

BRIEF CUTTING EDGE REPORT

Title: Bias measurement in, bias results out: how an assumption free height adjusted weight model outperforms body mass index

Megan M. Shuey;¹ Shi Huang;² Rebecca T. Levinson;¹ Eric Farber-Eger;¹ Katherine N. Cahill;¹ Joshua A. Beckman;¹ John R. Koethe;¹ Heidi J. Silver;^{1,3} Kevin D. Niswender;^{1,3} Nancy J. Cox;¹ Frank E. Harrell Jr.;² Quinn S. Wells^{1,4,5*}

¹Department of Medicine, Vanderbilt University Medical Center, Nashville, TN, USA;

²Department of Biostatistics, Vanderbilt University School of Medicine, Nashville, TN, USA

³Department of Veteran Affairs, Tennessee Valley Healthcare System, Nashville, TN, USA

⁴Department of Pharmacology, Vanderbilt University Medical Center, Nashville, TN, USA

⁵Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, TN, USA

Keywords: body mass index, prediction modeling, metabolic syndrome, cardiovascular disease

Running Title: BMI and height-adjusted weight prediction models

*Corresponding author

Quinn S. Wells, Pharm.D., M.D.

Departments of Medicine, Pharmacology, and Biomedical Informatics

Vanderbilt University Medical Center, 2525 West End Avenue, Suite 300

Nashville, TN, 37203, USA

Phone: (615) 936-1819

Email: quinn.s.wells@vumc.org

Word count: 1496

Funding: The dataset used for the analyses described was obtained from the Vanderbilt University Medical Center Synthetic Derivative, which is supported by institutional funding, the 1S10RR025141-01 instrumentation award, and by the CTSA grant UL1TR000445 from National Center for Advancing Translational Sciences/National Institutes of Health. The work was supported by an American Heart Association strategically focused research network award, 17SFRN33520017. In addition, FE Harrell's work on this paper was supported by CTSA award No. UL1 TR002243 from the National Center for Advancing Translational Sciences. Its contents are solely the responsibility of the authors and do not necessarily represent official views of the National Center for Advancing Translational Sciences or the National Institutes of Health.

Disclosures: J. Beckman declares potential conflicts with Amgen, Bayer, Glaxo Smith Kline, Janssen, Novartis, and Sanofi in the past twelve months. All other authors declare no conflicts of interest.

Author contributions: MMS, SH, FJH, and QSW were involved in the conception, design, analysis, and interpretation of data relating to the study as well as drafting of the initial

manuscript. MMS, SH, RTL, EFE, and QSW were involved in the data acquisition while KNC, JAB, JRK, HJS, NJC, and KDN were also involved in the interpretation of results. All authors were involved in the review and revision of the manuscript as well as approval for submission.

Study Importance Questions

What is already known about this subject?

- Body mass index, derived from collected height and weight measures, is an imperfect proxy measure of body fat composition often used in medical research.

What are the new findings in your manuscript?

- We demonstrate how BMI introduces complex non-uniform biases across outcome and height-weight space.

How might your results change the direction of research or the focus of clinical practice?

- Modeling height and weight as separate, non-linear, interacting variables improves clinical prediction across the complete spectrum of heights and weights for all clinical outcomes.

Abstract

Objective

Body mass index (BMI) is the most commonly used predictor of weight-related comorbidities and outcomes. However, the presumed relationship between height and weight intrinsic to BMI may introduce bias with respect to prediction of clinical outcomes. Using Vanderbilt University Medical Center's deidentified electronic health records and landmark methodology, we performed a series of analyses comparing the performance of models representing weight and height as separate interacting variables to models using BMI.

Methods

Model prediction was evaluated with respect to established weight-related cardiometabolic traits, metabolic syndrome and its components hypertension, diabetes mellitus, low high-density lipoprotein, and elevated triglycerides, as well as cardiovascular outcomes, atrial fibrillation, coronary artery disease, heart failure, and peripheral artery disease. Model performance was evaluated using likelihood ratio, R^2 , and Somers' Dxy rank correlation. Differences in model predictions were visualized using heatmaps.

Results

Regardless of outcome, the maximally flexible model had a higher likelihood ratio, R^2 , and Somers' Dxy rank correlation for event-free prediction probability compared to the BMI model. Performance differed based on the outcome and across the height and weight range.

Conclusions

Compared to BMI, modeling height and weight as independent, interacting variables results in less bias and improved predictive accuracy for all tested traits.

