# The CODATwins Project: the current status and recent findings of COllaborative project of Development of Anthropometrical measures in Twins

Karri Silventoinen (1) (2), Aline Jelenkovic (3) (4), Yoshie Yokoyama (5), Reijo Sund (1) (6), Masumi Sugawara (7), Mami Tanaka (8), Satoko Matsumoto (9), Leonie H Bogl (10) (11), Duarte L Freitas (12) (13), José Antonio Maia (13), Jacob v. B. Hjelmborg (14), Sari Aaltonen (10), Maarit Piirtola (10), Antti Latvala (4) (10), Lucas Calais-Ferreira (15), Vinicius C Oliveira (16), Paulo H Ferreira (17), Fuling Ji (18), Feng Ning (18), Zengchang Pang (18), Juan R Ordoñana (19) (20), Juan F Sánchez-Romera (19) (20), Lucia Colodro-Conde (19) (21), S Alexandra Burt (22), Kelly L Klump (22), Nicholas G Martin (23), Sarah E Medland (23), Grant W Montgomery (24), Christian Kandler (25), Tom A McAdams (26), Thalia C Eley (26), Alice M Gregory (27), Kimberly J Saudino (28), Lise Dubois (29), Michel Boivin (30), Mara Brendgen (31), Ginette Dionne (30), Frank Vitaro (32), Adam D Tarnoki (33) (34), David L Tarnoki (33) (34), Claire MA Haworth (35), Robert Plomin (26), Sevgi Y Öncel (36), Fazil Aliev (37), (38), Emanuela Medda (39), Lorenza Nisticò (39), Virgilia Toccaceli (39), Jeffrey M Craig (40) (41) (42), Richard Saffery (41) (42) Sisira H Siribaddana (43) (44), Matthew Hotopf (45) (46), Athula Sumathipala (43) (47), Fruhling Rijsdijk (26), Hoe-Uk Jeong (48), Timothy Spector (49), Massimo Mangino (49) (50), Genevieve Lachance (49), Margaret Gatz (51) (52), David A Butler (53), Wenjing Gao (54), Canqing Yu (54), Liming Li (54), Gombojav Bayasgalan (55), Danshiitsoodol Narandalai (56) (55), K Paige Harden (57), Elliot M Tucker-Drob (57), Kaare Christensen (14) (58), Axel Skytthe (14), Kirsten O Kyvik (59) (60), Catherine A Derom (61) (62), Robert F Vlietinck (61), Ruth JF Loos (63), Wendy Cozen (64) (65), Amie E Hwang (64) (65), Thomas M Mack (64) (65), Mingguang He (66) (67), Xiaohu Ding (66), Judy L Silberg (68), Hermine H Maes (69), Tessa L Cutler (70), John L Hopper (70) (71), Patrik KE Magnusson (52), Nancy L Pedersen (52), Anna K Dahl Aslan (52) (72), Laura A Baker (73), Catherine Tuvblad (73) (74), Morten Bjerregaard-Andersen (75) (76) (77), Henning Beck-Nielsen (77), Morten Sodemann (78), Vilhelmina Ullemar (52), Catarina Almqvist (52) (79), Qihua Tan (80), Dongfeng Zhang (81), Gary E Swan (82), Ruth Krasnow (83), Kerry L Jang (84), Ariel Knafo-Noam (85), David Mankuta (86), Lior Abramson (85), Paul Lichtenstein (52), Robert F Krueger (87), Matt McGue (87), Shandell Pahlen (88), Per Tynelius (89), Finn Rasmussen* (89), Glen E Duncan (90), Dedra Buchwald (90), Robin P Corley (91), Brooke M Huibregtse (92), Tracy L Nelson (93), Keith E Whitfield (94), Carol E Franz (95), William S Kremen (95) (96), Michael J Lyons (97), Syuichi Ooki (98), Ingunn Brandt (99), Thomas S Nilsen (99), Jennifer R Harris (99), Joohon Sung (71) (100), Hang A Park (101) (102), Jooyeon Lee (71), Soo Ji Lee (71) (103), Gonneke Willemsen (104), Meike Bartels (104), Catharina EM van Beijsterveldt (104), Clare H Llewellyn (105), Abigail Fisher (105), Esther Rebato (106), Andreas Busjahn (107), Rie Tomizawa (2), Fujio Inui (108) (2), Mikio Watanabe (2), Chika Honda (2), Norio Sakai (2), Yoon-Mi Hur (48), Thorkild IA Sørensen (109) (110), Dorret I Boomsma (104), Jaakko Kaprio (4) (10)

1. Department of Social Research, University of Helsinki, Helsinki, Finland.
2. Center for Twin Research, Osaka University Graduate School of Medicine, Osaka, Japan.
3. Department of Physiology, Faculty of Medicine and Nursing, University of the Basque Country, Leioa, Spain.
4. Department of Public Health, University of Helsinki, Helsinki, Finland.
5. Department of Public Health Nursing, Osaka City University, Osaka, Japan.
6. Institute of Clinical Medicine, University of Eastern Finland, Kuopio, Finland.
7. Department of Psychology, Ochanomizu University, Tokyo, Japan.
8. Center for Forensic Mental Health, Chiba University, Chiba, Japan.
9. Institute for Education and Human Development, Ochanomizu University, Tokyo, Japan.

10. Institute for Molecular Medicine Finland FIMM, Helsinki, Finland.
11. Leibniz Institute for Prevention Research and Epidemiology – BIPS, Bremen, Germany.
12. Department of Physical Education and Sport, University of Madeira, Funchal, Portugal.
13. CIFI2D, Faculty of Sport, Porto, University of Porto, Portugal.
14. The Danish Twin Registry, Department of Public Health, Epidemiology, Biostatistics & Biodemography, University of Southern Denmark Odense, Denmark.
15. Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, Australia.
16. Pós-Graduação em Reabilitação e Desempenho Funcional, Universidade Federal dos Vales do Jequitinhonha e Mucuri, Diamantina, Brazil.
17. Musculoskeletal Health Research Group, Faculty of Health Sciences, The University of Sydney, Sydney, Australia.
18. Department of Noncommunicable Diseases Prevention, Qingdao Centers for Disease Control and Prevention, Qingdao, China.
19. Department of Human Anatomy and Psychobiology, University of Murcia, Murcia, Spain.
20. IMIB-Arrixaca, Murcia, Spain.
21. QIMR Berghofer Medical Research Institute, Brisbane, Australia.
22. Michigan State University, East Lansing, Michigan, USA.
23. Genetic Epidemiology Department, QIMR Berghofer Medical Research Institute, Brisbane, Australia.
24. Institute for Molecular Bioscience, The University of Queensland, Brisbane, Australia.
25. Department of Psychology, University of Bremen, Bremen, Germany.
26. Social Genetic and Developmental Psychiatry Research Centre, Institute of Psychiatry, Psychology and Neuroscience, King’s College London.
27. Department of Psychology, Goldsmiths, University of London, London, UK .
28. Boston University, Department of Psychological and Brain Sciencies, Boston, MA, USA.
29. School of Epidemiology and Public Health, University of Ottawa, Ottawa, Ontario, Canada.
30. École de psychologie, Université Laval, Québec, Canada.
31. Département de psychologie, Université du Québec à Montréal, Montréal, Québec, Canada.
32. École de psychoéducation, Université de Montréal, Montréal, Québec, Canada.
33. Department of Radiology, Semmelweis University, Budapest, Hungary.
34. Hungarian Twin Registry, Budapest, Hungary.
35. MRC Integrative Epidemiology Unit, University of Bristol, Bristol, U.K.
36. Department of Statistics, Faculty of Arts and Sciences, Kırıkkale University, Kırıkkale, Turkey.
37. Psychology and African American Studies, Viginia Commonwealth University, Richmond, USA.
38. Faculty of Business, Karabuk University, Turkey.
39. Istituto Superiore di Sanità - Centre for Behavioural Sciences and Mental Health, Rome, Italy.
40. Centre for Molecular and Medical Research, Deakin University School of Medicine, Geelong, Australia.
41. Murdoch Childrens Research Institute, Royal Children's Hospital, Parkville, Victoria, Australia .
42. Department of Paediatrics, University of Melbourne, Parkville, Victoria, Australia.
43. Institute of Research & Development, Battaramulla, Sri Lanka.
44. Faculty of Medicine & Allied Sciences, Rajarata University of Sri Lanka Saliyapura, Sri Lanka.
45. Institute of Psychiatry Psychology and Neuroscience, King's College London, London, UK.
46. South London and Maudsley NHS Foundation Trust, London, UK.
47. Research Institute for Primary Care and Health Sciences, School for Primary Care Research (SPCR), Faculty of Health, Keele University, Staffordshire, UK.

