Evaluation of the Value of Waist Circumference and Metabolomics in the Estimation of Visceral Adipose Tissue

Sebastiaan C. Boone, Maarten van Smeden, Frits R. Rosendaal, Saskia le Cessie, Rolf H.H. Groenwold, J. Wouter Jukema, Ko Willems van Dijk, Hildo J. Lamb, Philip Greenland, Ian J. Neeland, Matthew A. Allison, Michael H. Criqui, Matthew J. Budoff, Lars L. Lind, Joel Kullberg, Håkan Ahlström, Dennis O. Mook-Kanamori, Renée de Mutsert

Correspondence to:
Sebastiaan Boone, Leiden University Medical Center (LUMC), Department of Clinical Epidemiology, PO Box 9600, 2300 RC Leiden, Department C7-P, Postal Zone C7-Q (email: s.c.boone@lumc.nl), Fax: +31 (0)71 526 6994, Tel: +31 (0)71 526 4037, ORCID: 0000-0002-2411-0699

Author affiliations:
Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, the Netherlands (Sebastiaan C. Boone, Maarten van Smeden, Frits R. Rosendaal, Saskia le Cessie, Rolf H.H. Groenwold, Dennis O. Mook-Kanamori, Renée de Mutsert); Department of Biomedical Data Sciences, section Medical Statistics and Bioinformatics, Leiden University Medical Center, Leiden, the Netherlands (Maarten van Smeden, Saskia le Cessie); Department of Cardiology, Leiden University Medical Center, Leiden, the Netherlands (J. Wouter Jukema); Netherlands Heart Institute, Utrecht, the Netherlands (J. Wouter Jukema); Department of Human Genetics, Leiden University Medical Center, Leiden, the Netherlands (Ko Willems van Dijk); Einthoven Laboratory for Experimental Vascular Medicine, Leiden University Medical Center, Leiden, the Netherlands (Ko Willems van Dijk); Department of Medicine, Division of Endocrinology, Leiden University Medical Center, Leiden, the Netherlands (Ko Willems van Dijk); Department of Radiology, Leiden University Medical Center, Leiden, the Netherlands (Hildø J. Lamb); Department of Preventive Medicine, Northwestern University, Chicago, Illinois, the United States of America (Philip Greenland); Division of Cardiovascular Disease, University Hospitals Harrington Heart and Vascular Institute, Cleveland, Ohio, the United States of America (Ian J. Neeland); Division of Cardiovascular Disease, Case Western Reserve University School of Medicine, Cleveland, Ohio, the United States of America (Ian J. Neeland); Department of Family Medicine and Public Health, Division of Preventive Medicine, University of California San Diego, San Diego, California, the United States of America (Matthew A. Allison, Michael H. Criqui); Department of Medicine, Lundquist Institute at Harbor-UCLA Medical Center, Torrance, California, the United States of America (Matthew J. Budoff); Department of Medical Sciences, Uppsala University, Uppsala, Sweden (Lars L. Lind); Department of Surgical Sciences, Section of Radiology, Uppsala University, Uppsala, Sweden (Joel Kullberg, Håkan Ahlström); Antaros Medical AB, Mölndal, Sweden (Joel Kullberg, Håkan Ahlström); Department of Public Health and Primary Care, Leiden University Medical Center, Leiden, the Netherlands (Dennis O. Mook-Kanamori).

© The Author(s) 2022. Published by Oxford University Press on behalf of the Johns Hopkins Bloomberg School of Public Health. This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journalpermissions@oup.com.
Funding: This work was supported by the Netherlands Cardiovascular Research Initiative: an initiative with support of the Dutch Heart Foundation (CVON2014-02 ENERGISE). The NEO study is supported by the participating Departments, the Division and the Board of Directors of the Leiden University Medical Centre, and by the Leiden University, Research Profile Area ‘Vascular and Regenerative Medicine’. The Multi-Ethnic Study of Atherosclerosis was supported by the National Heart, Lung and Blood Institute (HHSN268201500003I, N01-HC-95159, N01-HC-95160, N01-HC-95161, N01-HC-95162, N01-HC-95163, N01-HC-95164, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168 and N01-HC-95169); and the National Center for Advancing Translational Sciences (UL1-TR-000040, UL1-TR-001079, and UL1-TR-001420). The MESA Abdominal Body Composition Ancillary study was supported by the National Heart, Lung, and Blood Institute (R01-HL088451 to Dr. M.A.A.)

Conflict of interest statement: Dennis Mook-Kanamori was a part-time clinical research consultant for Metabolon, Inc.

Running head: VAT Estimation: Value of Waist and Metabolomics
DATA AVAILABILITY STATEMENT

The data underlying this article will be shared on reasonable request to the corresponding author.

ACKNOWLEDGEMENTS

We express our gratitude to all participants of the Netherlands Epidemiology in Obesity study, in addition to all participating general practitioners. We furthermore thank P.R. van Beelen and all research nurses for collecting the data and P.J. Noordijk and her team for sample handling and storage and I. de Jonge for data management. The authors thank the other investigators, the staff, and the participants of the Multi-Ethnic Study of Atherosclerosis study for their valuable contributions. A full list of participating Multi-Ethnic Study of Atherosclerosis investigators and institutions can be found at http://www.mesa-nhlbi.org. We also thank all participants of the Prospective Investigation of Vasculature in Uppsala Seniors study, as well as the staff and investigators involved in the collection and management of the data.
ABSTRACT

Visceral adipose tissue (VAT) is a strong prognostic factor for cardiovascular disease and a potential target for cardiovascular risk stratification. Because VAT is difficult to measure in clinical practice, we estimated prediction models with predictors routinely measured in general practice and VAT as outcome using ridge regression in 2,501 middle-aged participants from the Netherlands Epidemiology of Obesity study, 2008-2012. Adding waist circumference and other anthropometric measurements on top of the routinely measured variables improved the optimism-adjusted $R^2$ from 0.50 to 0.58 with a decrease in the root-mean-square error (RMSE) from 45.6 to 41.5 cm$^2$ and with overall good calibration. Further addition of predominantly lipoprotein-related metabolites from the Nightingale platform did not improve the optimism-corrected $R^2$ and RMSE. The models were externally validated in 370 participants from the Prospective Investigation of Vasculature in Uppsala Seniors (PIVUS, 2006-2009) and 1,901 participants from the Multi-Ethnic Study of Atherosclerosis (MESA, 2000-2007). Performance was comparable to the development setting in PIVUS ($R^2$: 0.63, RMSE: 42.4 cm$^2$, calibration slope: 0.94) but lower in MESA ($R^2$: 0.44, RMSE: 60.7 cm$^2$, calibration slope: 0.75). Our findings indicate that the estimation of VAT with routine clinical measurements can be substantially improved by incorporating waist circumference but not by metabolite measurements.

KEYWORDS:
added value, development, external validation, metabolomics, prediction, visceral adipose tissue
FOOTOTES/ABBREVIATIONS

CI, confidence interval
CITL, calibration-in-the-large
C-statistic, concordance statistic
HC, hip circumference
HDL-C, high-density lipoprotein cholesterol
LASSO, least absolute shrinkage and selection operator
MAE, mean absolute error
MRI, magnetic resonance imaging
MESA, Multi-Ethnic Study of Atherosclerosis
NEO, Netherlands Epidemiology of Obesity
PIVUS, the Prospective Investigation of Vasculature in Uppsala Seniors
R^2, explained variance
RMSE, root-mean-square error
SD, standard deviation
TBF, total body fat percentage
VAT, visceral adipose tissue
WC, waist circumference
WHR, waist-to-hip ratio
Abdominal obesity has been firmly established as an important prognostic factor for cardiometabolic changes (1) such as type 2 diabetes (2) and cardiovascular disease (3, 4). Visceral adipose tissue (VAT) is a major contributor to the increased cardiometabolic risk in individuals with abdominal obesity as it is causally associated with cardiovascular risk factors, coronary heart disease and type 2 diabetes as was demonstrated in several recent Mendelian randomization studies (5, 6). These associations make VAT a potentially interesting target to measure for cardiometabolic risk stratification in addition to existing cardiovascular risk models or to identify and motivate individuals that could benefit from preventive treatments such as lifestyle changes. However, direct measurements of VAT are only possible through imaging methods such as magnetic resonance imaging (MRI) or computed tomography. These measurements are expensive and time consuming compared with conventional clinical assessments such as a physical examination or blood sampling and therefore not well suited for large epidemiological studies or clinical practice.

