

## Article

# The Danish Twin Registry: An Updated Overview

Dorthe Almind Pedersen<sup>1,2</sup>, Lisbeth Aagaard Larsen<sup>1,2</sup>, Marianne Nygaard<sup>1,2</sup>, Jonas Mengel-From<sup>1,2</sup>, Matt McGue<sup>1,2,3</sup>, Christine Dalgård<sup>1,4</sup>, Lars Hvidberg<sup>1,2</sup>, Jacob Hjelmberg<sup>1,2</sup>, Axel Skytthe<sup>1,2</sup>, Niels V. Holm<sup>1,5</sup>, Kirsten Ohm Kyvik<sup>1,6,7</sup> and Kaare Christensen<sup>1,2,8,9</sup>

<sup>1</sup>The Danish Twin Registry, Department of Public Health, University of Southern Denmark, Odense, Denmark, <sup>2</sup>Epidemiology, Biostatistics and Biodemography, Department of Public Health, University of Southern Denmark, Odense, Denmark, <sup>3</sup>Department of Psychology, University of Minnesota, Minneapolis, Minnesota, USA, <sup>4</sup>Environmental Medicine, Department of Public Health, University of Southern Denmark, Odense, Denmark, <sup>5</sup>Department of Oncology, Odense University Hospital, Odense, Denmark, <sup>6</sup>Department of Clinical Research, University of Southern Denmark, Odense, Denmark, <sup>7</sup>Odense Patient data Explorative Network (OPEN), Odense University Hospital, Odense, Denmark and, <sup>8</sup>Department of Clinical Genetics, Odense University Hospital, Odense, Denmark and <sup>9</sup>Department of Clinical Biochemistry and Pharmacology, Odense University Hospital, Odense, Denmark

## Abstract

The Danish Twin Registry (DTR) was established in the 1950s, when twins born from 1870 to 1910 were ascertained, and has since been extended to include twins from birth cohorts until 2009. The DTR currently comprises of more than 175,000 twins from the 140 birth cohorts. This makes the DTR the oldest nationwide twin register and among the largest in the world. The combination of data from several surveys, including biological samples and repeated measurements on the same individuals, and data from Danish national registers provides a unique resource for a wide range of twin studies. This article provides an updated overview of the data in the DTR: First, we provide a summary of the establishment of the register, the different ascertainment methods and the twins included; then follows an overview of major surveys conducted in the DTR since 1994 and a description of the DTR biobank, including a description of the molecular data created so far; finally, a short description is given of the linkage to Danish national registers at Statistics Denmark and some recent examples of studies using the various data resources in the DTR are highlighted.

**Keywords:** Twin register; surveys; biobank; molecular; register-based research; Denmark

(Received 15 June 2019; accepted 1 July 2019; First Published online 23 September 2019)

## The Population in the Danish Twin Registry

Overall, four different ascertainment sources were used during the history of the Danish Twin Registry (DTR). These are described below. Detailed descriptions can be found in Hauge et al. (1968), Kyvik et al. (1995) and Skytthe et al. (2002).

### Birth Cohorts 1870–1930

The oldest twins in the DTR, born from 1870 to 1930, have been ascertained through an effort headed by Harvald and Hauge during the 1950s and 1960s. First, church records in every parish in Denmark were searched by the local clergy, who were asked to report twin births where both twins had survived to age 6 years. Both same-sex and opposite-sex twin pairs were registered for birth cohorts 1870–1910, whereas for birth cohorts 1911–1930, only same-sex twin pairs were registered. Subsequently, the twins were located by searching through population registers and other public sources.

### Birth Cohorts 1931–1952

The nationwide ascertainment of twins born from 1931 to 1952 was conducted from 1996 to 1998 by Skytthe and Christensen using the Danish Civil Registration System (CRS; Pedersen, 2011). Children born on the same date in the same parish with the same surname at birth were identified. A one-page questionnaire was sent to sets of children fulfilling these criteria to obtain confirmation of twin status. Approximately half of the responders were twins. To be selected using the CRS approach, the twins had to be alive on April 2, 1968, when the CRS was established. Several secondary sources were used for ascertainment, for example, same-sex male twin pairs and opposite-sex twin pairs had been identified through a conscription register if at least one male twin had been to conscription examination, and twins in the Copenhagen area were identified from the CRS by having the same date of birth and the same address.

### Birth Cohorts 1953–1982

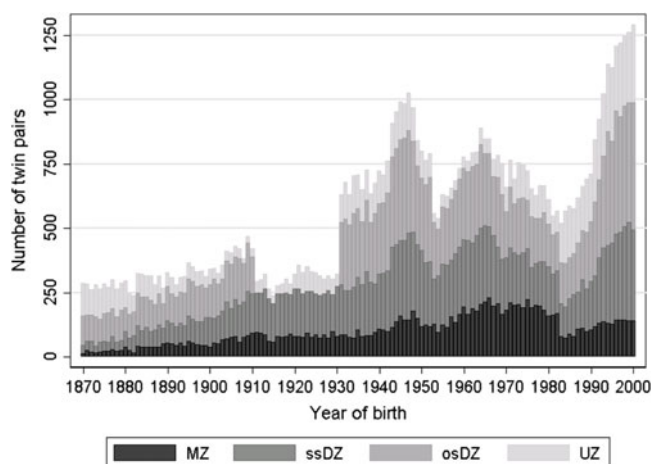
Twins born from 1953 to 1982 were ascertained at the beginning of the 1990s by Kyvik, mostly through links between parents and children in the CRS from 1953 and onwards. These links between mother and children were used to identify children born within three consecutive days by the same mother. In the early years of this period, the cohort was not complete. From April 1968 and

**Author for correspondence:** Kaare Christensen, Email: [kchristensen@health.sdu.dk](mailto:kchristensen@health.sdu.dk)

**Cite this article:** Pedersen DA, Larsen LA, Nygaard M, Mengel-From J, McGue M, Dalgård C, Hvidberg L, Hjelmberg J, Skytthe A, Holm NV, Kyvik KO, and Christensen K. (2019) The Danish Twin Registry: An Updated Overview. *Twin Research and Human Genetics* 22: 499–507, <https://doi.org/10.1017/thg.2019.72>

**Table 1.** Basic description of the population in the Danish Twin Registry

Birth cohorts	Number of twins, <i>N</i>	Males, <i>n</i> (%)	Females, <i>n</i> (%)	Number with known zygosity, <i>n</i> (%)	Number included in at least one major DTR survey, <i>n</i> (%)
1870–1910	26,474	12,938 (48.9)	13,536 (51.1)	20,638 (78.0)	787 (3.0)
1911–1930	12,300	5897 (47.9)	6403 (52.1)	9924 (80.7)	3944 (32.1)
1931–1952	34,492	18,759 (54.4)	15,553 (45.1)	29,106 (84.4)	15,171 (44.0)
1953–1967	21,796	11,399 (52.3)	10,384 (47.6)	20,346 (93.3)	16,016 (73.5)
1968–1982	20,418	10,415 (51.0)	9733 (47.7)	17,306 (84.8)	13,497 (66.1)
1983–2000	32,396	16,237 (50.1)	15,713 (48.5)	24,168 (74.6)	0 (0.0)
Total	147,876	75,645 (51.2)	71,322 (48.2)	121,488 (82.2)	49,419 (33.4)

