View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Carolina Digital Repository

OXFORD

*JNCI J Natl Cancer Inst* (2015) 107(11): djv219

doi:10.1093/jnci/djv219
First published online August 20, 2015
Article

ARTICLE

# Height and Breast Cancer Risk: Evidence From Prospective Studies and Mendelian Randomization

Ben Zhang, Xiao-Ou Shu, Ryan J. Delahanty, Chenjie Zeng, Kyriaki Michailidou, Manjeet K. Bolla, Qin Wang, Joe Dennis, Wanqing Wen, Jirong Long, Chun Li, Alison M. Dunning, Jenny Chang-Claude, Mitul Shah, Barbara J. Perkins, Kamila Czene, Hatef Darabi, Mikael Eriksson, Stig E. Bojesen, Børge G. Nordestgaard, Sune F. Nielsen, Henrik Flyger, Diether Lambrechts, Patrick Neven, Hans Wildiers, Giuseppe Floris, Marjanka K. Schmidt, Matti A. Rookus, Katja van den Hurk, Wim L. A. M. de Kort, Fergus J. Couch, Janet E. Olson, Emily Hallberg, Celine Vachon, Anja Rudolph, Petra Seibold, Dieter Flesch-Janys, Julian Peto, Isabel dos-Santos-Silva, Olivia Fletcher, Nichola Johnson, Heli Nevanlinna, Taru A. Muranen, Kristiina Aittomäki, Carl Blomqvist, Jingmei Li, Keith Humphreys, Judith Brand, Pascal Guénel, Thérèse Truong, Emilie Cordina-Duverger, Florence Menegaux, Barbara Burwinkel, Frederik Marme, Rongxi Yang, Harald Surowy, Javier Benitez, M. Pilar Zamora, Jose I. A. Perez, Angela Cox, Simon S. Cross, Malcolm W. R. Reed, Irene L. Andrulis, Julia A. Knight, Gord Glendon, Sandrine Tchatchou, Elinor J. Sawyer, Ian Tomlinson, Michael J. Kerin, Nicola Miller, Georgia Chenevix-Trench, kConFab Investigators, Australian Ovarian Study Group, Christopher A. Haiman, Brian E. Henderson, Fredrick Schumacher, Loic Le Marchand, Annika Lindblom, Sara Margolin, Maartje J. Hooning, John W. M. Martens, Madeleine M. A. Tilanus-Linthorst, J. Margriet Collée, John L. Hopper, Melissa C. Southey, Helen Tsimiklis, Carmel Apicella, Susan Slager, Amanda E. Toland, Christine B. Ambrosone, Drakoulis Yannoukakos, Graham G. Giles, Roger L. Milne, Catriona McLean, Peter A. Fasching, Lothar Haeberle, Arif B. Ekici, Matthias

**Received:** August 4, 2014; **Revised:** February 3, 2015; **Accepted:** July 15, 2015

© The Author 2015. Published by Oxford University Press. All rights reserved. For Permissions, please e-mail: journals.permissions@oup.com.

W. Beckmann, Hermann Brenner, Aida Karina Dieffenbach, Volker Arndt, Christa Stegmaier, Anthony J. Swerdlow, Alan Ashworth, Nick Orr, Michael Jones, Jonine Figueroa, Montserrat Garcia-Closas, Louise Brinton, Jolanta Lissowska, Martine Dumont, Robert Winqvist, Katri Pylkäs, Arja Jukkola-Vuorinen, Mervi Grip, Hiltrud Brauch, Thomas Brüning, Yon-Dschun Ko, Paolo Peterlongo, Siranoush Manoukian, Bernardo Bonanni, Paolo Radice, Natalia Bogdanova, Natalia Antonenkova, Thilo Dörk, Arto Mannermaa, Vesa Kataja, Veli-Matti Kosma, Jaana M. Hartikainen, Peter Devilee, Caroline Seynaeve, Christi J. Van Asperen, Anna Jakubowska, Jan Lubiński, Katarzyna Jaworska-Bieniek, Katarzyna Durda, Ute Hamann, Diana Torres, Rita K. Schmutzler, Susan L. Neuhausen, Hoda Anton-Culver, Vessela N. Kristensen, Grethe I. Grenaker Alnæs, the DRIVE Project, Brandon L. Pierce, Peter Kraft, Ulrike Peters, Sara Lindstrom, Daniela Seminara, Stephen Burgess, Habibul Ahsan, Alice S. Whittemore, Esther M. John, Marilie D. Gammon, Kathleen E. Malone, Daniel C. Tessier, Daniel Vincent, Francois Bacot, Craig Luccarini, Caroline Baynes, Shahana Ahmed, Mel Maranian, Catherine S. Healey, Anna González-Neira, Guillermo Pita, M. Rosario Alonso, Nuria Álvarez, Daniel Herrero, Paul D. P. Pharoah, Jacques Simard, Per Hall, David J. Hunter, Douglas F. Easton, Wei Zheng

**Affiliations of authors:**Division of Epidemiology, Department of Medicine, Vanderbilt Epidemiology Center, Vanderbilt-Ingram Cancer Center (BZ, XOS, RJD, CZ, WW, JL, WZ) and Department of Biostatistics (CL), Vanderbilt University School of Medicine, Nashville, TN; Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care(KM, MKB, QW, JD, PDPP, DFE) and Department of Oncology (AMD, MS, BJP, CL, CB, SA, MM, CSH, PDPP, DFE), University of Cambridge, Cambridge, UK; Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden (KC, HD, ME); Copenhagen General Population Study (SEB, BGN, SFN), Department of Clinical Biochemistry (SEB, BGN, SFN), and Department of Breast Surgery (HF), Herlev Hospital, Copenhagen, Denmark; Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark (SEB, BGN); Vesalius Research Center (VRC), VIB, Leuven, Belgium (DL); Laboratory for Translational Genetics, Department of Oncology, University of Leuven, Leuven, Belgium (DL); University Hospitals Leuven and Department of Oncology, Leuven, Belgium (PN, HW, GF); Netherlands Cancer Institute, Amsterdam, the Netherlands (MKS, MAR); Division Research, Department of Donor Studies, Sanquin Blood Supply, Amsterdam, the Netherlands (KVDH, WLAMDK); Department of Laboratory Medicine and Pathology (FJC) and Department of Health Sciences Research (JEO, EH, CV, SS), Mayo Clinic, Rochester, MN; Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany (JCC, AR, PS); Department of Cancer Epidemiology/Clinical Cancer Registry and Institute for Medical Biometrics and Epidemiology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany (DFJ); Department of Non-Communicable Disease Epidemiology, London School of Hygiene and Tropical Medicine, London, UK (JP, IDSS); Breakthrough Breast Cancer Research Centre, the Institute of Cancer Research, London, UK (OF, NJ); Department of Obstetrics and Gynecology (HN, TAM), Department of Clinical Genetics (KA), and Department of Oncology (CB), University of Helsinki and Helsinki University Central Hospital, Helsinki, Finland; Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden (PH, KH, JB); Human Genetics Division, Genome Institute of Singapore, Singapore, Singapore (JL); Inserm (National Institute of Health and Medical Research), CESP (Center for Research in Epidemiology and Population Health), Environmental Epidemiology of Cancer, Villejuif, France (PG, TT, ECD, FM); University Paris-Sud, Villejuif, France (PG, TT, ECD, FM); Department of Obstetrics and Gynecology (BB, FM, RY, HS) and National Center for Tumor Diseases (FM), University of Heidelberg, Heidelberg, Germany; Molecular Epidemiology Group, German Cancer Research Center, Heidelberg, Germany (BB, RY, HS); Human Genetics Group, Spanish National Research Center (CNIO), Human Genotyping Center, Madrid, Spain (JB); Oncology Service, Hospital La Paz, Madrid, Spain (MPZ); Surgery Service, Hospital Monte Naranco, Oviedo, Spain (JIAP); Department of Oncology (AC, MWRR) and Department of Neuroscience (SSC), University of Sheffield, Sheffield, UK; Department of Molecular Genetics (ILA), Lunenfeld-Tanenbaum Research Institute of Mount Sinai Hospital (GG, ST), University of Toronto, Toronto, Ontario, Canada; Prosserman Centre for Health Research, Lunenfeld-Tanenbaum Research Institute of Mount Sinai Hospital, Division of Epidemiology, Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada (JAK); Research Oncology, Division of Cancer Studies, Guy's Hospital, King's College London, London, UK (EJS); Wellcome Trust Centre for Human Genetics and Oxford Biomedical Research Centre, University of Oxford, UK (IT); Clinical Science Institute, University Hospital Galway, Galway, Ireland (MJK, NM); QIMR Berghofer Medical Research Institute, Brisbane, Australia (GCT, Australian Ovarian Cancer Study Group); Peter MacCallum Cancer Center, Melbourne, Australia (kConFab Investigators, Australian Ovarian Cancer Study Group); Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA (CAH, BEH, FS); Epidemiology Program, Cancer Research Center, University of Hawaii, Honolulu, HI (LLM); Department of Molecular Medicine and Surgery (AL) and Department of Oncology-Pathology (SM), Karolinska Institutet, Stockholm, Sweden; Family Cancer Clinic, Department of Medical Oncology (MJH, JWMM), Department of Surgical Oncology (MMATL), and Department of Clinical Genetics (JMC), Erasmus MC Cancer Institute, Rotterdam, the Netherlands; Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, the University of Melbourne, Melbourne, Australia (JLH, CA, GGG, RLM); Department of Pathology, the University of Melbourne, Melbourne, Australia (MCS, HT); Cancer Epidemiology Centre, Cancer Council Victoria, Melbourne, Australia (GGG, RLM); Anatomical Pathology, the Alfred Hospital, Melbourne, Australia (CM); Department of Molecular Virology, Immunology and Medical Genetics, Comprehensive Cancer Center, the Ohio State University, Columbus, OH (AET); Roswell Park Cancer Institute, Buffalo, NY (CBA); Molecular Diagnostics Laboratory, IRRP, National Centre for Scientific Research "Demokritos," Athens, Greece (DY); Department of Gynecology and Obstetrics, University Hospital Erlangen, Friedrich-Alexander University Erlangen-Nuremberg, Comprehensive Cancer Center Erlangen-EMN, Erlangen, Germany (PAF, LH, MWB); David Geffen School of Medicine, Department of Medicine Division of Hematology and Oncology, University of California at Los Angeles, CA (PAF); Institute of Human Genetics, University Hospital Erlangen, Friedrich-Alexander University Erlangen-Nuremberg, Erlangen, Germany (ABE); Division of Clinical Epidemiology and Aging Research, German Cancer Research Center (DKFZ), Heidelberg, Germany (HB, AKD, VA); Saarland Cancer Registry, Saarbrücken, Germany (CS); Division of Breast Cancer Research (AA, NO, AJS) and Division of Genetics and Epidemiology (MJ, AJS), Institute of Cancer Research, London, UK; Division of Cancer