Several studies have developed alternatives to direct measurements of VAT including the hypertriglyceridemic waist phenotype (7), the visceral adiposity index (8) and the lipid accumulation product (9), or models that estimate cross-sectional VAT surface area (10, 11). Most of these indices include waist circumference along with variables such as triglyceride concentrations and age (7-13). Although waist circumference is generally considered the best single predictor of VAT (14), it is not measured routinely in clinical settings (15). One reason that waist circumference is not measured more frequently may be that its added value for the prediction of VAT and other cardiometabolic outcomes on top of more routine measurements has not been established yet.

A potential source of new candidate predictors for VAT is the field of metabolomics, which concerns the measurement and analysis of metabolites, small molecular compounds that are
involved in metabolism. Contemporary technical developments enable the quantification of large numbers of metabolites at steadily decreasing costs, making the clinical use of metabolomics increasingly feasible. Although several studies have evaluated the associations of metabolites with VAT (16-19), there are no well-described attempts to use metabolite measurements in prediction models for VAT.

In this study, using three cohorts from the Netherlands, Sweden, and the United States of America, we developed and externally validated a prediction model for VAT containing routinely measured clinical predictors. We then investigated whether including waist circumference and measurements from a well-established lipid-focused metabolomics platform improved this model.

METHODS

Throughout the entire study, we adhered to the TRIPOD guidelines for the reporting of prediction studies (20).

Development study design and study population

The Netherlands Epidemiology of Obesity (NEO) study is a prospective, population-based cohort study aimed at investigating pathways leading to obesity-related conditions (21). Between 2008 and 2012 a total of 6,671 men and women aged 45-65 years were included, with an oversampling of persons with overweight or obesity. The study was approved by the medical ethics committee of the Leiden University Medical Center and all participants gave written informed consent.

Participants were invited for a baseline visit at the NEO study centre of the Leiden University Medical Center after an overnight fast of at least 10 hours. Participants came to the research
site in the morning and underwent a physical examination, including anthropometry and blood sampling. The Nightingale metabolomics platform was used to measure 224 metabolites of which 144 were used in this study (details provided in the Web Appendix). Approximately 35% of the participants (n = 2,580) without contra-indications for MRI were randomly selected to undergo direct assessment of the amount of VAT by MRI. For this study, we excluded participants with incomplete or poor-quality VAT measurements (n = 11), who were non-fasting during blood sampling at the baseline visit (n = 3), missing any of the predictors for the clinical models (n = 32), missing all metabolomics measurements (n = 17) and those with metabolite concentrations >10 standard deviations (SD) from the mean (n = 16) as they were deemed to be biologically implausible, leaving a study population of 2,501 participants (Web Figure 1).

External validation study designs and study populations

For external validation, we used two separate datasets from Sweden and the United States of America. First, the Prospective Investigation of Vasculature in Uppsala Seniors (PIVUS) is a population-based prospective cohort study that started in 2001 and included a random sample of 1,016 individuals aged 70 years from the city of Uppsala in Sweden (22). All living participants were invited back for a second measurement at age 75 of which 370 individuals had valid abdominal MRI measurements with quantification of VAT, metabolomics measurements with the Nightingale platform and no missing values in the clinical predictors. Second, the Multi-Ethnic Study of Atherosclerosis (MESA) is a population-based prospective cohort study which included 6,814 participants from different ethnicities (White, Black, Chinese and Hispanic) across the United States of America (23). Starting in July 2000, the participants underwent a baseline visit and a series of follow-up exams. During exams 2 and
3, a subset of the MESA cohort underwent abdominal computed tomography scans, from which VAT area was quantified. A total of 1,901 participants with valid VAT measurements and no missing values in the clinical predictors were available for analysis (Web Figure 1). In MESA, metabolomics measurements were performed using an untargeted NMR approach. Because only half of the metabolomics measurements from the Nightingale platform used in the NEO study overlapped with the MESA metabolomics data, we could not validate the metabolite-containing models and only validated the clinical models in the MESA cohort.

Data collection
Fasting blood samples were obtained in all three cohorts. A detailed description of the data collection of both predictors and outcomes in all participating cohorts is included in the Web Appendix.

Statistical analyses
Model development. Because of substantial heteroscedasticity when using the original cross-sectional VAT area, we used the natural logarithm of VAT area as outcome and back transformed all predicted values. To create the prediction models for VAT, all models were estimated using linear ridge regression, while we also used least absolute shrinkage and selection operator (LASSO) regression for models containing metabolomics measurements. As tuning parameter, we used the mean tuning parameter that minimized the mean squared error across 10 repeats of 10-fold cross-validation.

Based on knowledge from previously published studies, we considered the following variables for our core model: age, sex, height, body weight, body mass index (BMI), concentrations of
total cholesterol, triglycerides, high-density lipoprotein cholesterol (HDL-C) and glucose and the total cholesterol/HDL-C ratio (7-11, 24, 25). All these variables are measured routinely during a standard cardiovascular risk assessment by general practitioners in the Netherlands (26). We then sequentially added the following variables that are not routinely yet clinically measurable and have been associated with VAT, starting with the ones that are least expensive or time-consuming to measure: waist circumference (WC), hip circumference (HC), waist-to-hip ratio (WHR), serum insulin concentration and total body fat percentage (TBF). We selected an extended clinical model based on improvements in model performance, hereafter referred to as the extended clinical model. We then fitted ridge and LASSO regression models forcing in the predictors from this extended clinical model so they cannot be dropped from the models while further adding 144 metabolite variables from the Nightingale metabolomics platform but allowing shrinkage on all variables. All models were developed in the total sample as well as stratified by sex because men are known to have more VAT than women (25).

As performance measures we calculated the explained variance ($R^2$), root-mean-square error (RMSE), mean absolute error (MAE), calibration slope and calibration-in-the-large, the mean predicted versus observed values and c-statistics at VAT cut-offs of 70 cm$^2$ (27, 28), 100 cm$^2$ (27, 28) and 130 cm$^2$ (29, 30) from the predicted values on the logarithmic development scale as well as after back transforming the values to the original scale in cm$^2$. More detailed descriptions of these performance measures are included in the Web Appendix. Additionally, we created calibration plots with lowess curves and assessed the added predictive value of metabolites with back-to-back histograms and scatterplots of the predicted VAT of the metabolite containing versus the extended clinical models. We internally validated all models using bootstrapping with 1000 repeats for the models with only clinical predictors and 500
repeats for the metabolite containing models, repeating all modelling steps including the estimation of the tuning parameters for the ridge and LASSO models and calculated bootstrap optimism-corrected estimates of the performance measures (31).

**External validation.** We externally validated the core and extended clinical models in both PIVUS and MESA while we validated the metabolite-containing model in PIVUS. External validation was performed by calculating the expected amount of VAT with the regression equations developed in the NEO study and evaluating the same performance metrics as in the development study. Additionally, we evaluated model performance after recalibration by updating the model intercepts to account for differences in VAT distribution between the development and validation cohorts. All validations were performed in the total available study populations as well as stratified by sex, and additionally by ethnicity in MESA.

