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

Karri Silventoinen,1,2 Aline Jelenkovic,1,3 Reijo Sund,1 Chika Honda,2 Sari Aaltonen,1,4 Yoshie Yokoyama,5 Adam D. Tarnoki,6 David L. Tarnoki,6 Feng Ning,7 Fuling Ji,7 Zengchang Pang,7 Juan R. Ordoñana,8,9 Juan F. Sánchez-Romera,9,10 Lucia Colodro-Conde,8,9,11 S. Alexandra Burt,12 Kelly L. Klump,12 Sarah E. Medland,11 Grant W. Montgomery,11 Christian Kandler,13 Tom A. McAdams,14 Thalia C. Eley,14 Alice M. Gregory,15 Kimberly J. Saudino,16 Lise Dubois,17 Michel Boivin,18 Claire M. A. Haworth,19 Robert Plomin,14 Sevgi Y. Öncel,20 Razieh Alavi,20,21 Richard Mortensen,24,25 Richard Saffery,24,25 Massimo Mangino,30 Genevieve Lachance,30 Margaret Gatz,31 David A. Butler,32 Gombojav Bayasgalan,33 Danshiitsoodol Narandalai,33,34 Duarte L. Freitas,35 José Antonio Maia,36

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

Karri Silventoinen,1,2 Aline Jelenkovic,1,3 Reijo Sund,1 Chika Honda,2 Sari Aaltonen,1,4 Yoshie Yokoyama,5 Adam D. Tarnoki,6 David L. Tarnoki,6 Feng Ning,7 Fuling Ji,7 Zengchang Pang,7 Juan R. Ordoñana,8,9 Juan F. Sánchez-Romera,9,10 Lucia Colodro-Conde,8,9,11 S. Alexandra Burt,12 Kelly L. Klump,12 Sarah E. Medland,11 Grant W. Montgomery,11 Christian Kandler,13 Tom A. McAdams,14 Thalia C. Eley,14 Alice M. Gregory,15 Kimberly J. Saudino,16 Lise Dubois,17 Michel Boivin,18 Claire M. A. Haworth,19 Robert Plomin,14 Sevgi Y. Öncel,20 Razieh Alavi,20,21 Richard Mortensen,24,25 Richard Saffery,24,25 Massimo Mangino,30 Genevieve Lachance,30 Margaret Gatz,31 David A. Butler,32 Gombojav Bayasgalan,33 Danshiitsoodol Narandalai,33,34 Duarte L. Freitas,35 José Antonio Maia,36

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

Karri Silventoinen,1,2 Aline Jelenkovic,1,3 Reijo Sund,1 Chika Honda,2 Sari Aaltonen,1,4 Yoshie Yokoyama,5 Adam D. Tarnoki,6 David L. Tarnoki,6 Feng Ning,7 Fuling Ji,7 Zengchang Pang,7 Juan R. Ordoñana,8,9 Juan F. Sánchez-Romera,9,10 Lucia Colodro-Conde,8,9,11 S. Alexandra Burt,12 Kelly L. Klump,12 Sarah E. Medland,11 Grant W. Montgomery,11 Christian Kandler,13 Tom A. McAdams,14 Thalia C. Eley,14 Alice M. Gregory,15 Kimberly J. Saudino,16 Lise Dubois,17 Michel Boivin,18 Claire M. A. Haworth,19 Robert Plomin,14 Sevgi Y. Öncel,20 Razieh Alavi,20,21 Richard Mortensen,24,25 Richard Saffery,24,25 Massimo Mangino,30 Genevieve Lachance,30 Margaret Gatz,31 David A. Butler,32 Gombojav Bayasgalan,33 Danshiitsoodol Narandalai,33,34 Duarte L. Freitas,35 José Antonio Maia,36