ARTICLE

Epidemiology and Genetics, National Cancer Institute, Gaithersburg, MD, (JF, LB); Division of Genetics and Epidemiology and Breakthrough Breast Cancer Research Centre at The Institute of Cancer Research, London, UK (MGC); Department of Cancer Epidemiology and Prevention, M. Sklodowska-Curie Memorial Cancer Center & Institute of Oncology, Warsaw, Poland (JL); Centre Hospitalier Universitaire de Québec Research Center and Laval University, Quebec City, Quebec, Canada (JS, MD); Laboratory of Cancer Genetics and Tumor Biology, Department of Clinical Chemistry and Biocenter Oulu, NordLab Oulu/Oulu University Hospital (RW, KP), Department of Oncology (AJV), and Department of Surgery (MG), University of Oulu, Oulu, Finland; Dr. Margarete Fischer-Bosch-Institute of Clinical Pharmacology, Stuttgart, University Tübingen, German Cancer Consortium (DKTK) and German Cancer Research Center (DKFZ), Heidelberg, Germany (HB, The GENICA Network); Institute for Prevention and Occupational Medicine of the German Social Accident Insurance, Institute of the Ruhr University Bochum (IPA), Bochum, Germany (TB, The GENICA Network); Department of Internal Medicine, Evangelische Kliniken Bonn gGmbH, Johanniter Krankenhaus, Bonn, Germany (YDK, The GENICA Network); Molecular Genetics of Breast Cancer, Deutsches Krebsforschungszentrum (DKFZ), Heidelberg, Germany (The GENICA Network); IFOM, Fondazione Istituto FIRC di Oncologia Molecolare, Milan, Italy (PP); Unit of Medical Genetics (SM) and Unit of Molecular Bases of Genetic Risk and Genetic Testing (PR), Department of Preventive and Predictive Medicine, Fondazione IRCCS Istituto Nazionale dei Tumori (INT), Milan, Italy (SM); Division of Cancer Prevention and Genetics, Istituto Europeo di Oncologia (IEO), Milan, Italy (BB); Department of Radiation Oncology and Department of Obstetrics and Gynaecology (TD), Hannover Medical School, Hannover, Germany; N. N. Alexandrov Research Institute of Oncology and Medical Radiology, Minsk, Belarus (NA); Imaging Center, Department of Clinical Pathology (AM, VMK, JMH) and Cancer Center (AM,VK, VMK, JMH), Kuopio University Hospital, Kuopio, Finland; Institute of Clinical Medicine, Pathology and Forensic Medicine, University of Eastern Finland, Kuopio, Finland (AM, VMK, JMH); Central Finland Hospital District, Jyväskylä Central Hospital, Jyväskylä, Finland (VK); Department of Human Genetics and Department of Pathology (PD) and Department of Clinical Genetics (CJVA), Leiden University Medical Center, Leiden, the Netherlands (PD); Family Cancer Clinic, Department of Medical Oncology, Erasmus MC-Daniel den Hoed Cancer Centre, Rotterdam, the Netherlands (CS); Department of Genetics and Pathology, Pomeranian Medical University, Szczecin, Poland (AJ, JL, KJB, KD); Molecular Genetics of Breast Cancer, German Cancer Research Center, Heidelberg, Germany (UH, DT); Institute of Human Genetics, Pontificia University Javeriana, Bogota, Colombia (DT); Division of Molecular Gyneco-Oncology, Department of Gynaecology and Obstetrics and Center of Familial Breast and Ovarian Cancer and Center for Integrated Oncology, University Hospital of Cologne, Cologne, Germany (RKS); Center for Molecular Medicine Cologne, University of Cologne, Cologne, Germany (RKS); Department of Population Sciences, Beckman Research Institute of City of Hope, Duarte, CA (SLN); Department of Epidemiology, University of California Irvine, Irvine, CA (HAC); Department of Genetics, Institute for Cancer Research, Oslo University Hospital, Radiumhospitalet, Oslo, Norway (VNK, GIGA); Institute of Clinical Medicine and Department of Clinical Molecular Biology, University of Oslo, Oslo, Norway (VNK); Department of Health Studies, the University of Chicago, Chicago, IL (BLP, HA); Program in Genetic Epidemiology and Statistical Genetics, Department of Epidemiology, Harvard School of Public Health, Boston, MA (PK, SL, DJH); Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA (UP); Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, MD (DS); Department of Public Health and Primary Care, University of Cambridge, Strangeways Research Laboratory, 2 Worts Causeway, Cambridge, UK (SB); Department of Medicine and Department of Human Genetics and the University of Chicago Comprehensive Cancer Center, the University of Chicago, Chicago, IL (HA); Department of Health Research and Policy and Stanford Cancer Institute, Stanford University School of Medicine, Stanford, CA (ASW, EMJ); Cancer Prevention Institute of California, Fremont, CA (EMJ); Department of Epidemiology, University of North Carolina, Chapel Hill, NC (MDG); Division of Public Health Sciences, Program in Epidemiology, Fred Hutchinson Cancer Research Center, Seattle, WA (KEM); McGill University and Génome Québec Innovation Centre, Montréal, QC, Canada (DCT, DV, FB); Human Genotyping-CEGEN Unit, Human Cancer Genetics Program, Spanish National Cancer Research Centre, Madrid, Spain (AGN, GP, MRA, NA, DH).

**Correspondence to:** Wei Zheng, MD, PhD, Vanderbilt Epidemiology Center, Vanderbilt University School of Medicine, 2525 West End Avenue, Eighth Floor, Nashville, TN 37203-1738 (e-mail: wei.zheng@vanderbilt.edu).

## Abstract

**Background:** Epidemiological studies have linked adult height with breast cancer risk in women. However, the magnitude of the association, particularly by subtypes of breast cancer, has not been established. Furthermore, the mechanisms of the association remain unclear.

**Methods:** We performed a meta-analysis to investigate associations between height and breast cancer risk using data from 159 prospective cohorts totaling 5 216 302 women, including 113 178 events. In a consortium with individual-level data from 46 325 case patients and 42 482 control subjects, we conducted a Mendelian randomization analysis using a genetic score that comprised 168 height-associated variants as an instrument. This association was further evaluated in a second consortium using summary statistics data from 16 003 case patients and 41 335 control subjects.

**Results:** The pooled relative risk of breast cancer was 1.17 (95% confidence interval [CI] = 1.15 to 1.19) per 10 cm increase in height in the meta-analysis of prospective studies. In Mendelian randomization analysis, the odds ratio of breast cancer per 10 cm increase in genetically predicted height was 1.22 (95% CI = 1.13 to 1.32) in the first consortium and 1.21 (95% CI = 1.05 to 1.39) in the second consortium. The association was found in both premenopausal and postmenopausal women but restricted to hormone receptor–positive breast cancer. Analyses of height-associated variants identified eight new loci associated with breast cancer risk after adjusting for multiple comparisons, including three loci at 1q21.2, *DNAJC27*, and *CCDC91* at genome-wide significance level $P < 5 \times 10^{-8}$.

**Conclusions:** Our study provides strong evidence that adult height is a risk factor for breast cancer in women and certain genetic factors and biological pathways affecting adult height have an important role in the etiology of breast cancer.

ARTICLE

Breast cancer is a leading cause of cancer morbidity and mortality among women worldwide (1). Adult height has been found to be positively related to breast cancer risk in many epidemiological studies (2–30), reporting mostly a linear dose-response relationship. Results from previous studies, however, have been inconsistent, particularly with regard to the magnitude of the association and the association by subtypes of breast cancer. For example, relative risks of breast cancer associated with per 10 cm increase in adult height ranged from 1.08 to 1.38 in previous cohort studies. Furthermore, it remains unclear whether adult height is causally related to breast cancer risk through shared underlying genetic factors and biological pathways or serves only as a surrogate measure of certain environmental and lifestyle exposures that contribute to breast cancer risk. Answers to these questions may provide additional insight into breast tumorigenesis and strengthen the basis for classifying height as a breast cancer risk factor.

Mendelian randomization analysis can be used to minimize potential biases encountered in conventional observational studies and to determine the causal association of a

given exposure with disease risk (31). The causal association can also be manifested by common genetic and biological pathways that determine two sequentially developed phenotypes, such as adult height and breast cancer risk. Adult height is a classic quantitative trait determined, to a large extent, by genetic factors (32). Since 2007, genome-wide association studies (GWAS) have identified single-nucleotide polymorphisms (SNPs) in approximately 180 loci related to adult height (33–38). SNPs identified to date by GWAS explain approximately 10% of height variation in populations of European ancestry (38). The alleles associated with adult height should be randomly assigned to offspring from parents during mitosis, a process analogous to a random assignment of subjects to an exposure of interest in randomized clinical trials. Thus, a genetic score summarizing the effects of these height-associated SNPs can serve as an instrumental variable in a Mendelian randomization analysis of adult height and breast cancer risk (39).

Here we comprehensively assessed epidemiologic evidence from conventional observational studies regarding the association between height and breast cancer risk by performing a meta-analysis of 159 prospective cohorts including more than five million women of European ancestry. To determine the nature of the association, we conducted two Mendelian randomization analyses using data from two large consortia totaling 62 328 breast cancer case patients and 83 817 control subjects.

## Methods

### Meta-Analysis of Prospective Studies

We searched electronic databases to identify prospective studies that investigated the association between height and breast cancer risk among women of European ancestry published before December 2014 (Supplementary Figure 1, available online). We combined relative risks (RRs) of breast cancer with per 10 cm increase in height from each of the included studies using a random effects meta-analysis (40). We also performed subgroup meta-analyses based on method of height assessment (measured or self-reported), as well as menopausal, estrogen receptor (ER), and progesterone receptor (PR) status. We used the Cochran's Q statistic to test for heterogeneity (41) and the $I^2$ statistic to quantify heterogeneity across studies (42). Potential publication bias was assessed using Begg's and Egger's approaches (43,44). Sensitivity analyses were performed to evaluate the robustness of the results. We considered *P* values of less than .10 in tests of heterogeneity and publication bias and *P* values of less than .05 in the meta-analyses to be statistically significant. All tests were two-sided, with the exception of tests of heterogeneity and publication bias. Details of literature searches, study inclusion criteria, and meta-analysis are presented in the Supplementary Methods (available online).

### Mendelian Randomization Analysis

Our Mendelian randomization analysis was conducted using data from two consortia, the Breast Cancer Association Consortium (BCAC), and the Discovery, Biology, and Risk of Inherited Variants in Breast Cancer (DRIVE) Project (Supplementary Figure 1, available online). In BCAC, we included individual-level data for 46 325 breast cancer case patients and 42 482 control subjects of European ancestry from 39 studies (Supplementary Table 1). In DRIVE, only summary statistics data were available to our study, and these data were obtained from 16 003 breast cancer case patients and 41 335 control subjects of European ancestry from 11 studies (Supplementary Table 2, available online). Details of the methodology used by the BCAC and DRIVE have been published elsewhere (45–48) and are available on these websites (http://ccge.medschl.cam.ac.uk/research/consortia/icogs/, accessed August 8, 2015, and http://gameon.dfci.harvard.edu/, accessed August 8, 2015). Study descriptions and methods for SNP selection, genotyping, and imputation are presented in the Supplementary Methods (available online).

We examined associations between the 168 SNPs and adult height (in cm) in the BCAC using general linear models with adjustment for age and principal components. A weighted height genetic score (wHGS) was constructed for our primary analysis by using the 168 SNPs with the formula: $\mathrm{wHGS} = 9.101559(\sum_{i=1}^{168}\beta_i SNP_i)$, where 9.101559 is the coefficient to rescale the original wHGS to a mean of 336 risk alleles (Supplementary Table 3, available online), $\beta_i$ is the regression coefficient of the *i*th SNP for height, and $SNP_i$ is the dosage of the effect alleles (0,1, or 2) of the *i*th SNP. We converted all effect alleles to correspond to taller height in the SNP-based analyses and construction of the wHGS. Associations of breast cancer risk with the wHGS and each of the 168 SNPs were evaluated using unconditional logistic regression models to derive odds ratios (ORs) and 95% confidence intervals (CIs), adjusting for age and principal components. All analyses were performed for each study separately, and summary statistics were obtained using a fixed-effects meta-analysis.

In the first Mendelian randomization analysis, we estimated the potential causal association between height (X) and breast cancer risk (Y) by using the wHGS (G) as an instrumental variable. Specifically, the causal effect ($\beta_{YX}$) was calculated by using the Wald estimator: $\beta_{YX} = \frac{\beta_{YG}}{\beta_{XG}}$ (49), and the standard error for the causal effect ($SE_{YX}$) was derived using the delta method: $SE_{YX} = \sqrt{\left(\left(\frac{S_{YG}}{\beta_{XG}}\right)^2 + \frac{(S_{XG}\beta_{YG})^2}{\beta_{XG}^4} - \frac{2rS_{XG}S_{YG}\beta_{YG}}{\beta_{XG}^3}\right)}$ (50), where $\beta_{YG}$ is the natural log-scale OR of breast cancer risk associated with the wHGS, $\beta_{XG}$ is the regression coefficient of the wHGS on height, $S_{YG}$ and $S_{XG}$ are the corresponding standard errors, and *r* is the correlation between $\beta_{YG}$ and $\beta_{XG}$. Sensitivity analyses were performed (Supplementary Table 3, available online), and the strength of instrumental variables was evaluated using *F* statistic (51). The second Mendelian randomization analysis was conducted using the inverse-variance weighted method for summary statistics data to further evaluate the association (52). Specifically, the causal effect ($\beta_{YX}$) was estimated using a fixed-effects meta-analysis model: $\beta_{YX} = \frac{\sum_{i=1}^{168}\left(\frac{\beta_{XGi}\beta_{YGi}}{S_{YGi}^2}\right)}{\sum_{i=1}^{168}\left(\frac{\beta_{XGi}}{S_{YGi}}\right)^2}$, with its standard error ($SE_{YX}$) estimated using formula: $SE_{YX} = \sqrt{\frac{1}{\sum_{i=1}^{168}\left(\frac{\beta_{XGi}}{S_{YGi}}\right)^2}}$, where $\beta_{XGi}$ is the regression coefficient of the ith SNP on height obtained from approximately 110 500 women included in a GWAS of adult height published previously (38), and $\beta_{YGi}$ and $S_{YGi}$ are the natural log-scale odds ratio of breast cancer risk associated with the ith SNP and the corresponding standard error, obtained from the DRIVE Project. Details of methodology for statistical analyses are presented in the Supplementary Methods (available online).

