National Institute on Aging

FNIH
Foundation for the National Institutes of Health

Future Scenarios of Algorithm Building

With Illustration in the NIA/FNIH
Sarcopenia Definitions and Outcomes Project

Thomas G Travison, PhD
Hebrew SeniorLife Institute for Aging Research
Harvard Medical School

State of Play in Non-alcoholic Steatohepatitis (NASH)

- Clinical gold standard (liver biopsy & histology) invasive and expensive

- Phenotype of NASH may be described by combination of pathological + clinical characteristics
  - Circulating analytes
  - Metabolic / prediabetic illnesses
  - Proteomics, metabalomics
  - Imaging

Objective

- Develop model-based / algorithmic equation or decision rule integrating some or many factors sensitive and specific for presence of NASH

- ‘Dual purpose’ of
  - Individual level diagnosis
  - Population-level research / monitoring (‘on average’)

The dual-purpose problem

- Classical methods (e.g. linear / nonlinear regression) adept at describing central tendency
  - 'Typical' ~ average

- Not everyone is typical; indeed the atypical patient may be the most important or informative

Tension between population trends and inter-individual variation

ORIGINAL RESEARCH

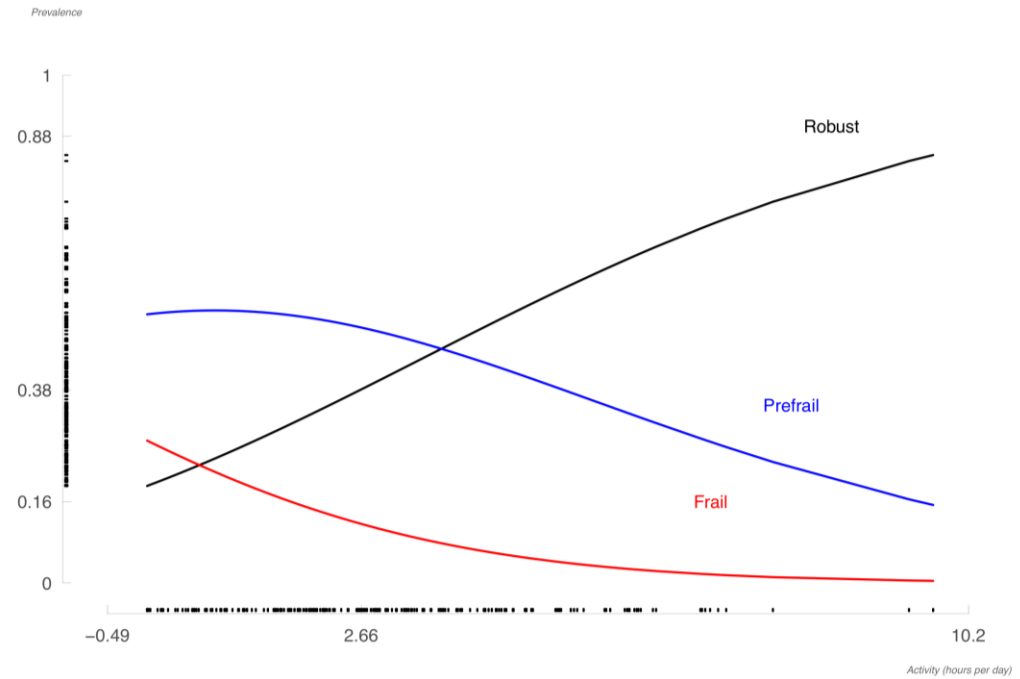
American Heart Association | American Stroke Association

Frailty, Physical Activity, and Mobility in Patients With Cardiac Implantable Electrical Devices

Daniel B. Kramer, MD, MPH; Timothy Tsai, MPH; Poorna Natarajan, MBBS; Elise Tewksbury, RN; Susan L. Mitchell, MD, MPH; Thomas G. Travison, PhD

