Humans with function-disrupting variants in the myostatin gene (MSTN) have increased skeletal muscle mass and strength, and less adiposity

Joseph L. Herman[‡1], Peter Dornbos[‡1], Karl Landheer[‡1], Benjamin J. Geraghty[‡1], Marc A. Egerman[2], Duc Phan[2], Mary Germino[4], Jason W. Mastaitis[3], Johnathon R. Walls[4], Luca A. Lotta[1], Gonçalo Abecasis[1], Aris Baras[1], Judith Y. Altarejos[3], Mark W. Sleeman[3], Regeneron Genetics Center[1#], Olle Melander[5,6], Malmö Diet and Cancer Study[#], Tea Shavlakadze[2], George D. Yancopoulos*[1,2,3,4], Jonas Bovijn[1]*, Jonathan Marchini[1]*, David J. Glass[2]*

1. Regeneron Genetics Center, Regeneron Pharmaceuticals, 777 Old Saw Mill River Road, Tarrytown, New York 10591

2. Aging and Age-Related Disorders, Regeneron Pharmaceuticals, 777 Old Saw Mill River Road, Tarrytown, New York 10591

3. Obesity, Muscle and Metabolism, Regeneron Pharmaceuticals, 777 Old Saw Mill River Road, Tarrytown, New York 10591

4. Regeneron Imaging Center, Regeneron Pharmaceuticals, 777 Old Saw Mill River Road, Tarrytown, New York 10591

5. Department of Clinical Sciences Malmö, Lund University, Malmö, Sweden

6. Department of Internal Medicine, Skåne University Hospital, Malmo, Sweden

[‡]these authors contributed equally to this work

[#]Full lists of authors for the banners are reported in the Supplementary Material

*corresponding authors, and equal contributors: david.glass@regeneron.com; jonathan.marchini@regeneron.com; jonas.bovijn@regeneron.com; george.yancopoulos@regeneron.com

This file includes:

Full lists of consortium members

Supplementary Tables 1-3; Supplementary Figures 1 to 9

Supplementary Data Captions (Supplementary Data 1 to 9 in Supplementary_Data.xlsx)

Regeneron Genetics Center

RGC Management & Leadership Team
Aris Baras, Gonçalo Abecasis, Adolfo Ferrando, Giovanni Coppola, Andrew Deubler, Luca A Lotta, John D Overton, Jeffrey G Reid, Alan Shuldiner, Katherine Siminovitch, Jason Portnoy, Marcus B Jones, Lyndon Mitnaul, Alison Fenney, Jonathan Marchini, Manuel Allen Revez Ferreira, Maya Ghoussaini, Mona Nafde, William Salerno, Cristen Willer, Lourdes Crane.

Sequencing & Lab Operations
John D Overton, Christina Beechert, Erin Fuller, Laura M Cremona, Eugene Kalyuskin, Hang Du, Caitlin Forsythe, Zhenhua Gu, Kristy Guevara, Michael Lattari, Alexander Lopez, Kia Manoochehri, Prathyusha Challa, Manasi Pradhan, Raymond Reynoso, Ricardo Schiavo, Maria Sotiropoulos Padilla, Chenggu Wang, Sarah E Wolf, Hang Du, Kristy Guevara.

Genome Informatics & Data Engineering
Jeffrey G Reid, Mona Nafde, Manan Goyal, George Mitra, Sanjay Sreeram, Rouel Lanche, Vrushali Mahajan, Sai Lakshmi Vasireddy, Gisu Eom, Krishna Pawan Punuru, Sujit Gokhale, Benjamin Sultan, Pooja Mule, Mudasar Sarwar, Muhammad Aqeel, Xiaodong Bai, Lance Zhang, Sean O'Keeffe, Razvan Panea, Evan Edelstein, Ayesha Rasool, William Salerno, Evan K Maxwell, Boris Boutkov, Alexander Gorovits, Ju Guan, Lukas Habegger, Alicia Hawes, Olga Krasheninina, Samantha Zarate, Adam J Mansfield, Lukas Habegger.

Analytical Genetics & Data Science
Gonçalo Abecasis, Manuel Allen Revez Ferreira, Joshua Backman, Kathy Burch, Adrian Campos, Liron Ganel, Sheila Gaynor, Benjamin Geraghty, Arkopravo Ghosh, Salvador Romero Martinez, Christopher Gillies, Lauren Gurski, Eric Jorgenson, Tyler Joseph, Michael Kessler, Jack Kosmicki, Adam Locke, Priyanka Nakka, Jonathan Marchini, Karl Landheer, Olivier Delaneau, Maya Ghoussaini, Anthony Marcketta, Joelle Mbatchou, Arden Moscati, Anita Pandit, Jonathan Ross, Carlo Sidore, Eli Stahl, Timothy Thornton, Sailaja Vedantam, Rujin Wang, Kuan-Han Wu, Bin Ye, Blair Zhang, Andrey Ziyatdinov, Yuxin Zou, Jingning Zhang, Kyoko Watanabe, Mira Tang, Frank Wendt, Suganthi Balasubramanian, Suying Bao, Kathie Sun, Chuanyi Zhang, Sean Yu, Aaron Zhang, David Corrigan, Dhruv Shidhaye, Chen Wang, Keyrun Adhikari, Alexander Lachmann, Mark Weiner.

Therapeutic Area Genetics
Adolfo Ferrando, Giovanni Coppola, Luca A. Lotta, Alan Shuldiner, Katherine Siminovitch, Brian Hobbs, Jon Silver, William Palmer, Rita Guerreiro, Amit Joshi, Antoine Baldassari, Cristen Willer, Sarah Graham, Ernst Mayerhofer, Erola Pairo Castineira, Mary Haas, Niek Verweij, George Hindy, Jonas Bovijn, Tanima De, Luanluan Sun, Olukayode Sosina, Arthur Gilly, Peter Dornbos, Juan Rodriguez-Flores, Moeen Riaz, Manav Kapoor, Gannie Tzoneva, Momodou W Jallow, Anna Alkelai, Ariane Ayer, Veera Rajagopal, Sahar Gelfman, Vijay Kumar, Jacqueline Otto, Jose Bras, Silvia Alvarez, Jessie Brown, Hossein Khiabanian, Joana Revez, Kimberly Skead, Valentina Zavala, Jae Soon Sul, Lei Chen, Sam Choi, Amy Damask, Nan Lin, Charles Paulding, Sameer Malhotra, Joseph Herman.

Research Program Management & Strategic Initiatives
Marcus B. Jones, Michelle G. LeBlanc, Nadia Rana, Jennifer Rico-Varela, Jaimee Hernandez, Larizbeth Romero, Ashley Paynter.

Senior Partnerships & Business Operations
Randi Schwartz, Lourdes Crane, Alison Fenney, Jody Hankins, Anna Han, Samuel Hart, Ryan Smith.

Business Operations & Administrative Coordinators
Ann Perez-Beals, Gina Solari, Johannie Rivera-Picart, Michelle Pagan, Sunilbe Siceron.

RGC contact: RGCcollaborations@regeneron.com

GHS-RGC DiscovEHR collaboration

GHS Author List:

1. Adam Buchanan. Geisinger Health System, Danville, PA, USA
    - E-mail: ahbuchanan@geisinger.edu
2. David J. Carey. Geisinger Health System, Danville, PA, USA
    - E-mail: djcarey@geisinger.edu
3. Christa L. Martin. Geisinger Health System, Danville, PA, USA
    - E-mail: clmartin1@geisinger.edu
4. Michelle Meyer. Geisinger Health System, Danville, PA, USA
    - E-mail: mmeyer@geisinger.edu
5. Kyle Retterer. Geisinger Health System, Danville, PA, USA
    - E-mail: kretterer1@geisinger.edu
6. David Rolston. Geisinger Health System, Danville, PA, USA
    - E-mail: ddrolston@geisinger.edu

Malmö Diet and Cancer Study

Olle Melander. E-mail: olle.melander@med.lu.se

Olle's affiliations:
1. Department of Clinical Science Malmö, Lund University, Malmö, Sweden
2. Department of Internal Medicine, Skane University Hospital, Malmö, Sweden

---

UCLA-RGC ATLAS Collaboration

UCLA Precision Health ATLAS
Alex Bui, Antonia Petruse, Arash Naeim, Chris Denny, Clara Lajonchere, Clara Magyar, Dan Geschwind, Maryam Ariannejad, Paul Boutros, Paul Spellman, Sarah Dry, Stan Nelson.