**Fig. 1.** Basic description of the zygosity by year of birth for twin pairs in the Danish Twin Registry. *Note:* MZ = monozygotic, ssDZ = same-sex dizygotic, osDZ = opposite-sex dizygotic, UZ = unknown zygosity.

onwards, all twin pairs where both twins were born alive were ascertained, since all live-born children receive a unique personal identification number (CPR number) and hence are included in the CRS. It was not possible to identify all twins born from 1953 to 1968 using the CRS, especially not in the birth cohorts 1953–1960. Therefore, additional twins were identified through secondary sources (e.g., the conscription register). Furthermore, the Danish Medical Birth Registry (MBR; Bliddal *et al.*, 2018) was used to identify births of stillborn twins from 1973 to 1982.

### Birth Cohorts 1983–2009

The youngest twins in the DTR, born from 1983 to 2009, were identified by Kyvik and Skytthe using the MBR where all births including stillbirths since 1973 are registered and where multiple births are marked with a code. Information about the twins and their mother, including CPR numbers, were extracted from the MBR and contact

information was retrieved from the CRS. The DTR is, with varying intervals, updated with new birth cohorts.

### Completeness of Ascertainment

There is complete ascertainment of all live-born twins since the establishment of the CRS in April 1968 and complete ascertainment of all twins, including stillborn, since 1973 when the MBR was initiated (Bliddal *et al.*, 2018; Pedersen, 2011). For different reasons, there is lower coverage for the earlier birth cohorts: For birth cohorts 1870–1930, only twins who survived to age 6 years are registered and, furthermore, no opposite-sex twin pairs were registered for birth cohorts 1911–1930. For birth cohorts 1931–1952, where the CRS was used as the primary source of ascertainment, the inclusion in DTR is conditioned on survival until April 1968. For twins born from 1931 to 1952, about 70 % of all twins have been ascertained and 90 % of those surviving to age 6 years. For twins born from 1953 to 1960, there is a low coverage due to incomplete links between parents and children in the CRS. The subsequent use of the conscription register helped identify additional male/male and male/female twin pairs. Therefore, there is lower coverage of female twins than male twins. A more detailed description of the completeness of ascertainment in the DTR can be found in Skytthe *et al.* (2002, 2011).

### Zygosity Assessment

The zygosity of all same-sex twins in the DTR has been determined by a questionnaire that was mailed to twins as part of the intake survey of all four birth cohort periods. The questionnaires included four questions about the degree of similarity in appearance between the co-twins and classified the twins as either monozygotic (MZ) or dizygotic (DZ); the four questions can be found in Kyvik *et al.* (1995). Zygosity was classified as unknown (UZ) in case of disagreement between the two twins, or inconsistent answers or no answer from either twin. If only one twin answered, zygosity classification was based on this answer alone. For the youngest twins born from 1989 to 2000, questionnaires were sent to the parents of the twins.

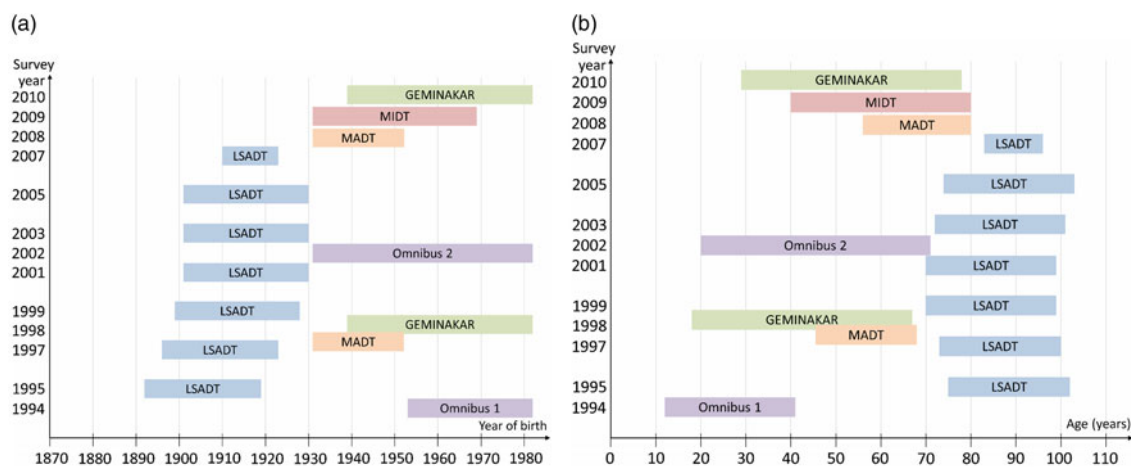
For some of the twins who participated in surveys where biological samples had been collected, zygosity was assessed using either serological or genetic markers. The questionnaire-based zygosity was found to be correct in 95 % of these cases (Christiansen *et al.*, 2003).

By May 2019, zygosity had been assessed for the twins born through 2000. Twins born from 2001 to 2009 have been identified, and zygosity assessments for these are planned in the end of 2019.

### Basic Description of the Twins in the DTR

By May 2019, the population in the DTR comprised of 175,518 individuals: 86,402 pairs of twins born from 1870 to 2009, and 874 sets of triplets and 23 sets of quadruplets born from 1931 to 2009. Since the DTR currently only includes zygosity for twins born through 2000, the remainder of this article focuses on twins born 1870–2000. A basic description of the twins in the DTR is shown in Table 1, and the distribution of the zygosity for twin pairs over birth cohorts is shown in Figure 1.

Figure 1 reflects the different ascertainment methods and zygosity assessments in the four birth cohort periods described above. As previously mentioned, no opposite-sexed twin pairs were registered for birth cohorts 1911–1930, and the smaller number of twins registered from 1953 to 1960 is due to low coverage



**Fig. 2.** (a) Overview of the birth cohorts included in major surveys in the Danish Twin Registry since 1994. (b) Overview of the ages at which twins are included in major surveys in the Danish Twin Registry since 1994.