- J Am Heart Assoc. 2017;6:e004659. DOI: 10.1161/JAHA.116.004659

Strong and consistently significant association at the mean

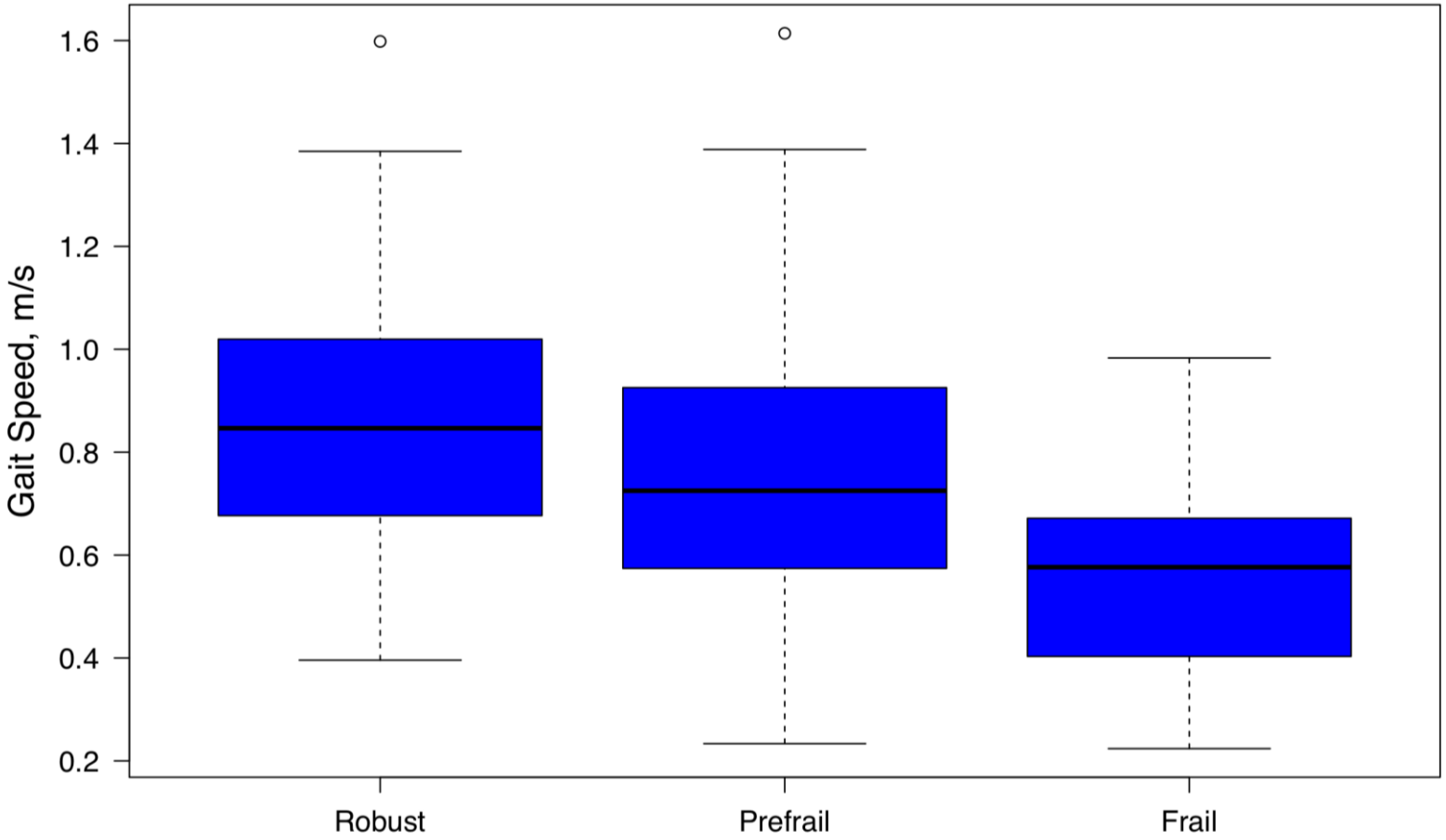


Multinomial Logistic Regression[a]

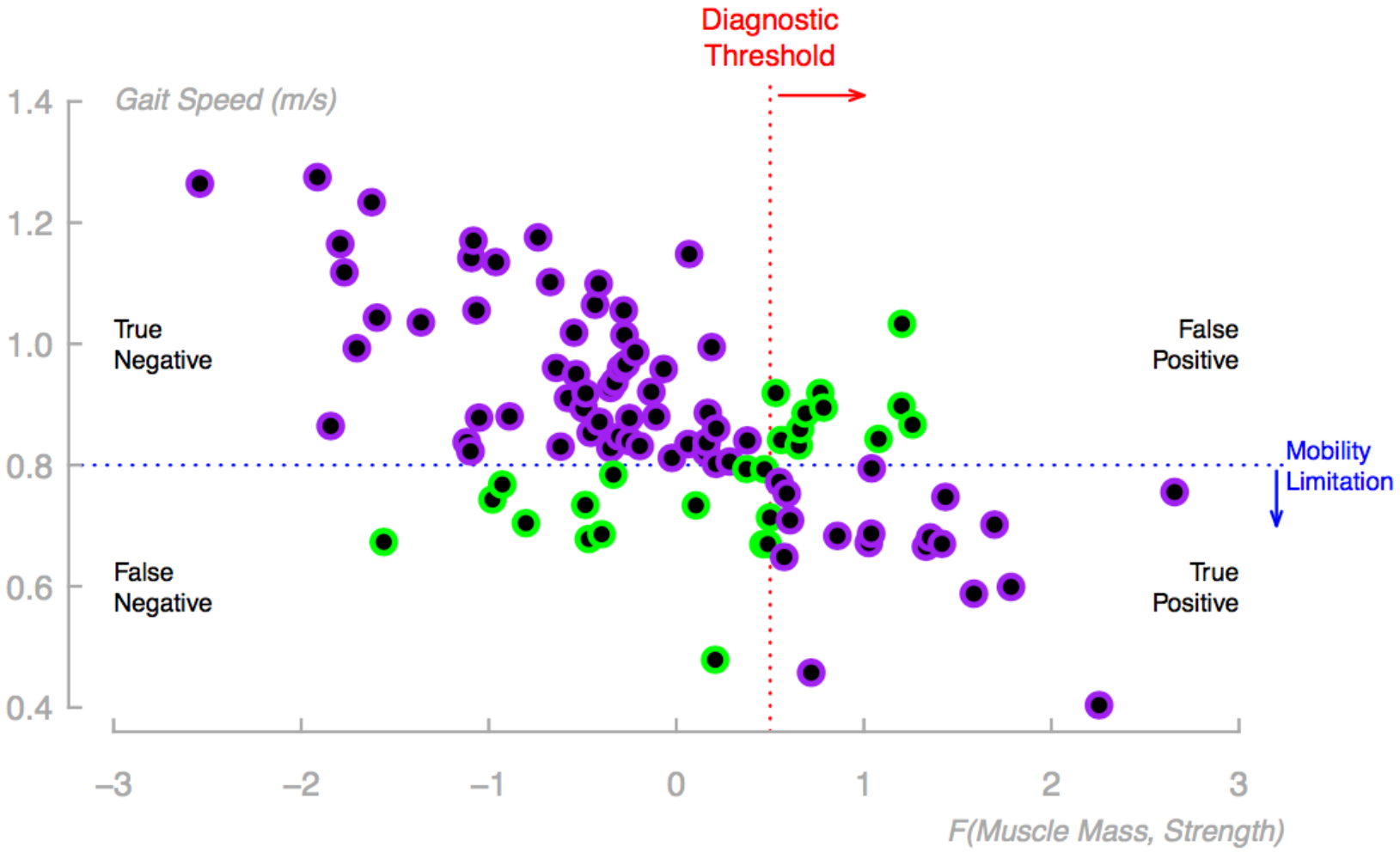
	Odds Ratio	95% Confidence Interval	p-val
Prefrail v. Robust	0.8	(0.68, 0.96)	0.015
Frail v. Robust	0.61	(0.44, 0.83)	0.002