UCLA Precision Health Data Discovery

Albert Duntugan, Alex Bui, Bogdan Pasaniuc, Chris Denny, Clara Lajonchere, Dan Geschwind, Paul Boutros, Paul Tung, Tim Chang, Yael Berkovich.

UCLA IT High Performance Computing
Albert Duntugan, Ankur Jain, Danielle Martinez, Ghouse Mohammed, Lora Illiev, Michael Broudy, Paul Boutros, Paul Tung, Roni Haas, Shaiful Alam, Taka Yamaguchi, Yael Berkovich, Yash Patel.

Contact: precisionhealth@mednet.ucla.edu

Penn Medicine Biobank

Penn Medicine BioBank Banner Author List:

PMBB Leadership Team
Daniel J. Rader, M.D., Marylyn D. Ritchie, Ph.D.
Contribution: All authors contributed to securing funding, study design and oversight. All authors reviewed the final version of the manuscript.

Patient Recruitment and Regulatory Oversight
JoEllen Weaver, Nawar Naseer, Ph.D., M.P.H., Giorgio Sirugo, M.D., P.h.D., Afiya Poindexter, Yi-An Ko, Ph.D., Kyle P. Nerz, Jenna Dever, Aidan Harvey, Sydney Linn
Contributions: JW manages patient recruitment and regulatory oversight of study. NN manages participant engagement, assists with regulatory oversight, and researcher access. GS assists with researcher access. AP, YK, KPN, JD, AH, and SH perform recruitment and enrollment of study participants.

Lab Operations
JoEllen Weaver, Meghan Livingstone, Fred Vadivieso, Stephanie DerOhannessian, Teo Tran, Julia Stephanowski, Salma Santos, Ned Haubein, P.h.D., Joseph Dunn
Contribution: JW, ML, FV, SD conduct oversight of lab operations. ML, FV, AK, SD, TT, JS, SS perform sample processing. NH, JD are responsible for sample tracking and the laboratory information management system.

Clinical Informatics
Anurag Verma, Ph.D., Colleen Morse Kripke, M.S. DPT, MSA, Marjorie Risman, M.S., Renae Judy, B.S., Colin Wollack, M.S.
Contribution: All authors contributed to the development and validation of clinical phenotypes used to identify study subjects and (when applicable) controls.

Genome Informatics
Anurag Verma Ph.D., Shefali S. Verma, Ph.D., Scott Damrauer, M.D., Yuki Bradford, M.S., Scott Dudek, M.S., Theodore Drivas, M.D., Ph.D.,
Contribution: AV, SSV, and SD are responsible for the analysis, design, and infrastructure needed to quality control genotype and exome data. YB performs the analysis. TD and AV provide variant and gene annotations and their functional interpretation of variants.

Contact: biobank@pennmedicine.upenn.edu

BELIEVE Study

The BELIEVE study Banner Author List:

Recruitment Coordinators and Field Supervisors: Israt Akter, Laboni Akter, Md Shehab Uddin Al-Abid, Arnab Aditya Das, Khairul Islam, Mir Shahadul Islam, Shoriful Islam, Mohammad Kamruzzaman, Md Taslim Uddin Miah, Monalisa Moni, Sabrina Monsur, Md Mostafa Monower, K M Thouhidur Rahman, Anjuman Ara Rahman, Mantaka Rahman, A H M Rezwan, Jayashree Saha, Sharraf Samin, Kazi Nazmus Saqeeb, Monjeline Sultana, Ishrat Tasmin.

Study Site Support Workers: Md Rezaul Karim Akanda, Jesmin Akhter, Ayasha Akter, Bakul Akter, Jharna Akter, Jesmin Akter, Juba Akter, Khadiza Akter, Khadija Akter, Lipi Akter, Maksuda Akter, Mousumi Akter, Mst Lovely Akter, Nahida Akter, Nasima Akter, Samima Akter, Sema Akter, Shahida Akter, Shahnaj Akter, Shamima Akter, Taslima Akter, A T M Zorjis Alam, Mahmuda Atique, Lutfa Begum, Mst Nazma Begum, Farjana Choudhury, Md Zakir Hossain Chowdhury, Mitali Paul Chowdhury, Robin Reza Chowdhury, Mukul Rani Debnath, Kaniz Fatema, Nahid Ferdash, Naima Ferdous, Md Rakib Al Hasan, Khandaker Hashanuzzaman, Shamima Haq, Md Riazul Haque, Alamgir Hossain, Md Ibrahim, Nusrat Jahan, Shahi Israt Jahan, Israt Jahan Jarin, Zohora Pradhan Jonaki, Asik Kabir, Tonema Kader, Md Mostafa Kamal, Sayed Kamruzzaman, Sadik Fatima Kanon, Nazmul Karim, Shamsul Karim, Tanuja Khanom, Shamim Ara Khatun, Badrun Nahar Lorin, Mst Sirajum Manira, Farhana Jahan Mary, Kazi Dilruba Mita, Lipi Mitra, Basudeb Mollik, Kamrun Nahar, Sudipta Nargis, Nusrat Alam Nawmee, Esrat Zahan Nesa, Mahmuda Akter Nipa, Mehenaz Parvin, Sanjida Parvin, Shahnaj Parvin, Suraiya Parvin, Kaniz Fatema Priya, Golam Mostafa Quadrey, Nayan Rabidash, Rulia Rahman, Madhabi Rani, Shahjalal Sarker, Smriti Sarker, Razia Sultana Shathi, Fatema Shelly, Ireen Sultana, Rovaiya Sultana, Israt Jahan Sumi, Sharmin Tamanna, Khadija Akter Topy, Umme Habiba Urmee, Suraya Yesmin, Shafia Zerin.

Data Management Team: Julianne Halley, Catherine Perry, Sarah Spackman, Charlotte van Coeverden, Matthew Walker.

Laboratory Team: Mahmuda Akther Akhi,Asia Akter, Labony Akter, Ms Mili Akter, Mst Shamima Akter, Sabina Akter, Salma Akter, Setu Akter, Tahmina Akter, Md Sabdar Ali, Mst Jesmin Ara, Edyta Bujnik, Jason Crawte, Apu Chandra Das, Many Das, Samantha Farrow, Nurjahan Fatema, Md Riyad Hasan, Md Saimul Islam, Soniya Jannat, Mst Amena Khatun, Most Nurnahar Khatun, William Mossman, Robyn Murdoch, Zannaton Naeem, Danh Nguyen-Murray, Evrikleia Ntasi, Silvia Alonso Rodriguez, Most Rezina Akter Roma, Tripty Roy, Most Abida Sultana, Sharmin Akter Samia.

Study Administration Team: Nazneen Ali, Laryssa Amado, Tanya Braune, Eilidh Cowan, Steve Ellis, Ellie Farrow, Richard Houghton, Md Zahidul Islam, Giulia Loffreda, Hannah Lombardi, Ank Michielsen, Niko Ovenden, Tamara Sabri, Karen Saunders neé Heasley, Md

Rafiqul Islam Rabbi, Hosne Ara Rekha, Valerie Rhenius, Md Khalid Sultan, Sophie Weston, Hannah Williams.