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

Karri Silventoinen,1,2 Aline Jelenkovic,1,3 Reijo Sund,1 Chika Honda,2 Sari Aaltonen,1,4 Yoshie Yokoyama,5 Adam D. Tarnoki,6 David L. Tarnoki,6 Feng Ning,7 Fuling Ji,7 Zengchang Pang,7 Juan R. Ordoñana,8,9 Juan F. Sánchez-Romera,9,10 Lucia Colodro-Conde,8,9,11 S. Alexandra Burt,12 Kelly L. Klump,12 Sarah E. Medland,11 Grant W. Montgomery,11 Christian Kandler,13 Tom A. McAdams,14 Thalia C. Eley,14 Alice M. Gregory,15 Kimberly J. Saudino,16 Lise Dubois,17 Michel Boivin,18 Claire M. A. Haworth,19 Robert Plomin,14 Sevgi Y. Öncel,20 Razieh Alavi,20,21 Richard Mortensen,24,25 Richard Saffery,24,25 Massimo Mangino,30 Genevieve Lachance,30 Margaret Gatz,31 David A. Butler,32 Gombojav Bayasgalan,33 Danshiitsoodol Narandalai,33,34 Duarte L. Freitas,35 José Antonio Maia,36
The genetics of human anthropometric traits has long attracted scientific interest. Height is a prototypical anthropometric phenotype because it is approximately normally distributed and does not change in adulthood except for slight shrinking in old age. By the late 19th century, Galton (1886) analyzed height of parents and offspring and inferred that ‘when dealing with the transmission of stature from parents to children, the average height of the two parents is all we need care to know about them’. Later, Pearson and Lee (1903) presented correlations of height between relatives, also suggesting genetic influence. The first heritability estimate of height was calculated by Fisher (1918) in his seminal paper presenting the statistical principles of quantitative genetics. Interest in the genetic influences on height was re-}

The genetics of human anthropometric traits has long attracted scientific interest. Height is a prototypical anthropometric phenotype because it is approximately normally distributed and does not change in adulthood except for slight shrinking in old age. By the late 19th century, Galton (1886) analyzed height of parents and offspring and inferred that ‘when dealing with the transmission of stature from parents to children, the average height of the two parents is all we need care to know about them’. Later, Pearson and Lee (1903) presented correlations of height between relatives, also suggesting genetic influence. The first heritability estimate of height was calculated by Fisher (1918) in his seminal paper presenting the statistical principles of quantitative genetics. Interest in the genetic influences on height was renewed when genetic linkage studies enabled research into genetic effects over the whole genome on quantitative traits (Perola et al., 2007). Later genome-wide association (GWA) studies allowed for the genome-wide identification of candidate genes. In 2010, a large scale GWA study identified 180 loci associated for height (Lango Allen et al., 2010), and...

For over 100 years, the genetics of human anthropometric traits has attracted scientific interest. In particular, height and body mass index (BMI, calculated as kg/m²) have been under intensive genetic research. However, it is still largely unknown whether and how heritability estimates vary between human populations. Opportunities to address this question have increased recently because of the establishment of many new twin cohorts and the increasing accumulation of data in established twin cohorts. We started a new research project to analyze systematically (1) the variation of heritability estimates of height, BMI and their trajectories over the life course between birth cohorts, ethnicities and countries, and (2) to study the effects of birth-related factors, education and smoking on these anthropometric traits and whether these effects vary between twin cohorts. We identified 67 twin projects, including both monozygotic (MZ) and dizygotic (DZ) twins, using various sources. We asked for individual level data on height and weight including repeated measurements, birth related traits, background variables, education and smoking. By the end of 2014, 48 projects participated. Together, we have 893,458 height and weight measures (52% females) from 434,723 twin individuals, including 201,192 complete twin pairs (40% monozygotic, 40% same-sex dizygotic and 20% opposite-sex dizygotic) representing 22 countries. This project demonstrates that large-scale international twin studies are feasible and can promote the use of existing data for novel research purposes.

**Keywords**: twins, height, BMI, heritability, international comparisons

---

**Received**: 30 March 2015; **Accepted**: 20 April 2015. First published online 27 May 2015.

**Address for Correspondence**: Karri Silventoinen, Population Research Unit, Department of Social Research, University of Helsinki, P.O. Box 18, FIN-00014 University of Helsinki, Finland. E-mail: karri.silventoinen@helsinki.fi
since then several large GWA studies have been published focusing on height on populations of European (Weedon et al., 2008), Asian (Cho et al., 2009, Hao et al., 2013; Okada et al., 2010), and African ancestry (N’Diaye et al., 2011). The latest GWA study for height published in 2014 found 697 genetic polymorphisms associated with height in populations of European ancestry (Wood et al., 2014). As a polygenic and normally distributed trait, height serves also to explore new methodological approaches to human genetics, such as assumption-free estimation of heritability from genome-wide identity-by-descent sharing between full siblings (Hemani et al., 2013; Visscher et al., 2006).