[a] Adjusted for Age and Body Mass Index

But! Dramatic inter-individual variation might preclude use of activity as a sensitive patient-level classifier of frailty status



Performance of a diagnostic cutpoint: composite predictor



Addressing the dual-purpose problem

- Potential strategy: use machine-learning / partitioning methods to divide population into subgroups at substantially differential risk
  - Deals less with ‘on average’ and more with rankings and individual inputs
  - Allows for high-dimension comparison of correlated and near-collinear candidate phenotypes
  - Deals intrinsically with nonlinearities and complex interactions

- Caveat: patient-level inference dramatically influenced by data quality, precision
  - Necessitates attention to specifics of data collection and aggregation

Illustration: NIA/FNIH Sarcopenia Definition and Outcomes Project

"Analyses of Datasets on Older Populations with High Prevalence of Mobility Disability to Develop Clinically Meaningful Diagnostic Cut-Points for Low Muscle Mass and/or Low Muscle Strength"

Sarcopenia Definition and Outcomes Consortium Investigators

PIs: Peggy Cawthon and Shalender Bhasin

Steering Committee and Analysts: Sheena Patel, Rosaly Correa-De-Araujo, Roger Fielding, Jay Magaziner, Todd Manini, Thomas Travison, Karol Pencina, Hao Zhu

Project Investigators: Shehzad Basaria, Ellen Binder, Todd Brown, Ralph D'Agostino, Susan Greenspan, Tamara Harris, Douglas Kiel, Steven Kritchevsky, Joe Massaro, Robert McLean, Anne Newman, Denise Orwig, Marco Pahor, Adam Santanasto, Michelle Shardell, Qian-Li Xue

Other Contributing Investigators: Steven R. Cummings, Robert Cumming, Kristine Ensrud, Vasant Harini, Joanne Jordan, Magnus K. Karlsson, Timothy Kwok, Östen Ljunggren, Dan Mellström, Claes Ohlsson, Eric S. Orwoll, Laura Schaap, Jean Woo, Marjolein Visser

NIH National Institute on Aging

Funding from the National Institute on Aging and the Foundation for the NIH is gratefully acknowledged (U01)

FNIH Foundation for the National Institutes of Health

Conceptual Problem

- Sarcopenia (age-related loss of muscle mass / strength / quality) difficult to measure
- Little consensus as to which aspects or in what combination constitute the phenotype
- Potential solution: establish predictive validity for patient-important endpoint (physical functioning / mobility)

Overall goals

- To develop diagnostic cut-points* for low muscle mass and strength for identifying those at risk of mobility limitation
  - *May be defined as a general function / combination of multiple inputs
- To establish utility of diagnostic cut-points for prediction of clinically important outcomes including mortality and falls
  - Once again ‘dual – purpose’

	Participants	N
Osteoporotic Fractures in Men (**MrOS**) Study	Ambulatory community dwelling men, 65+ y	5,835
Study of osteoporotic fractures (**SOF**)	Ambulatory community dwelling women, 65+ y	1,246
Health, Aging and Body Composition Study (**Health ABC**) Study	Non-disabled black and white men and women, 70-80 y	1,398
MrOS Sweden	Men in three Swedish communities, 70+ y	2,876
Mr&MsOS Hong Kong	Men and women residing in Hong Kong, 65+ y	4,000
Concord Health and Aging in Men Project (**CHAMP**)	Men living near Concord, Australia, 65+ y	1,529
Cardiovascular Health Study (**CHS**)	Community dwelling men and women, 65+ y	1,509
Johnston County Arthritis Study	White and black residents of rural Johnston County, North Carolina, 45+ y	438

Participant characteristics

Characteristic	Men	Women
N	14157	5723
Age, yrs	75	75
White race	79%	44%
Weight, kg	78.7	64.2
Walk speed, m/s	1.16	0.92
Gait speed (4 or 6m) < 0.8 m/s	8.7%	25.8%
Maximum grip strength, kg	39.6	22.1
BMI (kg/m^2)	26.6	26.3
Mobility complaints*	12%	26%
Percent fat (%)	27	38
ALM/height^2 (kg/m^2)	7.88	6.47
ALM/BMI	0.89	0.61

*any difficulty walking or stair climbing

Analytic Challenges and Approaches

- Variation of measurements across study populations.
  - Harmonization of body composition (DEXA) and other measurements across community-dwelling cohorts
  - Appropriate scaling of contributing variables (e.g. muscle / lean body mass)

- Uncertainty as to ‘best’ candidate measurements to be included in sarcopenia phenotype
  - Consensus development of candidate measurements of sarcopenia as functions of muscle mass and strength

- Multidimensionality of candidate predictive factors
  - Exploratory (ROC) and machine-learning approaches to derive candidate cutpoints

- Performance of candidate phenotypes against ‘hard’ endpoints
  - Sample-based estimation and regression modeling with eye to ‘dual purpose’ needs

Reduction in overall variation: Cross-calibration of DXA Measurements

Able to harmonize to 'NHANES' calibration on Hologic machines.