**RESULTS**

The development population in the NEO study consisted of 2,501 individuals, mostly of white ethnicity (96%), 1,311 (52.4%) were men (Table 1). Men had a larger waist circumference and more VAT, but less total body fat than women. Compared with PIVUS and MESA, NEO study participants were younger and had a higher BMI, waist and hip circumference and a poorer cardiometabolic profile with higher fasting glucose, cholesterol and the total cholesterol/HDL-C ratio (Table 1 and Web Table 1). The amount of VAT was similar in the NEO study and PIVUS but substantially higher in MESA.

**Clinical model development and validation**
The optimism-corrected performance measures for all models developed in the total population are summarized in Table 2 and Web Table 2. We observed a similar pattern on the logarithmic development scale as well as after back transforming the predicted values to the original scale. Adding waist circumference increased the R² from 0.50 (95% confidence interval (CI): 0.46, 0.53) to 0.55 (95%CI: 0.52, 0.58) on the original scale, while the RMSE decreased from 45.6 (95%CI: 43.6, 47.3) to 42.9 (95%CI: 41.1, 44.5) cm² and MAE from 33.3 (95%CI: 32.0, 34.4) to 31.5 (95%CI: 30.3, 32.6) cm². C-statistics increased from 0.91 (95%CI: 0.89, 0.99) to 0.92 (95%CI: 0.89, 0.94) for >70 cm², 0.86 (95%CI: 0.83, 0.88) to 0.88 (95%CI: 0.85, 0.89) for >100 cm² and 0.84 (95%CI: 0.81, 0.87) to 0.86 (0.83, 0.87) for >130 cm². Further adding hip circumference and WHR increased the R² to 0.58 (95% CI: 0.55, 0.61), while RMSE decreased to 41.5 (95%CI: 39.7, 42.8) cm² and MAE to 30.6 (95%CI: 29.4, 31.5) cm². Calibration slopes increased from 0.88 (95%CI: 0.84, 0.91) for the core model to 0.95 (95%CI: 0.92, 0.97) in the model that included the waist and hip circumference and the waist-to-hip ratio. We observed some overestimation of the amount of VAT for individuals at the upper end of the observed VAT distribution which improved when adding waist and hip circumference (Figure 1). The errors of the predicted values increased at higher predicted VAT amounts.

In the models developed in men and women separately, including waist and hip circumference and the waist-to-hip ratio also improved model performance, while further including total body fat or insulin did not. In men the R² was lower at 0.46 (95%CI: 0.43, 0.51) on the original scale for the model that included waist and hip circumference and the waist-to-hip ratio with a higher RMSE of 47.1 (95%CI: 44.6, 49.3) and MAE of 35.3 (95%CI: 33.5, 36.8) cm², while in women the R² was 0.58 (95%CI: 0.55, 0.63) with a RMSE of 34.4 (95%CI: 32.2, 36.2) and MAE of 25.3 (95%CI: 23.8, 26.5) cm² (Web Tables 3-4, Web Figures 2-3).
Overall, adding waist circumference resulted in the largest increase in model performance, with minor improvement from further addition of hip circumference and the waist-to-hip ratio. As all three can be easily measured together, we decided to use this model with the core predictors, waist and hip circumference and the waist-to-hip ratio as our extended clinical model on top of which to assess the added predictive value of metabolomics.

**Metabolite-extended model development and validation**

The optimism-corrected performance measures of the models in which the 144 metabolite predictors were added to the extended clinical model in the total development population are summarized in Table 2. The R² on the original scale increased from 0.58 (95%CI: 0.55, 0.61) to 0.60 (95% CI: 0.58, 0.63), RMSE decreased from 41.5 (95%CI: 39.7, 42.8) to 40.5 (95%CI: 38.4, 41.4) cm² and MAE decreased from 30.6 (95%CI: 29.4, 31.5) to 29.9 (95%CI: 28.4, 30.5) cm² for the ridge and LASSO models. Calibration characteristics were similar to the extended clinical model (Figure 2). The difference between the predictions from the extended clinical model and metabolite-extended models was small with no substantial differences in the size or distribution of the predicted values (Figure 3). The number of metabolites selected in the LASSO models varied by bootstrap sample from 25 to 62 with a median of 46 while no combination of metabolites was selected more than once. The most consistently selected metabolites included the amino acids and other non-lipoprotein compounds especially tyrosine, acetate and acetyl glycoprotein concentrations which were selected in all 500 bootstrap samples (Web Table 5). In men and women separately, we also observed small improvements after including metabolites (Web Tables 3-4, Web Figures 4-5).

The regression equations for the models developed in the complete population are summarized in Web Table 6.
External validation

Because the models that were developed in the complete NEO study population with sex included as a predictor performed similar to the models developed in men and women separately (Web Tables 3-4, 7-8), we decided to externally validate only the models developed in the complete NEO study population with sex included as a predictor.

The performance of the models with only clinical predictors was similar to the development setting in the total population of the PIVUS study with an $R^2$ of 0.65 (95%CI: 0.57, 0.71) on the logarithmic scale and 0.62 (95%CI: 0.50, 0.70) on the original scale for the extended clinical model that included waist, hip and waist-to-hip ratio (Table 3, Web Tables 9-10). RMSE and MAE were slightly higher at 42.9 (95%CI: 37.7, 48.0) and 31.0 (28.1, 43.2) cm$^2$ while c-statistics were practically identical to the NEO study. In MESA, the $R^2$ of both clinical models was lower while the extended clinical model still performed better with an $R^2$ of 0.40 (95%CI: 0.35, 0.45) on the logarithmic scale and 0.24 (95%CI: 0.19, 0.29) on the original scale compared with 0.28 (95%CI: 0.21, 0.34) and 0.07 (95%CI: -0.05, 0.16) for the core model. RMSE and MAE were substantially higher at 70.8 (95%CI: 68.1, 73.7) cm$^2$ and 53.8 (95%CI: 51.7, 55.9) cm$^2$ for the extended clinical model, while c-statistics were similar to the NEO study. Recalibration of the model intercept in MESA improved performance although the $R^2$ on the original scale remained lower at 0.44 (95%CI: 0.36, 0.51) and the RMSE and MAE remained higher at 60.7 (95%CI: 57.0, 65.0) and 43.6 (95%CI: 41.6, 45.6) cm$^2$. Calibration plots of both external cohorts show a similar pattern with a good calibration in PIVUS but an overall underestimation in MESA, which improved with recalibration of the model intercept (Figures 4-5).
The results in men and women were similar to the total population in MESA after recalibration of the intercept although the RMSE and MAE were lower in women than in men at 50.1 (95%CI: 45.4, 54.9) and 35.8 (95%CI: 33.6, 38.3) cm² for the extended clinical model in women versus 63.0 (95%CI: 56.5, 70.4) and 45.4 (95%CI: 42.7, 48.2) cm² in men (Web Tables 11-14, Web Figures 6-9). Conversely, in PIVUS the extended clinical model seemed to perform better in men (R² of 0.68 (95%CI: 0.55, 0.76), RMSE of 42.3 (95%CI: 36.9, 48.4) and MAE of 31.1 (95%CI: 27.2, 35.4) cm²) than in women (R² of 0.43 (95%CI: 0.20, 0.59), RMSE of 43.5 (95%CI: 35.8, 51.0) and MAE of 31.0 (26.8, 35.6) cm²).