Genetic studies of obesity and BMI (calculated as kg/m²) also have a long history. In an article published in 1923, Davenport showed that the tendency for obesity varies between families, and he interpreted this finding to suggest genetic effects on obesity (Davenport, 1923). After this initial paper, the evidence on the genetic effects on obesity accumulated, and in 1966 a review paper on previous family studies concluded that genetic factors played an important role in obesity (Seltzer & Mayer, 1966). After this review, interest in the genetics of BMI has rapidly increased because of the health consequences and related impact on public health of increased mean BMI over the world. The studies by Stunkard and colleagues demonstrating the importance of genetic factors underlying variation in BMI in studies based on twin (Stunkard et al., 1986a) and adoption data (Stunkard et al., 1986b; 1990) were a major achievement in this area. These findings corroborated earlier results reported on Finnish twins reared apart (Langvinainio et al., 1984). In 2007, the FTO gene was found to be associated with obesity in a case-control study of type 2 diabetes (Frayling et al., 2007), and it is now recognized to be the most promising candidate gene of obesity. The latest GWA study on BMI identified 97 loci explaining 2.7% of the variation of BMI, while all measurable variants accounted for around 20% of the variance (Locke et al., 2015).

After over a hundred years of research, we might assume that the heritability of height and BMI is already well known. However, surprisingly little research is available on the variation of heritability estimates of height and BMI between populations. Changes in mean height (Eveleth & Tanner, 2003) and BMI (Finucane et al., 2011) over time and changes in BMI across the human life span (Dahl et al., 2014) have been reported. According to basic principles of quantitative genetics, heritability estimates are not constant but rather are statistics describing the magnitude of genetic variation in a particular population and dependent on the underlying genetic make-up of the population under study and the environmental variation at play. Accordingly, these estimates may change over the life course and vary between study populations. A meta-analysis of nine twin studies found that the heritability of BMI increased over childhood and the effect of common environmental factors disappeared after mid-childhood (Silventoinen et al., 2010). Increasing heritability of height and BMI after early childhood was also found in a study of four twin cohorts (Dubois et al., 2012). However, these two studies did not reveal systematic variation in the heritability estimates between populations. A meta-analysis based on 88 independent heritability estimates of BMI reported inter-study variation in the heritability estimates, but meta-regression did not reveal any systematic patterns behind these differences (Elks et al., 2012). It is possible that this negative result was due to methodological limitations since many of the heritability estimates were based on data covering large age ranges, birth cohorts and social classes, and the authors did not have access to the original data. Twin studies for adult height (Silventoinen et al., 2003) and BMI (Schousboe et al., 2003) in seven European populations and Australia also found some variation in heritability estimates but were not able to find systematic patterns in these estimates. A study based on eight populations of adolescent twins found higher genetic variance of height and weight in Caucasian as compared to East Asian populations; however, because total variance for height and BMI was also higher in Caucasian populations, the heritability estimates were approximately equivalent (Hur et al., 2008). Thus, the previous meta-analyses have demonstrated the variation in the genetic components of height and BMI but have largely failed to identify factors behind the variation between populations.

The scant evidence on the variation of genetic and environmental contributions on height and BMI between populations may, however, reflect methodological limitations of previous studies rather than the lack of this type of variation. Previous studies conducted in Denmark (Rokholm et al., 2011a) and Sweden (Rokholm et al., 2011b) have demonstrated that genetic variation of BMI has increased over time in birth cohorts as the mean BMI increased; however, heritability estimates did not change. A Finnish study reported that environmental variation of height decreased especially in women from cohorts born at the beginning of the 20th century compared to those born after the World War II, leading to higher heritability estimates of height (Silventoinen et al., 2000). There is also evidence that parental social position may modify the genetic architecture of BMI in childhood (Lajunen et al., 2012). International comparisons addressing the methodological limitations of previous studies may be able to demonstrate comparable variation in genetic and environmental effects between populations.

During the recent decade, possibilities for international comparisons in twin studies have improved because of the establishment of new twin cohorts and the increasing accumulation of data in established twin cohorts. Thus, the number of twins available internationally for research has greatly increased, expanding the ability to examine ethnic, economic, and cultural variation between twin cohorts. These new opportunities to answer research questions not possible to address before led to the start of a new international research project: COllaborative project...
of Development of Anthropometrical measures in Twins (CODATwins). The aims of this project are to analyze systematically: (1) the variation of heritability estimates of height, BMI and their trajectories over the life course between birth cohorts, ethnicities, and countries; and (2) to study the effects of birth-related factors, education, and smoking on these anthropometric traits and whether these effects vary between twin cohorts. Additionally, this project aims to gain practical knowledge on the feasibility and opportunities offered by pooling a large number of twin cohorts as suggested by the International Network of Twin Registries (INTR) consortium (Buchwald et al., 2014).