Introduction

Both underweight and overweight individuals are at increased risk for adverse health outcomes and mortality.¹⁻⁴ The most common proxy-measure of body fat in the clinical setting is body mass index (BMI),^{5,6} a derived value where height and weight are assumed to act according to the fixed relation of weight divided by the square of height (kg/m^2). While BMI has been consistently associated with multiple weight-related outcomes,^{7,8} the intrinsic assumptions of BMI may result in limitations and biases as a predictive variable.^{9,10} For example, recent work has demonstrated that BMI undervalues the predictive potential of height for blood pressure variation and the addition of height to measures such as waist circumference improves cardiometabolic risk prediction.¹¹⁻¹⁴

Because excess weight is known to increase risk for the components of metabolic syndrome as well as cardiovascular disease,^{4,15,16} we compared BMI to a maximally flexible, interacting model of the height/weight relationship with respect to prediction of these outcomes. Specifically, we evaluated model performance for metabolic syndrome, metabolic syndrome components, and a range of cardiovascular outcomes

Methods

Study population

All data were extracted from a de-identified copy of the Vanderbilt University Medical Center electronic health record on 08/2019.¹⁷ Measures of height and weight after 18 years of age were cleaned and units harmonized based on a previous method and BMIs were calculated.¹⁸

Subjects were included in landmark analyses based on a prespecified three-year qualification period that required four height and weight measures separated by approximately one year (1 year \pm 4 months). For each outcome, subjects were excluded if the first occurrence of the particular outcome was before or during the qualification period. Validated data extraction methods were used to define outcomes: low high-density lipoprotein (HDL < 40 mg/dL in males and < 50 mg/dL in females), elevated triglycerides (triglycerides > 150 mg/dL), hypertension, diabetes mellitus, atrial fibrillation (AF), coronary artery disease (CAD), heart failure (HF), and peripheral artery disease (PAD).^{19,20} Because waist circumference and fasting glucose are infrequently ascertained in the clinical setting, we used a modified definition of metabolic syndrome, defined as two or more of the following events: diabetes mellitus, hypertension, low HDL, or elevated triglycerides.

Statistical analyses

Descriptive statistics were presented as count and frequencies for categorical variables and median and interquartile range for continuous variables. Comparisons were made using Pearson's chi-squared or Wilcoxon signed-rank, as appropriate. Cox regression analyses were conducted to examine how weight and height at the end of the qualification period (i.e., t_3) impact the hazard of developing each outcome. Analyses utilized two models: a maximally flexible model representing height and weight (each log transformed) as separate, non-linear (restricted cubic spline with 3 knots) terms and allowing for interactions; and a log transformed BMI model. All analyses were adjusted for sex, age (with restricted cubic spline with 3 knots), and race. Model performances were evaluated by likelihood ratio, R square, and Somers' Dxy rank correlation (index of discrimination between predicted score and observed responses).

We estimated event-free probability at five years across a wide range of height-weight combinations (weight range: 50-200 kg by 5kg; height 160-200 cm by 5cm [total 270 predictions]). In predicting event-free probability at five years, patient demographic characteristics were set to the population's median age of 51.2 years of age, white race, and female sex. Heatmaps were presented to visualize predicted 5-year event-free probability and the discrepancy between the two models across a full range of heights and weights. All statistical analyses were performed with R (version 3.3.1).

Results

The demographics of included subjects are available in Tables S1 and S2. Metabolic syndrome event frequency was 16.7% and the frequencies for individual components were 5.5% for DM, 19.5%, for hypertension, 6.3% for low HDL, and 4.9% for elevated triglycerides. For cardiovascular outcomes, the event frequencies were 9.6% for AF, 21.8% for CAD, 8.8% for HF, and 2.8% for PAD

Performances of the two body composition models for each outcome are summarized in Table 1. Briefly, the maximally flexible height*weight model had a better log likelihood ratio, R^2 , and discrimination ability (Somers' Dxy) than BMI with a maximum difference in model performance of 47.046, 0.0008, and 0.003, respectively.

Figures 1 and 2 (columns 1 and 2) present heat maps displaying the pattern of predicted 5-year event-free probability for each outcome across the range of height and weight measures.