Within the separate ethnicities in MESA, the clinical models performed best in White and Chinese participants but only after intercept recalibration (Table 4, Web Tables 15-16). For White participants, we observed an R² of 0.58 (95%CI: 0.52, 0.63), RMSE of 58.1 (95%CI: 54.4, 61.8) and MAE of 43.6 (95%CI: 40.9, 46.2) cm², while the R² in Chinese was 0.56 (95%CI: 0.44, 0.64) with a RMSE of 38.6 (95%CI: 34.0, 43.1) and MAE of 29.1 (95%CI: 25.8, 32.1) cm² on the original scale for the recalibrated extended clinical model. In Black and Hispanic participants, R² was lower even after recalibration (0.19 (95%CI: -0.08, 0.40) for Black and 0.31 (95%CI: 0.11, 0.47) for Hispanic) with higher values for the RMSE and MAE. The clinical models strongly overestimated VAT in three to six Black, Hispanic and White individuals (estimated VAT for these individuals ranged from 484 to 1010 cm²), which were removed from the calibration plots to improve the plots’ interpretability (Web Figures 10-13).

The metabolite-extended models, which were only validated in PIVUS, had a lower R², higher RMSE and MAE and poorer calibration than the extended clinical model (Table 3, Figure 4, Web Tables 17-19). Especially the ridge model performed poorly with an R² of -3.32 (95%CI: -4.58, -2.41) and c-statistics around 0.5, indicating non-useful predictions.
DISCUSSION

We developed prediction models to estimate VAT area in a middle-aged, ethnically white cohort and externally validated these models in a Swedish cohort with an older population and a multi-ethnic cohort from the United States. We showed that adding waist circumference to measurements routinely performed at a general practitioner’s office improved VAT prediction, while further adding insulin concentrations, total body fat or Nightingale metabolomics measurements did not. Models that included waist circumference also performed consistently better in the external validation with performances similar to the development setting in PIVUS and White and Chinese participants from the MESA cohort. The overall performance in the MESA cohort was lower, likely due to the poorer performance in Black and Hispanic participants.

Several previous studies have created prediction models for VAT, often incorporating similar predictors as such as age, waist circumference and cholesterol concentrations, with the earliest attempts dating back over thirty years (10, 11, 32-34). In a review of VAT prediction models that used anthropometric measurements up to 2014, waist measurements were used in 25 out of 27 equations and the explained variance of these models ranged between 30-80% with comparatively large prediction errors (13) which is in line with our own findings. A direct comparison with these and more recent studies is difficult as most of the models were tested only in their development cohorts, internal validation was not or only partially performed, and reporting of recommended information for prediction studies was often incomplete (20). We found only one study that performed an external validation but information on measures such calibration or the RMSE in the external dataset were not reported (11). In contrast, we were able to demonstrate our models’ efficacy in two
geographically separated and demographically diverse cohorts. Another important difference with previous studies is our focus on the added value of predictors as criterion for inclusion in our models. In previous studies, waist circumference was included in models on statistical grounds such as by stepwise regression (11, 32, 33) or because it was considered to be the best single predictors of VAT (10). With our study we now clearly demonstrate that waist circumference improves VAT prediction on top of more routine clinical measurements such as BMI.

Given the strong association of VAT with cardiometabolic disorders such as atherosclerosis, type 2 diabetes and cardiovascular disease (2, 24, 35) measurements of VAT could be useful for the identification of individuals with an elevated cardiometabolic risk or to motivate preventive measures such as lifestyle changes in addition to traditional measures such as BMI or risk predictions from cardiovascular risk prediction models. Measurements of VAT could also help in identifying individuals with a normal BMI at risk of metabolic dysfunction as even in these individuals a higher VAT is associated with cardiometabolic changes (36). Because computed tomography or MRI-imaging are practically unfeasible, the International Atherosclerosis Society and the International Chair on Cardiometabolic Risk Working Group on Visceral Obesity recently emphasized the need for more research into ‘simple, clinically applicable tools’ to identify individuals with excess VAT (37). Together, the results from previous studies and our own suggest that waist circumference is an important predictor for VAT, which strengthens the call to measure waist circumference routinely in clinical practice (15).

Although we previously showed that measurements from the Nightingale metabolomics platform, including acetyl glycoproteins, branched-chain amino acids and lipoprotein measures, were associated with VAT after adjusting for many of the variables that we
included in our models here (16), adding these measurements did not improve predictive performance. The high correlations between the predominantly lipid metabolism related measurements in the Nightingale platform and the standard laboratory measures in our clinical models, indicative of a high co-linearity between the metabolites and clinical predictors, likely explain why the metabolomics measurements did not have a clear added value on top of the clinical measurements (Web Table 5). Our findings should therefore not discourage further attempts to improve VAT prediction models. There are several other potential predictors, such as inflammatory markers other than acetyl glycoproteins (38), adipokines (39), sex hormones (40) and phosphatidylcholines and sphingomyelins that can be measured with other metabolomics platforms (18, 19, 41), that have been associated with VAT but whose role in VAT prediction has not yet been evaluated. As the costs of metabolomics measurements decrease, improvements in VAT prediction from other metabolomics platforms could result in a favorable cost-benefit ratio when weighed against the costs of performing imaging such as MRI. While higher age (42) and BMI (43) are associated with increases in VAT, our models seemed to generalize well to the validation cohorts whose participants were older and had a lower body weight than the participants from the NEO study. Further exploration of the role of ethnicity and sex in VAT prediction seems warranted given the heterogeneity in outcomes we observed within these subgroups. Similar differences in model performance were also observed in previous studies such as one in 622 middle-aged Japanese-Americans (11) where the explained variance was higher and mean-squared error was lower in women than in men. We found only two previous studies that examined the role of ethnicity in VAT prediction, one which estimated prediction formulas containing anthropometric variables separately in different ethnicities (10), while another included race as a predictor (34). Both studies reported high $R^2$ values ranging from
0.43 to 0.78, however the sample size in both studies was limited especially within the non-White subgroups and only (partial) internal validation was performed. Several descriptive studies have delineated ethnic differences in the distribution of VAT, such as a higher average VAT in White compared with Black individuals and differences in the association between body mass index and VAT in different ethnic subgroups (44-46), which could partially explain why our models, which were developed in an ethnically White cohort, performed poorer in Black and Hispanic participants in MESA.

Our study has several limitations including the use of cross-sectional VAT areas as outcome instead of total VAT volume. As a result, it is possible our models do not accurately estimate the VAT burden in an individual, although abdominal cross-sectional areas have been shown to correlate well ($\rho$: 0.70-0.99) with total VAT volume (47-49). In addition, VAT was measured using different imaging techniques and protocols in the NEO study, PIVUS, and MESA, which, together with measurement heterogeneity of the predictors, could explain some of the differences in performance between the cohorts (50). Strengths include our access to a large development cohort in which we performed a thorough development and internal validation. We also used a modelling approach that focuses on the added value of predictors on top of routinely measured ones instead of using purely statistical grounds for variable selection. Such statistical approaches can lead to models whose predictions transport poorly to a new setting (51) or that contain expensive or otherwise non-practically measurable predictors. Lastly, we were able to validate and demonstrate the efficacy of both the models with currently measurable predictors as well as those with metabolomics measurements in two cohorts that were demographically, ethnically and geographically distinct from the development cohort.
In conclusion, our study demonstrated that the prediction of VAT using routinely measured clinical variables can be improved by measuring and including waist circumference in the estimations but not by including metabolomics measurements from the Nightingale platform. We therefore encourage the routine measurement of waist circumference in clinical practice, for example during cardiovascular risk assessment.
REFERENCES:

1. Despres JP, Lemieux I. Abdominal obesity and metabolic syndrome. Nature. 2006;444(7121):881-7. doi:10.1038/nature05488
2. Freemantle N, Holmes J, Hockey A, Kumar S. How strong is the association between abdominal obesity and the incidence of type 2 diabetes? International journal of clinical practice. 2008;62(9):1391-6. doi:10.1111/j.1742-1241.2008.01805.x
3. Song X, Jousilaiti P, Stehouwer CD, et al. Cardiovascular and all-cause mortality in relation to various anthropometric measures of obesity in Europeans. Nutrition, metabolism, and cardiovascular diseases: NMCD. 2015;25(3):295-304. doi:10.1016/j.numecd.2014.09.004
4. Casanueva FF, Moreno B, Rodriguez-Azeredo R, et al. Relationship of abdominal obesity with cardiovascular disease, diabetes and hyperlipidaemia in Spain. Clinical endocrinology. 2010;73(1):35-40. doi:10.1111/j.1365-2265.2009.03727.x
5. Karlsson T, Rask-Andersen M, Pan G, et al. Contribution of genetics to visceral adiposity and its relation to cardiovascular and metabolic disease. Nature medicine. 2019;25(9):1390-5. doi:10.1038/s41591-019-0563-7
6. Dale CE, Fatemifar G, Palmer TM, et al. Causal Associations of Adiposity and Body Fat Distribution With Coronary Heart Disease, Stroke Subtypes, and Type 2 Diabetes Mellitus: A Mendelian Randomization Analysis. Circulation. 2017;135(24):2373-88. doi:10.1161/circulationaha.116.026560
7. Lemieux I, Pascot A, Couillard C, et al. Hypertriglyceridemic waist: A marker of the atherogenic metabolic triad (hyperinsulinemia; hyperapolipoprotein B; small, dense LDL) in men? Circulation. 2000;102(2):179-84.
8. Amato MC, Giordano C, Galia M, et al. Visceral Adiposity Index: a reliable indicator of visceral fat function associated with cardiometabolic risk. Diabetes care. 2010;33(4):920-2. doi:10.2337/dc09-1825
9. Kahn HS. The "lipid accumulation product" performs better than the body mass index for recognizing cardiovascular risk: a population-based comparison. BMC cardiovascular disorders. 2005;5:26. doi:10.1186/1471-2261-5-26
10. Eastwood SV, Tillin T, Wright A, et al. Estimation of CT-derived abdominal visceral and subcutaneous adipose tissue depots from anthropology in Europeans, South Asians and African Caribbeans. PloS one. 2013;8(9):e75085. doi:10.1371/journal.pone.0075085
11. Wander PL, Hayashi T, Sato KK, et al. Design and validation of a novel estimator of visceral adipose tissue area and comparison to existing adiposity surrogates. Journal of diabetes and its complications. 2018;32(11):1062-7. doi:10.1016/j.jdiacomp.2018.09.004
12. Oh SK, Cho AR, Kwon YJ, Lee HS, Lee JW. Derivation and validation of a new visceral adiposity index for predicting visceral obesity and cardiometabolic risk in a Korean population. PloS one. 2018;13(9):e0203787. doi:10.1371/journal.pone.0203787
13. Scafoglieri A, Clarys JP, Cattrysse E, Bautmans I. Use of anthropometry for the prediction of regional body tissue distribution in adults: benefits and limitations in clinical practice. Aging and disease. 2014;5(6):373-93. doi:10.14366/ad.2014.0500373
14. Poulion MC, Despres JP, Lemieux S, et al. Waist circumference and abdominal sagittal diameter: best simple anthropometric indexes of abdominal visceral adipose tissue accumulation and related cardiovascular risk in men and women. The American journal of cardiology. 1994;73(7):460-8.
15. Ross R, Neeland IJ, Yamashita S, et al. Waist circumference as a vital sign in clinical practice: a Consensus Statement from the IAS and ICCR Working Group on Visceral Obesity. Nature Reviews Endocrinology. 2020;16(3):177-89. doi:10.1038/s41574-019-0310-7
16. Neeland IJ, Boone SC, Mook-Kanamori DO, et al. Metabolomics Profiling of Visceral Adipose Tissue: Results From MESA and the NEO Study. Journal of the American Heart Association. 2019;8(9):e010810. doi:10.1161/jaha.118.010810
17. Szymanska E, Bouwman J, Strassburg K, et al. Gender-dependent associations of metabolite profiles and body fat distribution in a healthy population with central obesity: towards metabolomics diagnostics. Omics. 2012;16(12):652-67. doi:10.1089/omi.2012.0062

18. Syme C, Czajkowski S, Shin J, et al. Glycerophosphocholine Metabolites and Cardiovascular Disease Risk Factors in Adolescents: A Cohort Study. Circulation. 2016;134(21):1629-36. doi:10.1161/circulationaha.116.022993

19. Martin FP, Montoliu I, Collino S, et al. Topographical body fat distribution links to amino acid and lipid metabolism in healthy obese women [corrected]. PloS one. 2013;8(9):e73445. doi:10.1371/journal.pone.0073445

20. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): The TRIPOD Statement. Annals of internal medicine. 2015;162(1):55-63. doi:10.7326/M14-0697 %J Annals of Internal Medicine

21. de Mutsert R, den Heijer M, Rabelink TJ, et al. The Netherlands Epidemiology of Obesity (NEO) study: study design and data collection. European journal of epidemiology. 2013;28(6):513-23. doi:10.1007/s10654-013-9801-3

22. Lind L, Fors N, Hall J, Marttala K, Stenborg A. A Comparison of Three Different Methods to Evaluate Endothelium-Dependent Vasodilation in the Elderly. 2005;25(11):2368-75. doi:10.1161/01.ATV.0000184769.22061.da

23. Bild DE, Bluemke DA, Burke GL, et al. Multi-Ethnic Study of Atherosclerosis: objectives and design. American journal of epidemiology. 2002;156(9):871-81. doi:10.1093/aje/kwf113

24. Fox CS, Massaro JM, Hoffmann U, et al. Abdominal visceral and subcutaneous adipose tissue compartments: association with metabolic risk factors in the Framingham Heart Study. Circulation. 2007;116(1):39-48. doi:10.1161/circulationaha.106.675355

25. Karastergiou K, Smith SR, Greenberg AS, Fried SK. Sex differences in human adipose tissues – the biology of pear shape. Biology of Sex Differences. 2012;3(1):13. doi:10.1186/2042-6410-3-13

26. NHG. Cardiovasculair risicomanagement (CVRM). https://www.nhg.org/standaarden/volledig/cardiovasculair-risicomanagement. Accessed 10-05-2020.

27. Han JH, Park HS, Kim SM, Lee SY, Kim DJ, Choi WH. Visceral adipose tissue as a predictor for metabolic risk factors in the Korean population. Diabetic medicine : a journal of the British Diabetic Association. 2008;25(1):106-10. doi:10.1111/j.1464-5491.2007.02317.x

28. Miyawaki T, Hirata M, Moriyama K, et al. Metabolic syndrome in Japanese diagnosed with visceral fat measurement by computed tomography. Proceedings of the Japan Academy, Series B. 2005;81(10):471-9. doi:10.2183/pjab.81.471

29. Hunter GR, Snyder SW, Kekes-Szabo T, Nicholson C, Berland L. Intra-abdominal adipose tissue values associated with risk of possessing elevated blood lipids and blood pressure. Obesity research. 1994;2(6):563-8.

30. Oka R, Kobayashi J, Yagi K, et al. reassessment of the cutoff values of waist circumference and visceral fat area for identifying Japanese subjects at risk for the metabolic syndrome. Diabetes research and clinical practice. 2008;79(3):474-81. doi:10.1016/j.diabres.2007.10.016

31. Harrell F. Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis. 2 ed: Springer International Publishing, Basel, Switzerland; 2015.