48. Department of Education, Mokpo National University, Jeonnam, South Korea.
49. Department of Twin Research and Genetic Epidemiology, King's College, London, UK.
50. NIHR Biomedical Research Centre at Guy’s and St Thomas’ Foundation Trust, London, UK.
51. Center for Economic and Social Research, University of Southern California, Los Angeles, CA, USA.
52. Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden.
53. Health and Medicine Division, The National Academies of Sciences, Engineering, and Medicine Washington, DC, USA.
54. Department of Epidemiology and Biostatistics, School of Public Health, Peking University, Beijing, China.
55. Healthy Twin Association of Mongolia, Ulaanbaatar, Mongolia.
56. Graduate School of Biomedical and Health Sciences, Hiroshima University, Hiroshima, Japan.
57. Department of Psychology, University of Texas at Austin, Austin, TX, USA.
58. Department of Clinical Biochemistry and Pharmacology and Department of Clinical Genetics, Odense University Hospital, Odense, Denmark.
59. Department of Clinical Research, University of Southern Denmark, Odense, Denmark.
60. Odense Patient data Explorative Network (OPEN), Odense University Hospital, Odense, Denmark.
61. Centre of Human Genetics, University Hospitals Leuven, Leuven, Belgium.
62. Department of Obstetrics and Gynaecology, Ghent University Hospitals, Ghent, Belgium.
63. The Charles Bronfman Institute for Personalized Medicine, The Mindich Child Health and Development Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA.
64. Department of Preventive Medicine, Keck School of Medicine of USC, University of Southern California, Los Angeles, California, USA.
65. USC Norris Comprehensive Cancer Center, Los Angeles, California, USA.
66. State Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Center, Sun Yat-sen University, Guangzhou, China.
67. Centre for Eye Research Australia, University of Melbourne, Melbourne, Australia.
68. Department of Human and Molecular Genetics, Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University, Richmond, Virginia, USA.
69. Department of Human and Molecular Genetics, Psychiatry & Massey Cancer Center, Virginia Commonwealth University, Richmond, Virginia, USA.
70. Twins Research Australia, Centre for Epidemiology and Biostatistics, The University of Melbourne, Melbourne, Victoria, Australia.
71. Department of Epidemiology, School of Public Health, Seoul National University, Seoul, Korea.
72. Institute of Gerontology and Aging Research Network – Jönköping (ARN-J), School of Health and Welfare, Jönköping University, Jönköping, Sweden.
73. Department of Psychology, University of Southern California, Los Angeles, CA, USA.
74. School of Law, Psychology and Social Work, Örebro University, Sweden.
75. Bandim Health Project, INDEPTH Network, Bissau, Guinea-Bissau.
76. Research Center for Vitamins and Vaccines, Statens Serum Institute, Copenhagen, Denmark.
77. Department of Endocrinology, Odense University Hospital, Odense, Denmark.
78. Department of Infectious Diseases, Odense University Hospital, Odense, Denmark.
79. Pediatric Allergy and Pulmonology Unit at Astrid Lindgren Children’s Hospital, Karolinska University Hospital, Stockholm, Sweden.
80. Epidemiology, Biostatistics and Biodemography, Department of Public Health, University of Southern Denmark, Odense, Denmark.
81. Department of Public Health, Qingdao University Medical College, Qingdao, China.

82. Stanford Prevention Research Center, Department of Medicine, Stanford University School of Medicine, Stanford, CA, USA.
83. Center for Health Sciences, SRI International, Menlo Park, CA, USA.
84. Department of Psychiatry, University of British Columbia, Vancouver, BC, Canada.
85. The Hebrew University of Jerusalem, Jerusalem, Israel.
86. Hadassah Hospital Obstetrics and Gynecology Department, Hebrew University Medical School, Jerusalem, Israel.
87. Department of Psychology, University of Minnesota, Minneapolis, MN, USA.
88. Department of Psychology, University of California, Riverside, Riverside, CA 92521, USA.
89. Department of Public Health Sciences, Karolinska Institutet, Stockholm, Sweden.
90. Washington State Twin Registry, Washington State University - Health Sciences Spokane, Spokane, WA, USA.
91. Institute for Behavioral Genetics, University of Colorado, Boulder, Colorado, USA.
92. Institute of Behavioral Science, University of Colorado, Boulder, Colorado, USA.
93. Department of Health and Exercise Sciencies and Colorado School of Public Health, Colorado State University, USA.
94. Psychology and Neuroscience, Duke University, Durham, NC, USA.
95. Department of Psychiatry, University of California, San Diego, CA, USA.
96. VA San Diego Center of Excellence for Stress and Mental Health, ,La Jolla, CA, USA.
97. Department of Psychological and Brain Sciences, Boston University, Boston, MA, USA.
98. Department of Health Science, Ishikawa Prefectural Nursing University, Kahoku, Ishikawa, Japan.
99. Division of Health Data and Digitalization, Norwegian Institute of Public Health, Oslo, Norway.
100. Institute of Health and Environment, Seoul National University, Seoul, South-Korea.
101. Department of Emergency Medicine, Hallym University Dongtan Sacred Heart Hospital, Hwaseong, Korea.
102. Department of Health Science, School of Public Health Seoul National University.
103. Institute of Health & Environment, Seoul National University, South Korea.
104. Netherlands Twin Register, Department of Biological Psychology, Vrije Universiteit, Amsterdam, Amsterdam, Netherlands.
105. Health Behaviour Research Centre, Department of Epidemiology and Public Health, Institute of Epidemiology and Health Care, University College London, London, UK.
106. Department of Genetics, Physical Anthropology and Animal Physiology, University of the Basque Country UPV/EHU, Leioa, Spain.
107. HealthTwiSt GmbH, Berlin, Germany.
108. Faculty of Health Science, Kio University, Nara, Japan.
109. Novo Nordisk Foundation Centre for Basic Metabolic Research (Section of Metabolic Genetics), Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark.
110. Department of Public Health (Section of Epidemiology), Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark.
111. Molecular Epidemiology Department, QIMR Berghofer Medical Research Institute, Brisbane, Australia.

