# The Incorporation of Support Vector Machines and Hip Geometric Structure Assessments in the Development of Hip Fracture Risk Prediction Model

*Zhao Chen[1], Peng Jiang[2], Chengcheng Hu[1], Leslie Arendell[1], Skye Nicholas[1], John Robbins[3] and Samy Missoum[2]*

[1]Mel and Enid Zuckerman College of Public Health, University of Arizona, Tucson, AZ. [2]Aerospace and Mechanical Engineering Department, University of Arizona, Tucson, AZ. [3]University of California at Davis School of Medicine, Sacramento, CA .
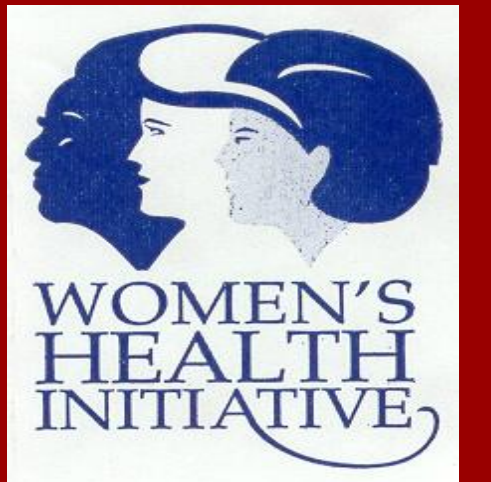
## Introduction

Hip fracture prediction models, such as FRAX, have been developed and widely used. Robbins and colleagues (2007) have shown that in the Women's Health Initiative (WHI), collectively, 11 predictors may generate a prediction with area under the curve (AUC) being 80% for 5-year hip fracture risk. The prediction property of this model was less accurate (AUC% = 71%) in the participants of the WHI bone mineral density (BMD) cohort, but when BMD was added into the model, AUC% reached 80%.

## Objective

❖ **Investigate the utility of support vector machines (SVM) technique and hip geometric structure assessments in predicting 10-year hip fracture risk in the participants of the WHI bone mineral density cohort.**

**Figure 1: SVM classifying WHI OS cohort by weight, age, and height**
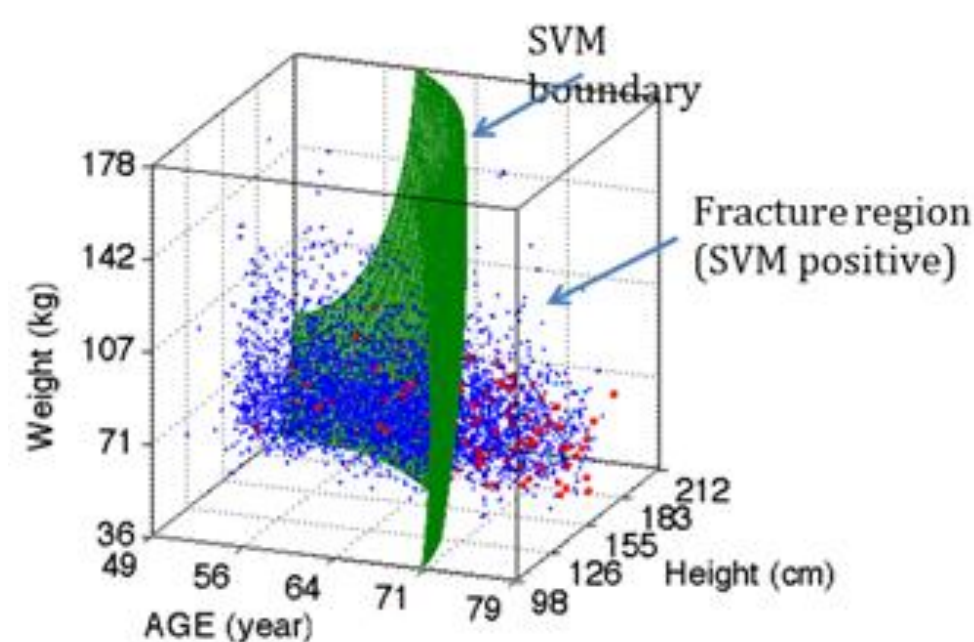


## Methods

**Study Participants:**

❖ Participants in the clinical trial (CT) arm (n= 5,016) and observational study (OS) arm (n=6,224) from the WHI BMD sub-cohort were selected for this analysis.

**Hip Fracture and other covariate assessment:**

❖ All hip fractures were adjudicated.

❖ Weight and height were measured in the clinic.

❖ Other covariates, including age, race/ethnicity, health status, fractured after age 55, physical activity, smoking status, a parent broke a hip, corticosteroid use, and diabetes treatment, were self-reported.