32. Seidell JC, Oosterlee A, Thijsens MA, et al. Assessment of intra-abdominal and subcutaneous abdominal fat: relation between anthropometry and computed tomography. Am J Clin Nutr. 1987;45(1):7-13. doi:10.1093/ajcn/45.1.7

33. Pinho CPS, Diniz ADS, de Arruda IKG, Leite A, Petribu MMV, Rodrigues IG. Predictive models for estimating visceral fat: The contribution from anthropometric parameters. PloS one. 2017;12(7):e0178958. doi:10.1371/journal.pone.0178958
34. Stanforth PR, Jackson AS, Green JS, et al. Generalized abdominal visceral fat prediction models for black and white adults aged 17–65 y: the HERITAGE Family Study. International Journal Of Obesity. 2004;28:925. doi:10.1038/sj.ijo.0802563
35. Després JP. Body Fat Distribution and Risk of Cardiovascular Disease. Circulation. 2012;126(10):1301-13. doi:10.1161/CIRCULATIONAHA.111.067264
36. Tatsumi Y, Nakao YM, Masuda I, et al. Risk for metabolic diseases in normal weight individuals with visceral fat accumulation: a cross-sectional study in Japan. BMJ open. 2017;7(1):e013831. doi:10.1136/bmjopen-2016-013831
37. Neeland IJ, Ross R, Despres JP, et al. Visceral and ectopic fat, atherosclerosis, and cardiometabolic disease: a position statement. The lancet. Diabetes & endocrinology. 2019;7(9):715-25. doi:10.1016/s2213-8587(19)30084-1
38. Yu JY, Choi WJ, Lee HS, Lee JW. Relationship between inflammatory markers and visceral obesity in obese and overweight Korean adults: An observational study. Medicine. 2019;98(9):e14740. doi:10.1097/md.0000000000014740
39. Christen T, Trompet S, Noordam R, et al. Sex differences in body fat distribution are related to sex differences in serum leptin and adiponectin. Peptides. 2018;107:25-31. doi:10.1016/j.peptides.2018.07.008
40. Mongraw-Chaffin ML, Anderson CA, Allison MA, et al. Association between sex hormones and adiposity: qualitative differences in women and men in the multi-ethnic study of atherosclerosis. The Journal of clinical endocrinology and metabolism. 2015;100(4):E596-600. doi:10.1210/jc.2014-2934
41. Boone S, Mook-Kanamori D, Rosendaal F, et al. Metabolomics: a search for biomarkers of visceral fat and liver fat content. Metabolomics : Official journal of the Metabolomic Society. 2019;15(10):139. doi:10.1007/s11306-019-1599-x
42. Borkan GA, Hults DE, Gerzof SG, Robbins AH, Silbert CK. Age changes in body composition revealed by computed tomography. J Gerontol. 1983;38(6):673-7. doi:10.1093/geronj/38.6.673
43. Enzi G, Gasparo M, Biondetti PR, Fiore D, Semisa M, Zurlo F. Subcutaneous and visceral fat distribution according to sex, age, and overweight, evaluated by computed tomography. Am J Clin Nutr. 1986;44(6):739-46. doi:10.1093/ajcn/44.6.739
44. Lear SA, Humphries KH, Kohll S, Chockalingam A, Frohlich JJ, Birmingham CL. Visceral adipose tissue accumulation differs according to ethnic background: results of the Multicultural Community Health Assessment Trial (M-CHAT). Am J Clin Nutr. 2007;86(2):353-9. doi:10.1093/ajcn/86.2.353
45. Despres JP, Couillard C, Gagnon J, et al. Race, visceral adipose tissue, plasma lipids, and lipoprotein lipase activity in men and women: the Health, Risk Factors, Exercise Training, and Genetics (HERITAGE) family study. Arteriosclerosis, thrombosis, and vascular biology. 2000;20(8):1932-8. doi:10.1161/01.atv.20.8.1932
46. Nazare J-A, Smith JD, Borel A-L, et al. Ethnic influences on the relations between abdominal subcutaneous and visceral adiposity, liver fat, and cardiometabolic risk profile: the International Study of Prediction of Intra-Abdominal Adiposity and Its Relationship With Cardiometabolic Risk/Intra-Abdominal Adiposity. The American Journal of Clinical Nutrition. 2012;96(4):714-26. doi:10.3945/ajcn.112.035758 %J The American Journal of Clinical Nutrition
47. Shen W, Punyanitya M, Wang Z, et al. Visceral adipose tissue: relations between single-slice areas and total volume. The American Journal of Clinical Nutrition. 2004;80(2):271-8. doi:10.1093/ajcn/80.2.271 %J The American Journal of Clinical Nutrition
48. So R, Matsuot S, Sasai H, et al. Best single-slice measurement site for estimating visceral adipose tissue volume after weight loss in obese, Japanese men. Nutrition & metabolism. 2012;9:56. doi:10.1186/1743-7075-9-56
49. Kobayashi J, Tadokoro N, Watanabe M, Shinomiya M. A novel method of measuring intra-abdominal fat volume using helical computed tomography. International journal of obesity and related metabolic disorders / journal of the International Association for the Study of the Obesity. 2002;26(3):398-402. doi:10.1038/sj.ijo.0801921

23
50. Luijken K, Wynants L, van Smeden M, et al. Changing predictor measurement procedures affected the performance of prediction models in clinical examples. Journal of clinical epidemiology. 2020;119:7-18. doi:https://doi.org/10.1016/j.jclinepi.2019.11.001

51. Smith G. Step away from stepwise. Journal of Big Data. 2018;5(1):32. doi:10.1186/s40537-018-0143-6
| Characteristic                        | Total (n = 2,501) | Men (n = 1,311) | Women (n = 1,190) |
|--------------------------------------|-------------------|-----------------|-------------------|
|                                      | Mean (SD)         | Median (IQR)    | Mean (SD)         | Median (IQR)    | Mean (SD)         | Median (IQR)    |
| Age, years                           | 56 (51–61)        | 56 (51–61)      | 56 (51–61)        |
| Male sex<sup>a</sup>                 | 1,311             | 52.4            |                   |
| White ethnicity<sup>a</sup>          | 2,394             | 95.9            | 1,265             | 96.7            | 1,129             | 95.0            |
| Height, m                            | 1.74 (0.10)       | 1.81 (0.07)     | 1.66 (0.06)       |
| Weight, kg                           | 89.2 (15.2)       | 96.0 (12.7)     | 81.7 (14.2)       |
| Body mass index<sup>b</sup>          | 29.5 (4.2)        | 29.4 (3.4)      | 29.5 (4.8)        |
| Waist circumference, cm              | 101 (12)          | 105 (10)        | 96 (13)           |
| Hip circumference, cm                | 109 (9)           | 108 (7)         | 110 (10)          |
| Waist-to-hip ratio                   | 0.92 (0.08)       | 0.97 (0.06)     | 0.87 (0.07)       |
| Body fat percentage                  | 33.8 (27.8–42.2)  | 28.5 (25.1–31.9) | 42.3 (38.1–45.4)  |
| Visceral adipose tissue, cm<sup>2</sup> | 114 (76–157)     | 134 (100–176)   | 89 (59–128)       |
| > 100 cm<sup>2</sup>a                | 1,478             | 59.1            | 978               | 74.6            | 500               | 42.0            |
| Laboratory measurements              |                   |                 |                   |
| Glucose, mmol/L                      | 5.5 (5.1–6.0)     | 5.6 (5.2–6.1)   | 5.4 (5.0–5.9)     |
| Insulin, IU                          | 10.5 (6.8–15.4)   | 11.5 (7.4–16.7) | 9.7 (6.3–14.3)    |
| Triglycerides, mmol/L                | 1.3 (0.9–1.8)     | 1.4 (1.0–2.0)   | 1.1 (0.8–1.5)     |
| HDL-cholesterol, mmol/L              | 1.4 (0.4)         | 1.3 (0.3)       | 1.6 (0.4)         |
| Total cholesterol, mmol/L            | 5.7 (1.1)         | 5.6 (1.1)       | 5.8 (1.1)         |
| Total cholesterol / HDL-C ratio      | 4.1 (3.3–5.0)     | 4.5 (3.8–5.5)   | 3.6 (3.0–4.4)     |