*Retired

Correspondence:
Karri Silventoinen
University of Helsinki, Department of Social Research

P.O. Box 18, FIN-00014 University of Helsinki, Finland
GSM: +358400-620726
E-mail: karri.silventoinen@helsinki.fi

# Abstract

The COllaborative project of Development of Anthropometrical measures in Twins (CODATwins) project is a large international collaborative effort to analyze individual-level phenotype data from twins in multiple cohorts from different environments. The main objective is to study factors that modify genetic and environmental variation of height, body mass index (BMI, $kg/m^2$) and size at birth, and additionally to address other research questions such as long-term consequences of birth size. The project started in 2013 and is open to all twin projects in the world having height and weight measures on twins with information on zygosity. Thus far, 54 twin projects from 24 countries have provided individual-level data. The CODATwins database includes 489,981 twin individuals (228,635 complete twin pairs). Since many twin cohorts have collected longitudinal data, there is a total of 1,049,785 height and weight observations. For many cohorts, we also have information on birth weight and length, own smoking behavior and own or parental education. We found that the heritability estimates of height and BMI systematically changed from infancy to old age. Remarkably, only minor differences in the heritability estimates were found across cultural–geographic regions, measurement time and birth cohort for height and BMI. In addition to genetic epidemiological studies, we looked at associations of height and BMI with education, birthweight and smoking status. Within family analyses examined differences within same-sex and opposite-sex dizygotic twins in birth size and later development. The CODATwins project demonstrates the feasibility and value of international collaboration to address gene-by-exposure interactions that require large sample sizes and address the effects of different exposures across time, geographical regions and socio–economic status.

Height and body mass index (BMI, kg/m$^2$) are among the most intensively studied traits in human genetics and public health. The first genetic study on height was published in the late 19th century (Galton, 1886) and on BMI in the early 20th century (Davenport, 1923), both utilizing information on familial resemblance. These anthropometric traits were also among the first traits studied in humans when new scientific innovations in molecular genetics became available (Silventoinen et al., 2015). With the current global obesity epidemic coexisting with severe undernutrition affecting growth in some populations (NCD Risk Factor Collaboration, 2016a) and well-known associations of growth and adult height with several health indicators (Silventoinen, 2015), questions about the roles of genetic and environmental factors in variations of these traits in rapidly changing environments are important.

In order to delineate conditions modifying the heritability of height and BMI, genotype-by-exposure (GxE) studies that include large, diverse samples are needed. To address GxE, there have been some previous studies comparing the heritability of height (Silventoinen et al., 2003) and BMI (Hur et al., 2008; Schousboe et al., 2003) between countries or over time periods when mean height (Silventoinen et al., 2000) and BMI have increased (Rokholm et al., 2011a; Rokholm et al., 2011b). In these studies, a significant difference in heritability indicates the modification of gene expression by environment (i.e., for GxE interaction) (Boomsma and Martin, 2002). There is also evidence that the heritability of BMI may change with age, based on both literature-based meta-analyses (Elks et al., 2012; Silventoinen et al., 2010) and pooled individual data (Dubois et al., 2012). Studies investigating factors affecting heritability require large sample sizes and broad ranges of exposures to obtain a comprehensive understanding of this variation. As our environment rapidly changes, with large differences within and between countries in rates of obesity and lifestyle factors, pooling twin data from different environments is critically important to analyze the dependence of heritability on environments.

## Objectives of the project

To obtain an accurate answer to the question about how the heritability of height and BMI vary over age and sex as well as time and space, the COllaborative project of the Development of Anthropometrical measures in Twins (CODATwins) project was started in June 2013. The idea behind the CODATwins project was to bring together all globally available twin data on height and weight.

Additionally, year of birth, sex, zygosity and age at measurement were collected, as well as data on birth weight and length, birth order, gestational age, ethnicity, own and parental education and own smoking status to analyze how these factors are related to height and BMI including the genetic architecture of these traits. The project is open for all twin cohorts that have collected data on height and weight from monozygotic (MZ) and dizygotic (DZ) twins. Twin cohorts were identified from several sources. The most important was the previous special issue of *Twin Research and Human Genetics* on twin registers (Hur and Craig, 2013), which was complemented with other sources and personal contacts. The first invitation letter was sent in September 2013 to the potential collaborators. Follow-up letters were sent in October 2013, January 2014 and September 2014. The first version of the harmonized database used in the first scientific papers was ready in January 2015 (Silventoinen et al., 2015). However, we have received several new cohorts and updates after this first manuscript. The current version of the CODATwins database includes a vast majority of height and weight measures known to have been collected from twins with information on zygosity.

The initial objective of the project was to analyze the heritability of height and BMI across different cultures and geographic regions. As the project progressed, we also analyzed how birth-related factors were associated with later physical development and education, differences in the physical development of same-sex and opposite-sex DZ twins, and associations between smoking and BMI. Throughout the project, we classified the twin cohorts to three cultural–geographic regions based on average BMI levels: East Asia, having the lowest BMI; Europe, having intermediate BMI; and North America and Australia having the highest BMI (NCD Risk Factor Collaboration, 2016a). These regions reflect different social and nutritional environments, including different obesogenic levels, which may affect not only the heritability of BMI, but also birth size and later height.

## Current status of the database

Figure 1 presents the geographic diversity of the 54 individual twin projects who have contributed data to the CODATwins database. The collaborators come from 24 countries. They mainly represent Europe, North America, Australia and, in to a lesser extent, East Asia; individual twin cohorts come from South Asia, Middle East, Africa and South America.

Table 1 presents the basic characteristics of the CODATwins database by country. The footnote indicates that there are 58 twin cohorts in the CODATwins database. Note that one study project can include more than one cohort. Together, the database includes nearly a half million twin individuals, including nearly a quarter million complete twin pairs. Half of the twins are females, but this overall proportion conceals variation in the sex ratios between twin cohorts. In many cohorts including adult twins, there are somewhat more women than men because of more active participation and lower mortality in women than men. However, we have three male-only cohorts including army veterans and conscripts, equalizing the overall sex ratio in the whole database. A larger proportion of twins are same-sex DZ twins (39%), compared to opposite-sex DZ twins (22%). This is mainly because some of the cohorts have collected data only on same-sex twin pairs, but also partly because of lower participation rates of opposite-sex compared to same-sex DZ twins in adult cohorts.

There are about 1 million height and weight measures after 6 months of age from the 489,981 twins available in the database. For around half of twins (N=252,624), we had only baseline data available, and for those having follow-up data, 112,691 twins had only one follow-up measure whereas 5695 had nine or more follow-up measures. The majority of measures are based on self-reports (63%) or parental reports (20%), and only a minority are measured values (17%). For nearly 150,000 twin individuals, we had additional information on birthweight, but only a minority of them were measured weights (7%) and the majority were parentally reported (78%) or self-reported (16%). For slightly less than half of these twin individuals, we have additional information on birth length (41%) and gestational age (44%). However, it is noteworthy that especially in the parental reports of height or length and weight, the reliability of the data probably varies between the cohorts. In some cohorts, parents were able to use the records of measures provided by medical doctors or registered nurses during child health check-ups whereas in others, they needed to rely on their own estimates and recall.

We also collected information on own smoking and own and parental education, but these measures were not mandatory for participation in the project and thus were available in only some of the cohorts, and in some cohorts were not available for all participants. Longitudinal information on smoking was collected if available, provided about a quarter of million assessments of smoking status. For own and parental education, only one measure was collected, preferring the most recent one (i.e., the highest attained educational level measure). We have information on maternal education for nearly 150,000 twin individuals, and for the large majority of them (96%) we have additional information on paternal education. Information on own education is available for about 200,000 twin individuals. However, because parental education is mainly available for children and own education for adults, we have only 63,138 twin individuals with all three educational measures.

## Major findings from studies on anthropometric traits

We started this project by detailing the growth development of twins. DZ twins were consistently taller and, in childhood and adolescence especially, they also had somewhat higher BMI than MZ twins. Slightly higher BMI variance was found for DZ twins than for MZ twins in childhood, but this difference disappeared in adolescence; for height, no zygosity differences in the variances were seen (Jelenkovic et al., 2015). First-born twins were heavier in infancy and also slightly taller than second-born twins over childhood and adolescence, but this difference disappeared in adulthood (Yokoyama et al., 2016). Thus, in further genetic modeling, we used different means of BMI and height for MZ and DZ twins. We also randomized the birth order within twin pairs since information on birth order was available for only 39% of the twins.