Similarly, heat maps were used to display discrepancies between model predictions ($\text{Predicted probability}_{\text{BMI}} - \text{Predicted probability}_{\text{height*weight}}$) (Figures 1 and 2, column 3). Colors indicate whether the BMI model predicts higher (blue) or lower (pink) probability than the height*weight model. In addition, for each outcome we present a histogram of prediction differences (Figures 1 and 2, column 4).

Discussion

We used electronic health record-derived data to conduct a comparative analysis of prediction models using either BMI or an unbiased height*weight interaction model. Our principal findings were that models with maximum flexibility outperform those using BMI across a wide range of cardiometabolic outcomes, and that discrepancies between models vary by outcome and location within the height-weight variable space.

Abnormal body composition is an important determinant of clinical outcomes, and accurately modeling the effect of height and weight on outcomes is critical for risk prediction. In this context, our findings have several important implications. First, although BMI is the most commonly used measure of body composition, it demonstrates inferior performance compared to a maximally flexible model across all outcomes, suggesting the assumed fixed relationship between weight and height (i.e., kg/m^2) inadequately represents the clinical impact of body composition. Second, BMI introduces complex non-uniform biases across outcome and height-weight space. For example, predicted risk for hypertension and diabetes mellitus is similar between both approaches with the exception of high body weight individuals with short stature where BMI significantly underestimates risk. In contrast, BMI introduces considerable error into

prediction of the features of atherogenic dyslipidemia (low HDL and high triglycerides), especially at higher weights, where the contribution of height is poorly modeled. Different patterns emerge for cardiovascular outcomes. BMI systematically overestimates the contribution of height for lower weight individuals and underestimates height for heavy individuals with respect to CAD, HF, and PAD risk. For instance, we observed a decrease in the predicted CAD event free risk in tall slender individuals, height > 180 cm and weight < 50 kgs, an observation that is missed by the BMI model. By comparison, BMI consistently overestimates the contribution of height for AF, though the bias is most pronounced for those at extremes of weight.

While the absolute magnitudes of discrepancies between flexible and BMI-based models were frequently modest, they are not clinically insignificant, as they are frequently not small compared to the absolute risk of the outcome in question. For example, for abnormal HDL, the BMI model overestimates the event free probability of individuals in tall patients with extreme obesity patients from 0.02 to 0.06 (2.0 to 6.0%). Considering the frequency of abnormal HDL in the total population is 6.3% this overestimate may be as frequent as the outcome alone. For abnormal HDL, a similar concern of underestimation arises in extremely heavy short patients.

The current analysis adds to the literature by systemically examining the limitations and biases of BMI across the height-weight space and a diverse set of cardiometabolic outcomes. Because BMI is calculated from height and weight, the use of BMI is a choice rather than an issue of data availability. The rationale for using BMI has often focused on the ease of calculation and interpretation, familiarity with its use among clinicians and scientists, and its established value as

a predictor. Here we show that allowing height and weight to “speak for themselves”, rather than be forced to exert effects through the BMI relationship, results in more accurate risk prediction across a range of conditions and outcomes. Thus, to the extent that more accurate risk prediction translates to improved patient care, future efforts should consider more flexible approaches to modeling height and weight.

There are limitations to this study. For example, known clinical predictors for the various outcomes were not included in models. However, this choice was made to allow characterization of how choice of body weight (i.e., BMI vs height*weight) impacts model performance across outcomes. As is common with use of electronic health record data there are always concerns related to data sparsity. While our study design did its best to minimize sparsity issues it remains possible that there is potential for confounding due to this, for example it is possible that a patient may have a particular outcome, however, it was missed either due to misclassification or its development outside of the follow-up period due to various reasons.

Conclusion

A data-driven, maximally flexible, log height-adjusted weight interaction model has better log likelihood for the prediction of weight-related outcomes than BMI. The prediction performance of these two models varies across the full spectrum of heights-weights and the absolute difference in model prediction may exceed the frequency of a given outcome. The scientific community should consider avoiding BMI when studying weight-related outcomes in favor of more flexible modeling strategies.