Scientific Investigators and Collaborators: Tahmeed Ahmed, Jim Ajioka, Khondker Abdul Awal, Aytalina Azarova, Arul Baradi, Adam S Butterworth, Evangelia Chatzidiakou, Sohel Reza Choudhury, Rajiv Chowdhury, Mayank Dalakoti, John Danesh, Emanuele Di Angelantonio, Camilla Faidutti, Joerg Feldmann, Richard Fenner, Meerjady Sabrina Flora, Simon Griffin, Louise Hair, Sharifuddin Hasnat, Sarah Hawkes, Shahid Akhter Hossain, Tuhin Haque, Tafsir Hassan, Md Mominul Islam, Shafiul Islam, Maria L C Iurilli, Syed Shariful Islam, Roderic L Jones, Stephen Kaptoge, Md Khalequzzaman, Md AlfazalKhan, Kamrul Hasan Khan, Nusrat Khan, Lawrence King, Joe Lavallée, Shammi Luhar, Abdul Malik, Fazila Tun-Nesa Malik, Nick Mascie-Taylor, Md Mostafa Monower, Ruchira Tabassum Naved, Md Sirajul Islam, Aliya Naheed, Anisur Rahman, Lisa Pennells, Olalekan Popoola, Mahbubur Rahman, Rubhana Raqib, Laurie Savage, Sara Shazad, Tahmina Shirin, Lalitha Sundaram, Stephen Sutton, Henry Taylor, Aloka Tulukdar, Kim Van Daalen, Angela Wood, Luisa Zuccolo.

Fellows and Trainees: Sadika Akhter, Tanvir Chowdhury, Nurul Huda, Samia Naz Isha, Riaz Hossain Khan, Samsad Rabbani Khan, Md Mostafa Monower, Aliva Salmeen, Zeeba Zahra Sultana, Animesh Talukder, Renesa Tarannum.

External Advisory Board (CAPABLE Initiative): Colin Baigent (Chair), Abbas Bhuiya, Shahida 'Lucky' Parvin.

Steering Committee & Data and Sample Access Committee (current membership): Sohel Reza Choudhury, John Danesh, Emanuele Di Angelantonio, Simon Griffin, Md Khalequzzaman, Md Alfazal Khan, Nick Mascie-Taylor.

Colorado Center for Personalized Medicine (CCPM)

Heather D. Anderson, Christina L. Aquilante, Kelsey Arbogast, Ian M. Brooks, Elizabeth E. Burke, Emily M. Casteel, Joanne B. Cole, Curtis R. Coughlin II, Jacob Crawford, Kristy Crooks, Erin Culver, Matthew J. Fisher, Teresa C. Frye, Hunter George,
Chris R. Gignoux, Elizabeth K. Gilliland, Casey S. Greene, Emily Hearst, Audrey E. Hendricks, Randi K. Johnson, Shelby Jones, Dave Kao, Gabrielle A. Knortz, Danielle Koffenberger, Santhanagopalan Krishnamoorthy, Lisa Ku, Elizabeth L. Kudron, Rashawnda Lacy, Ethan M. Lange, Joe A. Lesny, Meng Lin, James L. Martin.

Contact: CCPM-Biobank@ucdenver.edu

Mayo Clinic-RGC Project Generation (PG)

PG Leadership Team
Cerhan, James R., M.D.; Couch, Fergus J., Ph.D., Olson, Janet E., Ph.D.

Statistical Genetics and Bioinformatics
Larson, Nicholas B., Ph.D., M.S.; Fredericksen, Zachary S.

Laboratory Operations
Cicek, Mine, Ph.D.

Registry Principal Investigators
1. Alcohol Use Disorder (AUD): Biernacka, Joanna M., Ph.D., Karpyak, Victor M., M.D. Ph.D.
2. Alzheimer's Disease Research Center (ADRC): Vemuri, Prashanthi, Ph.D.; Ramanan, Vijay K., M.D., Ph.D.
3. Bipolar disorder registry: Biernacka, Joanna M., Ph.D., Frye, Mark A., M.D.
4. Brain: Eckel Passow, Jeanette E., Ph.D.
5. Breast - Mayo Florida: McLaughlin, Sarah A., M.D.
6. Breast - Mayo Rochester: Olson, Janet E., Ph.D.; Couch, Fergus J., Ph.D., Ruddy, Kathryn J., M.D.
7. Cardiovascular Disease Specimen Repository: Bielinski, Suzette J., Ph.D., M.Ed.
8. Chronic Kidney Disease: Lieske, John C., M.D.
9. Chronic Pain: Hooten, W. Michael, M.D.
10. Colorectal: Boardman, Lisa A., M.D.
11. COVID-19 Biobank: Kennedy, Richard B., Ph.D.; Cerhan, James R., M.D., Ph.D.; Badley, Andrew D., M.D.
12. Endometrial: Dowdy, Sean C., M.D.; Bakkum-Gamez, Jamie N., M.D.; Glaser, Gretchen E., M.D.
13. Lung: Yang, Ping, M.D., Ph.D.
14. Lymphoma: Cerhan, James R., M.D., Ph.D.
15. Mayo Clinic Biobank: Olson, Janet E., Ph.D.
16. Mayo Clinic Study of Aging (MCSA): Vemuri, Prashanthi, Ph.D.; Ramanan, Vijay K., M.D., Ph.D.
17. Mayo Mammography Health Study: Vachon, Celine M., Ph.D.
18. Multiple Myeloma (MM)/Smouldering MM (SMM): Dispenzieri, Angela, M.D.; Vachon, Celine M., Ph.D.
19. Neuroendocrine pancreatic tumor registry: Antwi, Samuel O., Ph.D.; Ann L. Oberg, Ph.D.; Kari G. Rabe, MS.
20. Mayo Clinic Biospecimen Resource for Ovarian Cancer Research: Kaufmann, Scott H.,M.D., Ph.D.; Goode, Ellen L., Ph.D.; William A. Cliby, M.D.
21. Biospecimen Resource for Pancreas Research: Antwi, Samuel O., Ph.D.; Ann L. Oberg, Ph.D.; Kari G. Rabe, MS.
22. Parkinson’s Disease: Ahlskog, J. Eric, M.D., Ph.D.; Bower, James H., M.D.
23. Polycystic kidney disease: Harris, Peter C., Ph.D.
24. Polyps: Boardman, Lisa A., M.D.
25. Prevalence of Asymptomatic Ventricular Dysfunction: Pereira, Naveen L., M.D.
26. PRISM Mammography study: Couch, Fergus J., Ph.D.; Vachon, Celine M., Ph.D.; Olson, Janet E., Ph.D.
27. Prostate: Cicek, Mine, Ph.D.
28. Prostate Family: Cicek, Mine, Ph.D.
29. 2Radiation Oncology Registry: Ma, Daniel J., M.D.; Mutter, Robert W., M.D.
30. Renal: Eckel Passow, Jeanette E., Ph.D.

31. Vascular Diseases Biorepository: Pereira, Naveen L., M.D.

Management
Colborn, Lisa K., M.B.A.; Danielsen, Andrew J.; Harrington, Jonathan J.; Kushwaha, Jennifer M.