At birth, genetic factors explained a smaller proportion of the variation of weight, length and ponderal index (PI, $kg/m^3$) than shared and unique environmental factors; the heritability estimates, however, somewhat increased when the results were adjusted for gestational age (Yokoyama et al., 2018). The heritability estimates of height increased from early childhood until adulthood because of decreasing shared environmental variation (Jelenkovic et al., 2016a). For BMI, the heritability estimates first decreased from infancy to early childhood and then increased being between 0.7 and 0.8 at most ages during late childhood and adolescence; these differences were due to shared environmental variation being highest in early childhood (Silventoinen et al., 2016). The heritability of BMI was highest in early adulthood and then decreased until old age due to increasing unique environmental variation (Silventoinen et al., 2017).

For height, we did not find systematic differences in the heritability estimates or the variances between the geographic–cultural regions in childhood and adolescence (Jelenkovic et al., 2016a) or in adulthood (Jelenkovic et al., 2016b). The variances and heritability estimates of adult height were also roughly similar across birth cohorts from the late 19th to late 20th centuries (Jelenkovic et al., 2016b). For BMI in childhood and adolescence, we found that the variation was highest in North America and Australia and lowest in East Asia, thus also corresponding to the mean BMI differences between the regions (Silventoinen et al., 2016). Similar differences between the cultural–geographic regions were also found in the variances of adult BMI. The adult BMI variance increased from the 1940s to the 2010s along with increasing mean BMI. The genetic and environmental variances, however, showed about the same differences leading to only minor and inconsistent differences in the heritability estimates of BMI over time or between the cultural–geographic regions (Silventoinen et al., 2017). The heritability estimates of height and BMI were about the same in males and females. However, we found sex-specific genetic effects already in childhood, increasing during adolescence and being highest in adulthood for both height (Jelenkovic et al., 2016a; Jelenkovic et al., 2016b) and BMI (Silventoinen et al., 2016; Silventoinen et al., 2017).

We utilized the discordant twin pair design to analyze the effect of differences in intrauterine environment for the twins on later physical development. Within twin pairs, a lighter and shorter twin at birth was also shorter over childhood and adolescence than the co-twin, and this effect was still seen in adulthood (Jelenkovic et al., 2018a). BMI showed similar effects, and a smaller co-twin at birth had lower BMI from early childhood to adulthood; however, the effect sizes somewhat attenuated in adulthood (Jelenkovic et al., 2017).

The comparison of opposite-sex and same-sex DZ twins is uniquely suited to study sex differences that may arise in utero through a putative masculinization process of female twins who have a male co-twin (compared to same-sex twin pairs) that may arise from being exposed in utero to sex hormones of the opposite sex. However, our data did not support these effects for anthropometric traits. When we studied the sex of the co-twin, we found that boys having a female co-twin had a slightly greater birthweight and longer gestational age than DZ boys having a same-sex co-twin; in girls, no differences were seen between opposite-sex and same-sex DZ twins (Jelenkovic et al., 2018b). In adulthood, both men and women having an opposite-sex co-twin were slightly taller than those DZ twins having a same-sex co-twin, whereas BMI showed no differences between these twin-type groups (Bogl et al., 2017).

## Major findings from studies on education and smoking

We have harmonized the different educational classifications in the individual datasets by transforming them into educational years. In our first study, we found only minor differences between MZ and DZ twins in own or parental education (Silventoinen et al., 2017). Because of large differences between countries and birth cohorts in educational levels, we decided to focus, in further studies, on relative education (i.e., education years adjusted for birth year and twin cohort). We used the discordant twin pair design to analyze how differences in birthweight between co-twins, which may reflect differences in the intrauterine environment, affect differences in education in adulthood. We found that the lighter co-twin at birth had shorter education than the heavier co-twin, but the differences were very small and somewhat inconsistent between birth cohorts and zygosities (Jelenkovic et al., 2018c). In another study, we analyzed how parental education modifies the genetic and environmental variation of BMI from infancy to old age in the cultural–geographic regions (Silventoinen et al., 2019). We found that the mean BMI and the genetic variance of BMI were greater in those whose parents had low education when compared to the offspring of highly educated parents. These associations were strongest in North America and Australia and weakest or non-existent in East Asia. These results suggest that the interplay between genetic predisposition, childhood social environment and macro-social context is important for socio–economic differences in BMI.

We wished to have smoking data as widely available as possible and therefore collected information on the current smoking status of twins themselves and harmonized it to three categories: never smokers, current smokers and former smokers. In our first study utilizing these measures, we examined twin pairs discordant for smoking, contrasting never, current and former smokers (Piirtola et al., 2018). As expected, the currently smoking twin had slightly lower BMI than the co-twin who had never smoked. Also as expected, the former smoking twins had a higher BMI than their current smoking co-twins. However, when comparing twins from MZ pairs discordant for former smoking and never smoking, we found only small differences, which suggests that the net effect of smoking initiation and subsequent quitting on weight trajectory is minor.

## Further study plans

There are numerous opportunities for further studies. Previous studies have presented literature-based meta-analyses of educational attainment (Branigan et al., 2013; de Zeeuw et al., 2015), but our database offers possibilities to study in much more detail how the genetic and environmental variation of educational years have changed over birth cohorts and vary between countries. We also plan to analyze how parental education modifies the genetic architecture of height and birthweight by using the same approach as in the previous study on parental education and BMI (Silventoinen et al., 2019). Further, we can analyze the associations of height and BMI with own education. By using the discordant twin pair design, as previously used by Piirtola et al. (2018), it is possible to analyze whether the association between smoking and education is causal or due to common genetic or common environmental factors.

We have not yet addressed longitudinal associations. Since around half of the participating twin cohorts have longitudinal measures of height and weight, the CODATwins database offers good opportunities for this type of research. We can analyze how genetic factors affect the tracking of height and BMI over childhood and weight change in adulthood. Finally, thus far most studies, including ours, have ignored the well-known skewness of BMI distribution, which may well be associated with both the increasing mean and variance of BMI (Pak et al., 2016). New methods to analyze the skewness of BMI distribution utilizing twin data are now available (Tsang et al., 2018), and the CODATwins database would allow for analyzing, in detail, the differences in the skewness of BMI distribution between different ages, measurement years and cultural–geographic regions.

## Discussion

The CODATwins project shows that conducting a large-scale international collaborative project of existing twin cohorts is feasible. Such large data sets can provide reliable answers to research questions that have not been resolved using small or moderate size cohorts. In addition, they can answer new research questions impossible to analyze in any single cohort, which rarely span multiple determinants such as age from birth to old age, time period and sufficient geo-cultural diversity. We have been able to collect an international database and answer several new research questions related to the genetic and environmental determinants of height and relative weight.