References

1. Laxy M, Teuner C, Holle R, Kurz C. The association between BMI and health-related quality of life in the US population: sex, age and ethnicity matters. *Int J Obes (Lond)*. 2018;42(3):318-326.
2. Folsom AR, Kaye SA, Sellers TA, et al. Body fat distribution and 5-year risk of death in older women. *JAMA*. 1993;269(4):483-487.
3. Singh PN, Lindsted KD, Fraser GE. Body weight and mortality among adults who never smoked. *Am J Epidemiol*. 1999;150(11):1152-1164.
4. Attard SM, Herring AH, Howard AG, Gordon-Larsen P. Longitudinal trajectories of BMI and cardiovascular disease risk: the national longitudinal study of adolescent health. *Obesity (Silver Spring)*. 2013;21(11):2180-2188.
5. Clinical guidelines on the identification, evaluation, and treatment of overweight and obesity in adults: executive summary. Expert Panel on the Identification, Evaluation, and Treatment of Overweight in Adults. *Am J Clin Nutr*. 1998;68(4):899-917.
6. Obesity: preventing and managing the global epidemic. Report of a WHO consultation. *World Health Organ Tech Rep Ser*. 2000;894:i-xii, 1-253.
7. Schneider HJ, Glaesmer H, Klotsche J, et al. Accuracy of anthropometric indicators of obesity to predict cardiovascular risk. *J Clin Endocrinol Metab*. 2007;92(2):589-594.
8. Global Burden of Metabolic Risk Factors for Chronic Diseases C. Cardiovascular disease, chronic kidney disease, and diabetes mortality burden of cardiometabolic risk factors from 1980 to 2010: a comparative risk assessment. *Lancet Diabetes Endocrinol*. 2014;2(8):634-647.

9. Jackson AS, Stanforth PR, Gagnon J, et al. The effect of sex, age and race on estimating percentage body fat from body mass index: The Heritage Family Study. *Int J Obes Relat Metab Disord.* 2002;26(6):789-796.
10. Garn SM, Leonard WR, Hawthorne VM. Three limitations of the body mass index. *Am J Clin Nutr.* 1986;44(6):996-997.
11. Michels KB, Greenland S, Rosner BA. Does body mass index adequately capture the relation of body composition and body size to health outcomes? *Am J Epidemiol.* 1998;147(2):167-172.
12. Mirzaei M, Khajeh M. Comparison of anthropometric indices (body mass index, waist circumference, waist to hip ratio and waist to height ratio) in predicting risk of type II diabetes in the population of Yazd, Iran. *Diabetes Metab Syndr.* 2018;12(5):677-682.
13. Chen X, Liu Y, Sun X, et al. Comparison of body mass index, waist circumference, conicity index, and waist-to-height ratio for predicting incidence of hypertension: the rural Chinese cohort study. *J Hum Hypertens.* 2018;32(3):228-235.
14. Lo K, Wong M, Khalechelvam P, Tam W. Waist-to-height ratio, body mass index and waist circumference for screening paediatric cardio-metabolic risk factors: a meta-analysis. *Obes Rev.* 2016;17(12):1258-1275.
15. Brown CD, Higgins M, Donato KA, et al. Body mass index and the prevalence of hypertension and dyslipidemia. *Obes Res.* 2000;8(9):605-619.
16. Sanada H, Yokokawa H, Yoneda M, et al. High body mass index is an important risk factor for the development of type 2 diabetes. *Intern Med.* 2012;51(14):1821-1826.

17. Roden DM, Pulley JM, Basford MA, et al. Development of a large-scale de-identified DNA biobank to enable personalized medicine. *Clin Pharmacol Ther.* 2008;84(3):362-369.
18. Jiang YBM, Kirby J, Wang X, Harris P, Denny J. Systematic and longitudinal approach to height, weight, and body mass index data cleaning for efficient reuse of EHR data for research. American Medical Informatics Association; 2015.
19. Beckman JA, Duncan MS, Damrauer SM, et al. Microvascular Disease, Peripheral Artery Disease, and Amputation. *Circulation.* 2019;140(6):449-458.
20. Aragam KG, Chaffin M, Levinson RT, et al. Phenotypic Refinement of Heart Failure in a National Biobank Facilitates Genetic Discovery. *Circulation.* 2018.