**Collection of a Collaborative Database**

We started the CODATwins project in May 2013 by identifying all twin projects in the world. The only criterion was the availability of data from both MZ and DZ twin pairs. The main sources used to identify the projects were a special issue of Twin Research and Human Genetics (Hur & Craig, 2013) and the participants of the INTR consortium (Buchwald et al., 2014, van Dongen et al., 2012); these sources were complemented by personal communications. Together we identified 67 eligible twin projects. We sent e-mail invitations to principle investigators of all these projects in September 2013 along with the study protocol. We asked the investigators to send us individual level data on height and weight including repeated measurements, birth-related traits (birth weight, birth length, birth order, and gestational age), background variables (twin identifier, sex, zygosity, ethnicity, birth year, and age at the time of measurements), education (own education for adults and mother’s and father’s education for children) and smoking for adults to the CODATwins data management center at the University of Helsinki. To those who did not respond, we sent reminders in October 2013, January 2014 and September 2014; with the final reminder, we sent the first year progress report including the list of all twin projects already collaborating with this project.

We did not receive a response from eight projects; internet searches (PubMed and Google) indicated that these projects had not been active in recent years and some of them may not even have ever been established. Eight projects declined: two because of lack of height and weight data, one because of lack of information on zygosity, and four because the delivery of the data was not possible to organize due to local regulations. One project informed that they are currently publishing their own results, but the data may become available later when the original articles have been published. Three projects that initially accepted the invitation have not sent data. Based on the correspondence, the main reason was the lack of resources to prepare the data file. By the end of 2014, 47 projects had sent data to the data management center. Additionally, one cohort is available through the remote access system but is not part of the pooled database. Figure 1 describes the accumulation of the CODATwins database.

**Structure of Database**

Table 1 presents the twin cohorts participating in the CODATwins project. Because one twin project can include several cohorts, there are 54 twin cohorts available representing 22 countries. From these cohorts, 35 are longitudinal. Figure 2 presents the number of height and weight measures by sex and age. Together there are 893,458 measures. Children are well represented, and 41% of the measures were conducted at 18 years of age or younger. Overall, about half of the measures are for females (52%); however, the cohorts vary considerably regarding the proportion of their samples that are females and some cohorts include only males while others mainly include females (Supplementary Table 1). Most of the height and weight measures were self-reported (63%) or parentally reported (21%) and only a minority was based on measured values (16%). The reason is that data in the largest cohorts were collected by questionnaires, and the collection of clinical measures was generally conducted in cohorts smaller in size. In 27 cohorts we had additional information on birth weight and in most of these cohorts also had data on birth length (Supplementary Table 1). Together, we have 122,321 birth weight measures in the database; 77% of these measures were parentally reported, 17% self-reported and 6% clinically measured.

In total, data are available for 434,723 twin individuals having at least one height and weight measure. Most of the twins are from Europe (60%) and North-America (30%), followed by Australia (6%), East Asia (3%), South Asia, and the Middle East (1%) and Africa (less than 0.1%); no twin cohort is available from Latin America. Figure 3 presents the number of complete twin pairs by birth year and zygosity. Together there are 201,192 complete twin pairs. Among these pairs, 40% are MZ twins, 40% same-sex DZ twins, and 20% opposite-sex DZ twins. A quarter of the twin pairs...
### TABLE 1
Number of Height and Weight Measures in the Twin Cohorts Participating in the CODATwins Project