Our main results concern the heritability estimates of height and relative weight. We found that the heritability estimates of both height and relative weight (BMI or PI) varied considerably by age (Jelenkovic et al., 2016a; Jelenkovic et al., 2016b; Silventoinen et al., 2016; Silventoinen et al., 2017; Yokoyama et al., 2018). Thus, it is likely that in some previous literature-based meta-analyses on the heritability of BMI (Elks et al., 2012; Min et al., 2013), the age ranges of original studies may have been too broad to capture this complexity, and the differences in the reported heritability estimates may reflect age differences between the cohorts. This further emphasizes the need for pooled analyses instead of relying only on meta-analyses of published results. On the other hand, we found only little evidence that the macro-environment modifies the heritability estimates of height or BMI. Height has increased all over the world during the 20th century (NCD Risk Factor Collaboration, 2016b), and it could be speculated that this has affected the genetic architecture of height as environmental stress has diminished. However, we found that the heritability estimates of adult height, and also total, genetic and environmental variances, were very similar between the cohorts born from the late 19th to late 20th centuries despite considerable differences in mean height across these environments (Jelenkovic et al., 2016b). On the other hand, we found that the variance of BMI was higher in more obesogenic environments measured by both the cultural–geographic regions with different mean BMI levels as well as when analyzing BMI measures from the 1940s to the 2010s, over which time the mean BMI has increased (Silventoinen et al., 2016; Silventoinen et al., 2017). However, even for BMI, we did not find any systematic differences in the heritability estimates across the measurement years or the cultural–geographic regions. This suggests that the heritability estimates of height and BMI are robust for differences in macro–environment. In the case of height, this seems to be because total variation is not sensitive to the change of environment. In contrast, BMI variation increased along with increasing mean BMI, but this was due to increases in both the genetic and environmental variations. This suggests that factors affecting increasing mean BMI may operate partly by amplifying the effect of genes on BMI variation.

In our studies not focusing on heritability, we found results both supporting and contradicting previous studies. Using the twin design, we demonstrated that intrauterine conditions affecting smaller birth size are associated with shorter height and lower BMI from early childhood to adulthood (Jelenkovic et al., 2017; Jelenkovic et al., 2018a), which are consistent with prior studies. On the other hand, our findings that birthweight was only weakly associated with adult education in discordant twin pairs (Jelenkovic et al., 2018c) and that males and females having opposite-sex co-twins showed no consistent differences as compared to those having same-sex co-twins in height and BMI (Bogl et al., 2017) are not consistent with previous hypotheses. Because there is a well-known tendency to publish positive results (Thornton and Lee, 2000), these types of large collaborative studies are important to validly test hypotheses and estimate effect sizes not inflated by publication bias.

This project provides a good estimation of the total number of twins in different cohorts potentially available for further collaborative studies, and thus demonstrates the opportunities, as well as certain limitations, of the currently available twin data. Height and weight are among the most commonly collected traits, and we are aware of only a few twin cohorts that are not part of the database. Thus, the total number of participants in any twin study in the world may not be much higher than the half million twin individuals assembled in the CODATwins database. Less than one-fifth of the height and weight values were based on direct measures, with the rest being self- or parentally reported. When studying physiological traits requiring clinical examination, such as blood pressure or cholesterol level, the number of participants available is likely to be much smaller. However, in the future, the linkage of twin cohorts to population-based biobanks together with health-care databases is an avenue to address more detailed biomedical and clinical research questions. This approach has already been exploded to estimate heritability by using an extended family design (Polubriaginof et al., 2018). The current twin data disproportionately represents Western populations and, to a lesser extent, East Asian populations. Thus, new data collections would be very important especially in the geographic regions currently having only limited twin data available.

In conclusion, the CODATwins project demonstrates the scientific value of an international collaboration that aims to pool individual phenotypic datasets. This allowed the analyses of macro-environmental effects on genetic and environmental variations and resulted in a tremendous increase in statistical power. A similar approach could be used to study many other traits not yet included in the CODATwins project. This would lead to new knowledge and make the maximal use of data already collected.

## Acknowledgements

This study was conducted within the CODATwins project (Academy of Finland #266592). K Silventoinen is supported by Osaka University's International Joint Research Promotion Program. This research was facilitated through access to Twins Research Australia, a national resource supported by a Centre of Research Excellence Grant (ID: 1079102), from the National Health and Medical Research Council. The Boston University Twin Project is funded by grants (#R01 HD068435 #R01 MH062375) from the National Institutes of Health to K. Saudino. Paulo Ferreira is funded by a National Medical Research Council Research Fellowship. California Twin Program was supported by The California Tobacco-Related Disease Research Program (7RT-0134H, 8RT-0107H, 6RT-0354H) and the National Institutes of Health (1R01ESO15150-01). The Carolina African American Twin Study of Aging (CAATSA) was funded by a grant from the National Institute on Aging (grant 1RO1-AG13662-01A2) to K. E. Whitfield. The CATSS-Study is supported by the Swedish Research Council through the Swedish Initiative for Research on Microdata in the Social And Medical Sciences (SIMSAM) framework grant no 340-2013-5867, grants provided by the Stockholm County Council (ALF-projects), the Swedish Heart-Lung Foundation and the Swedish Asthma and Allergy Association’s Research Foundation. Chinese National Twin Registry is funded by Special Fund for Health Scientific Research in the Public Welfare (Project No: 201502006), China. Colorado Twin Registry is funded by NIDA funded center grant DA011015, & Longitudinal Twin Study HD10333; Author Huibregtse is supported by 5T32DA017637 and 5T32AG052371. Danish Twin Registry is supported by the National Program for Research Infrastructure 2007 from the Danish Agency for Science, Technology and Innovation, The Research Council for Health and Disease, the Velux Foundation and the US National Institute of Health (P01 AG08761). Since its origin the East Flanders Prospective Survey has been partly supported by grants from the Fund of Scientific Research, Flanders and Twins, a non-profit Association for Scientific Research in Multiple Births (Belgium). Data collection and analyses in Finnish twin cohorts have been supported by ENGAGE – European Network for Genetic and Genomic Epidemiology, FP7-HEALTH-F4-2007, grant agreement number 201413, National Institute of Alcohol Abuse and Alcoholism (grants AA-12502, AA-00145, and AA-09203 to R J Rose, the Academy of Finland Center of Excellence in Complex Disease Genetics (grant numbers: 213506, 129680), and the Academy of Finland (grants 100499, 205585, 118555, 141054, 265240, 263278 and 264146 to J Kaprio). Gemini was supported by a grant from Cancer Research UK (C1418/A7974). Waves 1-3 of Genesis 12-19 were funded by the W T Grant Foundation, the University of London Central Research fund and a Medical Research Council Training Fellowship (G81/343) and Career Development Award (G120/635) to Thalia C. Eley. Wave 4 was supported by grants from the Economic and Social Research Council (RES-000-22-2206) and the Institute of Social Psychiatry (06/07 – 11) to Alice M. Gregory who was also supported at that time by a Leverhulme Research Fellowship (RF/2/RFG/2008/0145). Wave 5 was supported by funding to Alice M. Gregory from Goldsmiths, University of London. T.C. Eley is part funded by a program grant from the UK Medical Research Council (MR/M021475/1). This study presents independent research [part-] funded by the National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King’s College London. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health. The Minnesota Twin Registry (MTR) acknowledges support from NIH grant R01AG053217. Guangzhou Twin Eye Study is supported by National Natural Science Foundation of China (grant #81125007). Anthropometric measurements of the Hungarian twins were supported by Medexpert Ltd., Budapest, Hungary. Korean