ARTICLE

**Table 1.** Characteristics of prospective cohort studies included in the meta-analysis

| Author (reference) | Year | Study | Height* | Number of cohorts | Population† | Country | Cohort size | Number of events‡ | Years of recruitment | Follow-up§, y | Adjustment in model |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Tornberg (5) | 1988 | CSHS | Measured | 1 | European | Sweden | 46 570 | 1182 | 1963–1965 | 18–20 | Multivariable |
| Tretli (6) | 1989 | NMRS | Measured | 1 | European | Norway | 567 333 | 8427 | 1963–1975 | 6–18 | Multivariable |
| Vatten (7) | 1992 | NNHSS | Measured | 1 | European | Norway | 25 967 | 291 | 1974–1977 | 14 | Multivariable |
| De Stavola (8) | 1993 | Guernsey | Measured | 2 | European | UK | 6706 | 168 | 1968–1985 | 21/10 | Multivariable |
| Michels-Blanck (9) | 1996 | CPS-I | Self-reported | 1 | European | USA | 428 653 | 2226 | 1959–1960 | 13 | Age |
| Freni (10) | 1996 | NHANES-I | Measured | 1 | European | USA | 7622 | 182 | 1970–1975 | 12.9 | Multivariable |
| Tulinius (11) | 1997 | ICRFS | Measured | 1 | European | Iceland | 11 580 | 439 | 1967–1991 | 4–27 | Multivariable |
| Kaaks (12) | 1998 | DOM | Measured | 1 | European | The Netherlands | 11 480 | 275 | 1984–1986 | 10.6 | Multivariable |
| Sonnenschein (13) | 1999 | NYUWHS | Self-reported | 1 | European | USA | 8416 | 259 | 1985–1991 | 6.6 | Multivariable |
| van den Brandt (14) | 2000 | TPP | Self-reported | 7 | European | Multiple | 337 819 | 4385 | 1976–1987 | 3–7 | Multivariable |
| Weiderpass (15) | 2004 | NSWLHCS | Self-reported | 1 | European | Norway and Sweden | 99 717 | 733 | 1991–1992 | 6.7 | Multivariable |
| Ahlgren (16) | 2004 | CSHRR | Measured | 1 | European | Denmark | 117 415 | 3340 | N/A | 28.4 | Multivariable |
| Macinnis (17) | 2004 | MCCS | Measured | 1 | European | Australia | 13 598 | 357 | 1990–1994 | 9.1 | Multivariable |
| McCullough (18) | 2005 | CPS-II | Self-reported | 1 | European | USA | 430 236 | 4522 | 1982 | 20 | Multivariable |
| Baer (19) | 2006 | NHS-II | Self-reported | 1 | European | USA | 108 829 | 1315 | 1989 | 12 | Multivariable |
| Lacey (20) | 2009 | PLCO | Self-reported | 1 | European | USA | 70 575 | 2085 | 1993–2001 | 4.98 | Multivariable |
| Green (21) | 2011 | MWS | Self-reported | 1 | European | UK | 1 297 124 | 39 299 | 1996–2001 | 9·4 | Multivariable |
| White (22) | 2012 | MEC | Self-reported | 1 | European | USA | 19 815 | 835 | 1993–1996 | 8–11 | Multivariable |
| Canchola (23) | 2012 | CTS | Self-reported | 1 | European | USA | 52 642 | 2321 | 1995–1996 | 12.1 | Multivariable |
| Mellemkjaer (24) | 2012 | DCH | Measured | 1 | European | Denmark | 23 864 | 1209 | 1993–1997 | 11.8 | Multivariable |
| Kabat (25) | 2013 | CNBSS | Measured | 1 | European | Canada | 88 256 | 4224 | 1980–1985 | 16.2 | Multivariable |
| Wormser (26) | 2012 | ERFC | Measured | 121 | European | Multiple | 522 257 | 3926 | N/A | 13.7 | Multivariable |
| Ritte (27) | 2013 | EPIC | Measured | 1 | European | Multiple | 306 600 | 9307 | 1992–2000 | 10.8 | Multivariable |
| Kabat (28) | 2013 | WHI | Measured | 1 | European | USA | 144 701 | 6798 | 1993–1998 | 12.0 | Multivariable |
| Wiren (29) | 2014 | Me-Can | Measured | 7 | European | Multiple | 297 156 | 6161 | 1972–2005 | 12.7 | Multivariable |
| Kabat (30) | 2014 | NIH-AARP | Self-reported | 1 | European | USA | 192 514 | 9169 | 1995–1996 | 10.5 | Multivariable |

* Height was either measured or self-reported in the included study. CNBSS = the Canadian National Breast Screening Study; CPS-I = the American Cancer Society Cancer Prevention Study I; CPS-II = the American Cancer Society's Cancer Prevention Study II; CSHRR = the Copenhagen School Health Records Register; CSHS = Central Sweden Health Screening; CTS = the California Teachers Study Cohort; DCH = the Danish Diet, Cancer and Health Cohort; DOM = the Dutch Diagnostisch Onderzoek Mammacarcinoom Cohort; EPIC = the European Prospective Investigation into Cancer and Nutrition Cohort; ERFC = the Emerging Risk Factors Collaboration; ICRFS = the Icelandic Cardiovascular Risk Factor Study; MCCS = the Melbourne Collaborative Cohort Study; MEC = the Multiethnic Cohort; Me-Can = the Metabolic syndrome and Cancer project; MWS = the Million Women Study; N/A = no data were available; NHANES-I = the first National Health and Nutrition Examination Survey; NHS-II = the Nurses' Health Study II; NIH-AARP = the National Institutes of Health-American Association of Retired Persons Diet and Health Study; NMRS = National Mass Radiography Service; NNHSS = Norway National Health Screening Service; NSWLHCS = the Norwegian and Swedish Women's Lifestyle and Health Cohort Study; NYUWHS = the New York University Women's Health Study; PLCO = the Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial; TPP = the Pooling Project; WHI = the Women's Health Initiative.
† Population as reported for >80% of the included study.
‡ Events were fatal breast cancer cases in CPS-I, CPS-II, and ERFC.
§ Mean or median or range of duration of follow-up.

Analyses were performed using SAS (version 9.3), R (version 3.0.0), and PLINK (version 1.07). All tests were two-sided, and *P* values of less than .05 were considered statistically significant unless stated otherwise.

## Results

### Meta-Analysis of Association Between Height and Breast Cancer Risk

We identified 26 articles (5–30), containing information from 159 prospective cohorts that were eligible for inclusion in our meta-analyses (Table 1). Of these, 22 articles reported results from individual cohorts, while the remaining four articles provided results from combined analyses of two to 121 cohorts. After excluding overlapping cohorts, 5 216 302 participants of European ancestry were included in our analyses, including 113 178 women with breast cancer. Figure 1 presents relative risks (RRs) of breast cancer associated with per 10 cm increase in height for each of the published studies, all studies combined, and study subgroups. The pooled relative risk of breast cancer was 1.17 (95% confidence interval [CI] =1.15 to 1.19, $P < .001$) per 10 cm increase in height for all studies combined, with strong evidence of heterogeneity across studies ($P_{heterogeneity} < .001$,

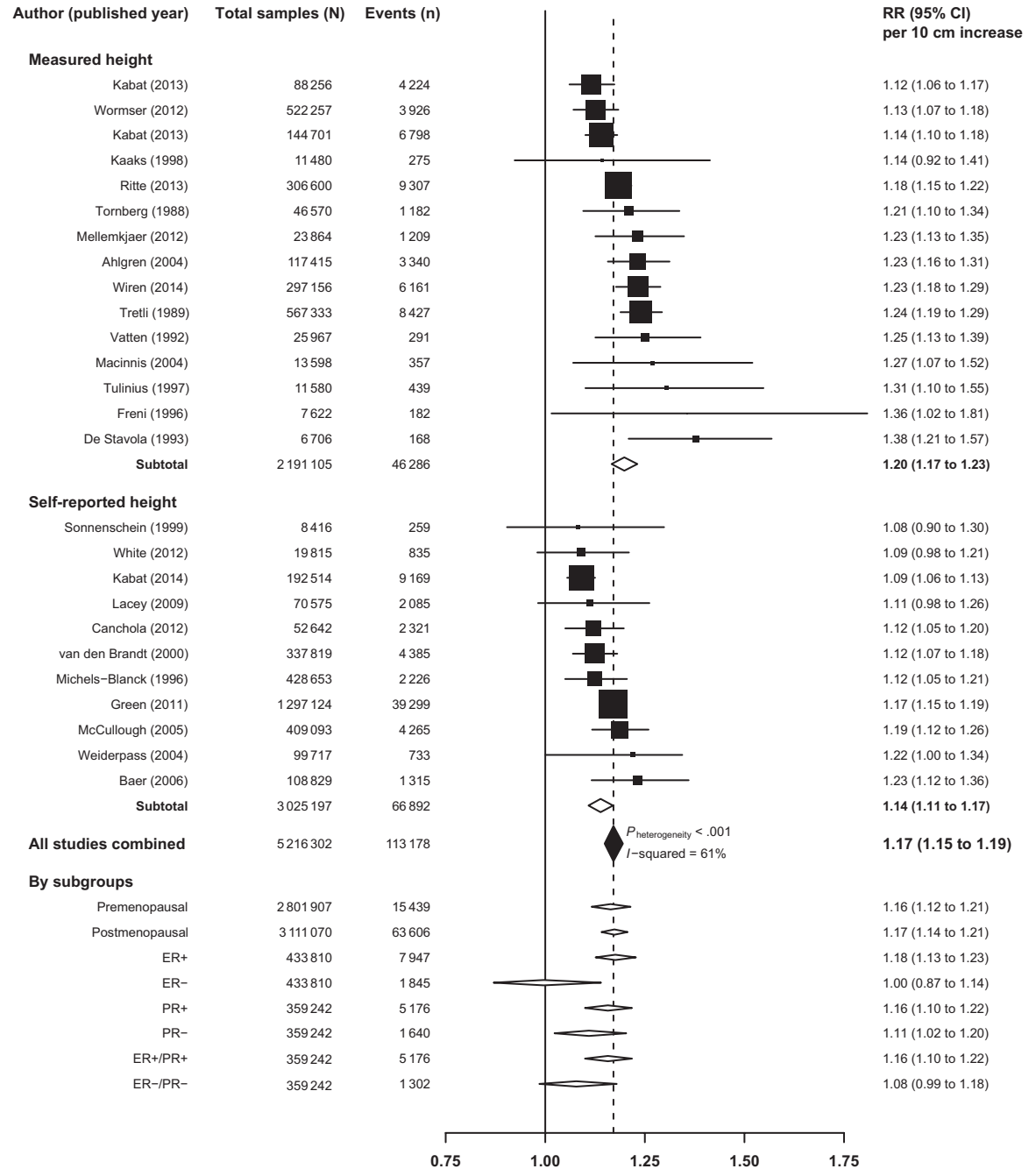


**Figure 1.** Meta-analysis of associations between height and risk of breast cancer in prospective cohort studies. All tests for meta-analyses were two-sided. CI = confidence interval; ER = estrogen receptor; PR = progesterone receptor; RR = relative risk.

ARTICLE

$I^2$ = 61%). Removal of the smallest, the largest, the first significant, or the study with adjustment only for age did not change the pooled risk estimate. There was little evidence of publication bias ($P$ > .33).

The association between height and breast cancer risk was stronger in the meta-analysis of studies with measured height than in the studies with self-reported height (RR = 1.20, 95% CI = 1.17 to 1.23 vs RR = 1.14, 95% CI = 1.11 to 1.17, $P_{heterogeneity}$ = .001). The association was similar among postmenopausal and premenopausal women (RR = 1.17, 95% CI = 1.14 to 1.21 vs RR = 1.16, 95% CI = 1.12 to 1.21, $P_{interaction}$ = .79). Statistically significant associations of height with breast cancer risk were found for ER-positive case patients (RR = 1.18, 95% CI = 1.13 to 1.23), PR-positive case patients (RR = 1.16, 95% CI = 1.10 to 1.22), and ER/PR-positive case patients (RR = 1.16, 95% CI = 1.10 to 1.22). The association was not statistically significant for ER-negative and ER/PR-negative breast cancer and was only nominally significant for PR-negative breast cancer ($P$ = .01).

## Mendelian Randomization Analyses of Association Between Height and Breast Cancer Risk

Of the 168 height-associated variants, 146 showed an association with height at $P$ values of less than .05 in the BCAC in the same direction and with comparable effect sizes as reported in previous GWAS (Supplementary Table 4, available online). The wHGS, constructed using all the 168 SNPs, ranged from 257 to 407 (mean = 336, standard deviation [SD] = 16.17) in BCAC participants, explaining approximately 8.99% of the height variation. A clear relation between the wHGS and height was found in the study (Table 2). The wHGS was associated, at a $P$ value of .02 or lower, with age at menarche, parity, age at first live birth, use of menopausal hormone therapy, weight, and body mass index (BMI), and these associations were no longer statistically significant after adjusting for height. No association was observed between wHGS and other risk factors for breast cancer. The mean wHGS was higher in case patients than in control subjects (336.02 vs 335.98, $P$ < .001). The wHGS was positively associated with breast cancer risk ($P = 6.97\times10^{-7}$) with an odds ratio of 1.22 (95% CI = 1.13 to 1.32) by an increment of the wHGS corresponding to a 10 cm increase in height (Figure 2). There was little evidence of heterogeneity across studies ($P_{heterogeneity}$ = .27, $I^2$ = 11%). This positive association between breast cancer risk and wHGS remained essentially unchanged after adjustment for breast cancer risk factors, including age at menarche, age at menopause, parity, family history of breast cancer, age at the first live birth, breast feeding, and use of oral contraceptive or postmenopausal hormone (data not shown). A 13% elevated risk of breast cancer was found to be associated with per 10 cm increment in measured or self-reported height in the BCAC (Table 3). Adjustment for measured/self-reported height eliminated the association between the wHGS and breast cancer risk ($P$ = .20). On the other hand, the significant association between measured/self-reported height remained unchanged (OR = 1.13, 95% CI = 1.10 to 1.17, $P$ < .0001) after adjusting for wHGS.