Abbreviations: SD, standard deviation; HDL-cholesterol, high-density lipoprotein cholesterol; IQR, interquartile range; Total cholesterol / HDL-C ratio, total cholesterol / high density lipoprotein cholesterol ratio

<sup>a</sup> Values are expressed as numbers and percentages

<sup>b</sup> Weight (kg)/height (m)<sup>2</sup>
Table 2. Development and Internal Validation: Bootstrap-Optimism Corrected Estimates\(^a\) of Model Performance in the Netherlands Epidemiology of Obesity Study (n = 2,501), the Netherlands, 2008-2012

| Models | Predictors | Logarithmic scale | | Original scale | C-statistics |
|---|---|---|---|---|---|
| | | R\(^2\) | RMSE | MAE | CITL | Slope | R\(^2\) | RMSE | MAE | CITL | Slope | > 70 | > 100 | > 130 |
| 1 | Core\(^b\) | 0.60 | 0.37 | 0.29 | 0.00 | 1.00 | 0.50 | 45.6 | 33.3 | 6.1 | 0.88 | 0.91 | 0.87 | 0.84 |
| 2 | Core + WC | 0.64 | 0.35 | 0.28 | 0.00 | 1.00 | 0.55 | 42.9 | 31.5 | 5.5 | 0.90 | 0.92 | 0.88 | 0.86 |
| 3 | Core + WC + HC | 0.65 | 0.35 | 0.27 | 0.00 | 1.00 | 0.57 | 42.1 | 30.9 | 5.4 | 0.90 | 0.92 | 0.89 | 0.86 |
| 4 | Core + WC + HC + WHR | 0.66 | 0.34 | 0.27 | 0.00 | 1.00 | 0.58 | 41.5 | 30.6 | 5.5 | 0.95 | 0.92 | 0.89 | 0.86 |
| 5 | Core + WC + HC + WHR + Insulin | 0.66 | 0.34 | 0.27 | 0.00 | 1.00 | 0.57 | 42.2 | 30.5 | 5.4 | 0.91 | 0.92 | 0.89 | 0.87 |
| 6 | Core + WC + HC + WHR + TBF | 0.66 | 0.34 | 0.27 | 0.00 | 1.00 | 0.58 | 41.2 | 30.4 | 5.4 | 0.94 | 0.92 | 0.89 | 0.86 |
| 7 | Core + WC + HC + WHR + Insulin + TBF | 0.66 | 0.34 | 0.27 | 0.00 | 1.00 | 0.57 | 41.8 | 30.4 | 5.3 | 0.91 | 0.92 | 0.89 | 0.87 |
| Ridge | Core + WC + HC + WHR + Metabolites | 0.68 | 0.33 | 0.26 | 0.00 | 0.99 | 0.60 | 40.5 | 29.9 | 5.0 | 0.93 | 0.93 | 0.89 | 0.87 |
| LASSO | Core + WC + HC + WHR + Metabolites | 0.68 | 0.33 | 0.26 | 0.00 | 0.99 | 0.60 | 40.5 | 29.9 | 5.1 | 0.95 | 0.93 | 0.89 | 0.87 |

Abbreviations: CITL, calibration-in-the-large; C-statistics, concordance statistics; HC, hip circumference; LASSO, least absolute shrinkage and selection operator; MAE, mean absolute error; R\(^2\), explained variance; RMSE, root-mean-square error; TBF, total body fat percentage; WC, waist circumference; WHR, waist-to-hip ratio.

\(^a\) The estimates are reported for the development scale (natural logarithm of VAT) and the original scale (in cm\(^2\)) after back transforming the predicted values.

\(^b\) The core predictors include age, sex, height, body weight, body mass index, concentrations of total cholesterol, triglycerides, high-density lipoprotein cholesterol and glucose and the total cholesterol/HDL-C ratio.
### Table 3. Estimates of Model Performance of the Core, Extended Clinical and Metabolite-extended Models in the Prospective Investigation of the Vasculature in Uppsala Seniors (n = 370), Uppsala, Sweden, 2006-2009, and the Multi-Ethnic Study of Atherosclerosis (n = 1,901), United States, 2000-2007

| Outcome measures | MESA | Clinical Models | PIVUS | Metabolite-extended Models |
|------------------|------|----------------|-------|---------------------------|
|                  | Core | Extended | Core | Extended | Ridge | LASSO |
| **Logarithmic scale** |      |          |      |            |       |       |
| Mean observed     | 4.94 | 4.66     | 4.70 | 4.70       | 4.70  | 4.70  |
| Mean predicted    | 4.94 | 4.66     | 4.77 | 4.70       | 4.66  | 4.70  |
| R²               | 0.55 | 0.66     | 0.66 | 0.66       | 0.68  | 0.51  |
| RMSE             | 0.38 | 0.34     | 0.33 | 0.32       | 0.33  | 0.39  |
| MAE              | 0.30 | 0.27     | 0.26 | 0.25       | 0.28  | 0.30  |
| CITL             | 0.00 | 0.00     | -0.07| 0.00       | 0.00  | 0.00  |
| slope            | 0.87 | 1.00     | 0.97 | 0.97       | 0.92  | 0.96  |
| **Original scale** |      |          |      |            |       |       |
| Mean observed     | 162.2| 122.1    | 127.2| 127.2      | 127.2 | 127.2 |
| Mean predicted    | 159.3| 116.5    | 131.1| 122.4      | 157.8 | 117.1 |
| R²               | 0.17 | 0.58     | 0.62 | 0.63       | 0.44  | 0.44  |
| RMSE             | 74.0 | 41.5     | 42.9 | 42.4       | 60.7  | 40.5  |
| MAE              | 47.9 | 30.6     | 31.0 | 30.1       | 43.6  | 29.9  |
| CITL             | 2.8  | 5.5      | -3.9 | 4.8        | 41.6  | 4.3   |
| slope            | 0.57 | 0.95     | 0.88 | 0.94       | 0.98  | 0.95  |
| **C-statistics** |      |          |      |            |       |       |
| >70 cm²          | 0.93 | 0.92     | 0.91 | 0.91       | 0.94  | 0.93  |
| >100 cm²         | 0.90 | 0.89     | 0.90 | 0.90       | 0.91  | 0.91  |
| >130 cm²         | 0.88 | 0.86     | 0.86 | 0.86       | 0.90  | 0.87  |

Abbreviations: CITL, calibration-in-the-large; MAE, mean absolute error; LASSO, least absolute shrinkage and selection operator; MESA, the Multi-Ethnic Study of Atherosclerosis; PIVUS, the Prospective Investigation of the Vasculature in Uppsala Seniors; R², explained variance; RMSE, root-mean-square error.

a Estimates are reported for the development scale (natural logarithm of VAT) and the original scale (in cm²) after back transforming the predicted values. We report the initial results (first column under each subheader) and the performance after intercept recalibration (second column under each subheader).

b Core predictors include age, sex, height, body weight, body mass index, total cholesterol, triglycerides, high-density lipoprotein cholesterol, glucose and the total cholesterol/HDL-C ratio.