Twin-Family Register was supported by the Global Research Network Program of the National Research Foundation (NRF 2011-220-E00006). Longitudinal Israeli Study of Twins was funded by the Starting Grant no. 240994 from the European Research Council (ERC) to Ariel Knafo. The Michigan State University Twin Registry has been supported by Michigan State University, as well as grants R01-MH081813, R01-MH0820-54, R01-MH092377-02, R21-MH070542-01, R03-MH63851-01, 1R01-MH118848-01 from the National Institute of Mental Health (NIMH), R01-HD066040 from the Eunice Kennedy Shriver National Institute for Child Health and Human Development (NICHD), and 11-SPG-2518 from the MSU Foundation. The content of this manuscript is solely the responsibility of the authors and does not necessarily represent the official views of the NIMH, the NICHD, or the National Institutes of Health. The Murcia Twin Registry is supported by Fundación Séneca, Regional Agency for Science and Technology, Murcia, Spain (08633/PHCS/08, 15302/PHCS/10 & 19479/PI/14) and Ministry of Science and Innovation, Spain (PSI2009-11560 & PSI2014-56680-R). The NAS-NRC Twin Registry acknowledges financial support from the National Institutes of Health grant number R21 AG039572. Netherlands Twin Register acknowledges the Netherlands Organization for Scientific Research (NWO) and MagW/ZonMW grants 904-61-090, 985-10-002, 912-10-020, 904-61-193,480-04-004, 463-06-001, 451-04-034, 400-05-717, Addiction-31160008, Middelgroot-911-09-032, Spinozapremie 56-464-14192; VU University's Institute for Health and Care Research (EMGO+ ); the European Research Council (ERC - 230374), the Avera Institute, Sioux Falls, South Dakota (USA). Osaka University Aged Twin Registry is supported by grants from JSPS KAKENHI JP (23593419, 24792601, 26671010, 24590695, 26293128, 16K15385, 16K15978, 16K15989, 16H03261). PETS was supported by grants from the Australian National Health and Medical Research Council (grant numbers 437015 and 607358 to JC, and RS), the Bonnie Babes Foundation (grant number BBF20704 to JMC), the Financial Markets Foundation for Children (grant no. 032-2007 to JMC), and by the Victorian Government's Operational Infrastructure Support Program. Madeira data comes from the following project: Genetic and environmental influences on physical activity, fitness and health: the Madeira family study Project reference: POCI/DES/56834/2004 Founded by the Portuguese agency for research (The Foundation for Science and Technology [FCT]). TwinsUK receives funding from the Wellcome Trust, Medical Research Council and European Union. TwinsUK and M. Mangino are supported by the National Institute for Health Research (NIHR)-funded BioResource, Clinical Research Facility and Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust in partnership with King's College London. The Quebec Newborn Twin Study acknowledges financial support from the Fonds Québécois de la Recherche sur la Société et la Culture, the Fonds de la Recherche en Santé du Québec, the Social Science and Humanities Research Council of Canada, the National Health Research Development Program, the Canadian Institutes for Health Research, Sainte-Justine Hospital's Research Center, and the Canada Research Chair Program (Michel Boivin). South Korea Twin Registry is supported by National Research Foundation of Korea (NRF-371-2011-1 B00047). We acknowledge The Swedish Twin Registry for access to data. The Swedish Twin Registry is managed by Karolinska Institutet and receives funding through the Swedish Research Council under the grant no 2017-00641. The Twins Early Development Study (TEDS) is supported by a program grant (G0901245) from the UK Medical Research Council and the work on obesity in TEDS is supported in part by a grant from the UK Biotechnology and Biological Sciences Research Council (31/D19086). Currently TEDS is supported by MRC grant 'MR/M021475/1'. The Texas Twin Project is currently funded by grant R01HD083613 from the National Institutes of Health. S.Y. Öncel and F. Aliev are supported by

Kırıkkale University Research Grant: KKU, 2009/43 and TUBITAK grant 114C117. The University of Southern California Twin Study is funded by a grant from the National Institute of Mental Health (R01 MH58354). Washington State Twin Registry (formerly the University of Washington Twin Registry) was supported in part by grant NIH RC2 HL103416 (D. Buchwald, PI). Vietnam Era Twin Study of Aging was supported by National Institute of Health grants NIA R01 AG018384, R01 AG018386, R01 AG022381, and R01 AG022982, and, in part, with resources of the VA San Diego Center of Excellence for Stress and Mental Health. The Cooperative Studies Program of the Office of Research & Development of the United States Department of Veterans Affairs has provided financial support for the development and maintenance of the Vietnam Era Twin (VET) Registry. The content of this manuscript is solely the responsibility of the authors and does not necessarily represent the official views of the NIA/NIH, or the VA. The West Japan Twins and Higher Order Multiple Births Registry was supported by Grant-in-Aid for Scientific Research (B) (grant number 15H05105) from the Japan Society for the Promotion of Science.

## References

Bogl, L.H., Jelenkovic, A., Vuoksimaa, E., Ahrenfeldt, L., Pietiläinen, K.H., Stazi, M.A., Fagnani, C., D'Ippolito, C., Hur, Y.M., Jeong, H.U., Silberg, J.L. et al. 2017, "Does the sex of one's co-twin affect height and BMI in adulthood? A study of dizygotic adult twins from 31 cohorts", Biol Sex Differ, vol. 8, no. 14.

Boomsma, D.I. & Martin, N.G. 2002, "Gene-environment interactions" in Biological Psychiatry, eds. H. D'haenen, J.A. den Boer & P. Wilner, John Wiley & Sons Ltd, New York, pp. 181-187.

Branigan, A.R., McCallum, K.J. & Freese, J. 2013, "Variation in the heritability of educational attainment: an international meta-analysis", Social Forces, vol. 92, pp. 109-140.

Davenport, C.B. 1923, Body-build and its inheritance, Carnegie Institution of Washington, No. 329, Washington.

de Zeeuw, E.L., de Geus, E.J.C. & Boomsma, D.I. 2015, "Meta-analysis of twin studies highlights the importance of genetic variation in primary school educational achievement", Trends in Neuroscience and Education, vol. 4, pp. 69-76.

Dubois, L., Ohm Kyvik, K., Girard, M., Tatone-Tokuda, F., Perusse, D., Hjelmborg, J., Skytthe, A., Rasmussen, F., Wright, M.J., Lichtenstein, P. & Martin, N.G. 2012, "Genetic and environmental contributions to weight, height, and BMI from birth to 19 years of age: an international study of over 12,000 twin pairs ", PloS One, vol. 7, pp. e30153.

Elks, C.E., den Hoed, M., Zhao, J.H., Sharp, S.J., Wareham, N.J., Loos, R.J. & Ong, K.K. 2012, "Variability in the heritability of body mass index: a systematic review and meta-regression ", Frontiers in Endocrinology, vol. 3, no. doi: 10.3389/fendo.2012.00029.

Galton, F. 1886, "Regression towards mediocrity in heriditary stature", Journal of Antropological Institute, vol. 15, pp. 246-262.

Hur, Y.M. & Craig, J.M. 2013, "Twin registries worldwide: an important resource for scientific research ", Twin Research and Human Genetics, vol. 16, pp. 1-12.

Hur, Y.M., Kaprio, J., Iacono, W.G., Boomsma, D.I., McGue, M., Silventoinen, K., Martin, N.G., Luciano, M., Visscher, P.M., Rose, R.J., He, M., Ando, J., Ooki, S., Nonaka, K., Lin, C.C., Lajunen, H.R., Cornes, B.K., Bartels, M., van Beijsterveldt, C.E., Cherny, S.S. & Mitchell, K. 2008, "Genetic influences on the difference in variability of height, weight and body mass index between Caucasian and East Asian adolescent twins", International Journal of Obesity, vol. 32, pp. 1455-1467.

Jelenkovic, A., Hur, Y.M., Sund, R., Yokoyama, Y., Siribaddana, S.H., Hotopf, M., Sumathipala, A., Rijsdijk, F., Tan, Q., Zhang, D. et al. 2016b, "Genetic and environmental influences on adult human height across birth cohorts from 1886 to 1994", Elife, vol. e20320.

Jelenkovic, A., Mikkonen, J., Martikainen, P., Latvala, A., Yokoyama, Y., Sund, R., Vuoksimaa, E., Rebato, E., Sung, J., Kim, J., Lee, J. et al. 2018c, "Association between birth weight and educational attainment: an individual-based pooled analysis of nine twin cohorts", J Epidemiol Community Health, vol. 72, pp. 832-837.