Table 3 presents associations of breast cancer risk with adult height as predicted using the wHGS as the instrument in Mendelian randomization analysis. As a comparison, results derived from the meta-analysis of prospective studies and the meta-analysis of studies included in BCAC are also presented. A 10 cm increase in height as predicted by the wHGS was associated with an approximately 22% elevated risk of breast cancer for all women combined (OR = 1.22, 95% CI = 1.13 to 1.32, $P = 7.52\times10^{-7}$), compared with the meta-analysis of prospective cohort studies and the meta-analysis of BCAC case-control studies, which showed a 17% and a 13% elevated risk per 10 cm increment in height, respectively (Table 3). Odds ratios associated with genetically predicted height did not vary by menopausal status ($P_{interaction}$ = .86). The association, however, was restricted primarily to hormone receptor–positive breast cancer.

**Table 2.** Associations of the weighted height genetic score with height and traditional breast cancer risk factors

| Variable | Number of participants | Summary effect* | Standard error | $P$ |
|---|---|---|---|---|
| Height, cm† | | | | |
| All participants | 50706 | 0.11 | 0.002 | $<1\times10^{-500}$ |
| Control subjects | 20458 | 0.11 | 0.003 | $<1\times10^{-200}$ |
| Case patients | 30248 | 0.11 | 0.002 | $<1\times10^{-200}$ |
| Traditional risk factors | | | | |
| Age, y | 80455 | -0.0033 | 0.002 | .17 |
| Age at menarche, y | 53990 | 0.0015 | 0.0004 | $2.67\times10^{-4}$ |
| Menopausal status, post vs pre | 61686 | 0.00043 | 0.0008 | .61 |
| Age at menopause, y | 26921 | -0.0013 | 0.002 | .54 |
| Family history of breast cancer, yes vs no | 47417 | 0.00097 | 0.0007 | .19 |
| Parous, yes vs no | 62683 | -0.0012 | 0.0007 | .09 |
| Parity, numbers | 61837 | -0.00071 | 0.0003 | .02 |
| Age at first live birth, y | 44736 | 0.0047 | 0.001 | $8.50\times10^{-4}$ |
| Use of oral contraceptives, ever vs never | 28941 | 0.00073 | 0.0009 | .42 |
| Use of menopausal hormone therapy, ever vs never | 30983 | 0.0018 | 0.0008 | .02 |
| Breastfeeding, ever vs never | 43321 | -0.00076 | 0.0007 | .26 |
| Smoking, ever vs never | 39562 | -0.000068 | 0.0006 | .92 |
| Weight, kg | 51634 | 0.070 | 0.004 | $2.01\times10^{-82}$ |
| BMI, kg/m$^2$ | 47221 | -0.0074 | 0.001 | $1.21\times10^{-7}$ |

* Regression coefficient is presented for continuous variables and natural log-scale odds ratio for dichotomous variables, per unit increase of the weighted height genetic score. BMI = body mass index.

† There was no heterogeneity in the association of the weighted height genetic score with height among case patients and control subjects ($P$ = .72). All tests were two-sided.

ARTICLE

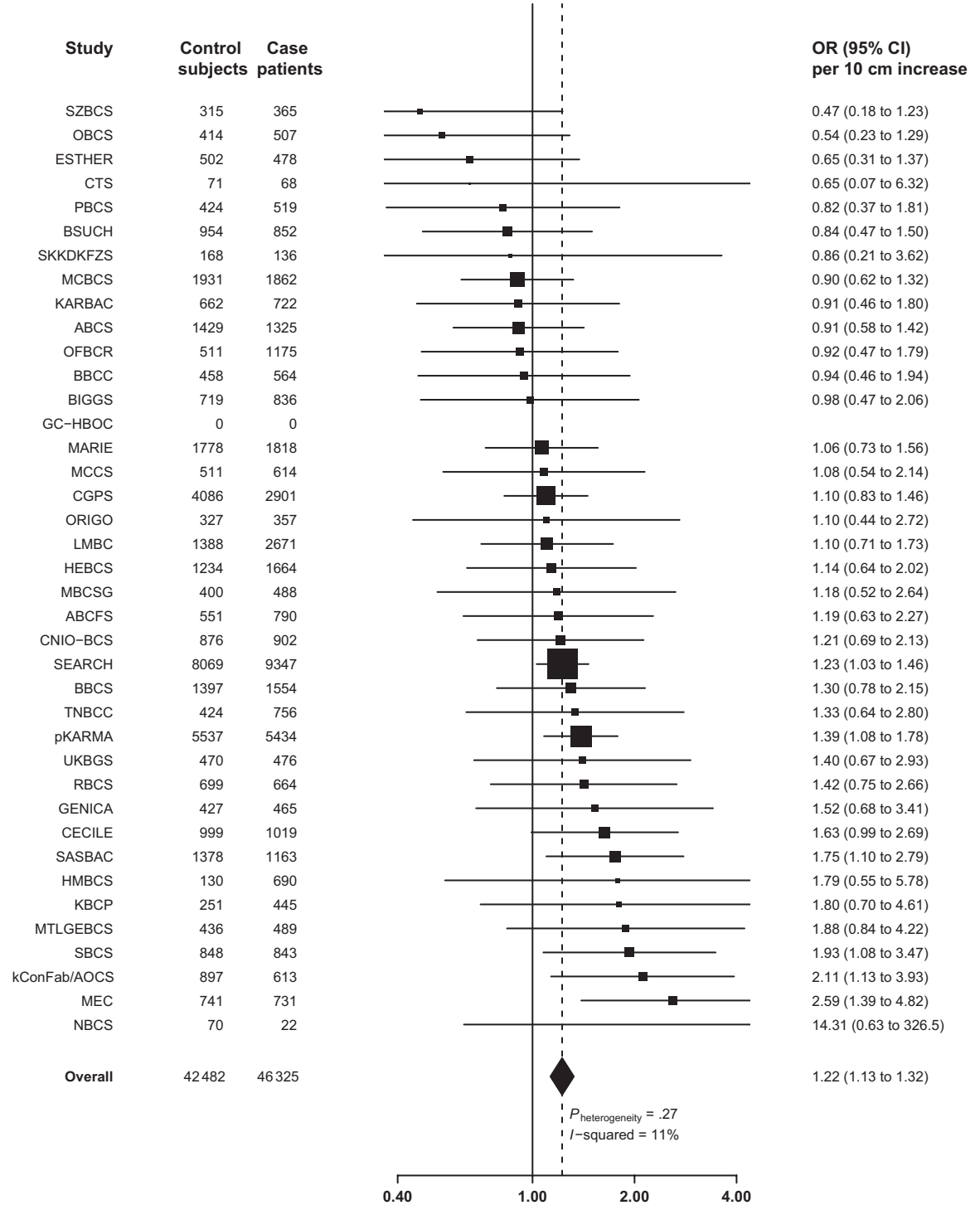


**Figure 2.** Association of the weighted height genetic score with breast cancer risk in the Breast Cancer Association Consortium. CI = confidence interval; OR = odds ratio.

For example, the odds ratios were 1.26 (95% CI = 1.14 to 1.38) and 1.02 (95% CI = 0.87 to 1.18), respectively, for ER-positive and ER-negative breast cancer ($P_{interaction}$ = .02), similar to the results obtained from meta-analyses of previous cohort studies and BCAC case-control studies. Sensitivity analyses were performed to evaluate the robustness of the associations between breast cancer risk and various height-associated genetic scores (all $F$ statistics > 3500) (Figure 3). All analyses yielded similar results, with little evidence of heterogeneity ($P$ > .66 for all tests). In the second Mendelian randomization analysis using summary statistics data from DRIVE for breast cancer risk and published GWAS for height, we found that a 10 cm increase in height as predicted by the 168 height-associated SNPs was associated with an approximately 21% elevated risk of breast cancer (OR = 1.21, 95% CI = 1.05 to 1.39, $P$ = .008), highly consistent with that observed in the analysis using individual-level data from BCAC. Sensitivity analyses similar to those performed in the analysis of data from BCAC described above yielded similar results (data not shown).

ARTICLE

**Table 3.** Association of height and breast cancer risk in women

| | Meta-analysis of prospective studies | | | Breast Cancer Association Consortium | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Observational estimate | | | Instrumental variable estimate | | | Observational estimate | | |
| Breast cancer group | N/events | RR (95% CI)* | *P* | Case patients/ control subjects | OR (95% CI)* | *P* | Case patients/ control subjects | OR (95% CI)* | *P* |
| All women combined | | | | | | | | | |
| All case patients | 5 216 302/113 178 | 1.17 (1.15 to 1.19) | <.001 | 46 325/42 482 | 1.22 (1.13 to 1.32) | <.001 | 30 248/20 458 | 1.13 (1.10 to 1.16) | <.001 |
| By menopausal status | | | | | | | | | |
| Premenopausal | 2 801 907/15 439 | 1.16 (1.12 to 1.21) | <.001 | 10 209/9053 | 1.29 (1.07 to 1.56) | .007 | 8959/6225 | 1.11 (1.05 to 1.17) | <.001 |
| Postmenopausal | 3 111 070/63 606 | 1.17 (1.14 to 1.21) | <.001 | 23 069/19 355 | 1.32 (1.17 to 1.49) | <.001 | 20 197/13 311 | 1.14 (1.10 to 1.18) | <.001 |
| $P_{interaction}$ | | | .79 | | | .86 | | | .35 |
| By ER status | | | | | | | | | |
| ER-positive | 433 810/7947 | 1.18 (1.13 to 1.23) | <.001 | 27 074/42 482 | 1.26 (1.14 to 1.38) | <.001 | 19 953/20 458 | 1.16 (1.12 to 1.20) | <.001 |
| ER-negative | 433 810/1845 | 1.00 (0.87 to 1.13) | .95 | 7288/42 482 | 1.02 (0.87 to 1.18) | .84 | 4810/20 458 | 1.05 (1.00 to 1.10) | .07 |
| $P_{interaction}$ | | | .02 | | | .02 | | | .002 |
| By PR status | | | | | | | | | |
| PR-positive | 359 242/5176 | 1.16 (1.10 to 1.22) | <.001 | 19 749/42 482 | 1.23 (1.11 to 1.37) | <.001 | 14 796/20 458 | 1.14 (1.10 to 1.18) | <.001 |
| PR-negative | 359 242/1640 | 1.11 (1.02 to 1.20) | .01 | 9726/42 482 | 1.07 (0.94 to 1.23) | .31 | 6787/20 458 | 1.09 (1.04 to 1.14) | <.001 |
| $P_{interaction}$ | | | .44 | | | .12 | | | .09 |
| By ER/PR status | | | | | | | | | |
| ER/PR-positive | 359 242/5176 | 1.16 (1.10 to 1.22) | <.001 | 18 948/42 482 | 1.25 (1.12 to 1.39) | <.001 | 14 162/20 458 | 1.15 (1.11 to 1.19) | <.001 |
| ER/PR-negative | 359 242/1302 | 1.08 (0.99 to 1.18) | .10 | 5848/42 482 | 1.03 (0.87 to 1.23) | .70 | 3779/20 458 | 1.07 (1.01 to 1.13) | .03 |
| $P_{interaction}$ | | | .22 | | | .07 | | | .03 |

* Results are presented for per 10 cm increase of height. All tests were two-sided. CI = confidence interval; ER = estrogen receptor; OR = odds ratio; PR = progesterone receptor; RR = relative risk.