Tables

Table 1: Performance Statistics for Event Free Prediction Using the Maximally Flexible Log Height-adjusted Weight Interaction or Log Body Mass Index models

Prediction Model	LR	DF	R ²	DI
Metabolic Syndrome				
log height-adjusted weight interaction	3074.44	15	0.0424	0.3060
logBMI	3053.02	10	0.0418	0.3050
Diabetes Mellitus				
log height-adjusted weight interaction	2735.044	15	0.0431	0.4319
logBMI	2687.998	10	0.0423	0.4289
High Density Lipoprotein				
log height-adjusted weight interaction	1102.552	15	0.0161	0.2645
logBMI	1084.594	10	0.0159	0.2636
Hypertension				
log height-adjusted weight interaction	2900.190	15	0.0538	0.3221
logBMI	2877.066	10	0.0534	0.3214
Triglycerides				
log height-adjusted weight interaction	1134.645	15	0.0184	0.2981
logBMI	1112.862	10	0.0181	0.2959
Atrial Fibrillation				
log height-adjusted weight interaction	5024.154	15	0.0591	0.4244
logBMI	4983.270	10	0.0586	0.4237
Coronary Artery Disease				

log height-adjusted weight interaction	3375.364	15	0.0481	0.2752
logBMI	3347.507	10	0.0477	0.2743

Heart Failure

log height-adjusted weight interaction	5486.422	15	0.0665	0.4390
logBMI	5470.308	10	0.0664	0.4380

Peripheral Artery Disease

log height-adjusted weight interaction	2486.560	15	0.0503	0.5049
logBMI	2468.629	10	0.0500	0.5030

Abbreviations: BMI, body mass index; DF, degrees of freedom; DI, discrimination index (Somers' Dxy); LR, likelihood ratio

