

The Norwegian Twin Registry

Thomas S. Nilsen, Ingunn Brandt, Per Magnus, and Jennifer R. Harris
Division of Epidemiology, Norwegian Institute of Public Health, Oslo, Norway

Norway has a long-standing tradition in twin research, but the data collected in several population-based twin studies were not coordinated centrally or easily accessible to the scientific community. In 2009, the Norwegian Twin Registry was established at the Norwegian Institute of Public Health (NIPH) in Oslo with the purpose of creating a single research resource for Norwegian twin data. As of today, the Norwegian Twin Registry contains 47,989 twins covering birth years 1895–1960 and 1967–1979; 31,440 of these twins consented to participate in health-related research. In addition, DNA from approximately 4,800 of the twins is banked at the NIPH biobank and new studies are continually adding new data to the registry. The value of the Norwegian twin data is greatly enhanced by the linkage opportunities offered by Norway's many nationwide registries, spanning a broad array of medical, demographic, and socioeconomic information.

■ **Keywords:** Twins, Norwegian Twin Registry, Registry Linkage

The Norwegian Twin Registry

The Norwegian Twin Registry (NTR) was established in 2009 as a merger of three major population-based Norwegian Twin Panels (Bergem, 2002; Harris et al., 2002, 2006). The NTR is housed at the Norwegian Institute of Public Health (NIPH), which currently runs 10 of Norway's national health registries and several large population-based cohort studies. The principal reasons for establishing the NTR are to promote research that exploits the value inherent in twin designs and to create a resource for studies in genetic epidemiology. Therefore, the registry makes data accessible to researchers for a wide range of studies. The registry infrastructure easily accommodates future expansions and enhancements of the twin data as described below. The purpose of this article is to describe the NTR today, including potentials for enriching the existing data through matches to other national registries and existing biological sample resources.

Since the 1960s, when twin research was initiated in Norway, substantial investments have been made in collecting population-based twin data for specific studies, but there were neither resources nor formal plans for developing a national twin data resource that would be made available to researchers. Only data from the NIPH panel had been fairly accessible (Harris et al., 2002, 2006). Establishment of the NTR was jointly funded by the University of Oslo, Oslo University Hospital, and NIPH. It does not have dedicated governmental funding and is currently supported through the Division for Epidemiology (NIPH) and through research studies using NTR data.

Panel I of the NTR covers cohorts born 1895–1945. A total of 37,000 pairs were identified for these birth cohorts through Statistics Norway, which also provided information about the twins' full names, and the date and place of birth of the twins and their parents. This information was then used to conduct linkages with national population registry files so that the national identity (NI) number of the twins could be assigned. This process identified a subset of 18,972 individuals (including 5,250 complete pairs) for whom the population registry files could provide validated NI numbers. The NI numbers make it possible to conduct linkages to other national registries. Importantly, a majority (10,000) of the same sex pairs from Panel I are also part of Panel II, and for these twins the NTR also contains information on zygosity plus other data as described below.

Panel II covers same sex twin births from 1915 through 1960 (and overlaps with Panel I for the years 1915–1945). Pairs where at least one twin had died before the age of 20 at the time of recruitment were excluded. Today, this panel contains 21,963 consenting twins, from an estimated 20,173 same sex pairs, comprising 9,183 complete pairs.

National identity numbers were introduced in Norway in 1964, based on the census of 1960; thus, twins deceased before 1960 were mostly lost to both Panel I and Panel II (Bergem, 2002; Iversen et al., 2001).

RECEIVED 27 June 2012; ACCEPTED 9 August 2012. First published online 5 September 12.

ADDRESS FOR CORRESPONDENCE: Jennifer R. Harris, Division of Epidemiology, Norwegian Institute of Public Health, P.O. Box 4404 Nydalen, Oslo, Norway. E-mail: jennifer.harris@fhi.no

TABLE 1
Overview of the Number of Individuals and Pairs in the Three Panels that Comprise the NTR

	Individuals with national identity number	Individuals with Consent	Complete pairs (male/female)			
			MZ	DZ	DZU	Unknown
Panel I 1895 –1945	8244*	0	0	0	371	892 (632/260)
Panel II 1915 –1960	26091	21963	3840 (1752/2088)	5343 (2545/2798)	0	53
Panel III 1967 –1979	13651	9477	1599 (681/918)	1359 (589/770)	1284	1
Total	47989	31440	5439 (2433/3006)	6702 (3134/3568)	1655	946

Note: *Unique for panel I—otherwise included in panel II count.
 Number of twins in NTR is not equal to the sum of previously published numbers of twins from the respective twin panels. This is due to three factors: (1) Files received from two of the three panels deviate from published numbers, (2) considerable overlap between two of the panels, and (3) legal issues concerning consent which have impact on number of twins available for research. The twin panels were established under different legal circumstances and with different purposes, thus making a legal integration of all our twin data challenging with certain issues still under pending decision from the Norwegian Data Protection Agency. Thus the number of twins available for research may vary.

Panel III covers birth cohorts from 1967 through 1979 (Harris et al., 2002). Twins were identified through the national Medical Birth Registry, which is complete for all births in Norway since the beginning of mandatory registration of all pregnancies in 1967 (Irgens, 2000). Approximately 15,000 twins were born in this period, of which 9,477 (4,242 complete pairs) consenting twins are now part of the NTR. Consent was granted via the return of a completed questionnaire or specific consent form. Today, studies will invariably need to obtain specific informed consent for the details of the data collection, storage and use.