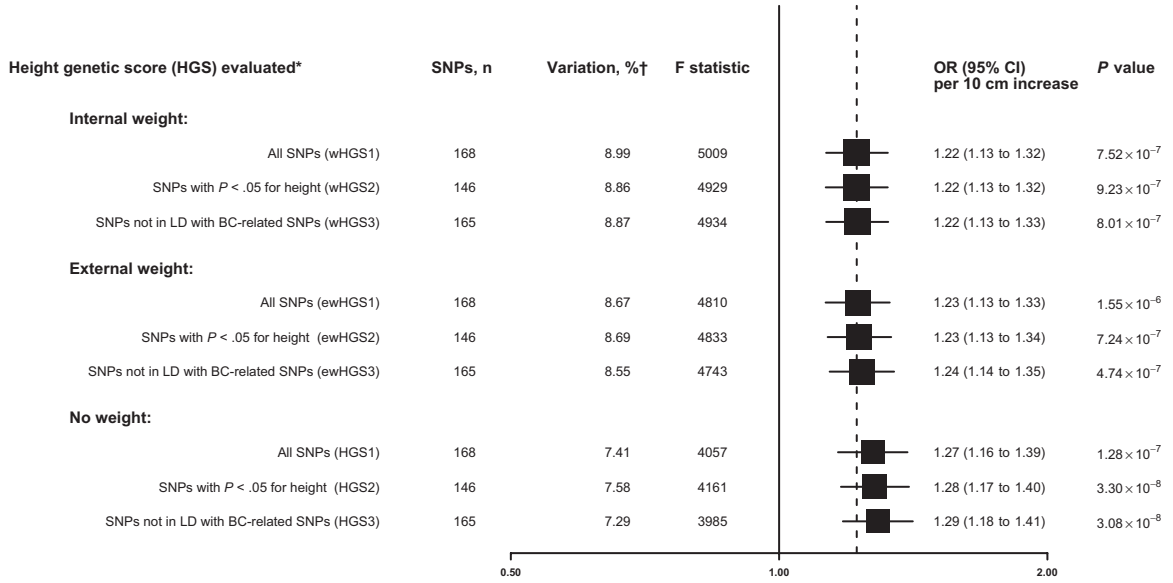


**Figure 3.** Sensitivity analyses for associations between genetically predicted height and breast cancer risk in the Breast Cancer Association Consortium. *The details of the formula to construct height genetic score are presented in Supplementary Table 3 (available online). †Phenotypic variation of height explained by height genetic scores in the study population. BC = breast cancer; HGS = height genetic score; LD = linkage disequilibrium; SNP = single-nucleotide polymorphism.

### Height-Associated Variants and Breast Cancer Risk

Statistically significant associations with breast cancer at *P* values of less than .05 in the same direction as observed for height were found for 16 SNPs in BCAC and 18 SNPs in DRIVE, both higher than expected by chance ($P = .01$ for BCAC and $P = .002$ for DRIVE) (Supplementary Tables 4 and 5, available online). In the combined analysis of data from both consortia, 25 SNPs were associated with breast cancer risk at *P* value sof less than .05 in the same direction as observed for height (Supplementary Table 6, available online). In particular, the association for eight SNPs remained statistically significant after adjusting for multiple comparisons of 168 independent SNPs, significantly higher than expected by chance ($P < 10^{-15}$) (Table 4). The association for three SNPs, rs11205277 near the *SF3B4* gene at 1q21.2, rs4665736 in the *DNAJC27* gene at 2p23.3, and rs2638953 in the *CCDC91* gene at 12p11.22, reached the genome-wide significance level of *P* values of less than $5.0 \times 10^{-8}$ (Table 4 and Figure 4). These three loci have not been previously reported in GWAS to be associated with breast cancer risk.

## Discussion

The association between adult height and breast cancer risk in women has been investigated in many epidemiological studies (2–4). However, the magnitude of this association, particularly for subtypes of breast cancer, has not been established. In our meta-analysis of data from more than five million women, including approximately 110 000 breast cancer events, we estimated that a 10 cm increase in height was associated with a 17% elevated risk of breast cancer. The association was stronger in the meta-analysis of studies with measured height than in meta-analysis of studies with self-reported height. This association was confirmed in our Mendelian randomization analysis including 62 328 breast cancer case patients and 83 817 control subjects from two large consortia, in which a 21% to 22% elevated risk of breast cancer was associated with per 10 cm increase in genetically predicted height. The weaker association observed in the meta-analysis of previous cohort studies was expected because some of these conventional observational studies may have suffered from possible biases, including confounding biases and measurement errors. In both meta-analysis of prospective studies and Mendelian randomization analysis, the association between height and breast cancer risk was observed in both premenopausal and postmenopausal women but was limited primarily to hormone receptor–positive breast cancer. Using the Mendelian randomization approach, our study provides strong evidence for a possible causal association between adult height and breast cancer risk. Results from this study have clarified the nature of the height and breast cancer association and provided additional insight into the genetic and biological basis of breast cancer development.

The genetic score used in our Mendelian randomization analysis explains approximately 10% of the height variation in populations of European ancestry. The remaining 90% of height variation would be explained by both environmental factors and genetic variants not yet identified. Given the small height variation explained by the genetic score used in our study, we expected that the association between measured/reported height and breast cancer risk should be similar with or without adjusting for the height-associated genetic score. Indeed, this is what we observed in the study. It has been reported that in addition to genetic factors, adult height is influenced by energy intake and socioeconomic status during growth spurts (53). It has been suggested that certain nutritional factors during childhood and adolescence may be related to breast cancer risk (2–4). Height is also influenced by the timing of puberty, which is affected by endogenous estrogen, a hormone that plays a central role in breast cancer etiology (2–4). However, very few studies have collected sufficiently detailed data on childhood and adolescent nutrition and health status and pubertal development to clearly disentangle the association of breast cancer risk with adult height from these exposures that could also contribute to breast cancer risk (54). Therefore, it has been unclear whether height is just a simple surrogate measure of early life exposures of breast cancer risk factors, in

**Table 4.** Association of breast cancer risk with eight height-associated SNPs that remained statistically significant after adjusting for 168 multiple comparisons ($P < .0003$)

| SNP | Locus | Position* | Gene† | Alleles‡ (1/2) | BCAC§ | | | DRIVE§ | | | Combined (62 328/83 817) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | RAF | OR (95% CI) | $P_{trend}$ | RAF | OR (95% CI) | $P_{trend}$ | OR (95% CI) | $P_{meta}$ |
| rs11205277 | 1q21.2 | 148159496 | *SF3B4* | G/A | 0.43 | 1.06 (1.04 to 1.08) | $4.58\times10^{-10}$ | 0.45 | 1.03 (1.00 to 1.08) | .08 | 1.06 (1.04 to 1.07) | $2.02\times10^{-10}$ |
| rs4665736 | 2p23.3 | 25041103 | *DNAJC27* | T/C | 0.53 | 1.06 (1.04 to 1.09) | $3.11\times10^{-7}$ | 0.51 | 1.04 (1.01 to 1.08) | .02 | 1.06 (1.04 to 1.08) | $2.65\times10^{-8}$ |
| rs724016 | 3q23 | 142588260 | *ZBTB38* | G/A | 0.45 | 1.05 (1.03 to 1.07) | $2.88\times10^{-6}$ | 0.45 | 1.05 (1.01 to 1.08) | .009 | 1.05 (1.03 to 1.06) | $8.64\times10^{-8}$ |
| rs17081935 | 4q12 | 57518233 | *NOA1* | T/C | 0.19 | 1.05 (1.02 to 1.07) | $2.52\times10^{-4}$ | 0.19 | 1.02 (0.98 to 1.07) | .36 | 1.04 (1.02 to 1.06) | $2.73\times10^{-4}$ |
| rs3782089 | 11q13.1 | 65093395 | *SSSCA1* | C/T | 0.93 | 1.08 (1.04 to 1.12) | $5.76\times10^{-5}$ | 0.93 | 1.03 (0.96 to 1.11) | .39 | 1.07 (1.03 to 1.11) | $7.32\times10^{-5}$ |
| rs2638953 | 12p11.22 | 28425682 | *CCDC91* | C/G | 0.68 | 1.05 (1.03 to 1.08) | $2.90\times10^{-7}$ | 0.68 | 1.08 (1.04 to 1.12) | $1.74\times10^{-5}$ | 1.06 (1.04 to 1.08) | $4.95\times10^{-11}$ |
| rs3764419 | 17q11.2 | 26188149 | *ATAD5* | C/A | 0.62 | 1.04 (1.02 to 1.06) | $1.14\times10^{-4}$ | 0.61 | 1.05 (1.01 to 1.08) | .01 | 1.04 (1.02 to 1.06) | $3.93\times10^{-6}$ |
| rs4605213 | 17q21.33 | 46599746 | *NME1-NME2* | C/G | 0.35 | 1.03 (1.01 to 1.05) | .003 | 0.35 | 1.04 (1.01 to 1.08) | .03 | 1.03 (1.02 to 1.05) | $2.31\times10^{-4}$ |

* The chromosome physical position is based on the National Center for Biotechnology Information database, Build 36.3. BCAC = Breast Cancer Association Consortium; CI = confidence interval; DRIVE = Discovery, Biology, and Risk of Inherited Variants in Breast Cancer; OR = odds ratio; RAF = risk allele frequency; SNP = single-nucleotide polymorphism.

† The closest gene(s).

‡ Alleles (1/2); Allele1 (risk allele) associated with increased height in initial study; Allele2, reference allele.

§ Results (RAF, OR, 95% CI, and *P* value) were derived from BCAC and DRIVE. Three loci (1q21.2, 2p23.3 and 12P11.22) were associated with breast cancer risk at genome-wide significant level $P < 5 \times 10^{-8}$. All tests were two-sided.

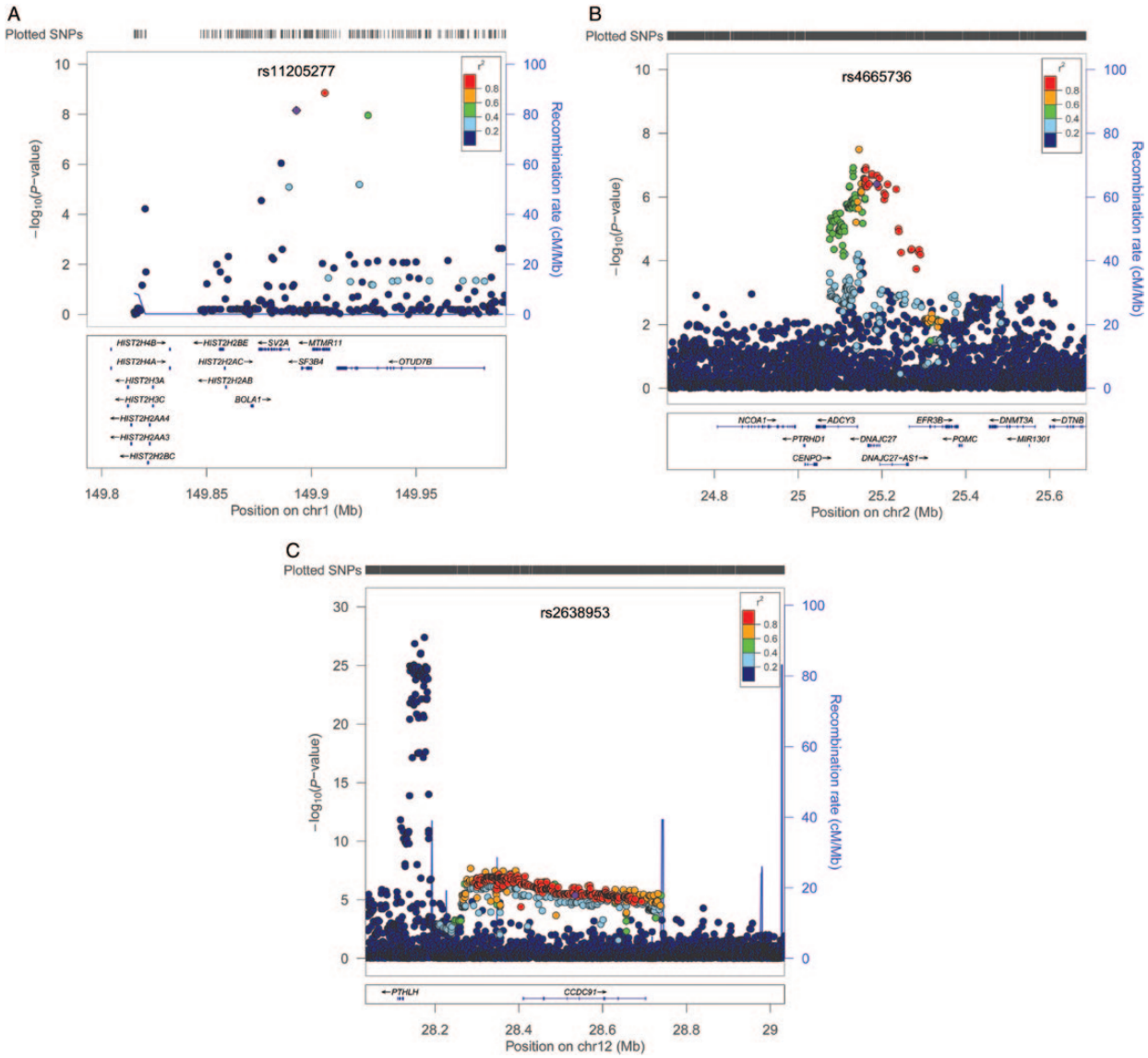


**Figure 4.** Regional association plots of the three new loci associated with breast cancer risk in the Breast Cancer Association Consortium. The **three plots** represent: **(A)** 1q21.2, **(B)** 2p23.3, and **(C)** 12p11.22. For each plot, the -$\log_{10}$ (P values) (y-axis) of single-nucleotide polymorphisms (SNPs) are shown according to their chromosomal positions (x-axis) in National Center for Biological Information (NCBI) Build 37. **Blue lines** represent the estimated recombination rates from the HapMap Project (NCBI Build 37). **Arrows** indicate genomic locations of genes within the LD block centered on the index SNPs in the NCBI Build 37 human assembly. The color of SNPs represents their LD ($r^2$, the 1000 Genomes Project Europeans), with the index SNP shown as a **purple diamond** at each locus.

which height per se is not causally related to breast cancer risk. It is also possible that the height–breast cancer association is causal, in which genetic and/or environmental factors determine height and subsequently contribute to breast cancer risk through the shared underlying biology. Using the Mendelian randomization approach, our study provides strong evidence for a possible causal association between adult height and breast cancer risk, suggesting that factors, both genetic variants and environmental exposures, that determine adult height, collectively, may be causally related to breast cancer risk. The primary goal of Mendelian randomization analyses is to minimize possible biases commonly encountered in conventional observational studies in order to provide strong evidence for causal influence. Like any other Mendelian randomization study, we cannot estimate the relative contribution of genetic variants and environmental exposures to the association between adult height and breast cancer risk in our study, particularly because many additional genetic variants related to height have not yet been identified. Given the incomplete understanding of the genetic component for complex traits, such as adult height, body weight, and blood lipids, no Mendelian randomization study conducted to date has attempted to determine relative contribution of genetics and environment in the association between these traits and disease risk (69–74).