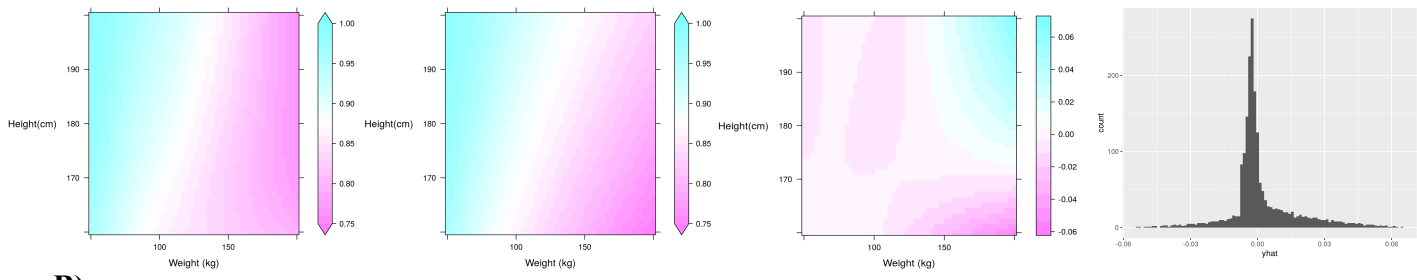
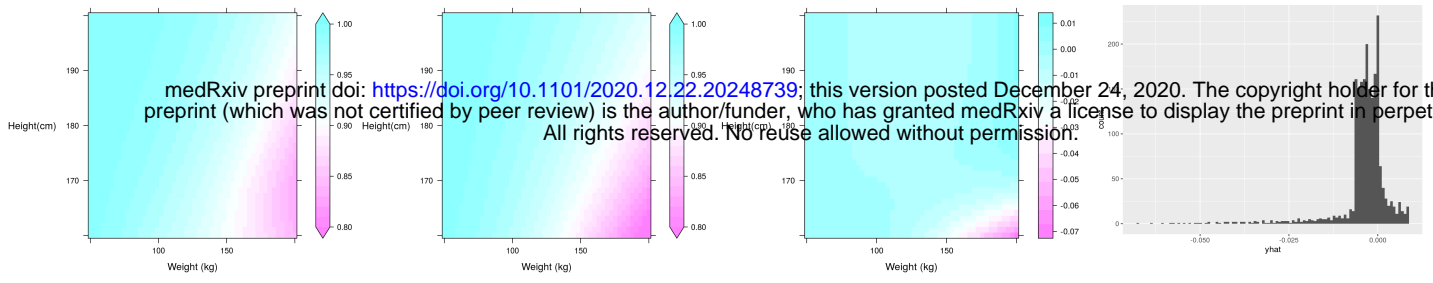
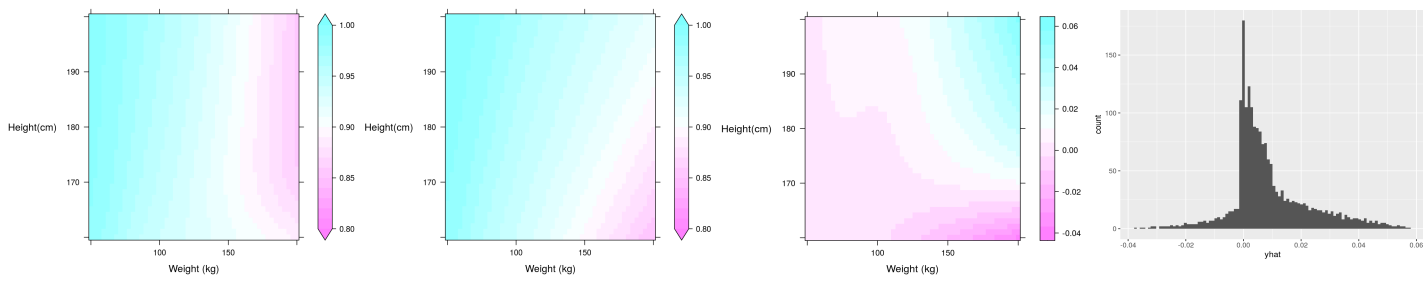
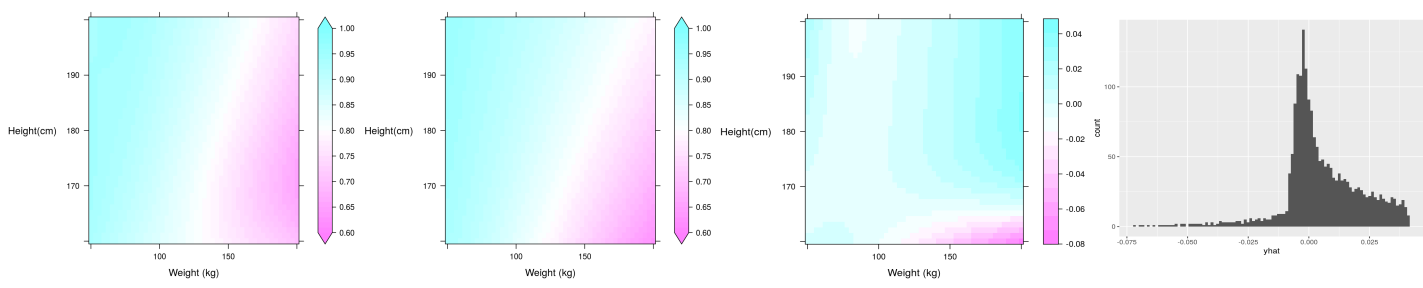
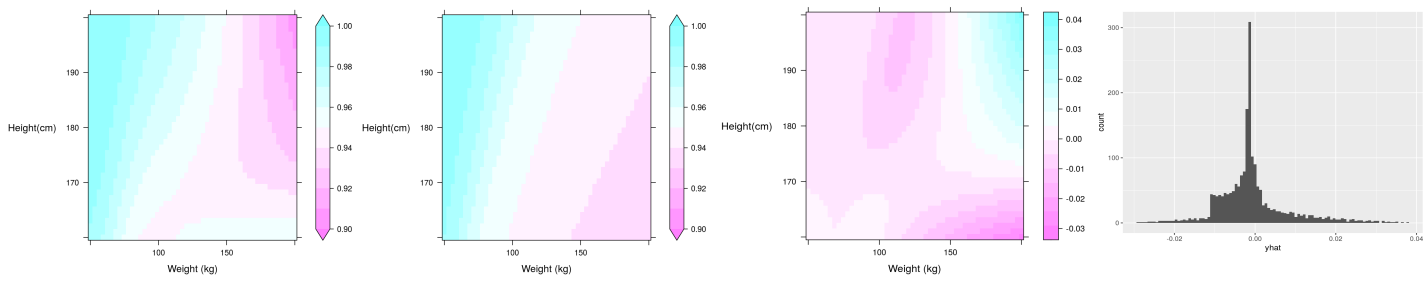
Figure Legends

Figure 1: Comparison of predictive models for metabolic dysregulation and its components.

A) Metabolic dysregulation; B) Diabetes mellitus; C) High density lipoprotein; D) Hypertension; E) Triglycerides. Graphics from left to right for each panel are: height-adjusted weight model prediction of 5-year free event probability across heights and weights; BMI model prediction of 5-year free event probability across heights and weights; difference in model prediction ($\text{Predicted probability}_{\text{BMI}} - \text{Predicted probability}_{\text{height*weight}}$); distribution of the difference in the prediction difference.

Figure 2: Comparison of predictive models for cardiovascular diseases.