**Table 2.** Overview of the major twin surveys in the Danish Twin Registry

Survey	Year of survey	Participants	Males/Females	Monozygotic	Same-sex dizygotic	Opposite-sex dizygotic	Unknown/other zygocities	Complete pairs
Omnibus I	1994	29,433	14,181/15,252	8670	10,424	9282	1057	12,494
Omnibus II	2002	34,943	15,905/19,038	9123	12,776	11,913	1131	11,751
LSADT	1995	2401	859/1542	756	1391	224	30	480
	1997	2172	773/1399	696	1339	129	8	487
	1999	2709	1144/1565	931	1714	63	1	583
	2001	2448	1090/1358	853	1562	33	0	525
	2003	1844	815/1029	651	1176	17	0	347
MADT	2005	1372	610/762	490	872	10	0	236
	2007	121	35/86	58	63	0	0	44
MIDT	1998	4314	2198/2116	1471	1389	1448	6	1884
	2008–2011	2400	1134/1266	820	763	816	1	816
GEMINAKAR	2008–2011	10,276	4784/5492	1983	3948	4004	341	3035
	1998	1512	729/783	622	628	262	0	756
Omnibus 1	2010–2012	1145	535/610	504	444	197	0	513

Note: LSADT = Longitudinal Study of Aging Danish Twins, MADT = Middle Age Danish Twins, MIDT = Middle age Danish Twins.

particularly for female twins. The larger number of twin pairs with unknown zygosity in birth cohorts 1983–2000 is due to lower response rate to the intake questionnaire among the youngest twins compared to the older cohorts. Furthermore, a large increase in number of twin births since the 1980s is seen, which is mainly due to increased use of assisted reproductive technology (Herskind et al., 2005).

### Data Security

Data in the DTR are stored at a secure server. Directly identifiable data such as name, CPR-number and address are stored separately from the survey data. Generally, researchers will only have access to deidentified data in a project folder on a secure server via a secure virtual private network connection, and the access is limited to data necessary for answering a specific research question.

### Major Surveys Conducted Since 1994 in the DTR

During the history of the DTR, several surveys have been conducted. Below is a short description of major twin surveys since 1994. Figure 2(a) and (b) provides an overview of these major surveys regarding the year they were initiated and the birth cohorts and ages they include. Table 2 shows the number of twins included in each survey. Some of the surveys also included a few triplets and quadruplets, as noted in the descriptions of the surveys below. These are included in the numbers in Table 2.

In relation to the establishment of the four cohorts in the DTR, the questionnaires confirming twin status and assessing twin zygosity also included other questions. For example, the questionnaire for the 2003 assessment of the zygosity of twins born from 1983 to 2000 also included questions about infertility treatment and childhood conditions such as asthma, febrile seizures, epilepsies and type I diabetes.

### **General Twin Surveys (Omnibus 1 and 2)**

Two general twin surveys comprising various twin health projects were undertaken. In both surveys, a questionnaire was developed based on the interests of the researchers involved, and the responses were used both to study health conditions using the entire cohort of twins and to identify specific subgroups of twins for more detailed follow-up studies.

The first Omnibus (Omnibus 1) survey was undertaken in 1994 to illustrate the relative importance of genes and environment for the development of common diseases and health conditions in young twins. A seven-page questionnaire was sent to twins born from 1953 to 1982. A total of 29,433 individuals (29,173 twins and 260 triplets/quadruplets) completed the questionnaire (Leboeuf-Yde & Kyvik, 1998).

In the second Omnibus (Omnibus 2) survey undertaken in 2002, a larger questionnaire was mailed to twins born from 1931 to 1982. The survey addressed issues of health and health-related behavior, fertility, socioeconomic characteristics and subjective wellbeing in young and middle-aged twins. A total of 34,943 individuals (34,771 twins and 172 triplets/quadruplets) completed the questionnaire (Skytthe *et al.*, 2006).

A total of 21,260 individuals participated in both Omnibus 1 and 2.

### **The Longitudinal Study of Aging Danish Twins (LSADT)**

A cohort sequential study of elderly Danish twins was initiated in 1995. All twins aged 75 years and older, alive and residing in Denmark in January 1995 were selected for participation in the study and subsequently followed at 2 years' intervals up to and including 2005. Additional cohorts were added in 1997, 1999 and 2001; twins aged 73–76 years were enrolled in 1997 and twins aged 70+ years at the beginning of 1999 and 2001 were enrolled in these follow-up surveys. Across all intake waves, a total of 4731 subjects completed the intake assessment. Twins were selected for participation regardless of whether their co-twin was alive. Thus, only 1152 intact pairs completed an intake assessment.

The study was initiated to investigate the relative influence of genetic and environmental factors on the health of the elderly in different domains of health and functioning. All waves of Longitudinal Study of Aging Danish Twins (LSADT) were interview-based examinations with comprehensive questionnaires including questions regarding self-reported health and medication, lifestyle and activities of daily living, socio-demographic factors, cognitive tests and measurements of symptoms of depression. From 1999 and onwards, a short physical performance test was included involving measurements of chair stand, handgrip strength and lung function. Furthermore, biological material in the form of blood spots or buccal swaps was collected. In addition, whole blood samples were collected in 1997 and for a subgroup of those twin pairs still alive in 2007, a smaller questionnaire was filled out and a whole blood sample was collected. A more detailed description of the LSADT can be found in McGue and Christensen (2007).

### **The Middle Age Danish Twins (MADT) Study**

In 1998, a twin study including middle-aged twins born in the period 1931–1952 was initiated as a parallel to the LSADT in order to investigate how midlife health, lifestyle and functioning contribute to differences in late-life health and mortality. The participants were selected through a sampling of 240 twins from each birth

cohort from 1931 to 1952. From each birth cohort, 240 twins from 120 twin pairs were selected, including 40 pairs of MZ twins, 40 pairs of same-sex DZ twins and 40 pairs of opposite-sex DZ twins, and an equal distribution of men and women. A total of 4314 twins (including two of three from three triplets sets) completed a personal interview by means of a structured questionnaire similar to that of LSADT, and biological material in the form of blood spots or buccal swaps was collected.

In 2008–2011, a follow-up assessment of the Middle Age Danish Twin (MADT) cohort was conducted as a part of a broader project (the Danish Twin Registry Infrastructure Study). This follow-up assessment involved a mailed questionnaire, which the participants filled out in advance and handed in at an in-person interview and examination that took place at one of the five centers located across Denmark. Furthermore, a whole blood sample was collected. A total of 2400 twins completed the questionnaire and/or the examination. The majority (2380 twins) completed both questionnaire and examination (Skytthe *et al.*, 2013).

### **The Middle age Danish Twins (MIDT) Study**

The Danish Twin Registry Infrastructure Study carried out from 2008 to 2011 also comprised twins born 1931–1969 who were not included in the MADT survey. Priority was given to the oldest birth cohorts, twin pairs where both twins were alive and twins who had previously shown willingness to participate in research projects. The Middle age Danish Twin (MIDT) study, with a few exceptions, shared the same questionnaire, examination and biological samples as the MADT follow-up study. A total of 10,276 twins (including 43 triplets/quadruplets) completed the questionnaire and/or examination. The majority (10,111 twins) completed both questionnaire and examination (Skytthe *et al.*, 2013).