**HSA (T Beck method):**

❖ Hip structural analysis (HSA) program was applied to archived DXA scans to estimate BMD and structural parameters (BMD [g/cm2], cross-sectional area [cm$^2$], outer diameter [cm], section modulus [cm$^3$], cortical thickness [cm], buckling ratio) at 3 femur cross-sections (narrow neck; intertrochanter; shaft)

**Support Vector Machines (SVMs):**

❖ SVM is a powerful classification technique used to extract, in a high dimensional space, complex nonlinear "decision functions" between two classes of data. It does not have any model assumptions and is often referred as a machine learning approach. (Figure 1)

**Logistic Regression:**

❖ Logistic regression was used to build a prediction model for 10-year hip fracture risk. The model was fitted on the OS cohort with all 11 predictors identified in Robbins et al. (2007) and used to predict 10-year hip fracture risk for the CT subjects.
❖ Area under the receiver operating characteristic curve (AUC) was calculated based on the predicted probabilities and actual observations.

The same procedure was repeated for logistic models expanded to include baseline hip BMD and other hip geometry variables.

**Imputation:**
❖ Any continuous missing data were imputed using the median of all observed values . For missing categorical predictors a separate category called "missing" was assigned .

## Results

**Table 1: Demographics**

| | Clinical Trial Cohort | | Observational Study Cohort | |
| --- | --- | --- | --- | --- |
| | N=5016 (%) | | N=6224 (%) | |
| | Fracture | No Fracture | Fracture | No Fracture |
| Age, y | 110 | 4906 | 158 | 6066 |
| Mean (SD) | 68.74 (6.47) | 62.67 (7.25) | 68.54 (6.45) | 63.56 (7.41) |
| Height, cm | 110 | 4899 | 158 | 6009 |
| Mean (SD) | 161.84 (6.27) | 161.75 (6.15) | 162.15 (7.29) | 161.46 (6.45) |
| Weight, kg | 110 | 4905 | 158 | 6044 |
| Mean (SD) | 70.17 (12.42) | 75.89 (15.67) | 67.67 (12.18) | 72.67 (16.89) |
| Race/ethnicity[a] | 110 | 4900 | 158 | 6053 |
| Non-Hispanic white | 102 (92.73) | 3803 (77.61) | 144 (91.14) | 4617 (76.28) |
| African American | 2 (1.82) | 751 (15.33) | 6 (3.80) | 813 (13.43) |
| Hispanic | 4 (3.64) | 272 (5.55) | 4 (2.53) | 464 (7.67) |
| Self-reported health[b] | 105 | 4762 | 157 | 6007 |
| Good | 44 (41.90) | 1823 (38.28) | 68 (43.31) | 2116 (35.23) |
| Fracture on or after age 55 | 71 | 2788 | 131 | 4616 |
| Yes | 13 (18.31) | 448 (16.07) | 45 (34.35) | 792 (17.16) |

[a] Other race/ethnicities groups (American Indians, Asians, and other) were ≤8%of the study population.
[b] Most participants reported their health was good versus to poor, fair, very good, or excellent.