c The extended clinical model includes the core predictors, waist circumference, hip circumference and the waist-to-hip ratio.
Table 4. Estimates of Model Performance of the Core and Extended Clinical Models Before and After Recalibration in White (n = 764), Chinese (n = 249), Black (n = 397) and Hispanic (n = 491) Participants of the Multi-Ethnic Study of Atherosclerosis Cohort (n = 1,901), United States, 2000-2007c

| Outcome measures | White Coreb | White Extendedc | Chinese Core | Chinese Extended | Black Core | Black Extended | Hispanic Core | Hispanic Extended |
|------------------|-------------|-----------------|--------------|-----------------|------------|---------------|--------------|------------------|
| Logarithmic scale |             |                 |              |                 |            |               |              |                  |
| Mean observed    | 5.00        | 5.00            | 5.00         | 5.00            | 4.71       | 4.71          | 4.71         | 4.71             |
| Mean predicted   | 4.63        | 5.00            | 4.67         | 5.00            | 4.36       | 4.71          | 4.40         | 4.71             |
| R²               | 0.30        | 0.65            | 0.40         | 0.69            | 0.11       | 0.58          | 0.27         | 0.64             |
| RMSE             | 0.51        | 0.36            | 0.47         | 0.34            | 0.49       | 0.34          | 0.44         | 0.31             |
| CITL             | 0.37        | 0.00            | 0.33         | 0.00            | 0.35       | 0.00          | 0.32         | 0.00             |
| Calibration slope| 0.98        | 0.98            | 1.01         | 1.01            | 0.96       | 0.96          | 0.95         | 0.95             |
| Original scale   |             |                 |              |                 |            |               |              |                  |
| Mean observed    | 174.1       | 174.1           | 174.1        | 174.1           | 125.6      | 125.6         | 125.6        | 125.6            |
| Mean predicted   | 116.6       | 168.0           | 120.0        | 166.7           | 85.2       | 121.4         | 89.5         | 122.8            |
| R²               | 0.12        | 0.46            | 0.22         | 0.58            | 0.07       | 0.53          | 0.20         | 0.56             |
| RMSE             | 84.6        | 66.6            | 79.5         | 58.1            | 56.1       | 39.7          | 51.8         | 38.6             |
| MAE              | 65.5        | 46.4            | 61.7         | 43.6            | 44.6       | 29.7          | 40.0         | 29.1             |
| CITL             | 57.5        | 6.1             | 54.1         | 7.4             | 40.4       | 4.1           | 36.1         | 2.8              |
| Calibration slope| 1.06        | 0.74            | 1.22         | 0.88            | 1.18       | 0.83          | 1.10         | 0.80             |
| C-statistics     |             |                 |              |                 |            |               |              |                  |
| >70 cm²          | 0.93        | 0.93            | 0.94         | 0.94            | 0.86       | 0.86          | 0.86         | 0.82             |
| >100 cm²         | 0.93        | 0.93            | 0.92         | 0.92            | 0.88       | 0.88          | 0.89         | 0.89             |
| >130 cm²         | 0.92        | 0.92            | 0.92         | 0.92            | 0.87       | 0.87          | 0.88         | 0.88             |

Abbreviations: CITL, calibration-in-the-large; C-stat., concordance statistics; MAE, mean absolute error; R², explained variance; RMSE, root-mean-square error

a Estimates are reported for the development scale (natural logarithm of VAT) and the original scale (in cm²) after back transforming the predicted values. We report the initial results (first column under each subheader for the core and extended clinical models) and the performance after intercept recalibration (second column under each subheader for the core and extended clinical models)
b Core predictors include age, sex, height, body weight, body mass index, total cholesterol, triglycerides, high-density lipoprotein cholesterol, glucose and the total cholesterol/HDL-C ratio.
c The extended clinical model includes the core predictors, waist circumference, hip circumference and the waist-to-hip ratio.
Figure 1. Calibration plots of models with clinical predictors developed in the Netherlands Epidemiology of Obesity Study (n = 2,501), the Netherlands, 2008-2012. A-D) Calibration plots on the natural logarithm scale (development scale); A) Core; B) Core + WC; C) Core + WC + HC + WHR; D) Core + WC + HC + WHR + Insulin + TBF. E-H) Calibration plots after back transforming the predicted values to the original scale (in cm²). E) Core; F) Core + WC; G) Core + WC + HC + WHR; H) Core + WC + HC + WHR + Insulin + TBF. The black striped line indicates unity (x = y); the black continuous line is a lowess curve fitted to the data points. Light gray dots are women; dark gray dots are men. HC, hip circumference; TBF, total body fat; VAT, visceral adipose tissue; WC, waist circumference; WHR, waist-to-hip ratio.
Figure 2. Calibration plots of the extended clinical and metabolite-extended models developed in the Netherlands Epidemiology of Obesity Study (n = 2,501), the Netherlands, 2008-2012. A-C) Calibration plots on the natural logarithm scale (development scale); A) Extended clinical model with core predictors, WC, HC and WHR; B) Metabolite-extended, ridge; C) Metabolite-extended, LASSO. D-F) Calibration plots after back transforming the predicted values to the original scale (in cm²); D) Extended clinical model with core predictors, WC, HC and WHR; E) Metabolite-extended, ridge; F) Metabolite-extended, LASSO. The black striped line indicates unity (x = y); the black continuous line is a lowess curve fitted to the data points. Light gray dots are women; dark gray dots are men. HC, hip circumference; LASSO, least absolute shrinkage and selection operator; VAT, visceral adipose tissue; WC, waist circumference; WHR, waist-to-hip ratio.
Figure 3. Back-to-back histograms and scatterplots of predicted values of ln(VAT) of the extended clinical model (core + WC + HC + WHR) versus the metabolite-extended models in the Netherlands Epidemiology of Obesity Study (n = 2,501), the Netherlands, 2008-2012. A-B) Back-to-back histograms with predictions from the extended clinical model on the left-hand side and the metabolite extended models on the right-hand side; A) Metabolite-extended, ridge; B) Metabolite-extended, LASSO. C-D) Scatterplots with values predicted by the extended clinical model on the x-axis and values predicted by the extended models on the y-axis; C) Metabolite-extended, ridge; D) Metabolite-extended, LASSO. The striped line is the x = y line. HC, hip circumference; LASSO, least absolute shrinkage and selection operator; VAT, visceral adipose tissue; WC, waist circumference; WHR, waist-to-hip ratio.
A) Ridge

B) Lasso

C) Extended Clinical Model Predicted In(VAT), ln(cm²)

D) Metabolite Extended Model Predicted ln(VAT), ln(cm²)
Figure 4. Calibration plots of the external validation of the core and extended clinical models and the metabolite extended models in the Prospective Investigation of the Vasculature in Uppsala Seniors (n = 370), Uppsala, Sweden, 2006-2009. A-D) Calibration plots on the natural logarithm scale (development scale); A) Core; B) Extended; C) Metabolite-extended, ridge, D) Metabolite-extended, LASSO. E-H) Calibration plots of the predicted values back transformed to the original scale (in cm²); E) Core; F) Extended; G) Metabolite-extended, ridge; H) Metabolite-extended, LASSO. The black striped line indicates unity (x = y); the black continuous line is a lowess curve fitted to the data points. Light gray dots are women; dark gray dots are men. VAT, visceral adipose tissue.
Figure 5. Calibration plots of the external validation of the core and extended clinical models before and after intercept recalibration in the Multi-Ethnic Study of Atherosclerosis (n = 1,901), United States, 2000-2007. A-D) Calibration plots on the natural logarithm scale (development scale); A) Core; B) Core, recalibrated; C) Extended; D) Extended, recalibrated. E-H) Calibration plots after back transforming the predicted values to the original scale (in cm²); E) Core; F) Core, recalibrated; G) Extended; H) Extended, recalibrated. To improve plot interpretability, we restricted the axis range in E-H, excluding between one to nine individuals with large overestimations of VAT. The black striped line indicates unity (x = y); the black continuous line is a lowess curve fitted to the data points. Light gray dots are women; dark gray dots are men. VAT, visceral adipose tissue.
