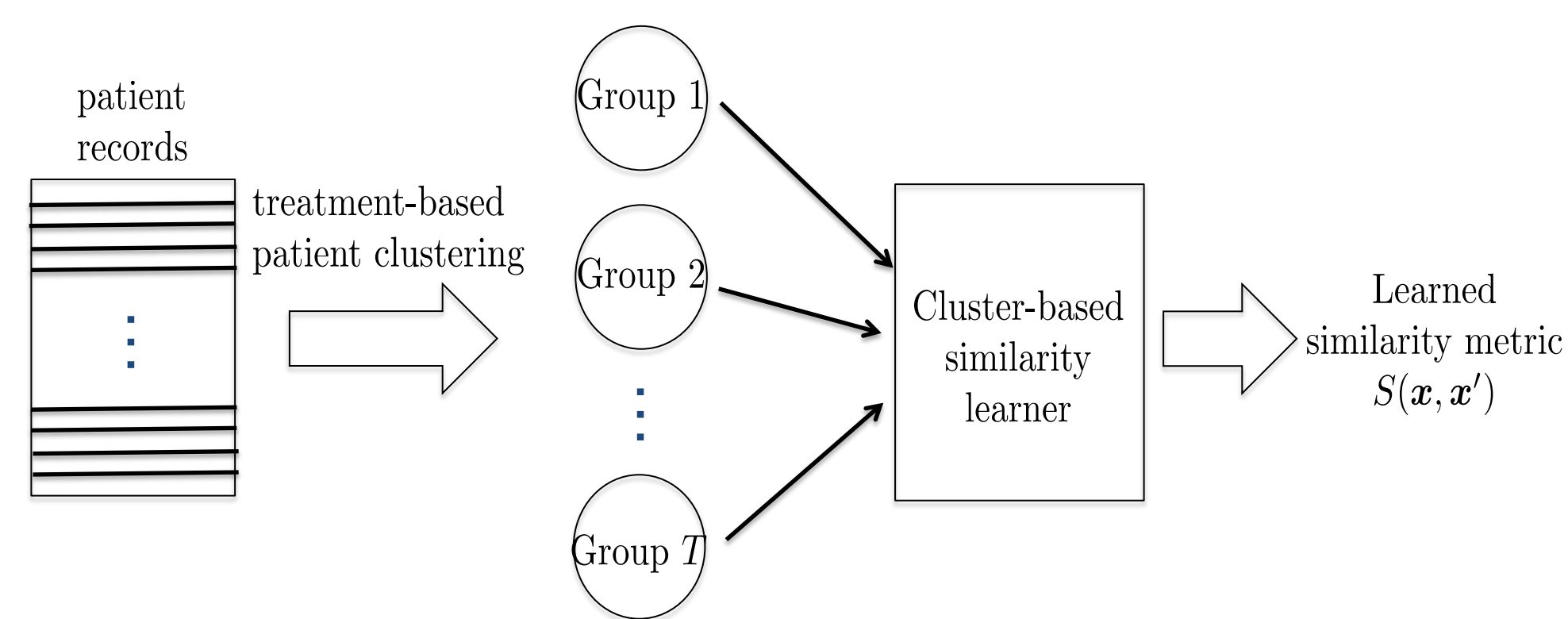
Step 1:

Structured database of experiential knowledge



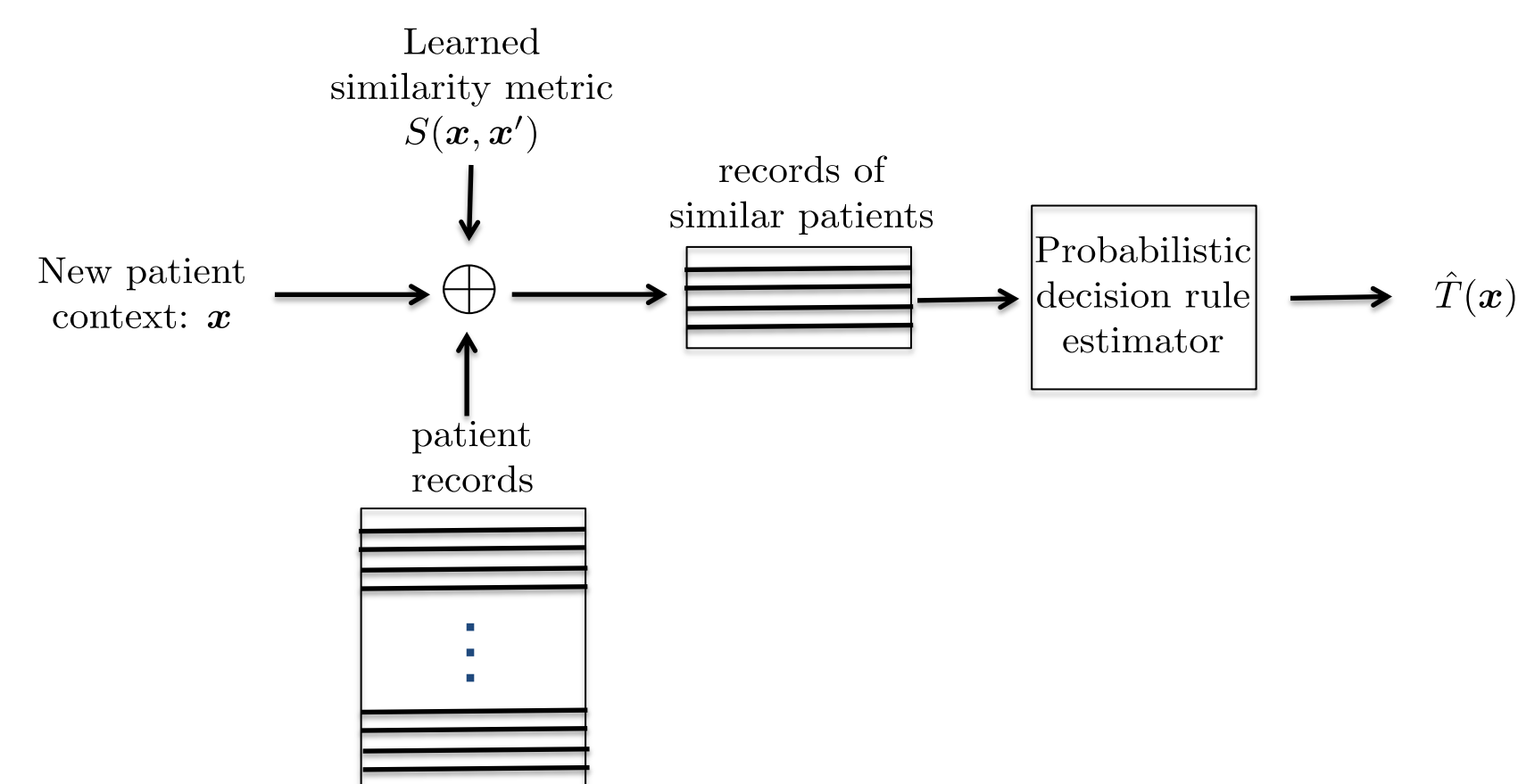
Step 2:

Similarity learning based on Rx decision point



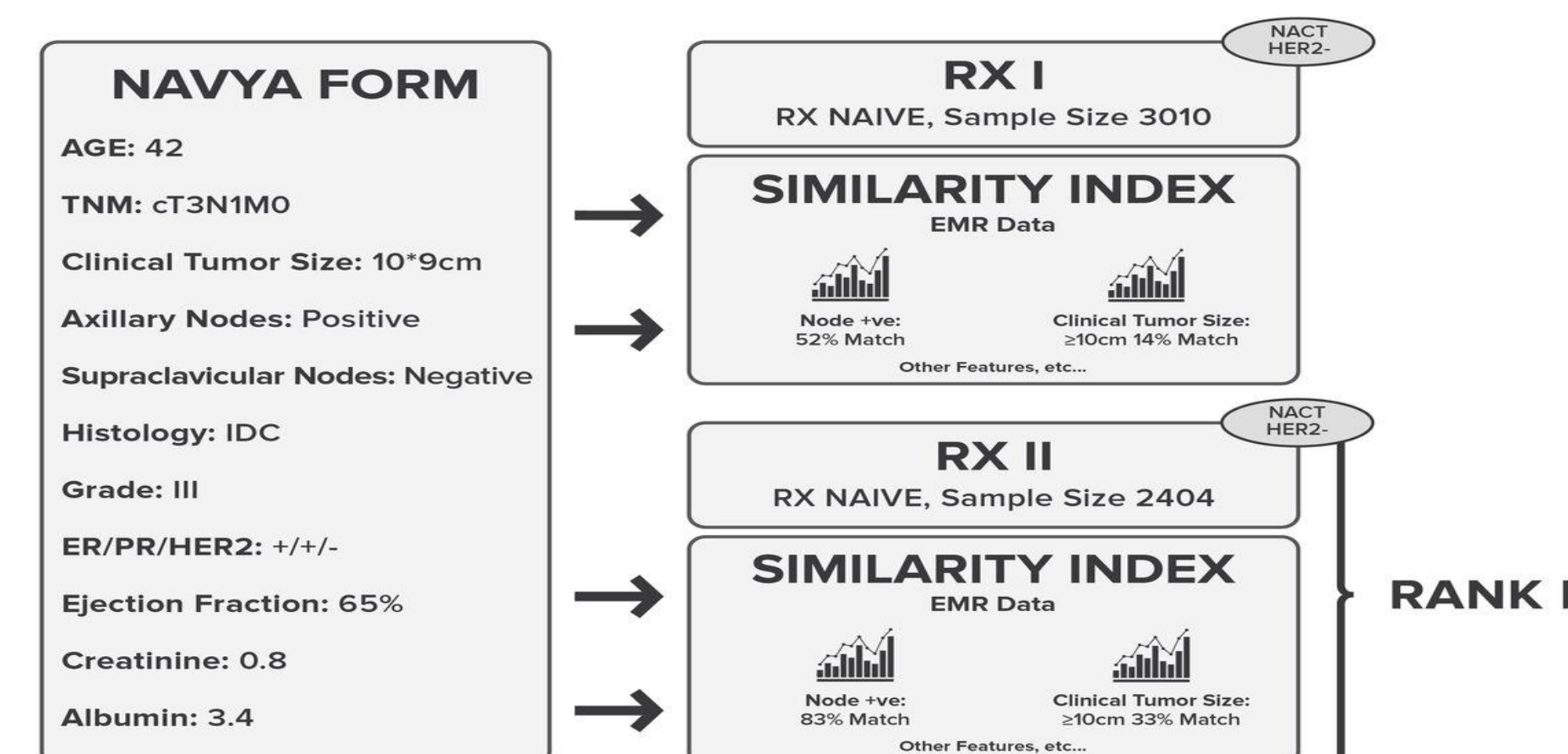
Step 3:

Patient record retrieval by learned similarity



Step 4:

Experience Engine Rx decision



Results

Baseline (weighted random guess) vs XE accuracy

- XE algorithms are significantly more accurate than baseline at predicting actual treatment decisions recommended by experts (Table 1)
- There is great variation in the number of treatment choices for each decision point
- Accuracy was higher for decision points with fewer treatment choices

XE accuracy for Standard (common, evidence-based) vs All treatments

- The most frequent treatments decisions recommended across all decision points were standard evidence based therapies
- XE algorithms are more accurate at predicting standard treatments than all treatments. Accuracy is significantly higher for decision points with a larger number of nonstandard treatment choices
- XE learned to weigh features relating to comorbidities and toxicities when recommending nonstandard treatments

Conclusion

- Experience Engine using machine learning on past expert decisions can predict treatments that experts are likely to recommend for a new patient
- Despite the limited dataset, the Experience Engine learned features that experts strongly consider when making decisions
- By including complex decisions that consider toxicities and morbidities, a rich new source of knowledge complementing evidence can be created
- The Experience Engine has the potential to analyze variations in decision making at expert practices, assess when to recommend nonstandard treatments, and serve as a training tool for new oncologists to make expert grade treatment decisions