<table>
<thead>
<tr>
<th>Cohort name</th>
<th>Main reference</th>
<th>Region</th>
<th>Number of height and weight measures</th>
<th>Number of longitudinal surveys</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Africa</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Guinea-Bissau Twin Study</td>
<td>(Bjerregaard-Andersen et al., 2013)</td>
<td>Guinea-Bissau</td>
<td>1,042</td>
<td>7</td>
</tr>
<tr>
<td><strong>Australia</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Australian Twin Registry</td>
<td>(Hopper et al., 2013)</td>
<td>Australia</td>
<td>2,536</td>
<td>1</td>
</tr>
<tr>
<td>Peri/Postnatal Epigenetic Twins Study (PETS)</td>
<td>(Loke et al., 2013)</td>
<td>Australia, Melbourne</td>
<td>571</td>
<td>2</td>
</tr>
<tr>
<td>Queensland Twin Register</td>
<td>(Liu et al., 2010)</td>
<td>Australia, Queensland province</td>
<td>55,479 &gt; 10</td>
<td></td>
</tr>
<tr>
<td><strong>East-Asia</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Guangzhou Twin Eye Study</td>
<td>(Zheng et al., 2013)</td>
<td>China, Guangzhou province</td>
<td>1,122</td>
<td>1</td>
</tr>
<tr>
<td>Japanese Twin Cohort</td>
<td>(Ooki, 2013)</td>
<td>Japan</td>
<td>34,405 &gt; 10</td>
<td></td>
</tr>
<tr>
<td>Mongolian Twin Registry</td>
<td>(Gombojav et al., 2013a)</td>
<td>Mongolia</td>
<td>166</td>
<td>1</td>
</tr>
<tr>
<td>Osaka University Aged Twin Registry</td>
<td>(Hayakawa et al., 2013)</td>
<td>Japan</td>
<td>1,289</td>
<td>4</td>
</tr>
<tr>
<td>South Korea Twin Registry</td>
<td>(Hur et al., 2013)</td>
<td>South Korea</td>
<td>2,278</td>
<td>1</td>
</tr>
<tr>
<td>Qingdao Twin Registry (adults)</td>
<td>(Duan et al., 2013)</td>
<td>China, Shandong province</td>
<td>986</td>
<td>1</td>
</tr>
<tr>
<td>Qingdao Twin Registry (children)</td>
<td>(Duan et al., 2013)</td>
<td>China, Shandong province</td>
<td>1,175</td>
<td>1</td>
</tr>
<tr>
<td>West Japan Twins and Higher Order Multiple Births Registry</td>
<td>(Yokoyama, 2013)</td>
<td>Japan</td>
<td>7,617 &gt; 10</td>
<td></td>
</tr>
<tr>
<td><strong>Europe</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adult Netherlands Twin Registry</td>
<td>(Willemsen et al., 2013)</td>
<td>Netherlands</td>
<td>37,638 &gt; 10</td>
<td></td>
</tr>
<tr>
<td>Berlin Twin Register</td>
<td>(Busjahn, 2013)</td>
<td>German, Berlin city</td>
<td>722</td>
<td>5</td>
</tr>
<tr>
<td>Bielefeld Longitudinal Study of Adult Twins</td>
<td>(Kandler et al., 2013)</td>
<td>German</td>
<td>2,366</td>
<td>1</td>
</tr>
<tr>
<td>Danish Twin Cohort</td>
<td>(Skythe et al., 2013)</td>
<td>Denmark</td>
<td>34,665</td>
<td>1</td>
</tr>
<tr>
<td>East Flanders Prospective Twin Survey</td>
<td>(Derom et al., 2013)</td>
<td>Belgium, East Flanders Province</td>
<td>803</td>
<td>1</td>
</tr>
<tr>
<td>Finnish Older Twin Cohort</td>
<td>(Kaprio, 2013)</td>
<td>Finland</td>
<td>68,683</td>
<td>4</td>
</tr>
<tr>
<td>FinnTwin12</td>
<td>(Kaprio, 2013)</td>
<td>Finland</td>
<td>16,211</td>
<td>4</td>
</tr>
<tr>
<td>FinnTwin16</td>
<td>(Kaprio, 2013)</td>
<td>Finland</td>
<td>24,438</td>
<td>5</td>
</tr>
<tr>
<td>Gemini Study</td>
<td>(van Jaarsveld et al., 2010)</td>
<td>UK</td>
<td>19,639 &gt; 10</td>
<td></td>
</tr>
<tr>
<td>Hungarian Twin Registry</td>
<td>(Littvay et al., 2013)</td>
<td>Hungary</td>
<td>825</td>
<td>1</td>
</tr>
<tr>
<td>Italian Twin Registry</td>
<td>(Brescianini et al., 2013)</td>
<td>Italy</td>
<td>18,834</td>
<td>5</td>
</tr>
<tr>
<td>Murcia Twin Registry</td>
<td>(Ordoñana et al., 2013)</td>
<td>Spain, Region of Murcia</td>
<td>4,392</td>
<td>3</td>
</tr>
<tr>
<td>Norwegian Twin Registry</td>
<td>(Nilsen et al., 2013)</td>
<td>Norway</td>
<td>20,188</td>