(Handles systematic, non-stochastic differences in measurement)

Classic to NHANES Pre-APEX 3.3

$FATMass_{NHANES\_PreAPEX\ 3.4} = FAT_{Classic} + 0.054 \times (Lean\ SoftTissueMass_{Classic} + BMC)$

$Lean\ SoftTissueMass_{NHANES\_PreAPEX\ 3.4} = TotalMass - BMC - FATMass_{NHANES\_PreAPEX\ 3.4}$

Classic to NHANES

$FATMass_{NHANES} = FATMass_{Classic} + 0.054 \times Lean\ SoftTissueMass_{Classic}$

$Lean\ SoftTissueMass_{NHANES} = TotalMass - BMC - FATMass_{NHANES}$

Shepherd et al JBMR 2012.

Data-driven approach for analytes

- Handles inter-cohort variation via statistical models of association, where systematic differences in assay methodology cannot be quantified by equations

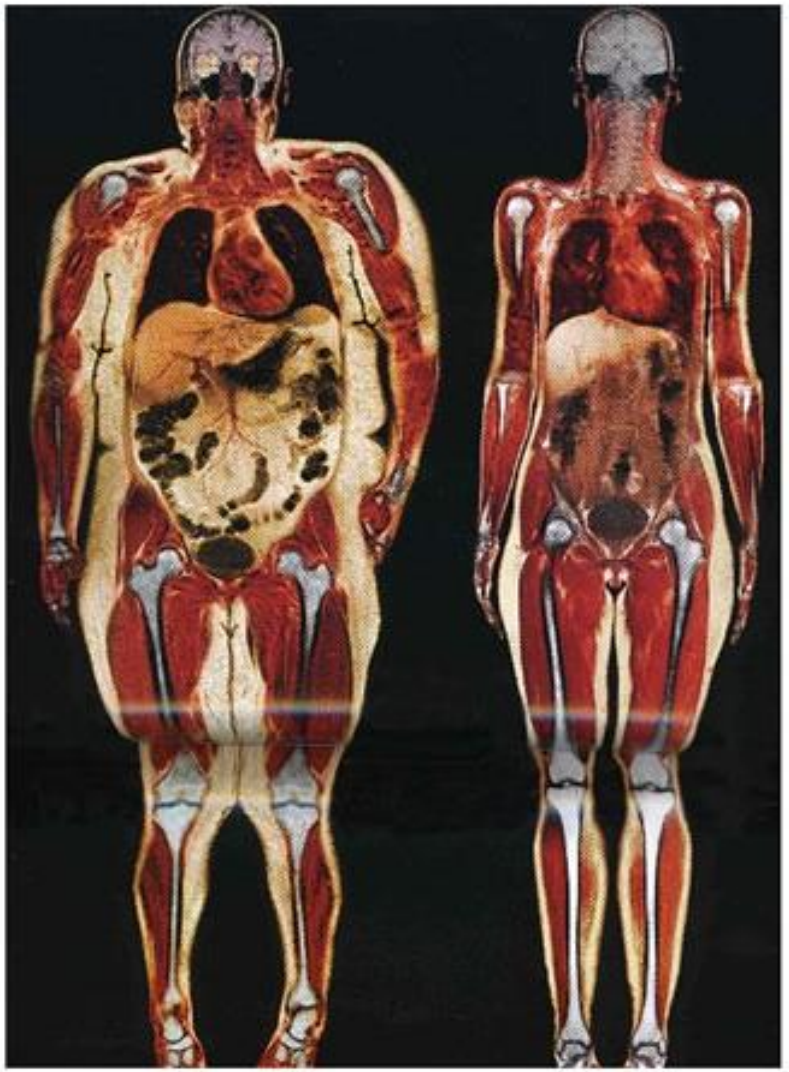
Harmonized Reference Ranges for Circulating Testosterone Levels in Men of Four Cohort Studies in the United States and Europe

Thomas G. Travison,[1] Hubert W. Vesper,[3] Eric Orwoll,[4] Frederick Wu,[5] Jean Marc Kaufman,[6] Ying Wang,[4] Bruno Lapauw,[6] Tom Fiers,[7] Alvin M. Matsumoto,[8] and Shalender Bhasin[2]

J Clin Endocrinol Metab, April 2017, 102(4):1161–1173

Defining muscle mass contribution – allometric scaling

- Absolute and proportionate scaled to:
  - Body mass
  - BMI
  - body surface area
  - height
  - height-squared or other
  - total & percent fat mass
- Consistency of relative lean mass or strength across populations and ranges of body size



Analytic Challenges and Approaches

- Variation of measurements across study populations.
  - Harmonization of body composition (DEXA) and other measurements across community-dwelling cohorts
- Uncertainty as to ‘best’ candidate measurements to be included in sarcopenia phenotype
  - Consensus development of candidate measurements of sarcopenia as functions of muscle mass and strength
- Multidimensionality of candidate predictive factors
  - Exploratory (ROC) and machine-learning approaches to derive candidate cutpoints
- Performance of candidate phenotypes against ‘hard’ endpoints
  - Sample-based estimation and regression modeling with eye to ‘dual purpose’ needs

