

Selecting a tool to predict hip fracture risk for use in a large integrated electronic health record system

Noa Dagan MD, MPH¹; Chandra Cohen-Stavi, MPA¹; Maya Leventer-Roberts MD, MPH¹; Ran Balicer, MD, PhD, MPH¹

¹ Clalit Research Institute, Chief Physician's Office, Clalit Health Services, Israel

BACKGROUND

- » Hip fractures are preventable events, associated with considerable morbidity, mortality and high costs
- » The three most studied tools for hip fracture prediction are Garvan, the WHO's FRAX, and QFracture, first published in 2007, 2008 and 2009, respectively
- » Comparison between the tools' performance is difficult, since previous external validations used different inclusion criteria
- » Applying these tools through automatic computation in an electronic health record (EHR) system, has yet to be established
- » Policy-makers face considerations of performance, local applicability, and required inputs in tool adoption decisions

Basic Features of the FRAX, Garvan and QFracture tools

	Number of risk factors	Prediction time	Age Range	Data Source for Model Development	Previous Independent Validations*
Garvan	5	5 & 10 years	50-95	Cohort, survey and physician-reported	6 studies (in 3 countries)
FRAX	11	10 years	50-90	Cohort, survey and patient-reported	26 studies (in 9 countries)
QFracture	26	1-10 years	30-100	EHR	3 studies (UK only)

*Marques A, Ferreira RJO, Santos E, et al. The accuracy of osteoporotic fracture risk prediction tools: a systematic review and meta-analysis. Ann Rheum Dis 2015;74:1958-1967.

OBJECTIVES

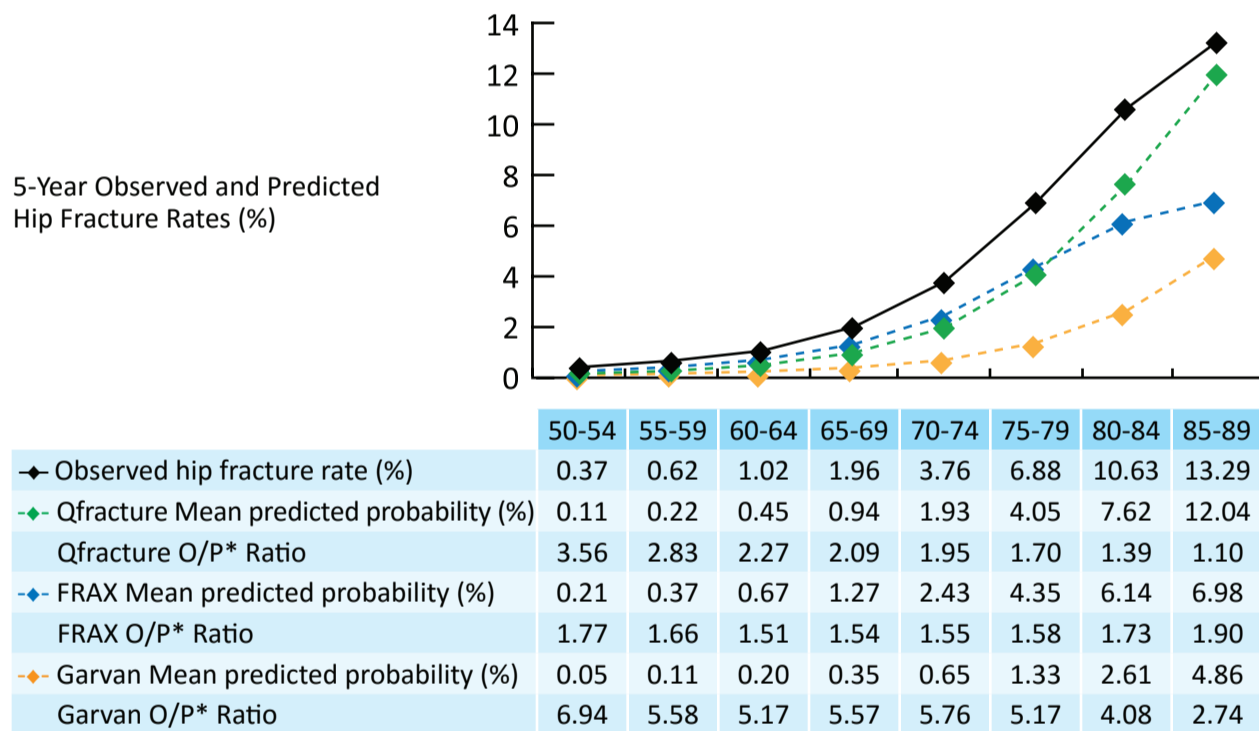
To test which of the three well-established hip-fracture prediction tools, Garvan, FRAX and QFracture, provides the greatest discrimination and calibration when cross-validated using a single integrated EHR system