<td>2</td>
</tr>
<tr>
<td>Portuguese Twin Cohort</td>
<td>(Maia et al., 2013)</td>
<td>Portugal, North of mainland, Acores, and Madeira Islands</td>
<td>7,189</td>
<td>1</td>
</tr>
<tr>
<td>Swedish Twin Cohorts</td>
<td>(Magnusson et al., 2013)</td>
<td>Sweden</td>
<td>110,117</td>
<td>3</td>
</tr>
<tr>
<td>Swedish Young Male Twins Study (adults)</td>
<td>(Rasmussen et al., 2006)</td>
<td>Sweden</td>
<td>5,702</td>
<td>3</td>
</tr>
<tr>
<td>Swedish Young Male Twins Study (children)</td>
<td>(Rasmussen et al., 2006)</td>
<td>Sweden</td>
<td>10,440 &gt; 10</td>
<td></td>
</tr>
<tr>
<td>Twins Early Development Study (TEDS)</td>
<td>(Haworth et al., 2013)</td>
<td>UK</td>
<td>59,108</td>
<td>7</td>
</tr>
<tr>
<td>TwinsUK</td>
<td>(Maayyeri et al., 2013)</td>
<td>UK</td>
<td>31,321</td>
<td>7</td>
</tr>
<tr>
<td>Young Netherlands Twin Registry</td>
<td>(van Beijsterveldt et al., 2013)</td>
<td>Netherlands</td>
<td>119,649</td>
<td>10</td>
</tr>
<tr>
<td><strong>South Asia and Middle East</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Longitudinal Israeli Study of Twins</td>
<td>(Avinun &amp; Knafo, 2013)</td>
<td>Israel</td>
<td>1,228</td>
<td>2</td>
</tr>
<tr>
<td>Sri Lanka Twin Registry</td>
<td>(Sumathipala et al., 2013)</td>
<td>Sri Lanka</td>
<td>2,485</td>
<td>1</td>
</tr>
<tr>
<td>Turkish Twin Study</td>
<td>(Oncel &amp; Aliev, 2013)</td>
<td>Turkey</td>
<td>584</td>
<td>1</td>
</tr>
<tr>
<td><strong>North America</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boston University Twin Project</td>
<td>(Saudino &amp; Asherson, 2013)</td>
<td>USA, Massachusetts</td>
<td>1,228</td>
<td>2</td>
</tr>
<tr>
<td>California Twin Program</td>
<td>(Cosen et al., 2013)</td>
<td>USA, California</td>
<td>27,237</td>
<td>1</td>
</tr>
<tr>
<td>Colorado Twin Registry</td>
<td>(Rhea et al., 2013)</td>
<td>USA, Colorado</td>
<td>8,671</td>
<td>5</td>
</tr>
<tr>
<td>Michigan State University Twin Registry</td>
<td>(Burt &amp; Klump, 2013)</td>
<td>USA, Michigan</td>
<td>22,172</td>
<td>2</td>
</tr>
<tr>
<td>Mid Atlantic Twin Registry</td>
<td>(Lilley &amp; Silberg, 2013)</td>
<td>USA, Virginia, North Carolina, South Carolina</td>
<td>11,801</td>
<td>1</td>
</tr>
<tr>
<td>Minnesota Twin Family Study</td>
<td>(Iacono &amp; McGue, 2002)</td>
<td>USA, Minnesota</td>
<td>3,269</td>
<td>3</td>
</tr>
<tr>
<td>Minnesota Twin Registry</td>
<td>(Krueger &amp; Johnson, 2002)</td>
<td>USA, Minnesota</td>
<td>10,122</td>
<td>1</td>
</tr>
<tr>
<td>NASA-NRC Twin Registry</td>
<td>(Gatz et al., 2004)</td>
<td>USA, WWII veterans</td>
<td>54,904</td>
<td>4</td>
</tr>
<tr>
<td>Quebec Newborn Twin Study</td>
<td>(Boivin et al., 2013)</td>
<td>Canada, Greater Montreal area</td>
<td>5,991</td>
<td>9</td>
</tr>
<tr>
<td>SRI-international</td>
<td>(Krasnow et al., 2013)</td>
<td>USA, California</td>
<td>1,092</td>
<td>1</td>
</tr>
<tr>
<td>Texas Twin Project</td>
<td>(Harden et al., 2013)</td>
<td>USA, Texas</td>
<td>565</td>
<td>1</td>
</tr>
<tr>
<td>University of British Columbia Twin Project</td>
<td>(Jang, 2013)</td>
<td>Canada, Greater Vancouver area</td>
<td>2,420</td>
<td>1</td>
</tr>
<tr>
<td>University of Southern California Twin Study</td>
<td>(Baker et al., 2013)</td>
<td>USA, Greater Los Angeles area</td>
<td>3,622</td>
<td>5</td>
</tr>
<tr>
<td>University of Washington Twin Registry</td>
<td>(Strachan et al., 2013)</td>
<td>USA, Washington State</td>
<td>27,452</td>
<td>3</td>
</tr>
<tr>
<td>Vietnam Era Twin Study of Aging</td>
<td>(Kremen et al., 2013)</td>
<td>USA, Vietnam era veterans</td>
<td>2,245</td>
<td>2</td>
</tr>
</tbody>
</table>
**FIGURE 2**
Number of height and weight measures by sex and age.