Jelenkovic, A., Sund, R., Hur, Y.M., Yokoyama, Y., Hjelmborg, J., Möller, S., Honda, C., Magnusson, P.K., Pedersen, N.L., Ooki, S. et al. 2016a, "Genetic and environmental influences on height from infancy to early adulthood: An individual-based pooled analysis of 45 twin cohorts", Scientific Reports, vol. 6, no. 28496.

Jelenkovic, A., Sund, R., Yokoyama, Y., Hur, Y.M., Ullemar, V., Almqvist, C., Magnusson, P.K., Willemsen, G., Bartels, M., Beijsterveldt, C.E.V. et al. 2018b, "Birth size and gestational age in opposite-sex twins as compared to same-sex twins: An individual-based pooled analysis of 21 cohorts", Sci Rep, vol. 19, no. 6300.

Jelenkovic, A., Yokoyama, Y., Sund, R., Honda, C., Bogl, L.H., Aaltonen, S., Ji, F., Ning, F., Pang, Z. et al. 2015, "Zygosity differences in height and body mass index of twins from infancy to old age: A study of the CODATwins project", Twin Res Hum Genet, vol. 18, pp. 557-570.

Jelenkovic, A., Yokoyama, Y., Sund, R., Hur, Y.M., Harris, J.R., Brandt, I., Nilsen, T.S., Ooki, S., Ullemar, V. et al. 2018a, "Associations between birth size and later height from infancy through adulthood: An individual based pooled analysis of 28 twin cohorts participating in the CODATwins project", Early Hum Dev, vol. 120, pp. 53-60.

Jelenkovic, A., Yokoyama, Y., Sund, R., Pietiläinen, K.H., Hur, Y.M., Willemsen, G., Bartels, M., van Beijsterveldt, T.C.E.M., Ooki, S. et al. 2017, "Association between birthweight and later body mass index: an individual-based pooled analysis of 27 twin cohorts participating in the CODATwins project", Int J Epidemiol, vol. 46, pp. 1488-1498.

Min, J., Chiu, D.T. & Wang, Y. 2013, "Variation in the heritability of body mass index based on diverse twin studies: a systematic review", Obes Rev, vol. 14, pp. 871-882.

NCD Risk Factor Collaboration 2016a, "Trends in adult body-mass index in 200 countries from 1975 to 2014: a pooled analysis of 1698 population-based measurement studies with 19·2 million participants", Lancet, vol. 387, pp. 1377-1396.

NCD Risk Factor Collaboration 2016b, "A century of trends in adult human height", Elife, vol. 26, no. e13410.

Pak, T.Y., Ferreira, S. & Colson, G. 2016, "Measuring and tracking obesity inequality in the United States: evidence from NHANES, 1971-2014", Popul Health Metr, vol. eCollection 2016, no. 10.1186/s12963-016-0081-5.

Piirtola, M., Jelenkovic, A., Latvala, A., Sund, R., Honda, C., Inui, F., Watanabe, M., Tomizawa, R. & Iwatani, Y. 2018, "Association of current and former smoking with body mass index: A study of smoking discordant twin pairs from 21 twin cohorts", PLoS One, vol. 13, no. e0200140.

Polubriaginof, F.C.G., Vanguri, R., Quinnies, K., Belbin, G.M., Yahi, A., Salmasian, H., Lorberbaum, T., Nwankwo, V., Li, L., Shervey, M.M., Glowe, P., Ionita-Laza, I., Simmerling, M., Hripcsak, G., Bakken, S., Goldstein, D., Kiryluk, K., Kenny, E.E., Dudley, J., Vawdrey, D.K., Tatonetti, N.P. 2018 "Disease heritability inferred from familial relationships reported in medical records". Cell, vol 173, pp. 1692-1704.

Rokholm, B., Silventoinen, K., Angquist, L., Skytthe, A., Kyvik, K.O. & Sørensen, T.I.A. 2011a, "Increased genetic variance of BMI with a higher prevalence of obesity", PloS One, vol. 6, pp. e20816.

Rokholm, B., Silventoinen, K., Tynelius, P., Gamborg, M., Sørensen, T.I.A. & Rasmussen, F. 2011b, "Increasing genetic variance of body mass index during the Swedish obesity epidemic ", PloS One, vol. 6, pp. e27135.

Schousboe, K., Willemsen, G., Kyvik, K.O., Mortensen, J., Boomsma, D.I., Cornes, B.K., Davis, C.J., Fagnani, C., Hjelmborg, J., Kaprio, J., De Lange, M., Luciano, M., Martin, N.G., Pedersen, N., Pietiläinen, K.H., Rissanen, A., Saarni, S., Sørensen, T.I.A., Van Baal, G.C. & Harris, J.R. 2003, "Sex differences in heritability of BMI: a comparative study of results from twin studies in eight countries", Twin Research, vol. 6, pp. 409-421.

Silventoinen, K. 2015, "Children's anthropometrics and later disease incidence" in The Oxford Handbook of Economics and Human Biology, eds. J. Komlos & I.R. Kelly, Oxford University Press, Oxford, pp. 1-20.

Silventoinen, K., Jelenkovic, A., Latvala, A., Sund, R., Yokoyama, Y., Ullemar, V., Almqvist, C., Derom, C.A., Vlietinck, R.F. , Loos, R.J.F. et al. 2017, "Education in twins and their parents across birth cohorts over 100 years: an individual-level pooled analysis of 42 twin cohorts", Twin Res Human Genet, vol. 20, pp. 395-405.

Silventoinen, K., Jelenkovic, A., Sund, R., Honda, C., Aaltonen, S., Yokoyama, Y., Tarnoki, A.D., Tarnoki, D.L., Ning, F., Ji, F., Pang, Z. et al. 2015, "The CODATwins Project: The cohort description of collaborative project of development of anthropometrical measures in twins to study macro-environmental variation in genetic and environmental effects on anthropometric traits.", Twin Res Hum Genet, vol. 18, pp. 348-360.

Silventoinen, K., Jelenkovic, A., Sund, R., Hur, Y., Yokoyama, Y., Honda, C., Hjelmborg, H. et al. 2016, "Genetic and environmental effects on body mass index from infancy to the onset of adulthood: an individual-based pooled analysis of 45 twin cohorts participating in the COllaborative project of Development of Anthropometrical measures in Twins (CODATwins) study", Am J Clin Nutr, vol. 104, pp. 371-379.

Silventoinen, K., Jelenkovic, A, Latvala, A, Yokoyama, Y., Sund, R., Sugawara, M., Tanaka, M., Matsumoto, S., Aaltonen, S., Piirtola, M., et al. 2019,"Parental education and genetics of body mass index from infancy to old age: a pooled analysis of 29 twin cohorts", Obesity, e-pub ahead of print.

Silventoinen, K., Jelenkovic, A., Sund, R., Yokoyama, Y., Hur, Y.M., Cozen, W., Hwang, A.E., Mack, T.M. et al. 2017, "Differences in genetic and environmental variation in adult body mass index by sex, age, time period, and region: an individual-based pooled analysis of 40 twin cohorts", Am J Clin Nutr, vol. 106, pp. 457-466.

Silventoinen, K., Kaprio, J., Lahelma, E. & Koskenvuo, M. 2000, "Relative effect of genetic and environmental factors on body height: differences across birth cohorts among Finnish men and women", American Journal of Public Health, vol. 90, pp. 627-630.

Silventoinen, K., Rokholm, B., Kaprio, J. & Sørensen, T.I.A. 2010, "The genetic and environmental influences on childhood obesity: a systematic review of twin and adoption studies", International Journal of Obesity, vol. 34, pp. 29-40.