### **The Importance of Genes, Familiar and Common Environment for the Development of Insulin Resistance, Abdominal Adiposity and Cardiovascular Risk Factors (GEMINAKAR)**

Twins were recruited during 1997–2000 covering the birth cohorts 1931–1952 and 1953–1982 to investigate the relative influence of genetic and environmental factors for variation in several quantitative biological traits related to insulin resistance, abdominal obesity and associated cardiovascular risk factors. Invitations were sent to 2585 randomly chosen twins pairs with at least one twin living within 100 km from one of two investigation sites, of whom 1098 complete twin pairs (42 %) were both willing and able to participate; an age, gender and zygosity-stratified sample of 756 complete pairs participated in a full day clinical examination program. The clinical examination included an oral glucose tolerance test, a maximal bicycle test, body composition determined by bio impedance, electrocardiograms (ECGs), blood pressure measurement and blood sampling for biobanking of DNA, serum and plasma samples. Also, waist and hip circumference were measured as well as weight and height. Furthermore, the participants completed a questionnaire regarding health, health-related behavior and disease, and they completed a comprehensive food frequency questionnaire designed to reflect the food intake during the previous month. DNA-based microsatellite markers were used to determine the zygosity of the twins. More details on recruitment to GEMINAKAR, inclusion and exclusion criteria can be found in Benyamin *et al.* (2007).

**Table 3.** Overview of the number of participants with specific samples from the major twin surveys in the biobank and existing molecular data from the biological samples

Survey	Year of survey	Participants with samples in the biobank	Biological samples in the biobank							Existing molecular data from biological samples		
			Blood spots	Buccal swabs/ Saliva	Plasma (EDTA)	Serum	PBMC/ DNA	Paxgene/ RNA	Other	Genome-wide SNP data	Genome-wide methylation data	Genome-wide gene expression data
LSADT	1997	689	617		689		687			554	310	
	1999	2319	2053	265 (swabs)								
	2001	2185	2023	163 (swabs)								
	2003	1734	1559	124 (swabs)					1709 (hair)			
	2005	1255	1154	105 (swabs)								
	2007	121			121		121	119	72 (viable cells)	96	86	119 (+58)
MADT	1998	4202	3972	246 (swabs)								
	2008–2011	2337		2303 (saliva)	2319	2317	2312	2212		1465	492	520 (+58)
MIDT	2008–2011	10,054		40 (saliva)	10,030	10,002	10,029					
GEMINAKAR	1998	1408				1407	1407			173	428	
	2010–2012	1139		1133 (saliva)	1136	1112	1136	1113	1138 (urine)		428	

Note: LSADT = Longitudinal Study of Aging Danish Twins, MADT = Middle Age Danish Twins, MIDT = Middle age Danish Twins.

The GEMINAKAR cohort was followed up during January 2010 to October 2013. A total of 1435 (95 % of the original 1512) twins were alive and still living in Denmark and invited.

A total of 1145 twins completed the questionnaire and/or examination and the majority (1135 twins) completed both questionnaire and examination. Participants were visited by a mobile examination unit at their home, at work or at any other place at their convenience. The participants completed three different questionnaires covering health and lifestyle, sleep and work, and diet and underwent a clinical examination including ECG, blood pressure measurement, body composition, spirometry, various anthropometric measures as well as blood, saliva and urine sampling for biobanking (Li et al., 2016).

### The DTR Biobank

The DTR biobank was established in 1997 when biological material was collected as a part of the second wave of the LSADT and the first wave of GEMINAKAR. Since then, the DTR biobank has grown and now comprises biological samples from more than 19,000 twins who have participated in the various DTR surveys. Table 3 gives an overview of the different biological samples collected in the major twin surveys since the establishment of the biobank as well as molecular data generated from the biological samples.

For financial and logistical reasons, LSADT 1999–2005 and MADT-98 included biological samples in the form of a blood spot for all non-proxy participants. If participants were reluctant to provide a blood sample, a buccal swap was used as an alternative

source of DNA. Whole blood was collected in the LSADT-97, LSADT-07, MADT-08, MIDT and GEMINAKAR surveys.

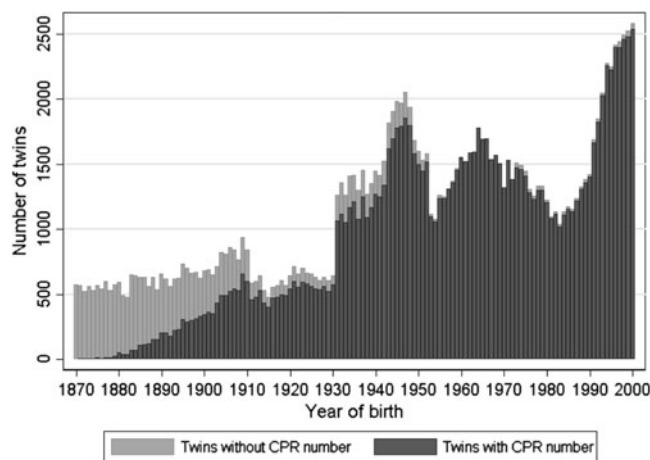
In several ancillary studies, the DTR was used to screen for the most informative twin pairs for a given phenotype — for example, twins with rheumatoid arthritis (Svendsen et al., 2013) and twins discordant for birth weight (Frost et al., 2012) — and included the relevant twins and collected biological material from these twins. An overview of ancillary twin studies with inclusion of biological material can be found in Skytthe et al. (2013).

### Storage of the Biological Samples

Several precautions have been taken in order to safely store the biological samples in the DTR biobank. The freezers in the biobank are housed in a secured area with alarm systems, controlling accessibility and warning systems in case of fire and flooding, cooling (and ventilation) systems with room temperature sensors and temperature sensor inside the freezers. All drains have also been sealed to prevent flooding. Furthermore, sample fractions from each participant are stored as duplets in separate freezers to prevent material from being damaged in case a freezer is malfunctioning.

### Molecular Data

**Genome-wide single nucleotide polymorphism (SNP) data** are available for a total of 554 LSADT-97, 96 LSADT-07, 1475 MADT-08 and 173 GEMINAKAR participants. Within the framework of the Interplay of Genes and Environment across Multiple Studies Consortium (Pedersen et al., 2019), 503 LSADT-97



**Fig. 3.** The number of twins with and without a person identification number (CPR number) in the Danish Twin Registry.

participants (including 115 MZ pairs, 133 SS-DZ pairs and 2 OS-DZ pairs) and 1465 MADT-08 participants (including 276 MZ pairs, 220 SS-DZ pairs and 231 OS-DZ pairs) were genotyped using the Illumina Infinium PsychArray. A subset of these data was recently included in a genome-wide association meta-analysis of intelligence (Savage *et al.*, 2018). In addition, 46 MADT-08, 46 LSADT-07 and 141 LSADT-97 participants were genotyped using the Illumina OmniExpress Array. The 141 LSADT-97 participants were selected for longevity, that is, surviving to at least 90 years of age (Deelen *et al.*, 2014) and only one of the twins in a pair was genotyped. Also, 80 LSADT-97 and 64 LSADT-07 participants were genotyped on the Illumina Human Omni2.5 Array in a study focusing on somatic structural genetic variants acquired over a 10 years' period (Magaard Koldby *et al.*, 2016). Finally, 173 participants within the GEMINAKAR cohort (one twin from 173 female MZ pairs) were genotyped using the Infinium II assay on the Illumina HumanHap300-Duo Genotyping BeadChips (Surakka *et al.*, 2012). The genome-wide SNP data for most of the LSADT-97 (539 of the 544) and MADT-08 (1465 of the 1475) participants have been imputed to the 1000 Genomes reference panel (Genomes Project Consortium *et al.*, 2015).