METHODS

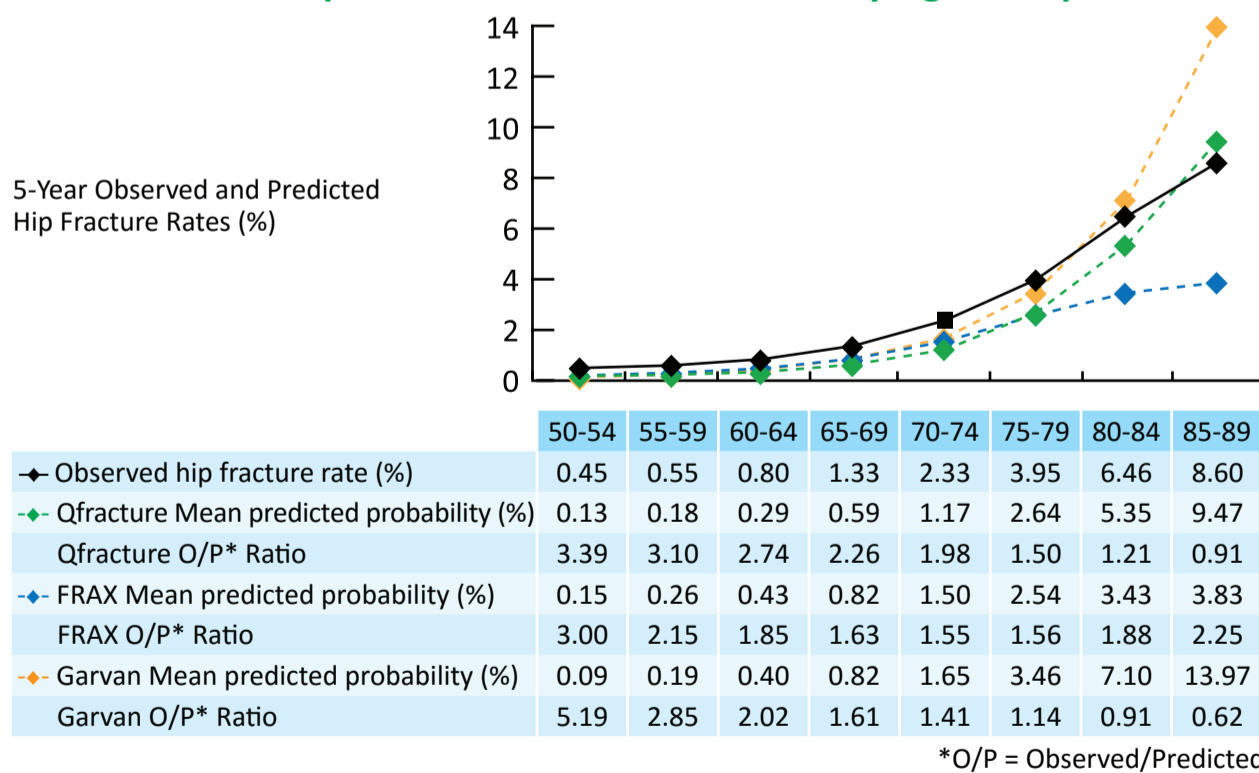
- » A retrospective cohort study among 989,291 Clalit (insurer and provider for >50% of the Israeli population) members aged 50-90 years
- » All variables were extracted from the comprehensive Clalit EHR system, which is universally deployed in all Clalit clinics
- » Five-year hip fracture estimated probabilities were calculated according to the Garvan, FRAX and QFracture tools, as of January 1, 2010 (index date)
- » Hip fracture events were measured in the 5-year follow-up between 2010-2014
- » Each of the tools' performance was evaluated for the following aspects:
 - Discrimination – AUC, sensitivity, specificity, PPV, NPV measures
 - Calibration – Observed-to-predicted ratio measure