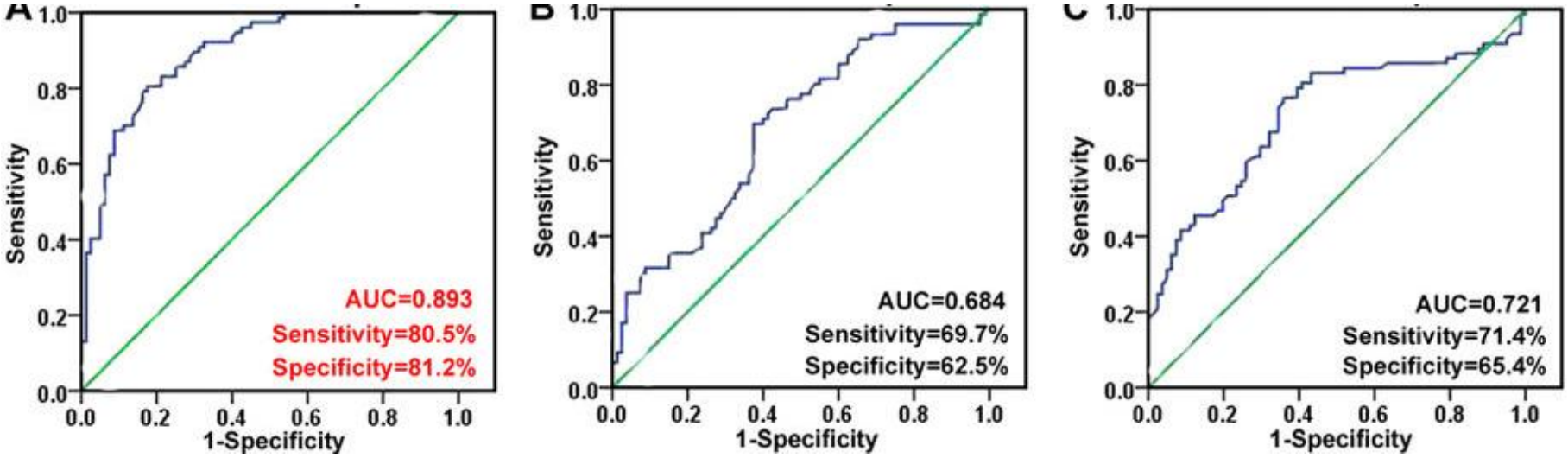
36 Candidate sarcopenia variables

	Variable
1	AGE
2	ALM
3	Maximum grip strength
4	ARM lean mass (ARMLM), kg
5	Appendicular fat mass (AFM), kg
6	Appendicular lean mass/Height
7	Appendicular lean mass/BMI
8	Appendicular lean mass/Total body fat
9	Appendicular lean mass/Weight
10	Appendicular lean mass/ height**2, (kg/m2)
11	Leg lean mass (LLM), kg
12	Leg lean mass/total body weight
13	Leg lean mass/Height
14	Leg lean mass/ Height**2
15	Leg lean mass/BMI
16	Leg lean mass/Total body fat
17	Maximum grip strength/ALM
18	Maximum grip strength/weight

	Variable
19	Maximum grip strength/BMI
20	Maximum grip strength/Height
21	Maximum grip strength/ Height**2
22	Maximum grip strength/Total body fat
23	ALM/Body Surface Area
24	Maximum grip strength/Body Surface Area
25	Maximum grip strength/LLM
26	LLM/Body Surface Area
27	ALM/Percent Fat
28	Maximum grip strength/Percent fat
29	ALM/AFM
30	Maximum grip strength/AFM
31	Maximum grip strength/ARMLM
32	Maximum grip strength/ARMLM**2
33	Height
34	Weight
35	Total body fat
36	BMI

Receiver operating curve (AUC)

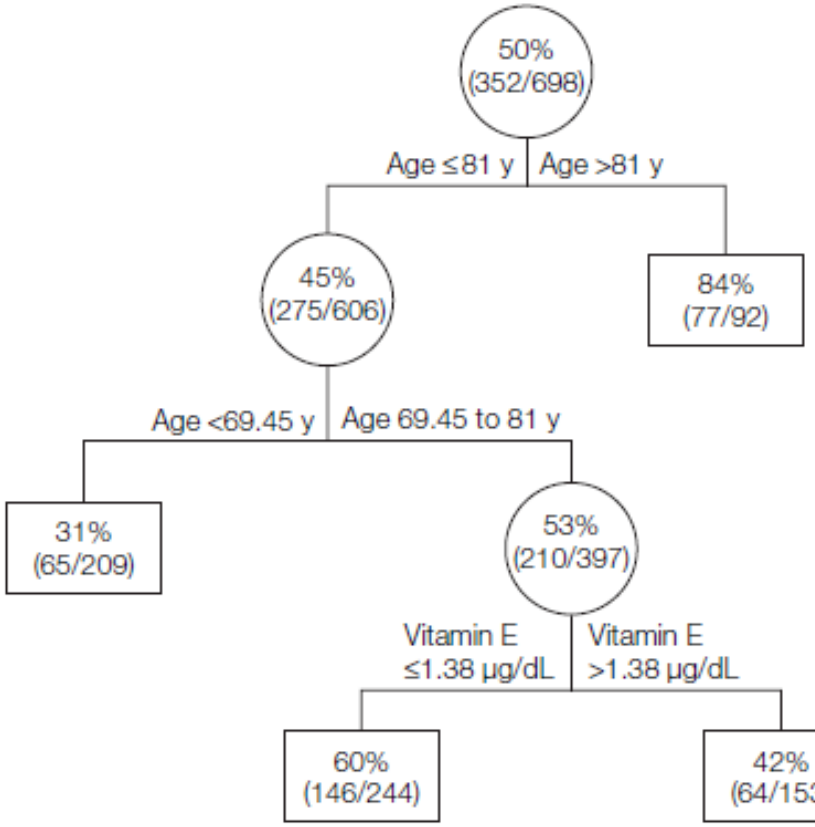
- Area under (ROC) curve used to compare the accuracy of “diagnosing” low walk speed (<0.80 m/sec) Traditional accuracy ranking:
  - 1.0-0.90=excellent; 0.80-0.89=good; 0.79-0.70=fair; <0.70=poor