South Asia Biobank Study

1. Prof. Anu Kasturiratne, Faculty of Medicine, University of Kelaniya
   - E-mail: anuradhani@kln.ac.lk
2. Dr. Khadija I Khawaja, Institute of Endocrinology and Metabolic Diseases, Services Institute of Medical Sciences, Lahore, Pakistan
   - E-mail: khadijairfan@yahoo.com
3. Dr. Malay Mridha, Centre for Non-Communicable Diseases and Nutrition, BRAC James P Grant School of Public Health, BRAC University, Dhaka, Bangladesh
   - E-mail: malay.mridha@bracu.ac.bd
4. Prof. John C Chambers, School of Public Health, Imperial College London, UK, Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, 308232, Singapore
   - E-mail: john.chambers@imperial.ac.uk

Contact: john.chambers@imperial.ac.uk

UTSW-Dallas Biobank

1. Julia Kozlitina, The Eugene McDermott Center for Human Growth and Development, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA.
   - E-mail: Julia.Kozlitina@UTSouthwestern.edu
2. Jonathan Cohen, The Center for Human Nutrition, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA
   - E-mail: Jonathan.Cohen@UTSouthwestern.edu

Contact: Julia.Kozlitina@UTSouthwestern.edu

Mount Sinai Million Health Discoveries Program

- Alexander W. Charney, MD
- **Affiliation:** Icahn School of Medicine at Mount Sinai
- **Email:** Alexander.Charney@mssm.edu

For any questions regarding this dataset, please contact Alexander Charney.

Supplementary Table 1: Anthropometric data for three homozygote carriers of *MSTN* missense mutations found in UK Biobank. Data reported is untransformed, with % deviation from covariate-adjusted expected values shown alongside. Results for body fat percentage, body fat-free mass, and fat free mass index (FFMI) are derived from bioimpedance analysis (BIA). Expected values computed using a linear model with sex, age, age squared, height, height squared, and sex interaction terms as predictors.

Genotype	Hom Ile225Thr	Hom Ile225Thr	Hom Arg65His
Sex	F	F	F
Age	55-60	50-55	65-70
BMI (kg m^{-2})	36.8	22.3	25.7
Body fat %	43.3 (+17.3%)	30.0 (-16.4%)	25.7 (-31.8%)
Body fat-free mass (kg)	55.4 (+24.0%)	41.1 (-8.0%)	47.1 (+12.3%)
FFMI (kg m^{-2})	21.9 (+24.0%)	16.7 (-8.0%)	20.6 (+12.3%)
Creatinine (μmol dm^{-3})	75.3 (+16.8%)	46.7 (-26.0%)	77.1 (+17.8%)
Creatinine:cystatin C ratio	78.7 (-5.8%)	69.7 (-20.3%)	107.4 (+37.7%)

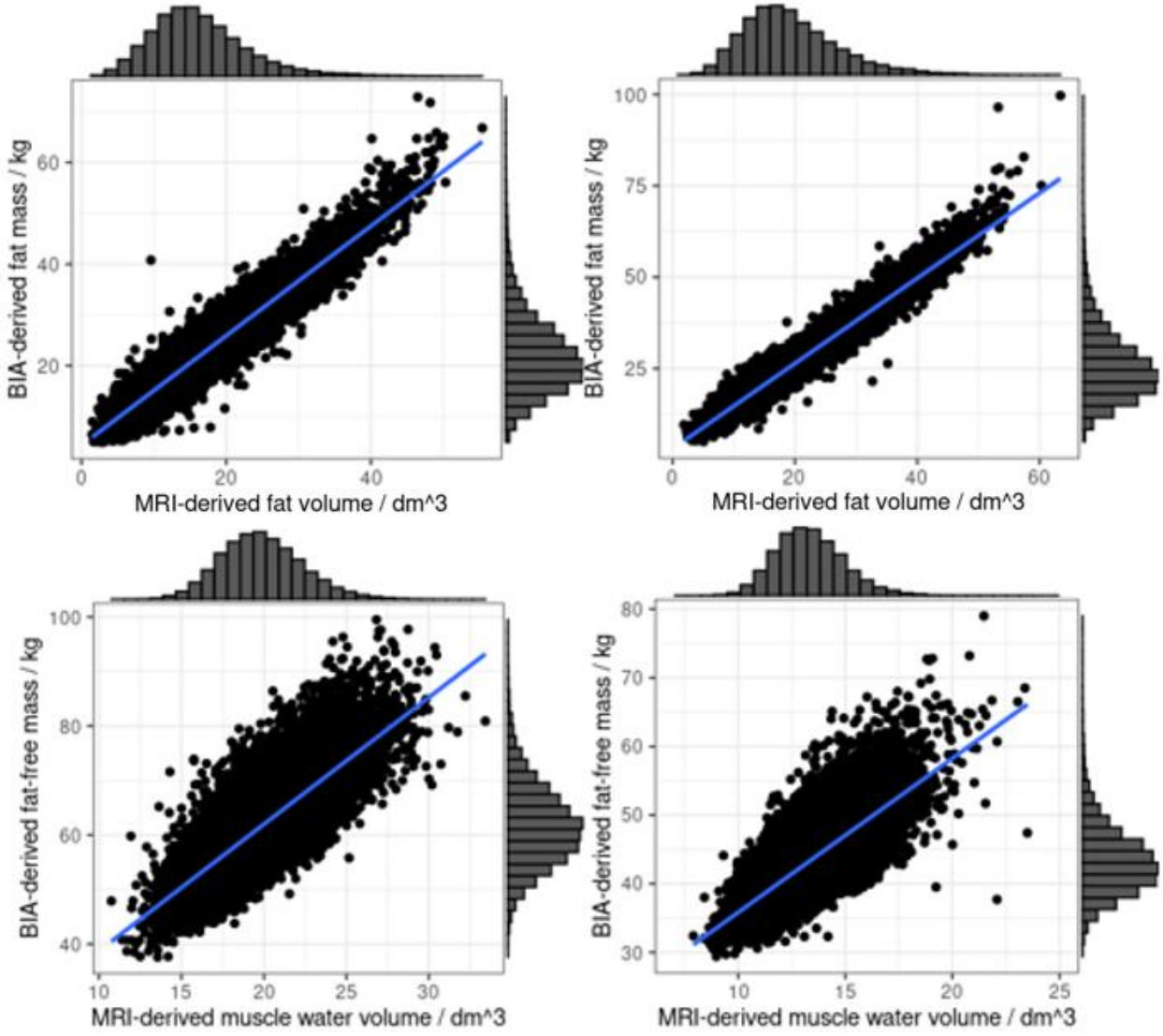
Supplementary Table 2: Five-fold cross-validation mean Dice scores (unitless) for the custom segmentation model of the four thigh muscle groups segmented for left and right legs, demonstrating strong concordance with manual annotations.

	Fold 1	*Fold 2*	*Fold 3*	*Fold 4*	*Fold 5*
Right quadriceps femoris	0.96	0.93	0.91	0.98	0.96
Left quadriceps femoris	0.97	0.95	0.93	0.95	0.96
Right medial	0.93	0.94	0.91	0.97	0.96
Left medial	0.93	0.94	0.96	0.96	0.92
Right posterior	0.93	0.94	0.92	0.96	0.95
Left posterior	0.93	0.92	0.94	0.94	0.93
Right sartorius	0.92	0.89	0.88	0.93	0.90
Left sartorius	0.94	0.89	0.92	0.85	0.87