RESULTS - CALIBRATION

Calibration of Hip Fracture Prediction Models by Age Groups - Females



Calibration of Hip Fracture Prediction Models by Age Groups - Males



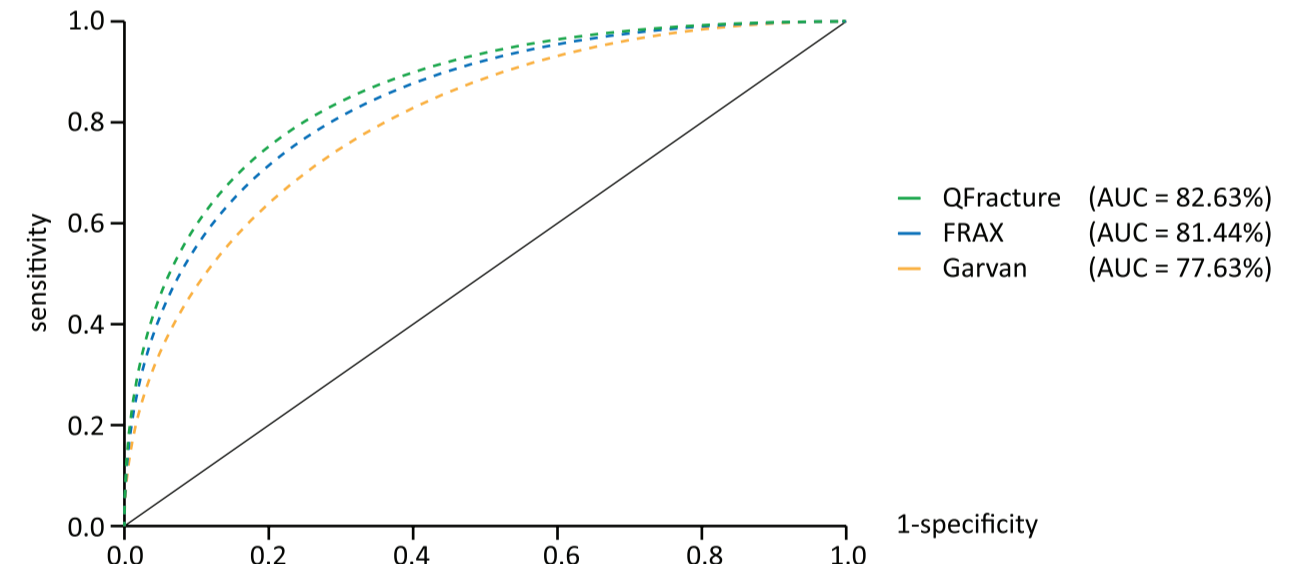
- » A majority of the observed-to-predicted ratios for hip fractures were >1, indicating underestimation of the risk by all three models
- » The QFracture and Garvan ratios presented a consistent downward trend with the increase in age groups, but were steadier across age groups for FRAX

RESULTS - DISCRIMINATION

Hip Fractures Prediction Performance Measures

	Garvan		FRAX		QFracture	
AUC (%)	77.63		81.44		82.63	
	Top 10% risk	Top 20% risk	Top 10% risk	Top 20% risk	Top 10% risk	Top 20% risk
Sensitivity (%)	36.90	56.67	43.85	66.02	45.25	68.65
Specificity (%)	90.74	81.01	90.91	80.98	90.97	81.34
PPV (%)	9.88	7.58	11.72	8.71	12.11	9.19
NPV (%)	98.12	98.55	98.33	98.86	98.37	98.95
Lift	3.69	2.83	4.38	3.25	4.52	3.43

Hip fracture prediction models - ROC curves



- » Both QFracture and FRAX had high discriminatory power, with QFracture performing slightly better
- » By targeting those with the top 20% hip fracture risk scores for prevention, QFracture, FRAX and Garvan would capture 68.7%, 66.0% and 56.7%, respectively, of those who suffered hip fractures
- » The specificity and NPV values were high and similar for all three models

CONCLUSIONS

- » This is the first study to directly compare three leading hip-fracture prediction tools and externally validate QFracture outside of the United Kingdom
- » Two international fracture prediction tools, QFracture and FRAX, performed well when externally validated in an EHR setting, although both tools would require local calibration
- » When EHR implementation is the chosen method for hip-fractures prediction, and all of the QFracture input variables are available within the system, QFracture may be the preferred tool, as it offers the best discrimination
- » For manual physician use, or when some of the QFracture input variables are not available within the EHR, the simpler FRAX which performs almost as well, may be a more suitable tool