Silventoinen, K., Sammalisto, S., Perola, M., Boomsma, D.I., Cornes, B.K., Davis, C., Dunkel, L., De Lange, M., Harris, J.R., Hjelmborg, J.V., Luciano, M., Martin, N.G., Mortensen, J., Nistico, L., Pedersen, N.L., Skytthe, A., Spector, T.D., Stazi, M.A., Willemsen, G. & Kaprio, J. 2003, "Heritability of adult body height: a comparative study of twin cohorts in eight countries", Twin Research, vol. 6, pp. 399-408.

Thornton, A. & Lee, P. 2000, "Publication bias in meta-analysis: its causes and consequences ", Journal of Clinical Epidemiology, vol. 53, pp. 207-216.

Tsang, S., Duncan, G.E., Dinescu, D. & Turkheimer, E. 2018, "Differential models of twin correlations in skew for body-mass index (BMI)", PLoS One, vol. 13, no. e0194968.

Yokoyama, Y., Jelenkovic, A., Hur, Y.M., Sund, R., Fagnani, C., Stazi, M.A., Brescianini, S., Ji, F., Ning, F., Pang, Z., Knafo-Noam, A., Mankuta, D., Abramson, L. et al. 2018, "Genetic and environmental factors affecting birth size variation: a pooled individual-based analysis of secular trends and global geographical differences using 26 twin cohorts", Int J Epidemiol, vol. 47, pp. 1195-1206.

Yokoyama, Y., Jelenkovic, A., Sund, R., Sung, J., Hopper, J.L., Ooki, S., Heikkilä, K., Aaltonen, S., Tarnoki, A.D., Tarnoki, D.L. et al. 2016, "Twin's birth-order differences in height and body mass index from birth to old age: A pooled study of 26 twin cohorts participating in the CODATwins project", Twin Res Hum Genet, vol. 19, pp. 112-124.

## Figure legends

Table 1. The current status of the CODATwins database by country.

| Country | Twin individuals[1] | | Complete twin pairs[1] | | | Number of measures[2] | | Number of twin individuals[3] | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | N | % of females | N | % of MZ | % of OSDZ | BMI | Smoking status | Birth weight | Maternal education | Own education |
| Australia[4] | 26435 | 61 | 12116 | 47 | 22 | 58586 | 35578 | 2354 | 2179 | 23403 |
| Belgium[5] | 803 | 52 | 379 | 64 | 11 | 803 | 803 | 803 | 747 | 803 |
| Brazil[6] | 211 | 76 | 97 | 79 | 12 | 211 | NA | NA | NA | 206 |
| Canada[7] | 2792 | 59 | 1387 | 46 | 21 | 7441 | NA | 1171 | NA | NA |
| China[8] | 22531 | 50 | 11264 | 54 | 20 | 22531 | 1006 | 19738 | 943 | 1006 |
| Denmark[9] | 65459 | 52 | 27261 | 27 | 33 | 108790 | NA | NA | NA | NA |
| Finland[10] | 40224 | 51 | 18735 | 29 | 17 | 110118 | 79965 | 10038 | 7621 | 36611 |
| Germany[11] | 2833 | 72 | 1395 | 66 | 10 | 2941 | 478 | NA | NA | 1793 |
| G-Bissau[12] | 253 | 53 | 108 | 15 | 58 | 1042 | NA | 253 | NA | NA |
| Hungary[13] | 825 | 66 | 387 | 59 | 13 | 825 | 805 | 259 | NA | 303 |
| Israel[14] | 995 | 49 | 489 | 23 | 35 | 1228 | NA | 818 | 745 | NA |
| Italy[15] | 17361 | 56 | 8630 | 44 | 25 | 18834 | 5651 | 1904 | 5767 | 11417 |
| Japan[16] | 9994 | 52 | 4932 | 54 | 18 | 52342 | 1023 | 9276 | 4900 | 431 |
| Mongolia[17] | 164 | 50 | 82 | 43 | 22 | 164 | NA | 138 | NA | NA |
| Netherlands[18] | 44169 | 53 | 22023 | 37 | 31 | 157287 | NA | 33779 | 34338 | 5873 |
| Norway[19] | 13941 | 53 | 5254 | 46 | 0 | 20188 | 20004 | 7900 | 8327 | 9228 |
| Portugal[20] | 432 | 51 | 216 | 40 | 26 | 432 | NA | NA | 430 | NA |
| S-Korea[21] | 4513 | 55 | 2248 | 59 | 18 | 5862 | 2649 | 779 | 3203 | 1346 |
| Spain[22] | 2258 | 57 | 1000 | 35 | 27 | 4392 | 4358 | 588 | NA | 2222 |
| Sri Lanka[23] | 2485 | 56 | 933 | 45 | 27 | 2485 | NA | NA | NA | 2469 |
| Sweden[24] | 74709 | 51 | 33768 | 33 | 19 | 185454 | 73828 | 5769 | 5310 | 29921 |
| Turkey[25] | 584 | 46 | 288 | 37 | 27 | 584 | 584 | NA | 584 | NA |
| UK[26] | 32545 | 61 | 15865 | 39 | 24 | 112198 | NA | 21262 | 3495 | NA |
| USA[27] | 123465 | 43 | 59778 | 42 | 19 | 175047 | 45396 | 31430 | 70316 | 77014 |
| Total | 489981 | 51 | 228635 | 39 | 22 | 1049785 | 272128 | 148259 | 148905 | 204046 |

MZ = monozygotic twins; OSDZ = opposite-sex dizygotic twins; BMI = body mass index; NA = not available. [1]Only twins including information on height, weight, age at the time of measurement, sex and zygosity are included; [2]Longitudinal measures available for part of twins; [3]Only one measure available for each twin; [4]Australian Twin Registry, Peri/Postnatal Epigenetic Twins Study (PETS), Queensland Twin Register; [5]East Flanders Prospective Twin Survey; [6]Brazilian Twin Registry; [7]Quebec Newborn Twin Study, University of British Columbia Twin Project; [8]Chinese National Twin Cohort Study, Guangzhou Twin Eye Study, Qingdao Twin Registry (Children), Qingdao Twin Registry (Adults); [9]Danish Twin Cohort; [10]Finnish Older Twin Cohort, FinnTwin12, FinnTwin16; [11]Berlin Twin Register Health (TwiSt), Bielefeld Longitudinal Study of Adult Twins; [12]Guinea Bissau Twin Study; [13]Hungarian Twin Registry; [14]Longitudinal Israeli Study of Twins; [15]Italian Twin Registry; [16]Japanese Twin Registry, Ochanomizu University Twin Project, Osaka University Aged Twin Registry, West Japan Twins and Higher Order Multiple Births Registry; [17]Mongolian Twin Registry; [18]Netherlands Twin cohort (children), Netherlands Twin cohort (adults); [19]Norwegian Twin Registry; [20]Portugal Twin Cohort, Madeira Twin Family Study; [21]Korean Twin-Family Register, South Korea Twin Registry; [22]Murcia Twin Registry; [23]Sri Lanka Twin Registry; [24]Child and Adolescent Twin Study in Sweden (CATSS), TCHAD-Study, Swedish Twin Cohorts, Swedish Young Male Twins Study; [25]Turkish Twin Study; [26]Gemini Study, Genesis 12-19 study, Twins Early Developmental Study, TwinsUK; [27]Boston University Twin Project, California Twin Program, Carolina African American Twin Study of Aging, Colorado Twin Registry, Michigan Twins Study, Mid-Atlantic Twin Registry, Minnesota Twin Family Study, Minnesota Twin Registry, NAS-NRC Twin Cohort, SRI-International, Texas Twin Project, University of Southern California Twin Study, University of Washington Twin Registry, Vietnam Era Twin Registry

Figure 1. Geographic distribution of the CODATwins collaborators.