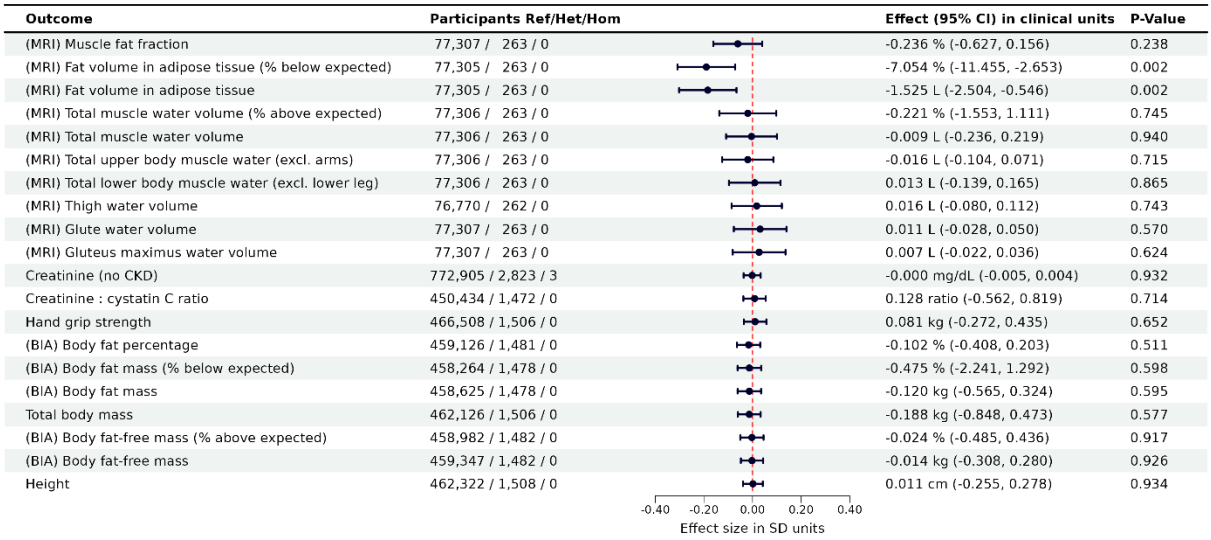
Supplementary Table 3: Five-fold cross-validation mean Dice scores (unitless) for the custom segmentation model of the five adipose tissue depots, demonstrating strong concordance with manual annotations.

	Fold 1	*Fold 2*	*Fold 3*	*Fold 4*	*Fold 5*
Upper body subcutaneous fat	0.89	0.93	0.96	0.84	0.96
Abdominal subcutaneous fat	0.92	0.90	0.93	0.86	0.93
Gluteofemoral subcutaneous fat	0.91	0.90	0.95	0.91	0.94
Visceral fat	0.92	0.91	0.93	0.89	0.85
Mediastinal fat	0.83	0.90	0.87	0.63	0.56

Supplementary Figure 1. Concordance between MRI-and BIA-derived measures of fat and fat-free mass for subjects in the UK Biobank imaging cohort, shown for men (left) and women (right).
Linear model fits were:
Men: $FM_{BIA} = 3.57 + 1.13 * FV_{MRI}$ ($R^2 = 0.90$) and $FFM_{BIA} = 15.46 + 2.33 * WV_{MRI}$ ($R^2 = 0.62$)
Women: $FM_{BIA} = 3.04 + 1.17 * FV_{MRI}$ ($R^2 = 0.95$) and $FFM_{BIA} = 13.39 + 2.24 * WV_{MRI}$ ($R^2 = 0.56$)
where FM = fat mass, FV = fat volume, FFM = fat-free mass, and WV = muscle water volume. MRI fat volumes are computed using the total across all fat depots, and total muscle water volume was computed using the sum over muscle groups segmented by TVS. The expected slope for fat mass versus fat volume given a fat density of 0.9 kg/dm^3 would be 0.9, and the slope of fat-free mass versus muscle water volume would be 2 in the case that muscle constitutes approximately 50% of BIA-derived fat-free mass. The higher estimated slope values versus these expectations are likely due to exclusion of arms and lower legs from the MRI-derived measurements, since these are partially or fully outside the field of view in most subjects. As an approximate estimate, the volume of muscle contained within lower legs may constitute around 1.3 dm^3 (Fuchs *et al.*, 2023) and the volume in shoulders and arms may constitute approximately 2 dm^3 (Holzbaur *et al.*, 2007); excluding these volumes will lead to an average underestimate of around 18% of the total muscle volume in UKB.

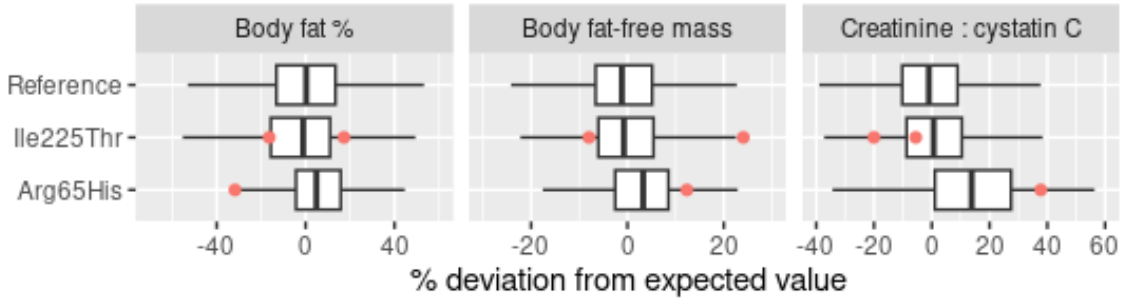


Supplementary Figure 2. *MSTN* Arg88Gln single variant associations with body composition, quantitative biomarkers, and MRI-derived fat and muscle metrics. This variant was not included in the set of predicted function-disrupting variants as defined by the 5-algorithm consensus described in the Methods. Effect size estimates are plotted in standard deviation units with numerical values shown alongside in clinical units; p-values are computed from gene burden test statistics using a chi-squared test (see Methods for additional details). Error bars indicate 95% confidence intervals (±1.96 SE); arrows indicate the error bar extends beyond axis limits. Summary statistics are included in Supplementary Data 6.



Outcome	Participants Ref/Het/Hom	Effect (95% CI) in clinical units	P-Value
(MRI) Muscle fat fraction	77,307 / 263 / 0	-0.236 % (-0.627, 0.156)	0.238
(MRI) Fat volume in adipose tissue (% below expected)	77,305 / 263 / 0	-7.054 % (-11.455, -2.653)	0.002
(MRI) Fat volume in adipose tissue	77,305 / 263 / 0	-1.525 L (-2.504, -0.546)	0.002
(MRI) Total muscle water volume (% above expected)	77,306 / 263 / 0	-0.221 % (-1.553, 1.111)	0.745
(MRI) Total muscle water volume	77,306 / 263 / 0	-0.009 L (-0.236, 0.219)	0.940
(MRI) Total upper body muscle water (excl. arms)	77,306 / 263 / 0	-0.016 L (-0.104, 0.071)	0.715
(MRI) Total lower body muscle water (excl. lower leg)	77,306 / 263 / 0	0.013 L (-0.139, 0.165)	0.865
(MRI) Thigh water volume	76,770 / 262 / 0	0.016 L (-0.080, 0.112)	0.743
(MRI) Glute water volume	77,307 / 263 / 0	0.011 L (-0.028, 0.050)	0.570
(MRI) Gluteus maximus water volume	77,307 / 263 / 0	0.007 L (-0.022, 0.036)	0.624
Creatinine (no CKD)	772,905 / 2,823 / 3	-0.000 mg/dL (-0.005, 0.004)	0.932
Creatinine : cystatin C ratio	450,434 / 1,472 / 0	0.128 ratio (-0.562, 0.819)	0.714
Hand grip strength	466,508 / 1,506 / 0	0.081 kg (-0.272, 0.435)	0.652
(BIA) Body fat percentage	459,126 / 1,481 / 0	-0.102 % (-0.408, 0.203)	0.511
(BIA) Body fat mass (% below expected)	458,264 / 1,478 / 0	-0.475 % (-2.241, 1.292)	0.598
(BIA) Body fat mass	458,625 / 1,478 / 0	-0.120 kg (-0.565, 0.324)	0.595
Total body mass	462,126 / 1,506 / 0	-0.188 kg (-0.848, 0.473)	0.577
(BIA) Body fat-free mass (% above expected)	458,982 / 1,482 / 0	-0.024 % (-0.485, 0.436)	0.917
(BIA) Body fat-free mass	459,347 / 1,482 / 0	-0.014 kg (-0.308, 0.280)	0.926
Height	462,322 / 1,508 / 0	0.011 cm (-0.255, 0.278)	0.934