**FIGURE 3**
Number of complete twin pairs by birth year and zygosity.
(25%) were born in the 1980s and 1990s. The numbers of twin individuals and complete twin pairs by cohorts are presented in Supplementary Table 1.

Discussion

We have successfully launched a large international twin collaboration, and our database now includes slightly over 200,000 complete twin pairs with height and weight measures from 22 countries. The vast majority of established twin cohorts responded positively to our request for individual level data. For some of the cohorts who did not participate, the reason was the lack of suitable data or that the cohort was no longer active. The value of pooling either summary data in GWA studies for height (Wood et al., 2014) and BMI (Locke et al., 2015) or pooling individual data for psychiatric conditions (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014) is well recognized. This project demonstrates that the same strategy can be used in classical twin research as well.

However, this project also revealed certain limitations with respect to available twin data. While European countries, especially in the northern and western parts of Europe, North America, and Australia are well represented, there is much less data on twins from other parts of the world. Our final database is heavily weighted toward European-origin populations following the Westernized lifestyle. The exception is East Asia, with several twin cohorts available from China, Japan, and South Korea and one from Mongolia. Even though many of these non-Western cohorts are not very large, these cohorts do provide an invaluable resource for studying the potential genetic variations in anthropometric phenotypes. It was unfortunate that there are few twin cohorts from Southern Asia, Africa and all of South America. As pointed out earlier, there is a real need and value to the creation of new twin cohorts in the developing world (Sung et al., 2006). Increasing collaboration between established twin projects can be helpful to stimulate new research activity and starting new twin projects (Buchwald et al., 2014).

In addition to the lack of representation of specific ethnic groups among the registry populations included, another limitation is that the populations represented are relatively affluent populations. Of the four countries officially classified as non-industrialized countries represented in this project, only Guinea-Bissau can be regarded as a real developing country. In contrast, China and Sri Lanka are moderately affluent societies and enjoy life expectancy nearly comparable to the United States, whereas Mongolia can be regarded as a middle-income country with life expectancy at the level of East European countries (Wang et al., 2012). Anthropometric data from twin pairs in diverse populations going through the demographic transition would be invaluable in understanding the influence of broad societal change on many phenotypes. However, it is noteworthy that we have substantial variation in birth cohorts; the oldest twins were born at the end of 19th century and around one-fifth of them before 1940. Major changes in the prevalence of obesity and standard of living during the 20th century allow for the testing of different hypotheses as demonstrated before for BMI in Denmark (Rokholm et al., 2011a) and Sweden (Rokholm et al., 2011b) and height in Finland (Silventoinen et al., 2003).

When considering further collaborative twin research projects, it is noteworthy that only 16% of the height and weight measures were based on clinical measure, whereas the majority was obtained by self- or parental report. Height and weight are some of few anthropometric traits possible to measure relatively reliably based on self-report. Data on even the most basic metabolic traits such as blood glucose, blood pressure, and blood lipids would require clinical assessments that are currently lacking in many twin samples. This shows that even when there are many large twin cohorts available, more data collection using clinical measures is still needed. Height and weight are widely available in twin cohorts, and there is also much less variation in the measurement protocols of these traits compared to other anthropometric traits, such as waist circumference, making harmonization straightforward; the biggest difference we found was the measurement units used for height (cm vs. foot and inch) and weight (kg and g vs. pound and ounce). However, it is noteworthy that even for height and weight there can be differences in the precision of equipment used for measuring weight (scale) and height (tape, anthropometer or stadiometer). When examining other traits, availability of the data and differences in measurement protocols will increase challenges to data harmonization.