Methods: Machine Learning

- Classification and regression tree (CART) analysis
  - Designed to identify subgroups that maximize outcome homogeneity
  - Produces mutually exclusive & exhaustive population subgroups whose members are most similar with respect to the outcome.

Example: Detecting complex interactions

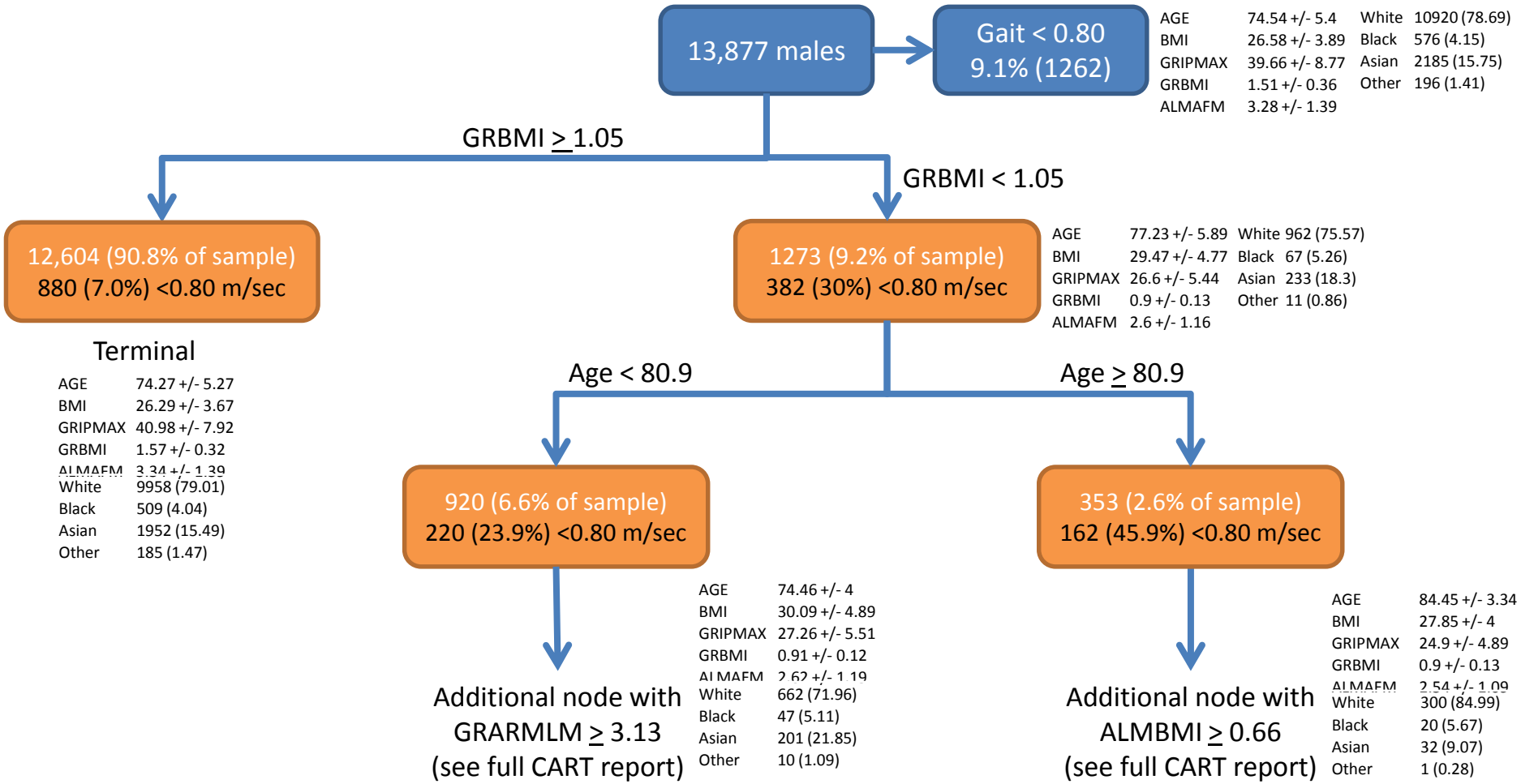


CART: decline in physical performance

Considered the following factors: α-tocopherol, vitamin B12, vitamin B6, folate, 25-hydroxyvitamin D, iron, age, sex, educational achievement, marital status, household composition, smoking, physical activity level, number of chronic conditions, body mass index, depression and cognitive function

Bartali et al., *JAMA* 2008.