Supplementary Figure 3: Distributions of deviations from expected values for fat-free mass and creatinine:cystatin c ratio for homozygote and heterozygote carriers of Arg65His and Ile225Thr. Homozygote missense carriers (red dots) are shown relative to heterozygote carriers (box plots) for Ile225Thr (second row, N=1,786 for body fat percentage and fat-free mass, N=1,793 for creatinine:cystatin c ratio) and Arg65His (third row, N=113 for body fat percentage and fat-free mass, N=114 for creatinine:cystatin c ratio). The distribution of homozygous reference individuals is shown on the top row for comparison (N= 458,758 for body fat percentage and fat-free mass, N=461,190 for creatinine:cystatin c ratio). Expected values computed using a linear model with sex, age, age squared, height, height squared and sex interaction terms as predictors. Boxes show medians and interquartile ranges (IQR); whiskers extend in each direction to the most extreme data point that is no further than 1.5 multiples of the IQR from the upper and lower quartile. Mean (SD) percent deviations for Arg65His heterozygotes were +5.3% (17.7) for body fat percentage, +3.8% (9.75) for body fat-free mass, and +13.9% (20.2) for creatinine : cystatin C ratio. Mean (SD) percent deviations for Ile225Thr heterozygotes were -2.3% (20.6) for body fat percentage, +0.3% (9.23) for body fat-free mass, and +1.8% (16.0) for creatinine : cystatin C ratio. While homogygous effect sizes cannot be accurately estimated due to the low sample sizes, Z-scores for the homozygotes relative to the heterozygote distribution were calculated, and converted to one-sided p-values, to provide a measure of deviation of the homozygote(s) from the heterozygote distributions. For the single Arg65His homozygote, the Z-scores and corresponding p-values obtained by this procedure were 2.08 (p=0.018) for body fat percentage, 0.87 (p=0.19) for body fat-free mass, and 1.18 (p=0.12) for creatinine : cystatin C ratio.



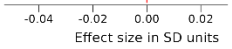
Supplementary Figure 4. Gene burden association analysis of *MSTN* with cardiometabolic disease outcomes (A) and quantitative biomarkers (B). Predicted loss-of-function and deleterious missense variants in *MSTN* with allele frequency below 0.5% were included in the gene burden tests. For the Malmo cohort the Ile225Thr variant has a frequency of 0.66%, but was included in the gene burden test for this cohort since the allele frequency is below 0.5% in all other cohorts. Effect sizes are reported in odds ratios (OR) for disease outcomes (**A**), and standard deviation (SD) and clinical units for quantitative traits (**B**). Error bars indicate 95% confidence intervals, calculated as ±1.96 SE from the effect estimate; for binary traits, the interval was calculated on the logistic scale and converted to OR units. p-values are computed from gene burden test statistics using a chi-squared test (see Methods for additional details). Arrows indicate the error bar extends beyond axis limits. Summary statistics are included in Supplementary Data 8.

A.

Outcome	Cases Ref/Het/Hom	Controls Ref/Het/Hom	OR (95% CI)	P-Value
Cardiomyopathy (Hypertrophic)	3,403 / 18 / 0	588,838 / 3,945 / 3	0.799 (0.525, 1.218)	0.298
Coronary Artery Disease (Loose)	152,007 / 1,012 / 1	667,381 / 4,210 / 3	0.972 (0.899, 1.051)	0.477
Liver Disease (Non Alcoholic)	51,863 / 331 / 1	741,063 / 4,942 / 3	0.973 (0.867, 1.092)	0.637
Liver Cirrhosis (Any)	14,601 / 96 / 0	741,063 / 4,942 / 3	1.001 (0.810, 1.236)	0.994
Type 2 diabetes	159,939 / 1,010 / 1	632,830 / 3,955 / 3	1.011 (0.933, 1.095)	0.793
Heart failure	82,813 / 610 / 1	592,427 / 3,956 / 3	1.028 (0.934, 1.131)	0.570
Cardiomyopathy (unspecified)	24,131 / 184 / 0	588,838 / 3,945 / 3	1.096 (0.930, 1.292)	0.272

B.

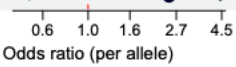
Outcome	Participants Ref/Het/Hom	Effect (95% CI) in clinical units	P-Value
Waist Hip Ratio adjusted for BMI	617,278 / 3,589 / 4	-0.002 ratio (-0.004, 0)	0.049
HbA1c (all)	807,828 / 4,929 / 4	-0.021 % (-0.043, 0.001)	0.056
Alanine Aminotransferase	895,752 / 5,664 / 5	-0.277 U/L (-0.643, 0.088)	0.137
Random glucose (all)	963,619 / 6,080 / 5	-0.232 mg/dL (-0.864, 0.4)	0.474
Triglycerides	863,359 / 5,560 / 5	-0.863 mg/dL (-3.274, 1.548)	0.486
Body Mass Index	985,795 / 6,044 / 6	-0.035 kg/m2 (-0.168, 0.097)	0.600
Aspartate Aminotransferase	913,721 / 5,827 / 5	-0.007 U/L (-0.268, 0.253)	0.956
Systolic Blood Pressure	1,105,210 / 6,921 / 6	0.011 mmHg (-0.359, 0.381)	0.954
LDL cholesterol (Friedewald)	809,201 / 5,244 / 5	0.139 mg/dL (-0.975, 1.253)	0.805
HDL cholesterol	821,563 / 5,313 / 5	0.126 mg/dL (-0.258, 0.51)	0.519



Supplementary Figure 5. Gene burden association analysis of *MSTN* with fertility-related quantitative traits (A) and fertility and polycystic ovary syndrome (PCOS; B). Burden test associations are reported for predicted loss-of-function and deleterious missense variants in *MSTN* with allele frequency below 0.5% and fertility-related traits. Effect sizes are reported in odds ratios (OR) for disease outcomes (**A**), and standard deviation (SD) and clinical units for quantitative traits (**B**). Error bars indicate 95% confidence intervals, calculated as ±1.96 SE from the effect estimate; for binary traits, the interval was calculated on the logistic scale and converted to OR units. p-values are computed from gene burden test statistics using a chi-squared test (see Methods for additional details). Arrows indicate the error bar extends beyond axis limits. PCOS indicates polycystic ovary syndrome. FSHB indicates follicle stimulating hormone subunit beta levels. Summary statistics are included in Supplementary Data 9.

A.

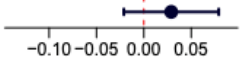
Outcome	Cases Ref/Het/Hom	Controls Ref/Het/Hom	OR (95% CI)	P−Value
PCOS (diagnosis)	965 / 3 / 0	189,029 / 1,152 / 3	0.607 (0.264, 1.397)	0.24
PCOS (diagnosis or Rotterdam)	1,128 / 8 / 0	189,029 / 1,152 / 3	1.205 (0.559, 2.597)	0.63
Infertility (female, any)	2,917 / 22 / 0	232,234 / 1,392 / 3	1.333 (0.823, 2.160)	0.24
Infertility (sex−combined)	3,189 / 25 / 0	431,548 / 2,537 / 3	1.393 (0.879, 2.210)	0.16
Infertility (male, any)	272 / 3 / 0	199,314 / 1,145 / 0	2.486 (0.516, 11.983)	0.26



Odds ratio (per allele)