Adult height is the result of various growth and development processes that are determined by many biological pathways. Among them, the insulin-like growth factor (IGF) signaling pathway is of particular interest. Multiple genetic variants in the IGF signal pathway have been identified by GWAS to be related to height (55). In fact, it is known that IGFs, particularly IGF1, are major regulators of growth in utero and during childhood and adolescence (56). IGF1 also plays an important role in carcinogenesis through promotion of epithelial cell proliferation and inhibition of apoptosis (57). Circulating IGF1 levels were found

ARTICLE

to be higher among taller compared with shorter women and were positively associated with breast cancer risk in a recent pooled analysis of 17 prospective studies (58). In contrast, low levels of IGF1 due to mutations in the *GHR* gene were associated with severe short stature and absence of breast cancer (or overall cancer) in individuals with Laron dwarfism in a prospective study (59). In addition to the IGF signal pathway, multiple other biological pathways have also been identified by GWAS to be associated with adult height, including Hedgehog, MARK, TGF-β, WNT, BMP, and mTOR (38,55,60). Some of these may also be involved in the pathogenesis of breast cancer (61–66). To our knowledge, to date no genetic variants in these pathways have been conclusively associated with breast cancer risk. Our study suggests that height and breast cancer susceptibility share some common genes and biological pathways, and thus focused search in future studies for variants in genes and biological pathways established for height may help to identify additional genetic risk variants for breast cancer.

By analyzing height-associated SNPs, we identified eight variants associated with breast cancer risk after adjusting for multiple comparisons. In particular, the association with three loci previously not reported in relation to breast cancer risk reached the genome-wide significance level of *P* values of less than $5.0\times10^{-8}$. At locus 1q21.2, the risk-associated SNP rs11205277 is located in an intergenic region between *SV2A* and *SF3B4*. Two other genes, *MTMR11* and *OTUD7B,* are also included in the linkage disequilibrium (LD) block tagged by the SNP. At the 2p23.3 locus, the risk-associated variant rs4665736 lies in intron 2 of the *DNAJC27* gene. This SNP is related to the expression of the *ADCY3* and *DNAJC27* genes in peripheral blood samples (67). At the 12p11.22 locus, rs2638953 maps to intron 6 of the *CCDC91* gene. Data from the ENCODE Project suggest that rs2638953 and other highly correlated SNPs ($r^2 > 0.8$) in the LD block might have regulatory functions (68). Additional studies are warranted to fine-map and functionally characterize the regions identified in our study.

Our analysis based on height-associated genetic score is consistent with a Mendelian randomization. The instrumental variable (wHGS) was strongly associated with adult height, the exposure of interest. The large *F*-statistic value (>3500) indicated that wHGS is a very strong instrumental variable. Although wHGS was related to some known breast cancer risk factors, all of the observed associations were much weaker than adult height, and all of the association can be explained by height. Indeed, we have shown that the associations between the wHGS and these breast cancer risk factors all disappeared after adjusting for measured/reported height. One possible limitation for the Mendelian randomization analysis is that of the 168 height-associated SNPs included in our study, 145 were imputed in BCAC, which could lead to an overall less precise estimate for the genetic association with height, underestimating the association between genetically predicted height and breast cancer risk. However, this bias should not be substantial because the imputation $R^2$ was greater than 0.50 for all of the SNPs included in the analysis with a mean value of 0.88. Furthermore, results from BCAC were replicated in DRIVE, and the results from these two large consortia were very close. Our meta-analysis of prospective cohort studies may be subject to potential biases inherent in the original studies. To minimize these biases, we included only prospective studies with age or multivariable-adjusted relative risks and excluded non-European studies from our analysis to minimize heterogeneity. We also conducted subgroup meta-analyses and found that the association of breast cancer risk was stronger with measured height than self-reported height. Most studies participating in BCAC are case-control studies, with height information obtained after cancer diagnosis. This may have contributed to the lower risk estimates for the association between adult height and breast cancer risk in BCAC than those obtained from the meta-analysis of prospective studies and the Mendelian randomization analysis. We present results from the analyses of prospective cohort studies, Mendelian randomization, and case-control studies in parallel to illustrate a potential biased estimate of the association between height and breast cancer risk from conventional case-control studies.

To our knowledge, this is the largest Mendelian randomization analysis conducted to date for any cancer. With 62 328 breast cancer case patients and 83 817 control subjects, our study has excellent power to quantify the association with overall breast cancer and by breast cancer subtypes. Our study, with data from a large meta-analysis of prospective cohort studies and Mendelian randomization analysis, provides strong evidence that adult height is a risk factor for breast cancer in women and that the association between adult height and breast cancer risk is likely to be causal. Furthermore, our study revealed that there are shared underlying genetic pathways affecting both height and the pathogenesis of breast cancer.

## Funding

The work for this project at Vanderbilt University was supported primarily by US National Institutes of Health (NIH) grant R37CA070867 and funds from Ingram Professorship and Anne Potter Wilson endowments. This work was partly supported by the European Community's Seventh Framework Programme under grant agreement number 223175 (grant number HEALTH-F2-2009–223175) (Collaborative Oncological Gene-Environment Study [COGS]). Funding for the iCOGS infrastructure came from: the European Community's Seventh Framework Programme under grant agreement No. 223175 (HEALTH-F2-2009–223175) (COGS), Cancer Research UK (CR-UK; C1287/A10118, C1287/A 10710, C12292/A11174, C1281/A12014, C5047/A8384, C5047/A15007, C5047/A10692), the National Institutes of Health (CA128978), and Post-Cancer GWAS initiative (1U19 CA148537, 1U19 CA148065 and 1U19 CA148112 - the GAME-ON initiative), the Department of Defence (W81XWH-10-1-0341), the Canadian Institutes of Health Research (CIHR) for the CIHR Team in Familial Risks of Breast Cancer, Komen Foundation for the Cure, the Breast Cancer Research Foundation, and the Ovarian Cancer Research Fund.

The Genetic Associations and Mechanisms in Oncology Discovery, Biology, and Risk of Inherited Variants in Breast Cancer (GAME-ON DRIVE) Consortium was supported by National Cancer Institute (NCI) Grant number U19 CA148065. The work of the Breast Cancer Family Registry (BCFR) included in DRIVE was supported by grant UM1 CA164920 from the National Cancer Institute. The content of this manuscript does not necessarily reflect the views or policies of the National Cancer Institute or any of the collaborating centers in the BCFR, nor does mention of trade names, commercial products, or organizations imply endorsement by the US government or the BCFR.

The Australian Breast Cancer Family Study (ABCFS), NC-BCFR, and Ontario Familial Breast Cancer Registry (OFBCR) work was supported by grant UM1 CA164920 from the National Cancer Institute. The ABCFS was also supported by the National Health and Medical Research Council of Australia, the New South Wales Cancer Council, the Victorian Health Promotion Foundation (Australia), and the Victorian Breast Cancer Research Consortium. John. L. Hopper is a National Health and Medical Research Council (NHMRC) Australia Fellow and a Victorian Breast Cancer Research Consortium Group

Leader. Melissa C. Southey is an NHMRC Senior Research Fellow and a Victorian Breast Cancer Research Consortium Group Leader. The Amsterdam Breast Cancer Study (ABCS) study was supported by the Dutch Cancer Society (grants NKI 2007–3839; 2009 4363); BBMRI-NL, which is a Research Infrastructure financed by the Dutch government (NWO 184.021.007); and the Dutch National Genomics Initiative. The work of the Bavarian Breast Cancer Cases and Controls (BBCC) was partly funded by ELAN-Fond of the University Hospital of Erlangen. The Bavarian British Breast Cancer Study (BBCS) is funded by Cancer Research UK and Breakthrough Breast Cancer and acknowledges National Health Service (NHS) funding to the National Institute for Health Research (NIHR) Biomedical Research Centre, and the National Cancer Research Network (NCRN). The Breast Cancer Association Consortium (BCAC) is funded by CR-UK (C1287/A10118 and C1287/A12014). Meetings of BCAC have been funded by the European Union European Cooperation in Science and Technology (COST) programme (BM0606). Douglas F. Easton is a Principal Research Fellow of CR-UK. In the Breast Cancer in Galway Genetic Study (BIGGS), Elinor Sawyer is supported by NIHR Comprehensive Biomedical Research Centre, Guy's & St. Thomas' NHS Foundation Trust in partnership with King's College London, United Kingdom. Ian Tomlinson is supported by the Oxford Biomedical Research Centre. The Breast Cancer Study of the University of Heidelberg (BSUCH) study was supported by the Dietmar-Hopp Foundation, the Helmholtz Society, and the German Cancer Research Center (DKFZ).

The CECILE Breast Cancer Study was funded by Fondation de France, Institut National du Cancer (INCa), Ligue Nationale contre le Cancer, Ligue contre le Cancer Grand Ouest, Agence Nationale de Sécurité Sanitaire (ANSES), Agence Nationale de la Recherche (ANR). The Copenhagen General Population Study (CGPS) was supported by the Chief Physician Johan Boserup and Lise Boserup Fund, the Danish Medical Research Council, and Herlev Hospital. The Spanish National Cancer Centre Breast Cancer Study (CNIO-BCS) was supported by the Genome Spain Foundation, the Red Temática de Investigación Cooperativa en Cáncer, and grants from the Asociación Española Contra el Cáncer and the Fondo de Investigación Sanitario (PI11/00923 and PI081120). The Human Genotyping-CEGEN Unit (CNIO) is supported by the Instituto de Salud Carlos III. The California Teachers Study (CTS) was initially supported by the California Breast Cancer Act of 1993 and the California Breast Cancer Research Fund (contract 97-10500) and is currently funded through the National Institutes of Health (R01 CA77398). Collection of cancer incidence data was supported by the California Department of Public Health as part of the statewide cancer reporting program mandated by California Health and Safety Code Section 103885. Hoda Anton-Culver receives support from the Lon V Smith Foundation (LVS39420). The Esther Breast Cancer Study (ESTHER) study was supported by a grant from the Baden Württemberg Ministry of Science, Research and Arts. Additional case patients were recruited in the context of the Verlauf der diagnostischen Abklärung study, which was supported by a grant from the German Cancer Aid (Deutsche Krebshilfe). The German Consortium for Hereditary Breast & Ovarian Cancer (GC-HBOC) was supported by Deutsche Krebshilfe (107 352). The Gene Environment Interaction and Breast Cancer in Germany (GENICA) was funded by the Federal Ministry of Education and Research (BMBF) Germany grants 01KW9975/5, 01KW9976/8, 01KW9977/0, and 01KW0114, the Robert Bosch Foundation, Stuttgart, Deutsches Krebsforschungszentrum (DKFZ), Heidelberg, Institute for Prevention and Occupational Medicine of the German Social Accident Insurance (IPA), Bochum, as well as the Department of Internal Medicine, Evangelische Kliniken Bonn gGmbH, Johanniter Krankenhaus, Bonn, Germany. The Helsinki Breast Cancer Study (HEBCS) was financially supported by the Helsinki University Central Hospital Research Fund, Academy of Finland (266528), the Finnish Cancer Society, the Nordic Cancer Union, and the Sigrid Juselius Foundation. The Hannover-Minsk Breast Cancer Study (HMBCS) was supported by a grant from the Friends of Hannover Medical School and by the Rudolf Bartling Foundation. Financial support for Karolinska Breast Cancer Study (KARBAC) was provided through the regional agreement on medical training and clinical research (ALF) between Stockholm County Council and Karolinska Institutet, the Swedish Cancer Society, The Gustav V Jubilee Foundation and and Bert von Kantzows Foundation. The Kuopio Breast Cancer Project (KBCP) was financially supported by the special Government Funding (EVO) of Kuopio University Hospital grants, Cancer Fund of North Savo, the Finnish Cancer Organizations, and by the strategic funding of the University of Eastern Finland. The Kathleen Cuningham Foundation Consortium for research into Familial Breast Cancer (kConFab) is supported by a grant from the National Breast Cancer Foundation, and previously by the National Health and NHMRC, the Queensland Cancer Fund, the Cancer Councils of New South Wales, Victoria, Tasmania, and South Australia, and the Cancer Foundation of Western Australia. The kConFab Clinical Follow Up Study was funded by the NHMRC (145684, 288704, 454508). Financial support for the Australian Ovarian Cancer Study (AOCS) was provided by the United States Army Medical Research and Materiel Command (DAMD17-01-1-0729), the Cancer Council of Tasmania and Cancer Foundation of Western Australia, and the NHMRC [199600]. Georgia Chenevix-Trench and Penelope M. Webb are supported by the NHMRC. Leuven Multidisciplinary Breast Centre (LMBC) is supported by the 'Stichting tegen Kanker' (232–2008 and 196–2010). Diether Lambrechts is supported by the Flemish Fund for Scientific Research and the KULPFV/10/016-SymBioSysII. The Mammary Carcinoma Risk Factor Investigation (MARIE) study was supported by the Deutsche Krebshilfe e.V. (70-2892-BR I, 106332, 108253, 10841), the Hamburg Cancer Society, the German Cancer Research Center, and the Federal Ministry of Education and Research (BMBF) Germany (01KH0402). Milan Breast Cancer Study Group (MBCSG) is supported by grants from the Italian Association for Cancer Research (AIRC) and by funds from the Italian citizens who allocated the 5/1000 share of their tax payment in support of the Fondazione IRCCS Istituto Nazionale Tumori, according to Italian laws (INT-Institutional strategic projects "5x1000"). The Mayo Clinic Breast Cancer Study (MCBCS) Clinic Breast Cancer Study (MCBCS) was supported by the NIH grants CA128978, CA116167, CA176785, an NIH Specialized Program of Research Excellence (SPORE) in Breast Cancer (CA116201), the Breast Cancer Research Foundation, and a generous gift from the David F. and Margaret T. Grohne Family Foundation and the Ting Tsung and Wei Fong Chao Foundation. Melbourne Collaborative Cohort Study (MCCS) cohort recruitment in the study was funded by VicHealth and Cancer Council Victoria. The MCCS was further supported by Australian NHMRC grants 209057, 251553, and 504711 and by infrastructure provided by Cancer Council Victoria. The Multi-ethnic Cohort (MEC) was support by NIH grants CA63464, CA54281, CA098758, and CA132839. The work of Montreal Gene-Environment Breast Cancer Study (MTLGEBCS) was supported by the Quebec Breast Cancer Foundation, the Canadian Institutes of Health Research for the "CIHR Team in Familial Risks of Breast Cancer" program (grant number CRN-87521), and the Ministry of Economic Development, Innovation and Export Trade (grant number PSR-SIIRI-701). The Norwegian Breast Cancer Study was supported by grants from the Norwegian Research council, 155218/V40, 175240/S10 to ALBD, FUGE-NFR 181600/V11 to VNK and a Swizz Bridge Award to ALBD. The Oulu Breast Cancer Study (OBCS) was supported by research grants from the Finnish Cancer Foundation, the Academy of Finland, the University of Oulu, and the Oulu University Hospital.