Sarcopenia Results: MEN

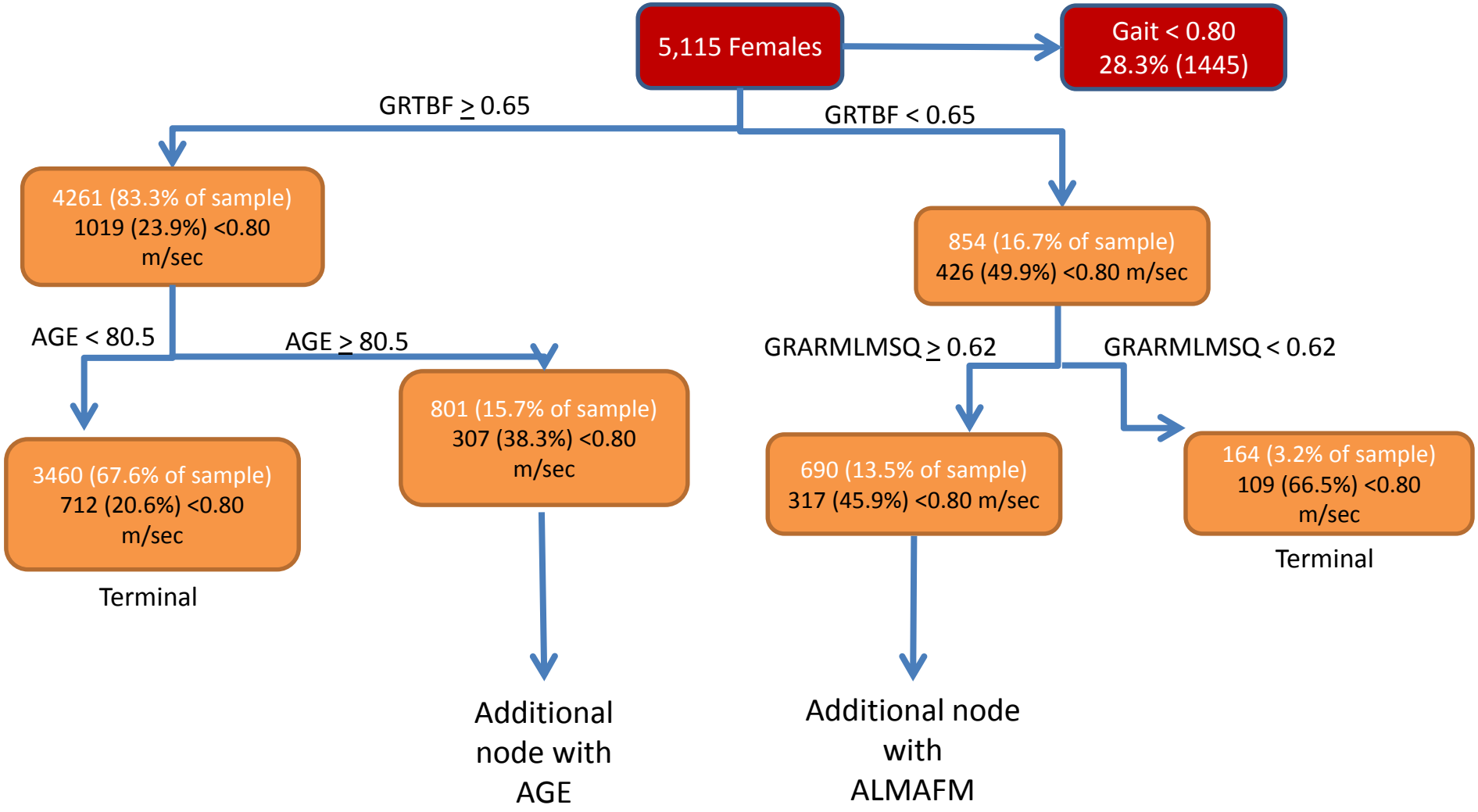


Men

- Those with GRBMI values less than 1.05 were much older and had much higher BMI and much lower grip than those with GRBMI>1.05

- Age <83.5 further discriminates slowness, although the group that the age cut-point applies to (about 10%) is a fairly small subset of the participants

Results: Women



Women

- Grip strength/total body fat is the first discriminator
  - splits the sample such that 83% has higher strength to total body fat and 17% had lower strength to total body fat