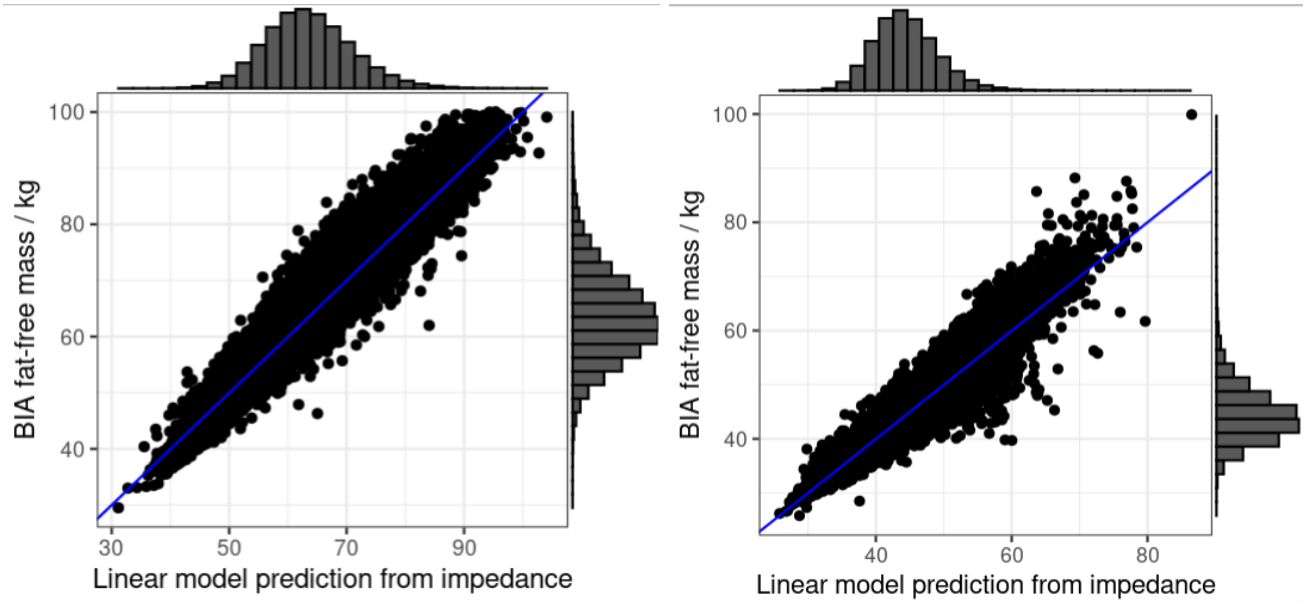
B.

Outcome	Participants Ref/Het/Hom	Effect (95% CI)	P−Value
FSHB	50,129 / 316 / 0	−3.00 % change (−9.31, 3.74))	0.37
FSHB (females)	27,033 / 170 / 0	−5.79 % change (−13.13, 2.10)	0.15
FSHB (males)	23,096 / 146 / 0	−2.67 % change (−12.82, 8.67)	0.63
Age of menarche	245,507 / 1,486 / 2	−0.027 years (−0.103, 0.050)	0.5
Age of menopause	155,650 / 968 / 2	0.299 years (−0.010, 0.608)	0.058
Number of live births	466,619 / 2,754 / 3	0.013 count (−0.030, 0.055)	0.56
Number of live births (female)	252,235 / 1,521 / 3	−0.005 count (−0.059, 0.048)	0.84
Number of live births (male)	211,704 / 1,219 / 0	0.039 count (−0.029, 0.107)	0.26