ARTICLE

The OFBCR was supported by grant UM1 CA164920 from the National Cancer Institute. The Leiden University Medical Centre Breast Cancer Study (ORIGO) study was supported by the Dutch Cancer Society (RUL 1997-1505) and the Biobanking and Biomolecular Resources Research Infrastructure (BBMRI-NL CP16). The NCI Polish Breast Cancer Study (PBCS) was funded by Intramural Research Funds of the National Cancer Institute, Department of Health and Human Services. The Karolinska Mammography Project for Risk Prediction of Breast Cancer - prevalent cases (pKARMA) study was supported by Märit and Hans Rausings Initiative Against Breast Cancer. The Rotterdam Breast Cancer Study (RBCS) was funded by the Dutch Cancer Society (DDHK 2004–3124, DDHK 2009–4318). The Singapore and Sweden Breast Cancer Study (SASBAC) study was supported by funding from the Agency for Science, Technology and Research of Singapore (A*STAR), the US National Institute of Health, and the Susan G. Komen Breast Cancer Foundation. The Sheffield Breast Cancer Study (SBCS) was supported by Yorkshire Cancer Research S295, S299, and S305PA. Study of Epidemiology and Risk factors in Cancer Heredity (SEARCH) is funded by a program grant from Cancer Research UK [C490/A10124] and supported by the UK National Institute for Health Research Biomedical Research Centre at the University of Cambridge. Städtisches Klinikum Karlsruhe Deutsches Krebsforschungszentrum Study (SKKDKFZS) is supported by the DKFZ. The IHCC-Szczecin Breast Cancer Study (SZBCS) was supported by Grant PBZ_KBN_122/P05/2004. The Triple Negative Breast Cancer Consortium Study (TNBCC) was supported by: a Specialized Program of Research Excellence (SPORE) in Breast Cancer (CA116201), a grant from the Breast Cancer Research Foundation, a generous gift from the David F. and Margaret T. Grohne Family Foundation and the Ting Tsung and Wei Fong Chao Foundation, the Stefanie Spielman Breast Cancer fund and the OSU Comprehensive Cancer Center, DBBR (a CCSG Share Resource by National Institutes of Health Grant P30 CA016056), the Hellenic Cooperative Oncology Group research grant (HR R_BG/04) and the Greek General Secretary for Research and Technology (GSRT) Program, Research Excellence II, the European Union (European Social Fund), and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF), ARISTEIA. The UK Breakthrough Generations Study (UKBGS) is funded by Breakthrough Breast Cancer and the Institute of Cancer Research (ICR). ICR acknowledges NHS funding to the NIHR Biomedical Research Centre.

## Notes

Acknowledgements: We thank all the individuals who took part in these studies and all the researchers, clinicians, technicians, and administrative staff who have enabled this work to be carried out. We thank Ms. Bethanie Rammer, who edited the manuscript.

In particular, we thank: Andrew Berchuck (Ovarian Cancer Association Consortium [OCAC]); Rosalind A. Eeles, Ali Amin Al Olama, Zsofia Kote-Jarai, and Sara Benlloch (Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome [PRACTICAL]); Antonis Antoniou, Lesley McGuffog, and Ken Offit (The Consortium of Investigators of Modifiers of BRCA1/2 [CIMBA]); Andrew Lee, Ed Dicks, Craig Luccarini, and the staff of the Centre for Genetic Epidemiology Laboratory; the staff of the CNIO genotyping unit; Sylvie LaBoissière and Frederic Robidoux and the staff of the McGill University and Génome Québec Innovation Centre; and the staff of the Copenhagen DNA laboratory; and Julie M Cunningham, Sharon A. Windebank, Christopher A. Hilker, Jeffrey Meyer, and the staff of Mayo Clinic Genotyping Core Facility.

The authors also wish to thank: Maggie Angelakos, Judi Maskiell, and Gillian Dite (ABCFS); Annegien Broeks, Sten Cornelissen, Richard van Hien, Frans Hogervorst, Senno Verhoef, Laura van 't Veer, Emiel Rutgers, Ellen van der Schoot, and Femke Atsma (ABCS); Matthias Rübner, Silke Landrith, Alexander Hein, Michael Schneider, and Sonja Oeser (BBCC); Eileen Williams, Elaine Ryder-Mills, and Kara Sargus (BBCS); Niall McInerney, Gabrielle Colleran, Andrew Rowan, and Angela Jones (BIGGS); Peter Bugert and Medical Faculty Mannheim (BSUCH); staff and participants of the Copenhagen General Population Study, Dorthe Uldall Andersen, Maria Birna Arnadottir, Anne Bank, and Dorthe Kjeldgård Hansen for the excellent technical assistance, and The Danish Breast Cancer Group (DBCG) for the tumor information (CGPS); Charo Alonso, and Primitiva Menendez (CNIO-BCS); Leslie Bernstein, James Lacey, Sophia Wang, Huiyan Ma, Yani Lu, and Jessica Clague DeHart at the Beckman Research Institute of City of Hope, Dennis Deapen, Rich Pinder, and Eunjung Lee at the University of Southern California, Pam Horn-Ross, Peggy Reynolds, Christina Clarke Dur, David Nelson at the Cancer Prevention Institute of California; and Argyrios Ziogas, and Hannah Park at the University of California Irvine (CTS); Hartwig Ziegler, Sonja Wolf, and Volker Hermann (ESTHER); Heide Hellebrand, Stefanie Engert (GC-HBOC); The GENICA Network (Dr. Margarete Fischer-Bosch-Institute of Clinical Pharmacology, Stuttgart, University of Tübingen, Germany [Wing-Yee Lo, Christina Justenhoven], Department of Internal Medicine, Evangelische Kliniken Bonn gGmbH, Johanniter Krankenhaus, Bonn, Germany [Christian Baisch], Institute of Pathology, University of Bonn, Germany [Hans-Peter Fischer], Molecular Genetics of Breast Cancer, Deutsches Krebsforschungszentrum, Heidelberg, Germany, and Institute for Prevention and Occupational Medicine of the German Social Accident Insurance, Institute of the Ruhr University Bochum, Bochum, Germany [Beate Pesch, Sylvia Rabstein, Anne Lotz], Institute of Occupational Medicine and Maritime Medicine, University Medical Center Hamburg-Eppendorf, Germany [Volker Harth]); Kirsimari Aaltonen, Karl von Smitten, Sofia Khan, Tuomas Heikkinen, and Irja Erkkilä (HEBCS); Peter Hillemanns, Hans Christiansen, and Johann H. Karstens (HMBCS); Eija Myöhänen and Helena Kemiläinen (KBCP); Heather Thorne and Eveline Niedermayr (kConFab/AOCS); Gilian Peuteman, Dominiek Smeets, Thomas Van Brussel, and Kathleen Corthouts (LMBC); Tracy Slanger, Elke Mutschelknauss, Katharina Buck, Alina Vrieling, Ursula Eilber, Sabine Behrens, Muhabbet Celik, and Til Olchers (MARIE); Bernard Peissel and Daniela Zaffaroni of the Fondazione IRCCS Istituto Nazionale dei Tumori (INT), Monica Barile and Irene Feroce of the Istituto Europeo di Oncologia (IEO), Loris Bernard, and the personnel of the Cogentech Cancer Genetic Test Laboratory (MBCSG); Martine Tranchant at CHU de Québec Research Center, and Marie-France Valois, Annie Turgeon, and Lea Heguy at McGill University Health Center, Royal Victoria Hospital, McGill University (MTLGEBCS); NBCS study group (NBCS); Meeri Otsukka and Kari Mononen (OBCS); Teresa Selander and Nayana Weerasooriya (OFBCR); Ellie Krol-Warmerdam, Jannet Blom, and Jan Molenaar (ORIGO); Louise Brinton, Mark Sherman, Stephen Chanock, Neonila Szeszenia-Dabrowska, Beata Peplonska, Witold Zatonski, Pei Chao, and Michael Stagner (PBCS); The Swedish Medical Research Counsel (pKARMA and SASBAC); Ans van den Ouweland, Anja Nieuwlaat, Ellen Crepin, and Petra Bos (RBCS); Sue Higham, Helen Cramp, Dan Connley, Ian brock, and Sabapathy Balasubramanian (SBCS); the SEARCH and EPIC teams (SEARCH); and Robert Pilarski, Charles Shapiro, the OSU Breast Cancer Tissue Bank, and the Human Genetics Sample Bank (TNBCC).

Author contributions: WZ conceived and directed the study. BZ performed literature searches, data extraction, quality assessment, and statistical analysis. BZ and WZ wrote the paper with significant contributions from XOS. BZ and RJD prepared the dataset for the Mendelian randomization analysis. WW, JL, and CL contributed to the discussion of statistical analysis. CZ assessed the quality of the data for the meta-analysis of prospective studies. DFE led BCAC and COGS. KM, MKB, QW, JD, AMD, JCC, AGN, PDPP, JS, and PH contributed significantly to BCAC and COGS. DJH directed the GAME-ON DRIVE Project. All authors contributed to collection of the data and biological samples in the original studies. All authors reviewed the manuscript and approved its submission for publication.

The authors declare no conflicts of interest. The sponsors of this study had no role in study design, data collection, analysis, interpretation, writing of the report, or the decision to publish. The authors had full access to the data and are responsible for the content of this manuscript.