**Genome-wide methylation data** created using the Illumina Infinium HumanMethylation450K BeadChip are available for 155 LSADT-97 twin pairs and 246 MADT-08 twin pairs. The methylation data have been used in multiple epigenetic studies (e.g., Bolund *et al.*, 2017; Debrabant *et al.*, 2018; Lund *et al.*, 2019; Starnawska *et al.*, 2017; Svane *et al.*, 2018). Methylation data have mainly been made for MZ twin pairs, but for 43 twin pairs (18 MZ and 25 DZ twin pairs) repeated methylation data were made both at baseline (LSADT-97) and follow-up (LSADT-07; Christiansen *et al.*, 2016; Tan *et al.*, 2016). DNA methylation data have also been used for studying DNA methylation clocks (Christiansen *et al.*, 2016; Jylhava *et al.*, 2019). Finally, methylation data using the Illumina Infinium EPIC BeadChip array are available from both baseline and follow-up for 214 MZ twin pairs (856 samples) in GEMINAKAR (unpublished).

**Genome-wide gene expression data** have been generated using whole blood from 119 LSADT-07 (including 16 MZ and 25 DZ twin pairs) and 520 MADT-08 (260 MZ twin pairs) participants and the Agilent SurePrint G3 Human GE 8 × 60K Microarray. Also, whole blood-based RNA sequencing, including microRNA sequencing, has been performed for 58

LSADT-07 (10 MZ and 14 DZ twin pairs and 10 unrelated individuals) and 58 MADT-08 (12 MZ and 12 DZ twin pairs and 10 unrelated individuals) participants. For further information, see Hasler *et al.* (2017).

**Circulating microRNAs (plasma) data** from 45 MZ twin pairs from LSADT-97 were generated using the 754 human miRNAs OpenArray system (Mengel-From, Feddersen *et al.*, 2018), and in a study of 48 LSADT-97 MZ twin pairs (selected for dementia diagnosis concordant and discordant plus twin pairs of different degrees of cognitive discordance) a panel of 48 miRNAs was quantified using the Fluidigm BioMark microfluidic chip system (Mengel-From, Ronne *et al.*, 2018).

**Data on mitochondrial DNA copy number** have been conducted on approximately 2300 twins, including 176 twins participating in the first round of GEMINAKAR, 671 LSADT-97 participants (including 291 complete twin pairs), 1321 MADT-08 and MIDT participants (comprising 513 complete twin pairs), and 149 twins with rheumatoid arthritis and co-twins (including 73 complete pairs). Further information can be found in Mengel-From *et al.* (2014) and Svendsen *et al.* (2019).

Another epigenetic dataset includes *X-inactivation pattern* in female twins for 151 MZ and 361 DZ twin pairs from LSADT-97 and GEMINAKAR (Kristiansen *et al.*, 2005; Mengel-From *et al.*, 2012). Additionally, X-inactivation pattern data for about 500 twins from MADT-08 are available (unpublished).

In addition, *telomere length* is available on a large number of twins. Telomere length measured by Southern blot method using the *HphI/MnII* enzyme is available on 548 twins from LSADT-97, whereas telomere length measured using the *HinfI/RsaI* enzyme is available on 482 twins from LSADT-97 and 80 twins from LSADT-07 (Kimura *et al.*, 2010). For the GEMINAKAR study, telomere length measured using the *HinfI/RsaI* enzyme is available on 820 twins measured both at baseline and at follow-up. These data have been used in a number of studies (Hjelmberg *et al.*, 2015; Kimura *et al.*, 2008; Soerensen *et al.*, 2012).

### Linkage to Nationwide Registers at Statistics Denmark

The Danish CRS was established for administrative purposes on April 2, 1968. All persons alive and living in Denmark by April 2, 1968, or born or emigrated hereafter are assigned a unique personal identification number (CPR number; Pedersen, 2011). This unique CPR number enables unambiguous linkage between all Danish nationwide administrative registers (Thygesen *et al.*, 2011; Thygesen & Erbsoll, 2011).

Figure 3 provides the number of twins in DTR with and without a CPR number for each birth cohort and hence gives an overview of the number of twins that can be linked to the Danish national registers. The presence of twins without a CPR number in birth cohorts from 1973 and onwards is a consequence of stillborn infants being included in the DTR from 1973 and not receiving a CPR number.

Twins with a CPR number, a 5% random sample of the Danish population, as well as first-degree relatives and spouses have been linked to several Danish national administrative registers at Statistics Denmark. These registers are shown in Figure 4. Detailed descriptions of most of the Danish national registers can be found in a special issue of Scandinavian Journal of Public Health (2011).

The linkage between datasets from these registers is constructed at Statistics Denmark. The datasets are stored at a server located at Statistics Denmark, and access to the individual-level data in deidentified form is given via a secure virtual private network connection. Only researchers affiliated with an authorized research

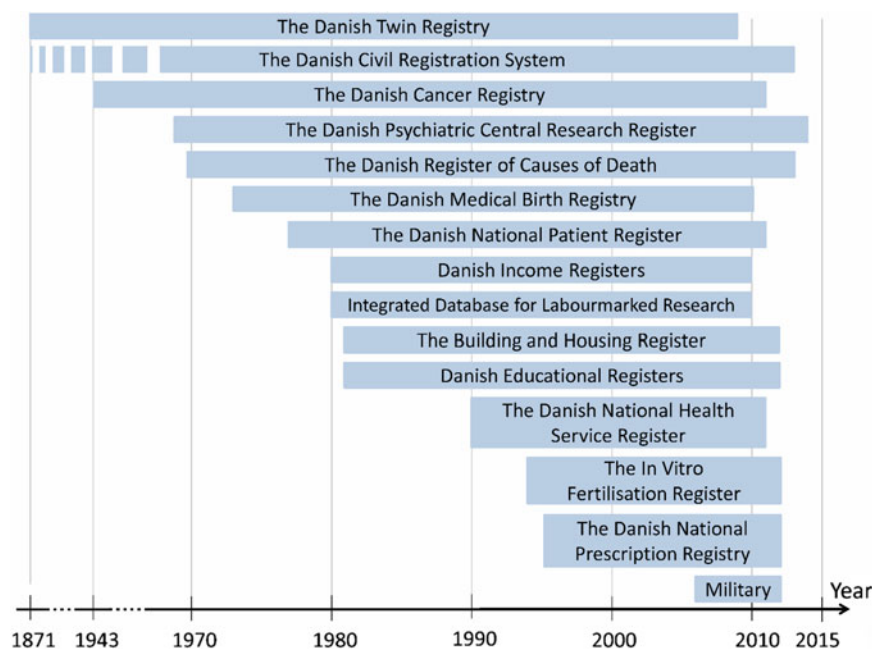


Fig. 4. The Danish Administrative Registers linked to the twins from the Danish Twin Registry at Statistic Denmark.

institution can establish remote access to these datasets and only permanent Danish research institutions of a certain size can be authorized. Researchers are only allowed to publish results on an aggregated level, preserving full anonymity of participants.