A) Atrial fibrillation; B) coronary artery disease; C) heart failure; D) peripheral artery disease. Graphics from left to right for each panel are: height-adjusted weight model prediction of 5-year free event probability across heights and weights; BMI model prediction of 5-year free event probability across heights and weights; difference in model prediction ($\text{Predicted probability}_{\text{BMI}} - \text{Predicted probability}_{\text{height*weight}}$); distribution of the difference in the prediction difference.

A)**B)****C)****D)****E)****Figure 1**

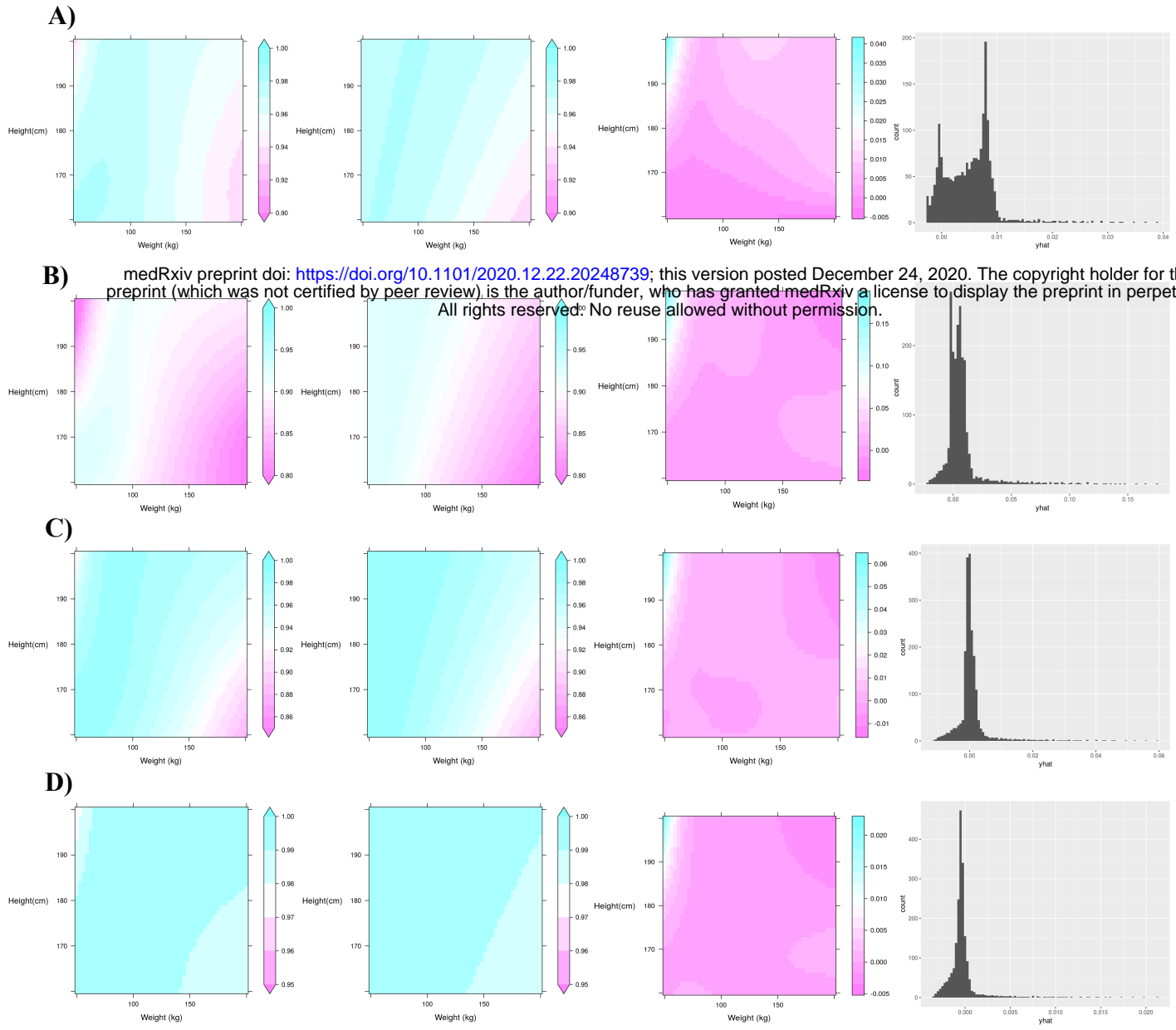


Figure 2