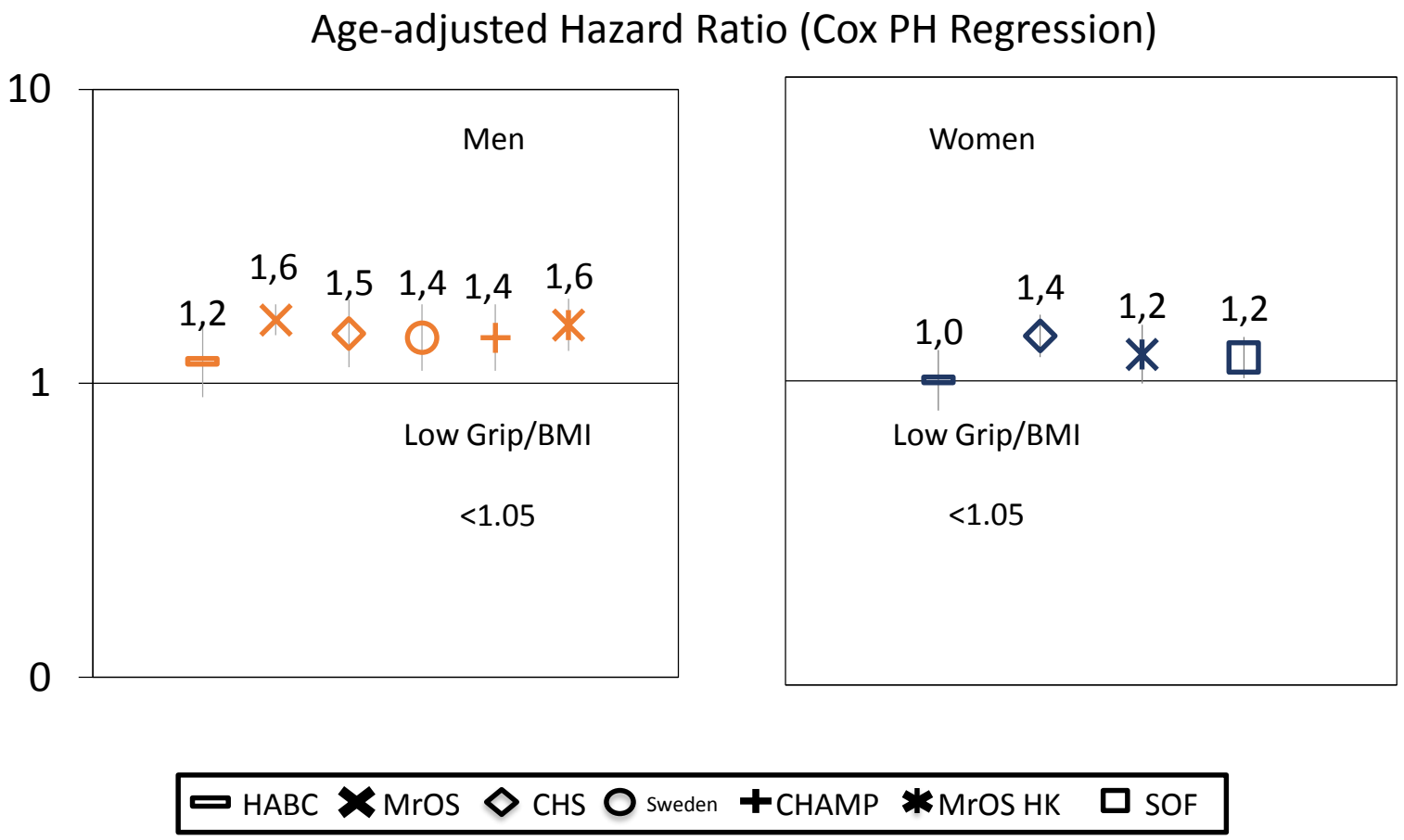
- GRARMLMSQ < 0.62 further discriminated women with low strength to total body fat (3.2% of sample)

- AGE $\geq$ 80.5 further discriminated women with high strength to total body fat (15.7% of sample)

Analytic Challenges and Approaches

- Variation of measurements across study populations.
  - Harmonization of body composition (DEXA) and other measurements across community-dwelling cohorts
- Uncertainty as to ‘best’ candidate measurements to be included in sarcopenia phenotype
  - Consensus development of candidate measurements of sarcopenia as functions of muscle mass and strength
- Multidimensionality of candidate predictive factors
  - Exploratory (ROC) and machine-learning approaches to derive candidate cutpoints
- Performance of candidate phenotypes against ‘hard’ endpoints
  - Sample-based estimation and regression modeling with eye to ‘dual purpose’ needs

10 y Mortality: Low Grip/BMI By Cohort



Conclusions / Lessons Learned

- Combination of classical / and computing-based approaches may yield dividends
  - Attention to definitions, harmonization, and data quality is critical
- The ‘dual-purpose’ problem is highly context dependent
  - Individual-level inference relies critically on calibration of phenotypes
- Modern methods cannot solve – and may exacerbate – tension between complexity and clinical utility
  - Heuristic underpinning of results may not be clear
    - Partially addressed by profiling individuals in specific risk categories
  - Association with other endpoints may not ‘validate’ results to sufficient satisfaction
  - Sarcopenia: Controversy over exclusion of direct measure of muscle in main results (To be continued…)

Acknowledgments

- DEXA Harmonization
  - Karol Pencina, Brigham and Women's Hospital
  - Kevin Wilson, Tom Kelly; Hologic, Inc.

- Allometric scaling
  - Todd Manini; University of Florida

- Machine Learning analyses:
  - Peggy Cawthon, Sheena Patel; UCSF, San Francisco Coordinating Center

We gratefully acknowledge funding for this project

National Institute on Aging
Foundation for the NIH

Funding for cohort studies included in this presentation was provided by NIA, NIAMS, NINDS, NCATS, NHLBI, NIH Roadmap, Research Grants Council (Hong Kong), The Chinese University of Hong Kong, Swedish Research Council, the Swedish Foundation for Strategic Research, the ALF/LUA research grant in Gothenburg, the Lundberg Foundation, the Torsten and Ragnar Söderberg's Foundation, Petrus and Augusta Hedlunds Foundation, the Västra Götaland Foundation, the Göteborg Medical Society, the Novo Nordisk Foundation, National Health and Medical Research Council (Australia), Ageing and Alzheimer's Institute (Australia)

NIH grant and contract numbers: AG051421 , AG027810, AG042124, AG042139, AG042140, AG042143, AG042145, AG042168, AR066160, TR000128, AR049439-01A1, AG005407, AR35582, AR35583, AR35584, AG005394, AG027574, AG027576, N01AG62101, N01AG62103, N01AG62106, N01-HC- 85079, N01-HC-85080, N01-HC-85081, N01-HC-85082, N01-HC-85083, N01-HC-85084, N01-HC-85085, N01-HC-85086; N01-HC-35129, N01 HC-15103, N01 HC-55222, N01-HC-75150, N01-HC-45133, N01-HC-85239, HHSN268201200036C, HL080295, HL087652, HL105756, HL103612, AG023629, AG15928, AG20098, AG027058