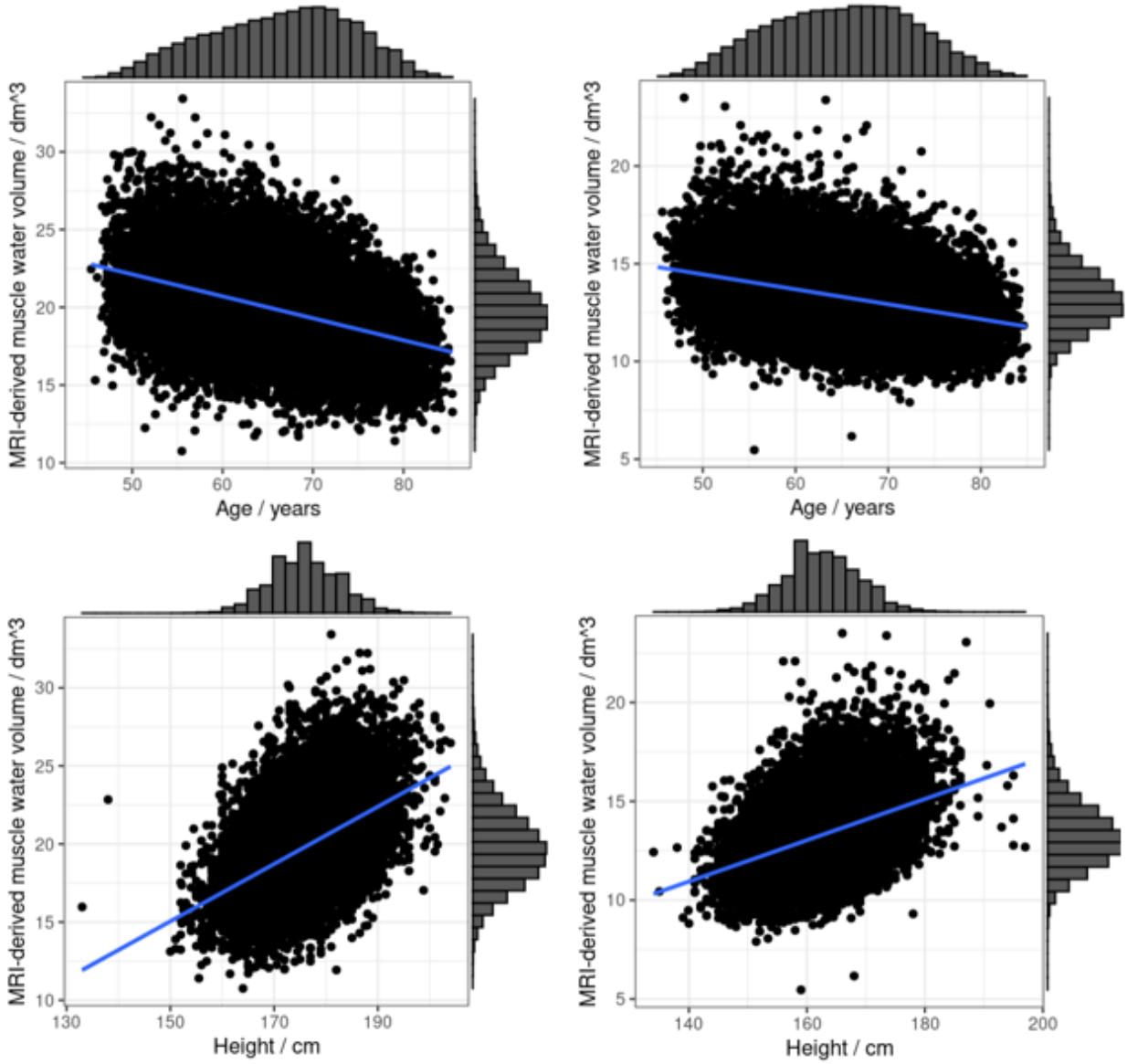


Per allele beta (95% CI) in SD units

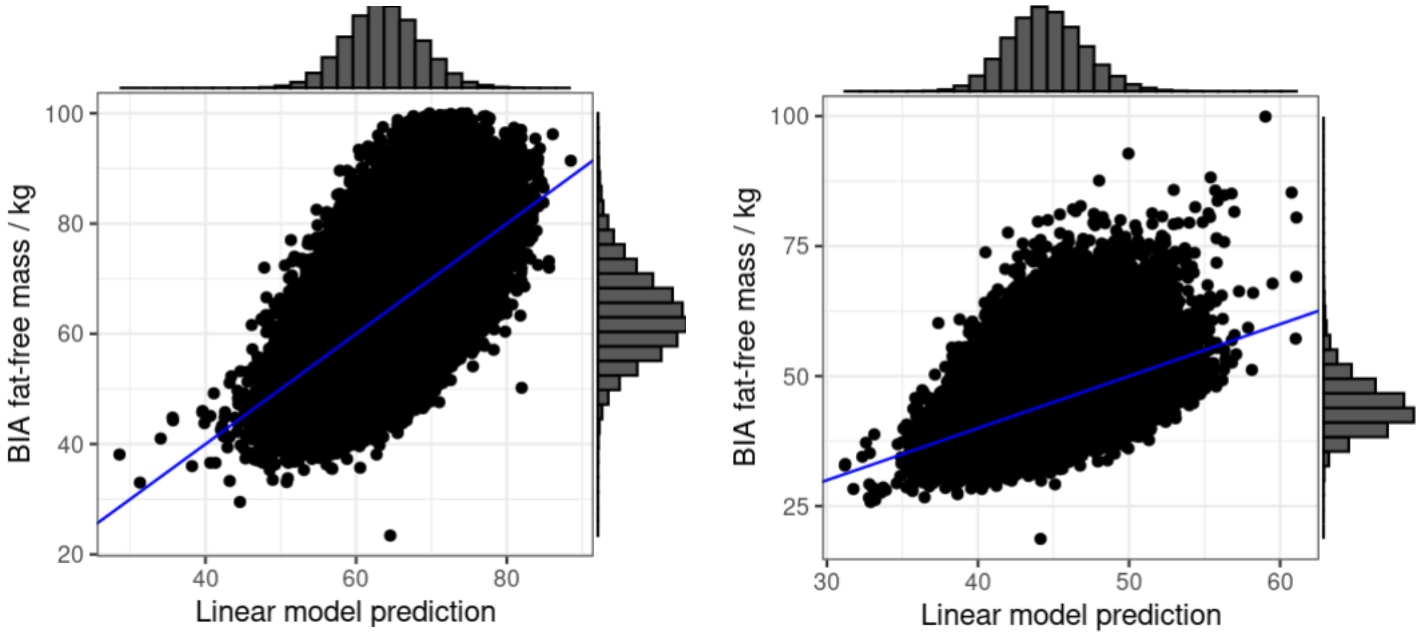
Supplementary Figure 6. Comparison between downloaded values of fat-free mass from UKB (field 23101) and values obtained from a linear fit of inverse whole-body impedance (UKB field 23106) of the form FFM ~ $1/R + h^2/R + age + age^2$, where R is impedance in Ohms and h is height, fitted separately for men (left) and women (right). Instance 0 values were used to fit the model, after filtering outliers with impedance < 360 Ohms. Identity lines are shown in blue. All terms in the model were estimated as significantly different from zero, and the R^2 for men and women were 0.93 and 0.91 respectively, indicating that the proprietary model used to derive these fat-free mass values from the impedance measurements is well approximated by an equation of this form, as noted in previous studies (Janssen *et al.*, 2000). A corollary of this is that estimated fat-free mass will have a positive genetic correlation with height.



Supplementary Figure 7: Correlations of total body muscle water volume with age and height, stratified by sex. Scatter plots show total muscle water volume versus age (top) for men (left) and women (right); and versus height (bottom) for men (left) and women (right). As expected, muscle water volume decreases with age (top rows), and increases in an approximately linear fashion with height (fit of linear model with single predictor shown in blue).



Supplementary Figure 8: Comparisons between BIA-derived fat-free mass values obtained from UK Biobank, and predicted values using a linear model with age, sex and height, plus squared and interaction terms, for men (left) and women (right). The overall R^2 for this model was 0.80, and 0.41 and 0.25 for men and women respectively. Percentage deviations from these predicted values were used to assess the effects on fat-free mass in excess of variation expected due to differences in height. Identity lines are shown in blue for reference. The fit for fat mass, defined as weight minus fat-free mass, was poor with this model (overall R^2=0.08), suggesting that estimated fat mass may be less sensitive to variations in body size.



Supplementary Data Captions (data provided in Supplementary_Data.xlsx)

For summary statistic tables, effect values are in standard deviation (SD) units for quantitative traits or odds ratios (ORs) for binary traits. Effect estimates and p-values for single-variant and gene burden tests were obtained via REGENIE v4.1, using default settings.

The lci_effect and uci_effect columns indicate lower and upper bounds for the 95% confidence intervals. For quantitative traits these are calculated as ±1.96 standard errors from the effect estimate; for binary traits, the confidence interval is computed on the logistic scale and converted into OR units.

The num_cases and num_controls columns indicate the total number of cases and controls included in the analysis. The case_ref, case_het, and case_alt columns indicate the case counts who respectively do not carry a variant in the burden test, are heterozygous for a variant included in the burden test, or homozygous for a variant included in the burden test.

The control_ref, control_het, and control_alt columns indicate the control counts who respectively do not carry a variant in the burden test, are heterozygous for a variant included in the burden test, or homozygous for a variant included in the burden test.

In all cases, our analyses were based on the GRCh38 genome assembly.

Supplementary Data 1. List of cohorts included for each phenotype where cross-cohort meta-analyses were available.
Supplementary Data 2. List of all variants predicted to induce loss of function of MSTN as detected in whole exome sequencing data from 11 cohorts (see Supplementary Data 1 for cohort list). Positions are based on the GRCh38 genome assembly.
Supplementary Data 3. Summary statistics for gene burden associations between MSTN and quantitative traits related to lean mass and fat mass, including (1) MSTN pLoF + missense (5/5), AAF<0.5%, (2) MSTN pLoF, AAF<0.5%, and (3) MSTN non-function-disrupting missense, AAF<0.5% (the latter includes all missense variants not present in the other two categories). AAF indicates alternative allele frequency. pLoF indicates putative loss-of-function. Also included are summary statistics for sex-stratified analyses.
Supplementary Data 4. Summary statistics for associations between MSTN gene burden variants and DXA-derived measures of lean and fat mass. AAF indicates alternative allele frequency. pLoF indicates putative loss-of-function.
Supplementary Data 5. Summary statistics for associations between individual MSTN missense variants and measures of lean and fat mass. Positions are based on the GRCh38 genome assembly.
Supplementary Data 6. Summary statistics for custom gene burden associations between MSTN missense masks and MRI-derived fat volume traits, including (1) MSTN pLoF + missense (5/5), AAF<0.5%, (2) MSTN any other missense, AAF<0.5%, and (3) MSTN any other missense, AAF<0.5% removing the Arg88Gln variant. AAF indicates alternative allele frequency. pLoF indicates putative loss-of-function. Also included are summary statistics for associations with the single Arg88Gln variant.

Supplementary Data 7. Summary statistics for associations between MSTN missense variants and cardiac wall thickness. Positions are based on the GRCh38 genome assembly.
Supplementary Data 8. Summary statistics for gene burden test associations across cardiometabolic traits. AAF indicates alternative allele frequency. pLoF indicates putative loss-of-function.
Supplementary Data 9. Summary statistics for gene burden test associations across infertility-related traits. Within the outcome column, FSHB indiciates follicle stimulating hormone subunit beta levels and PCOS indicates polycystic ovary syndrome. AAF indicates alternative allele frequency. pLoF indicates putative loss-of-function.

Supplementary References

1. Fuchs, C.J.*, et al.* Thigh muscles are more susceptible to age-related muscle loss when compared to lower leg and pelvic muscles. *Exp Gerontol* 175, 112159 (2023).
2. Holzbaur, K.R., Murray, W.M., Gold, G.E. & Delp, S.L. Upper limb muscle volumes in adult subjects. *J Biomech* 40, 742-749 (2007)
3. Janssen, I., Heymsfield, S.B., Baumgartner, R.N. & Ross, R. Estimation of skeletal muscle mass by bioelectrical impedance analysis. *J Appl Physiol (1985)* 89, 465-471 (2000).