The primary research area of interest in the DTR is human development and aging with a focus on health, and therefore the registers providing information about use of healthcare services are the most widely used in the DTR, in particular the Danish National Patient Register, which contains information about hospitalizations, diagnoses classified according to the 8th and 10th revision of the International Classification of Diseases (ICD-8/ICD-10), dates of admission and discharge, and operation codes (Lynge et al., 2011). The Danish healthcare system is tax based, and therefore the vast majority of health services in Denmark are free of charge and equal access to healthcare is provided for all Danish residents. The health institutions receive reimbursement for services, which are therefore registered in administrative registers. Hence, these registers provide nationwide coverage and are a unique source to investigate the use of healthcare services in large-scale studies virtually without selection bias.

### Recent Illustrative Examples of Studies Using the Various Data Resources in DTR

A study based on the MADT and LSADT survey data investigated whether twin similarity in cognitive ability, depression symptomatology and handgrip strength changed with age and found no evidence of declining twin similarities in any of the investigated outcomes (McGue & Christensen, 2013).

Linkage between the DTR and the Danish national registers has resulted in studies on twin-singleton differences, classic twin studies and twin-co-twin control studies. Two studies compared opposite-sex and same-sex twins on different outcomes to assess any potential impact of intra-uterine hormone transfer. The outcomes in the two studies were cancer diagnoses from the Danish Cancer Registry and academic performance using the education registers. Neither of the studies found any evidence of differences between opposite-sex and same-sex twins (Ahrenfeldt, Petersen

et al., 2015; Ahrenfeldt, Skytthe et al., 2015). Another study combined the male twin population with singleton men and their brothers to investigate the association between early cognitive ability as determined by conscription board examinations and later diagnosis of dementia. The study found that low cognitive ability in young adulthood was associated with increased risk of dementia before age 78 years. Intra-brother and intra-pair twin analyses showed that the association was partly explained by shared family factors (Osler et al., 2017).

Previously, a range of studies using the DTR found twins to be similar to singletons in various aspects, for example, overall mortality after age 6 years, fecundity, school tests (shortly described in Christensen et al., 2011). In line with this, a recent register-based study compared the overall survival prognosis for twins and singletons with congenital heart defects and found no significant differences (Herskind et al., 2017). Furthermore, a recent study combining register data from the Nordic countries found that twins and singletons have similar cancer incidence (Skytthe et al., 2019).

A study combining longitudinal measurements of cognition from the MADT and LSADT surveys with data from the Danish National Patient Register on exposure to anesthesia examined the possibility of a negative effect of anesthesia on cognition in elderly. The study found that having had a major surgery was only associated with a negligibly lower level of cognitive functioning and that the key factors were preoperative cognitive functioning and underlying disease status (Dokkedal et al., 2016).

Another study combined register data from the Nordic countries to investigate the familial risk and heritability of cancer using the national cancer registries and found a significant excess of familial risk overall and for specific cancers, among others prostate, melanoma, breast and ovary (Mucci et al., 2016). The studies on cancer follow up on the first cancer study in the world to involve twin data, conducted by Harvald and Hauge (1963). Further use of the combined Nordic data revealed how genetic influence on risk of major cancer sites varies with age for prostate cancer (Hjelmberg et al., 2014), breast cancer (Moller et al., 2016) and colon cancer (Graff et al., 2017). These data were also used to develop

methodology for the time-to-event analysis for studies involving twins (Scheike et al., 2015, 2014). Collected survey data on smoking status over decades made it possible, via the matched twin design, to study smoking exposure for risk of lung cancer throughout life, in particular to gauge the magnitude of direct causal effect of smoking and the interaction effect of genes with smoking on lung cancer (Hjelmborg et al., 2017).

A study using the molecular data for twins aged 30–82 years found that an epigenetic profile defining the DNA methylation age was highly correlated to chronological age, and that an increase in methylation age was associated with increased mortality risk. For the oldest old, intra-pair twin analyses showed a stronger association between DNA methylation age and mortality when controlling for familial factors. Hence, the study supports previous findings that DNA methylation age is a useful biomarker of aging (Christiansen et al., 2016).

## Conclusion

In the DTR, the combination of long and nearly complete follow-up of twins with extensive phenotypic and molecular information from surveys and registers provides a valuable resource for studies of the mechanisms of human development, aging and diseases.

**Acknowledgement.** We want to thank the following for their considerable financial contribution to the foundation and the continued support of the Danish Twin Registry: the Danish Research Council, University of Southern Denmark, the Danish National Research Foundation, the Danish Cancer Society, the Danish Diabetic Association, the Danish Heart Foundation, the Novo Nordisk Foundation, United States National Institute of Aging (P01AG08761), United States National Cancer Institute, and the National Program for Research Infrastructure 2007 from the Danish Agency for Science Technology and Innovation.