In addition to the anthropometric traits, we collected information on own education, parental education, and smoking. After reporting the main results for the anthropometric indicators, we will move to study how they are modified by education and smoking. Working with these variables is much more challenging compared to the anthropometric traits because of different classifications, varying educational systems, and large differences in mean levels of education between countries and birth cohorts. However, this variation also presents an opportunity because it allows for the study of these associations in very different environments and, for example, to study the relevance of absolute and relative education. In these future analyses, we can rely on work done to harmonize these variables in other contexts, such as the OECD classification of educational level (oecd.org) and the P3G consortium (p3g.org). This effort also demonstrates the potential of international collaborations of twin projects beyond calculating heritability estimates. For example, there are 10,410 adult MZ twin pairs discordant for BMI (more than 3 kg/m²) at least at one time point when measured at the same age in the database. Previous studies have demonstrated the high value of BMI discordant pairs for epigenetic research (Pietiläinen et al., 2008).
In conclusion, the CODATwins project demonstrates that large-scale international studies obtaining individual-level data from twin cohorts are feasible. Using the data from these twin cohorts creates novel opportunities for examining how genetic and environmental influences may vary across countries and regions. Future efforts in the CODATwins project will continue to extract from the substantial data already collected in the various twin projects in order to contribute to this objective.

Acknowledgments

This study was conducted within the CODATwins project (Academy of Finland #266592). Support for participating twin projects: the University of Southern California Twin Study is funded by a grant from the National Institute of Mental Health (R01 MH58354). 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 NAS-NRC Twin Registry acknowledges financial support from the National Institutes of Health grant number R21 AG039572. 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 (RF2/RFG/2008/0145). Wave 5 was supported by funding to Alice M. Gregory from Goldsmiths, University of London. Anthropometric measurements of the Hungarian twins were supported by Medexpert Ltd., Budapest, Hungary. South Korea Twin Registry is supported by National Research Foundation of Korea (NRF-2017R1A2B4003964). The 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). Korean Twin-Family Register was supported by the Global Research Network Program of the National Research Foundation (NRF 2011-220-E00006). The Colorado Twin Registry is funded by NIDA funded center grant DA011015 and Longitudinal Twin Study HD10333; Author Huibregtsen is supported by 5T32DA017637-10. The 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 Australian Twin Registry is supported by a Centre of Research Excellence (grant ID 1079102) from the National Health and Medical Research Council administered by the University of Melbourne. 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 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 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 (1R01ES015150-01). The Guangzhou Twin Eye Study is supported by National Natural Science Foundation of China (grant #81125007). 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. 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). K. Silventoinen is supported by Osaka University’s International Joint Research Promotion Program. S. Y. Oncel and F. Aliev are supported by Kirikkale University Research Grant; KKU, 2009/43 and TUBITAK grant 114C117. The Longitudinal Israeli Study of Twins was funded by the Starting Grant no. 240994 from the European Research Council (ERC) to Ariel Knafo. Data collection and research stemming from the Norwegian Twin Registry is supported, in part, from the European Union’s Seventh Framework Programmes ENGAGE Consortium (grant agreement HEALTH-F4-2007-201413, and BioSHaRE EU (grant agreement HEALTH-F4-2010-261433). The Murcia Twin Registry is...
supported by the Seneca Foundation, Regional Agency for Science and Technology, Murcia, Spain (08633/PHCS/08 & 15302/PHCS/10) and Ministry of Science and Innovation, Spain (PSI11560-2009). The Twins Early Development Study (TEDS) is supported by a program grant (G9001245) 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). The 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). The Boston University Twin Project is funded by grants (#R01 HD068435 #R01 MH062375) from the National Institutes of Health to K. Saudino. TwinSUk was funded by the Wellcome Trust; European Community’s Seventh Framework Programme (FP7/2007-2013). The study also receives support from the National Institute for Health Research (NIHR) BioResource Clinical Research Facility and Biomedical Research Centre based at Guy’s and St Thomas’ NHS Foundation Trust and King’s College London. The University of Washington Twin Registry is supported by the grant NIH RC2 HL103416 (D. Buchwald, PI). The 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 Avrera Institute, Sioux Falls, South Dakota (USA). Gemini was supported by a grant from Cancer Research UK (C1418/A7974).

Supplementary Material

To view supplementary material for this article, please visit http://dx.doi.org/10.1017/thg.2015.29.

References

Dubois, L., Ohm Kyvik, K., Girard, M., Tatone-Tokuda, F., Perusse, D., Hjelmborg, J., ... 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, 7, 1–12.