## References


1. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. *CA Cancer J Clin*. 2011;61(2):69–90.
2. Gunnell D, Okasha M, Smith GD, Oliver SE, Sandhu J, Holly JM. Height, leg length, and cancer risk: a systematic review. *Epidemiol Rev*. 2001;23(2):313–342.
3. Okasha M, McCarron P, Gunnell D, Smith GD. Exposures in childhood, adolescence and early adulthood and breast cancer risk: a systematic review of the literature. *Breast Cancer Res Treat*. 2003;78(2):223–276.
4. Friedenreich CM. Review of anthropometric factors and breast cancer risk. *Eur J Cancer Prev*. 2001;10(1):15–32.
5. Tornberg SA, Holm LE, Carstensen JM. Breast cancer risk in relation to serum cholesterol, serum beta-lipoprotein, height, weight, and blood pressure. *Acta Oncol*. 1988;27(1):31–37.
6. Tretli S. Height and weight in relation to breast cancer morbidity and mortality. A prospective study of 570,000 women in Norway. *Int J Cancer*. 1989;44(1):23–30.
7. Vatten LJ, Kvinnsland S. Prospective study of height, body mass index and risk of breast cancer. *Acta Oncol*. 1992;31(2):195–200.
8. De Stavola BL, Wang DY, Allen DS, et al. The association of height, weight, menstrual and reproductive events with breast cancer: results from two prospective studies on the island of Guernsey (United Kingdom). *Cancer Causes Control*. 1993;4(4):331–340.
9. Michels-Blanck H, Byers T, Mokdad AH, Will JC, Calle EE. Menstrual patterns and breast cancer mortality in a large U.S. cohort. *Epidemiology*. 1996;7(5):543–546.
10. Freni SC, Eberhardt MS, Turturro A, Hine RJ. Anthropometric measures and metabolic rate in association with risk of breast cancer (United States). *Cancer Causes Control*. 1996;7(3):358–365.
11. Tulinius H, Sigfusson N, Sigvaldason H, Bjarnadottir K, Tryggvadottir L. Risk factors for malignant diseases: a cohort study on a population of 22,946 Icelanders. *Cancer Epidemiol Biomarkers Prev*. 1997;6(11):863–873.
12. Kaaks R, Van Noord PA, Den Tonkelaar I, Peeters PH, Riboli E, Grobbee DE. Breast-cancer incidence in relation to height, weight and body-fat distribution in the Dutch "DOM" cohort. *Int J Cancer*. 1998;76(5):647–651.
13. Sonnenschein E, Toniolo P, Terry MB, et al. Body fat distribution and obesity in pre- and postmenopausal breast cancer. *Int J Epidemiol*. 1999;28(6):1026–1031.
14. van den Brandt PA, Spiegelman D, Yaun SS, et al. Pooled analysis of prospective cohort studies on height, weight, and breast cancer risk. *Am J Epidemiol*. 2000;152(6):514–527.
15. Weiderpass E, Braaten T, Magnusson C, et al. A prospective study of body size in different periods of life and risk of premenopausal breast cancer. *Cancer Epidemiol Biomarkers Prev*. 2004;13(7):1121–1127.
16. Ahlgren M, Melbye M, Wohlfahrt J, Sorensen TI. Growth patterns and the risk of breast cancer in women. *N Engl J Med*. 2004;351(16):1619–1626.
17. Macinnis RJ, English DR, Gertig DM, Hopper JL, Giles GG. Body size and composition and risk of postmenopausal breast cancer. *Cancer Epidemiol Biomarkers Prev*. 2004;13(12):2117–2125.
18. McCullough ML, Feigelson HS, Diver WR, Patel AV, Thun MJ, Calle EE. Risk factors for fatal breast cancer in African-American women and White women in a large US prospective cohort. *Am J Epidemiol*. 2005;162(8):734–742.
19. Baer HJ, Rich-Edwards JW, Colditz GA, Hunter DJ, Willett WC, Michels KB. Adult height, age at attained height, and incidence of breast cancer in premenopausal women. *Int J Cancer*. 2006;119(9):2231–2235.
20. Lacey JV Jr, Kreimer AR, Buys SS, et al. Breast cancer epidemiology according to recognized breast cancer risk factors in the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial Cohort. *BMC Cancer*. 2009;9:84.
21. Green J, Cairns BJ, Casabonne D, Wright FL, Reeves G, Beral V. Height and cancer incidence in the Million Women Study: prospective cohort, and meta-analysis of prospective studies of height and total cancer risk. *Lancet Oncol*. 2011;12(8):785–794.
22. White KK, Park SY, Kolonel LN, Henderson BE, Wilkens LR. Body size and breast cancer risk: the Multiethnic Cohort. *Int J Cancer*. 2012;131(5):E705–E716.
23. Canchola AJ, Anton-Culver H, Bernstein L, et al. Body size and the risk of postmenopausal breast cancer subtypes in the California Teachers Study cohort. *Cancer Causes Control*. 2012.
24. Mellemkjaer L, Christensen J, Frederiksen K, et al. Leg length, sitting height and postmenopausal breast cancer risk. *Br J Cancer*. 2012;107(1):165–168.
25. Kabat GC, Heo M, Kamensky V, Miller AB, Rohan TE. Adult height in relation to risk of cancer in a cohort of Canadian women. *Int J Cancer*. 2013;132(5):1125–1132.
26. Emerging Risk Factors Collaboration. Adult height and the risk of cause-specific death and vascular morbidity in 1 million people: individual participant meta-analysis. *Int J Epidemiol*. 2012;41(5):1419–1433.
27. Ritte R, Lukanova A, Tjonneland A, et al. Height, age at menarche and risk of hormone receptor-positive and -negative breast cancer: A cohort study. *Int J Cancer*. 2013;132(11):2619–2629.
28. Kabat GC, Anderson ML, Heo M, et al. Adult stature and risk of cancer at different anatomic sites in a cohort of postmenopausal women. *Cancer Epidemiol Biomarkers Prev*. 2013;22(8):1353–1363.
29. Wiren S, Haggstrom C, Ulmer H, et al. Pooled cohort study on height and risk of cancer and cancer death. *Cancer Causes Control*. 2014;25(2):151–159.
30. Kabat GC, Kim MY, Hollenbeck AR, Rohan TE. Attained height, sex, and risk of cancer at different anatomic sites in the NIH-AARP Diet and Health Study. *Cancer Causes Control*. 2014;25(12):1697–1706.
31. Smith GD, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol*. 2003;32(1):1–22.
32. Yang J, Benyamin B, McEvoy BP, et al. Common SNPs explain a large proportion of the heritability for human height. *Nat Genet*. 2010;42(7):565–569.
33. Gudbjartsson DF, Walters GB, Thorleifsson G, et al. Many sequence variants affecting diversity of adult human height. *Nat Genet*. 2008;40(5):609–615.
34. Weedon MN, Lettre G, Freathy RM, et al. A common variant of HMGA2 is associated with adult and childhood height in the general population. *Nat Genet*. 2007;39(10):1245–1250.
35. Sanna S, Jackson AU, Nagaraja R, et al. Common variants in the GDF5-UQCC region are associated with variation in human height. *Nat Genet*. 2008;40(2):198–203.
36. Lettre G, Jackson AU, Gieger C, et al. Identification of ten loci associated with height highlights new biological pathways in human growth. *Nat Genet*. 2008;40(5):584–591.
37. Weedon MN, Lango H, Lindgren CM, et al. Genome-wide association analysis identifies 20 loci that influence adult height. *Nat Genet*. 2008;40(5):575–583.
38. Lango AH, Estrada K, Lettre G, et al. Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature*. 2010;467(7317):832–838.
39. Palmer TM, Lawlor DA, Harbord RM, et al. Using multiple genetic variants as instrumental variables for modifiable risk factors. *Stat Methods Med Res*. 2012;21(3):223–242.
40. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials*. 1986;7(3):177–188.
41. Lau J, Ioannidis JP, Schmid CH. Quantitative synthesis in systematic reviews. *Ann Intern Med*. 1997;127(9):820–826.
42. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med*. 2002;21(11):1539–1558.
43. Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. *Biometrics*. 1994;50(4):1088–1101.
44. Egger M, Davey SG, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ*. 1997;315(7109):629–634.
45. Garcia-Closas M, Couch FJ, Lindstrom S, et al. Genome-wide association studies identify four ER negative-specific breast cancer risk loci. *Nat Genet*. 2013;45(4):392.
46. Ghoussaini M, Fletcher O, Michailidou K, et al. Genome-wide association analysis identifies three new breast cancer susceptibility loci. *Nat Genet*. 2012;44(3):312–318.
47. Michailidou K, Hall P, Gonzalez-Neira A, et al. Large-scale genotyping identifies 41 new loci associated with breast cancer risk. *Nat Genet*. 2013;45(4):353–361.
48. Siddiq A, Couch FJ, Chen GK, et al. A meta-analysis of genome-wide association studies of breast cancer identifies two novel susceptibility loci at 6q14 and 20q11. *Hum Mol Genet*. 2012;21(24):5373–5384.
49. Palmer TM, Sterne JA, Harbord RM, et al. Instrumental variable estimation of causal risk ratios and causal odds ratios in Mendelian randomization analyses. *Am J Epidemiol*. 2011;173(12):1392–1403.
50. Thomas DC, Lawlor DA, Thompson JR. Re: Estimation of bias in nongenetic observational studies using "Mendelian triangulation" by Bautista *et al*. *Ann Epidemiol*. 2007;17(7):511–513.
51. Lawlor DA, Harbord RM, Sterne JA, Timpson N, Davey SG. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Stat Med*. 2008;27(8):1133–1163.

52. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet Epidemiol*. 2013;37(7):658–665.
53. Silventoinen K. Determinants of variation in adult body height. *J Biosoc Sci*. 2003;35(2):263–285.
54. Renehan AG. Height and cancer: consistent links, but mechanisms unclear. *Lancet Oncol*. 2011;12(8):716–717.
55. He M, Xu M, Zhang B, et al. Meta-analysis of genome-wide association studies of adult height in East Asians identifies 17 novel loci. *Hum Mol Genet*. 2015;24(6):1791–1800.
56. Rosenfeld RG. Insulin-like growth factors and the basis of growth. *N Engl J Med*. 2003;349(23):2184–2186.
57. Pollak MN, Schernhammer ES, Hankinson SE. Insulin-like growth factors and neoplasia. *Nat Rev Cancer*. 2004;4(7):505–518.
58. Key TJ, Appleby PN, Reeves GK, Roddam AW. Insulin-like growth factor 1 (IGF1), IGF binding protein 3 (IGFBP3), and breast cancer risk: pooled individual data analysis of 17 prospective studies. *Lancet Oncol*. 2010;11(6):530–542.
59. Guevara-Aguirre J, Balasubramanian P, Guevara-Aguirre M, et al. Growth hormone receptor deficiency is associated with a major reduction in pro-aging signaling, cancer, and diabetes in humans. *Sci Transl Med*. 2011;3(70):70ra13.
60. Wood AR, Esko T, Yang J, et al. Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat Genet*. 2014;46(11):1173–1186.
61. Jiang J, Hui CC. Hedgehog signaling in development and cancer. *Dev Cell*. 2008;15(6):801–812.
62. Kasper M, Jaks V, Fiaschi M, Toftgard R. Hedgehog signalling in breast cancer. *Carcinogenesis*. 2009;30(6):903–911.
63. Benson JR. Role of transforming growth factor beta in breast carcinogenesis. *Lancet Oncol*. 2004;5(4):229–239.
64. Howard B, Ashworth A. Signalling pathways implicated in early mammary gland morphogenesis and breast cancer. *PLoS Genet*. 2006;2(8):e112.
65. Zhou J, Wulfkuhle J, Zhang H, et al. Activation of the PTEN/mTOR/STAT3 pathway in breast cancer stem-like cells is required for viability and maintenance. *Proc Natl Acad Sci U S A*. 2007;104(41):16158–16163.
66. Lindvall C, Bu W, Williams BO, Li Y. Wnt signaling, stem cells, and the cellular origin of breast cancer. *Stem Cell Rev*. 2007;3(2):157–168.
67. Westra HJ, Peters MJ, Esko T, et al. Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat Genet*. 2013;45(10):1238–1243.
68. Ward LD, Kellis M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res*. 2012;40(Database issue):D930-D934.
69. Nelson CP, Hamby SE, Saleheen D, et al.Genetically Determined Height and Coronary Artery Disease. *New Engl J Med*. 2015;372(17):1608–1618.
70. Do R, Willer CJ, Schmidt EM, et al. Common variants associated with plasma triglycerides and risk for coronary artery disease. *Nat Genet*. 2013;45(11):1345–1352.
71. Voight BF, Peloso GM, Orho-Melander M, et al. Plasma HDL cholesterol and risk of myocardial infarction: a mendelian randomisation study. *Lancet*. 2012;380(9841):572–580.
72. IL6R Genetics Consortium Emerging Risk Factors Collaboration, Sarwar N, Butterworth AS, et al. Interleukin-6 receptor pathways in coronary heart disease: a collaborative meta-analysis of 82 studies. *Lancet*. 2012;379(9822):1205–1213.
73. Interleukin-6 Receptor Mendelian Randomisation Analysis (IL6R MR) Consortium. The interleukin-6 receptor as a target for prevention of coronary heart disease: a mendelian randomisation analysis. *Lancet*. 2012;379(9822):1214–1224.
74. Vimaleswaran KS, Berry DJ, Lu C, et al. Causal relationship between obesity and vitamin D status: bi-directional Mendelian randomization analysis of multiple cohorts. *PLoS Med*. 2013;10(2):e1001383.