## References

- Ahrenfeldt, L., Petersen, I., Johnson, W., & Christiansen, K. (2015). Academic performance of opposite-sex and same-sex twins in adolescence: A Danish national cohort study. *Hormones and Behavior*, 69, 123–131.
- Ahrenfeldt, L. J., Skytthe, A., Moller, S., Czene, K., Adami, H. O., Mucci, L. A., ... Lindahl-Jacobsen, R. (2015). Risk of sex-specific cancers in opposite-sex and same-sex twins in Denmark and Sweden. *Cancer Epidemiology, Biomarkers & Prevention*, 24, 1622–1628.
- Benyamin, B., Sorensen, T. I. A., Schousboe, K., Fenger, M., Visscher, P. M., & Kyvik, K. O. (2007). Are there common genetic and environmental factors behind the endophenotypes associated with the metabolic syndrome? *Diabetologia*, 50, 1880–1888.
- Bliddal, M., Broe, A., Pottgard, A., Olsen, J., & Langhoff-Roos, J. (2018). The Danish Medical Birth Register. *European Journal of Epidemiology*, 33, 27–36.
- Bolund, A. C. S., Starnawska, A., Miller, M. R., Schlunssen, V., Backer, V., Borglum, A. D., ... Sigsgaard, T. (2017). Lung function discordance in monozygotic twins and associated differences in blood DNA methylation. *Clinical Epigenetics*, 9, 132.
- Christensen, K., Kyvik, K. O., Holm, N. V., & Skytthe, A. (2011). Register-based research on twins. *Scandinavian Journal of Public Health*, 39, 185–190.
- Christiansen, L., Frederiksen, H., Schousboe, K., Skytthe, A., von Wurmb-Schwark, N., Christiansen, K., & Kyvik, K. (2003). Age- and sex-differences in the validity of questionnaire-based zygosity in twins. *Twin Research*, 6, 275–278.
- Christiansen, L., Lenart, A., Tan, Q., Vaupel, J. W., Aviv, A., McGue, M., & Christiansen, K. (2016). DNA methylation age is associated with mortality in a longitudinal Danish twin study. *Aging Cell*, 15, 149–154.
- Debrabant, B., Soerensen, M., Christiansen, L., Tan, Q., McGue, M., Christiansen, K., & Hjelmborg, J. (2018). DNA methylation age and perceived age in elderly Danish twins. *Mechanisms of Ageing and Development*, 169, 40–44.
- Deelen, J., Beekman, M., Uh, H. W., Broer, L., Ayers, K. L., Tan, Q., ... Slagboom, P. E. (2014). Genome-wide association meta-analysis of human longevity identifies a novel locus conferring survival beyond 90 years of age. *Human Molecular Genetics*, 23, 4420–4432.
- Dokkedal, U., Hansen, T. G., Rasmussen, L. S., Mengel-From, J., & Christiansen, K. (2016). Cognitive functioning after surgery in middle-aged and elderly Danish twins. *Anesthesiology*, 124, 312–321.
- Frost, M., Petersen, I., Brixen, K., Beck-Nielsen, H., Holst, J. J., Christiansen, L., ... Christiansen, K. (2012). Adult glucose metabolism in extremely birthweight-discordant monozygotic twins. *Diabetologia*, 55, 3204–3212.
- Genomes Project Consortium, Auton, A., Brooks, L. D., Durbin, R. M., Garrison, E. P., Kang, H. M., ... Abecasis, G. R. (2015). A global reference for human genetic variation. *Nature*, 526, 68–74.
- Graff, R. E., Moller, S., Passarelli, M. N., Witte, J. S., Skytthe, A., Christiansen, K., ... Hjelmborg, J. B. (2017). Familial risk and heritability of colorectal cancer in the Nordic Twin Study of Cancer. *Clinical Gastroenterology and Hepatology*, 15, 1256–1264.
- Harvald, B., & Hauge, M. (1963). Heredity of cancer elucidated by a study of unselected twins. *JAMA*, 186, 749–753.
- Hasler, R., Venkatesh, G., Tan, Q., Flachsbarth, F., Sinha, A., Rosenstiel, P., ... Nebel, A. (2017). Genetic interplay between human longevity and metabolic pathways – A large-scale eQTL study. *Aging Cell*, 16, 716–725.
- Hauge, M., Harvald, B., Fischer, M., Gotlieb-Jensen, K., Juel-Nielsen, N., Raebild, I., ... Videbech, T. (1968). The Danish Twin Register. *Acta Geneticae Medicae et Gemellologiae*, 17, 315–332.
- Herskind, A. M., Basso, O., Olsen, J., Skytthe, A., & Christiansen, K. (2005). Is the natural twinning rate still declining? *Epidemiology*, 16, 591–592.
- Herskind, A. M., Larsen, L. A., Pedersen, D. A., & Christiansen, K. (2017). Comparison of late mortality among twins versus singletons with congenital heart defects. *American Journal of Cardiology*, 119, 1680–1686.
- Hjelmborg, J. B., Dalgard, C., Mangino, M., Spector, T. D., Halekoh, U., Moller, S., ... Aviv, A. (2015). Paternal age and telomere length in twins: The germ stem cell selection paradigm. *Aging Cell*, 14, 701–703.
- Hjelmborg, J., Korhonen, T., Holst, K., Skytthe, A., Pukkala, E., Kutschke, J., ... Kaprio, J. (2017). Lung cancer, genetic predisposition and smoking: The Nordic Twin Study of Cancer. *Thorax*, 72, 1021–1027.
- Hjelmborg, J. B., Scheike, T., Holst, K., Skytthe, A., Penney, K. L., Graff, R. E., ... Mucci, L. A. (2014). The heritability of prostate cancer in the Nordic Twin Study of Cancer. *Cancer Epidemiology, Biomarkers & Prevention*, 23, 2303–2310.
- Jylhava, J., Hjelmborg, J., Soerensen, M., Munoz, E., Tan, Q., Kuja-Halkola, R., ... Reynolds, C. A. (2019). Longitudinal changes in the genetic and environmental influences on the epigenetic clocks across old age: Evidence from two twin cohorts. *EBioMedicine*, 40, 710–716.
- Kimura, M., Hjelmborg, J. V., Gardner, J. P., Bathum, L., Brimacombe, M., Lu, X., ... Christiansen, K. (2008). Telomere length and mortality: a study of leukocytes in elderly Danish twins. *American Journal of Epidemiology*, 167, 799–806.
- Kimura, M., Stone, R. C., Hunt, S. C., Skurnick, J., Lu, X., Cao, X., ... Aviv, A. (2010). Measurement of telomere length by the Southern blot analysis of terminal restriction fragment lengths. *Nature Protocols*, 5, 1596–1607.
- Kristiansen, M., Knudsen, G. P., Bathum, L., Naumova, A. K., Sorensen, T. I., Brix, T. H., ... Orstavik, K. H. (2005). Twin study of genetic and aging effects on X chromosome inactivation. *European Journal of Human Genetics*, 13, 599–606.
- Kyvik, K. O., Green, A., & Beck-Nielsen, H. (1995). The new Danish Twin Register: Establishment and analysis of twinning rates. *International Journal of Epidemiology*, 24, 589–596.
- Leboeuf-Yde, C., & Kyvik, K. O. (1998). At what age does low back pain become a common problem? A study of 29,424 individuals aged 12–41 years. *Spine (Phila Pa 1976)*, 23, 228–234.
- Li, S., Kyvik, K. O., Duan, H., Zhang, D., Pang, Z., Hjelmborg, J., ... Dalgard, C. (2016). Longitudinal investigation into genetics in the conservation of metabolic phenotypes in Danish and Chinese twins. *PLoS One*, 11, e0162805.