**Figure 2: Predicted probability in clinical trial arm based on 11 predictors identified by Robbins and colleagues (2007) plus BMD**
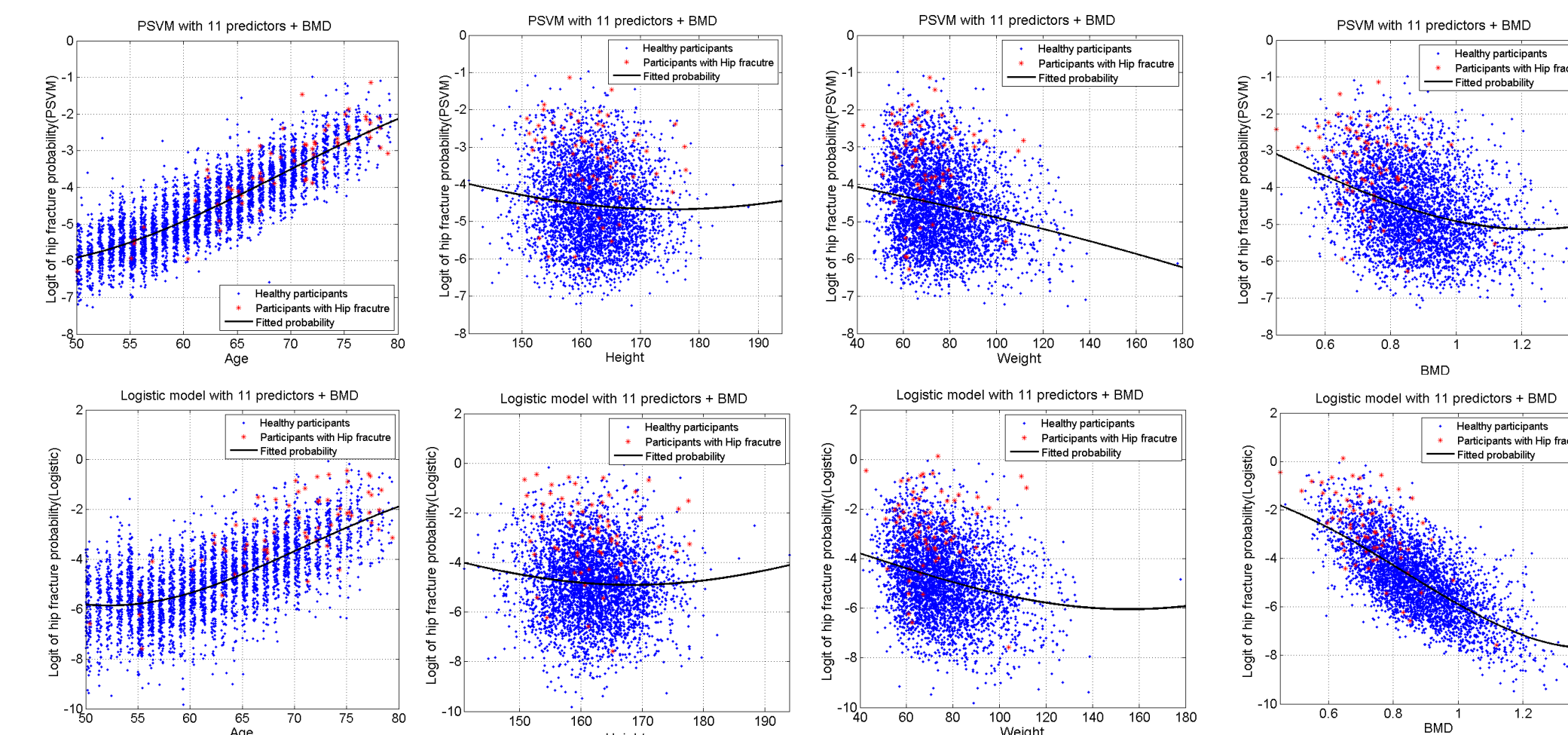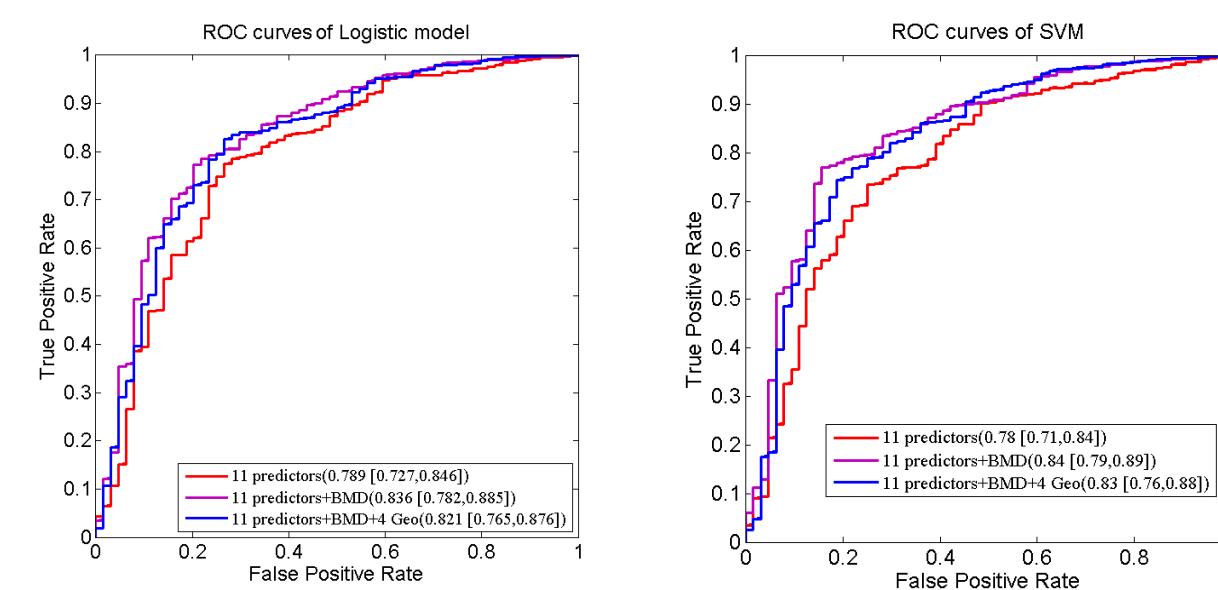


**Table 2: 10 year prediction with imputation**

| | Observational Study Cohort | | Clinical Trial Cohort | |
| --- | --- | --- | --- | --- |
| | #= 113/3892 | | #= 64/3768 | |
| | AUC | 95% CI | AUC | 95% CI |
| SVM with 11 predictors | 0.79 | [0.74, 0.83] | 0.78 | [0.72, 0.84] |
| SVM with 11 predictors + BMD | 0.84 | [0.80, 0.87] | 0.84 | [0.79, 0.89] |
| SVM with 11 predictors + BMD + 4 hip geometry parameters | 0.84 | [0.80, 0.88] | 0.83 | [0.76, 0.88] |

**Figure 3: ROC**



## Conclusion

❖ The preliminary results suggest that 10-years hip fracture risk may be well predicted by clinical risk factors with or without BMD.
❖ SVM and HSA parameters don't seem to add significant prediction value for 10-years hip fracture risk prediction.
❖ This study is limited by the number of hip fractures in the model building (observational study) and model validation (clinical trial) cohorts.

## Reference

Robbins et al. Factors Associated With 5-Year Risk of Hip Fracture in Postmenopausal Women. JAMA. 2007;298(20):2389-2398.

## Acknowledgements