- Lund, J. B., Li, S., Christensen, K., Mengel-From, J., Soerensen, M., Marioni, R. E., ... Tan, Q. (2019). Age-dependent DNA methylation patterns on the Y chromosome in elderly males. *Aging Cell*, e12907.
- Lynge, E., Sandegaard, J. L., & Rebolj, M. (2011). The Danish National Patient Register. *Scandinavian Journal of Public Health*, 39, 30–33.
- Magaard Koldby, K., Nygaard, M., Christensen, K., & Christiansen, L. (2016). Somaticly acquired structural genetic differences: A longitudinal study of elderly Danish twins. *European Journal of Human Genetics*, 24, 1506–1510.
- McGue, M., & Christensen, K. (2007). Social activity and healthy aging: A study of aging Danish twins. *Twin Research and Human Genetics*, 10, 255–265.
- McGue, M., & Christensen, K. (2013). Growing old but not growing apart: Twin similarity in the latter half of the lifespan. *Behavior Genetics*, 43, 1–12.
- Mengel-From, J., Feddersen, S., Halekoh, U., Heegaard, N. H. H., McGue, M., Christensen, K., ... Christiansen, L. (2018). Circulating microRNAs disclose biology of normal cognitive function in healthy elderly people – A discovery twin study. *European Journal of Human Genetics*, 26, 1378–1387.
- Mengel-From, J., Ronne, M. E., Carlsen, A. L., Skogstrand, K., Larsen, L. A., Tan, Q., ... Heegaard, N. H. H. (2018). Circulating, cell-free micro-RNA profiles reflect discordant development of dementia in monozygotic twins. *Journal of Alzheimer's Disease*, 63, 591–601.
- Mengel-From, J., Thinggaard, M., Christiansen, L., Vaupel, J. W., Orstavik, K. H., & Christensen, K. (2012). Skewed X inactivation and survival: A 13-year follow-up study of elderly twins and singletons. *European Journal of Human Genetics*, 20, 361–364.
- Mengel-From, J., Thinggaard, M., Dalgard, C., Kyvik, K. O., Christensen, K., & Christiansen, L. (2014). Mitochondrial DNA copy number in peripheral blood cells declines with age and is associated with general health among elderly. *Human Genetics*, 133, 1149–1159.
- Moller, S., Mucci, L. A., Harris, J. R., Scheike, T., Holst, K., Halekoh, U., ... Hjelmborg, J. B. (2016). The heritability of breast cancer among women in the Nordic Twin Study of Cancer. *Cancer Epidemiology, Biomarkers & Prevention*, 25, 145–150.
- Mucci, L. A., Hjelmborg, J. B., Harris, J. R., Czene, K., Havelick, D. J., Scheike, T., ... Kaprio, J. (2016). Familial risk and heritability of cancer among twins in Nordic countries. *JAMA*, 315, 68–76.
- Osler, M., Christensen, G. T., Garde, E., Mortensen, E. L., & Christensen, K. (2017). Cognitive ability in young adulthood and risk of dementia in a cohort of Danish men, brothers, and twins. *Alzheimer's & Dementia*, 13, 1355–1363.
- Pedersen, C. B. (2011). The Danish Civil Registration System. *Scandinavian Journal of Public Health*, 39, 22–25.
- Pedersen, N. L., Gatz, M., Finch, B. K., Finkel, D., Butler, D. A., Aslan, D. A., ... Whitfield, K. E. (2019). IGEMS: The Consortium on Interplay of Genes and Environment across Multiple Studies – An update. *Twin Research and Human Genetics*, [This issue].
- Savage, J. E., Jansen, P. R., Stringer, S., Watanabe, K., Bryois, J., de Leeuw, C. A., ... Posthuma, D. (2018). Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nature Genetics*, 50, 912–919.
- Scandinavian Journal of Public Health. (2011). Danish population-based registers for public health and health-related welfare research – A description of Danish registers and results from their application in research. *Scandinavian Journal of Public Health*, 39(7 Suppl), 8–10.
- Scheike, T. H., Hjelmborg, J. B., & Holst, K. K. (2015). Estimating twin pair concordance for age of onset. *Behavior Genetics*, 45, 573–580.
- Scheike, T. H., Holst, K. K., & Hjelmborg, J. B. (2014). Estimating twin concordance for bivariate competing risks twin data. *Statistics in Medicine*, 33, 1193–1204.
- Skytthe, A., Christiansen, L., Kyvik, K. O., Bodker, F. L., Hvidberg, L., Petersen, I., ... Christensen, K. (2013). The Danish Twin Registry: Linking surveys, national registers, and biological information. *Twin Research and Human Genetics*, 16, 104–111.
- Skytthe, A., Harris, J. R., Czene, K., Mucci, L., Adami, H. O., Christensen, K., ... Pukkala, E. (2019). Cancer incidence and mortality in 260,000 Nordic twins with 30,000 prospective cancers. *Twin Research and Human Genetics*, 22, 99–107.
- Skytthe, A., Kyvik, K., Bathum, L., Holm, N., Vaupel, J. W., & Christensen, K. (2006). The Danish Twin Registry in the new millennium. *Twin Research and Human Genetics*, 9, 763–771.
- Skytthe, A., Kyvik, K., Holm, N. V., Vaupel, J. W., & Christensen, K. (2002). The Danish Twin Registry: 127 birth cohorts of twins. *Twin Research*, 5, 352–357.
- Skytthe, A., Kyvik, K. O., Holm, N. V., & Christensen, K. (2011). The Danish Twin Registry. *Scandinavian Journal of Public Health*, 39, 75–78.
- Soerensen, M., Thinggaard, M., Nygaard, M., Dato, S., Tan, Q., Hjelmborg, J., ... Christiansen, L. (2012). Genetic variation in TERT and TERC and human leukocyte telomere length and longevity: A cross-sectional and longitudinal analysis. *Aging Cell*, 11, 223–227.
- Starnawska, A., Tan, Q., McGue, M., Mors, O., Borglum, A. D., Christensen, K., ... Christiansen, L. (2017). Epigenome-wide association study of cognitive functioning in middle-aged monozygotic twins. *Frontiers in Aging Neuroscience*, 9, 413.
- Surakka, I., Whitfield, J. B., Perola, M., Visscher, P. M., Montgomery, G. W., Falchi, M., ... Peltonen, L. (2012). A genome-wide association study of monozygotic twin-pairs suggests a locus related to variability of serum high-density lipoprotein cholesterol. *Twin Research and Human Genetics*, 15, 691–699.
- Svane, A. M., Soerensen, M., Lund, J., Tan, Q., Jylhava, J., Wang, Y., ... Hjelmborg, J. B. (2018). DNA methylation and all-cause mortality in middle-aged and elderly Danish twins. *Genes (Basel)*, 9(2).
- Svendsen, A. J., Kyvik, K. O., Houen, G., Junker, P., Christensen, K., Christiansen, L., ... Hjelmborg, J. V. (2013). On the origin of rheumatoid arthritis: The impact of environment and genes – A population based twin study. *PLoS One*, 8, e57304.
- Svendsen, A. J., Tan, Q., Jakobsen, M. A., Thyagarajan, B., Nygaard, M., Christiansen, L., & Mengel-From, J. (2019). White blood cell mitochondrial DNA copy number is decreased in rheumatoid arthritis and linked with risk factors. A twin study. *Journal of Autoimmunity*, 96, 142–146.
- Tan, Q., Heijmans, B. T., Hjelmborg, J. V., Soerensen, M., Christensen, K., & Christiansen, L. (2016). Epigenetic drift in the aging genome: A ten-year follow-up in an elderly twin cohort. *International Journal of Epidemiology*, 45, 1146–1158.
- Thygesen, L. C., Daasnes, C., Thaulow, I., & Bronnum-Hansen, H. (2011). Introduction to Danish (nationwide) registers on health and social issues: Structure, access, legislation, and archiving. *Scandinavian Journal of Public Health*, 39, 12–16.
- Thygesen, L. C., & Ersboll, A. K. (2011). Danish population-based registers for public health and health-related welfare research: Introduction to the supplement. *Scandinavian Journal of Public Health*, 39, 8–10.