Collectively, the three panels comprising the NTR span cohorts 1895–1960 and 1967–1979, and include a total of 14,742 complete pairs. Birth cohorts from 1961–1966 are missing because Panel II only recruited twins who were of consenting age (18 years), which meant birth cohorts before and including 1960, identified through Statistics Norway. Panel III relied on the Medical Birth Registry, which started in 1967 to identify twins nationwide.

Zygoty Classification

Panel I holds no information on zygoty except from opposite sex pairs. Panels II and III have determined zygoty by

questionnaire and subsequently conducted genetic marker analysis and DNA-based verification for a sub-sample of twins in Panel II and Panel III. These tests revealed high validity of the questionnaire-based method (Harris et al., 2006; Magnus et al., 1983). More specifically, for Panel II zygoty classification was determined by concordance of genetic markers from nine blood type systems and five serum group systems and four red cell enzymes systems and is reported in detail elsewhere (Berg, 1973). Tests of zygoty classification in Panel III were based on two multiplex panels of 12 micro-satellite markers each (for a total of 24 micro-satellite markers covering various autosomes). The alleles were scored by two individuals independently. Once discrepancies were resolved, the consensus data was reviewed for homozygoty between the two twins of each pair manually. Table 1 provides an overview of the number of twin individuals and the number of complete pairs across the three panels by zygoty.

Complete pairs are pairs where both twins have consented, although zygoty can be predicted with great accuracy using questionnaire data from one twin only. Twins with no consent and only NI are included in the base registry, but cannot be included in studies nor have their NI

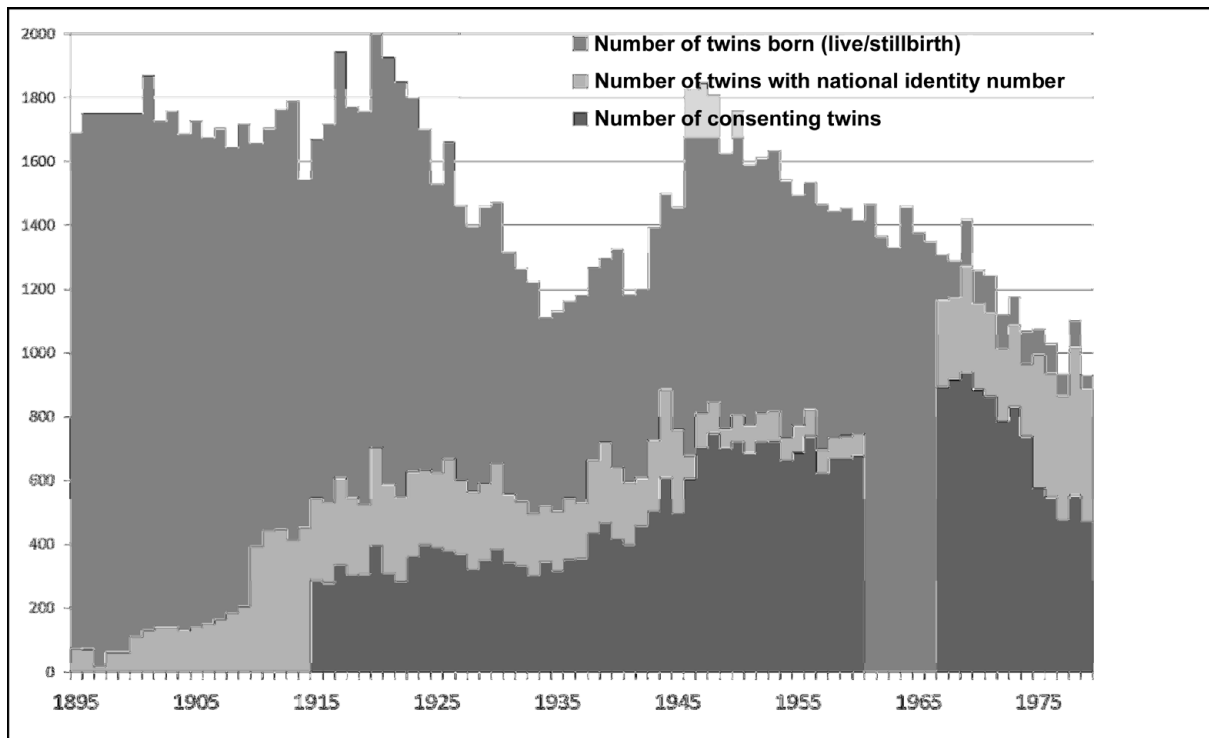


FIGURE 1

Structure of the twin data in the NTR showing the number of twins born (stillbirths and live births) on the y-axis by birth year on the x-axis.

Note. The large discrepancy between the number of twins born and number of twins with national identity number (introduced in 1964), is mainly due to the high number of stillbirths (5% in 1896–1900) and high general child mortality rates (8% dead \leq 1 year in 1901–1905), which were considerably higher for twins than singletons (Statistics Norway, 2011). A similar profile is reported for the Danish Twin Registry (Skytthe et al., 2002). Panel II twins (1915–1960) contain only same sex twins, thus accounting for a substantial number of the missing twins.

number linked to other registry data without prior consent. The exception to this rule applies to deceased twins, which can be included in studies, as the Personal Data Act in Norway generally only applies to living persons. Hence, the number of twins available for linkage is greater than the number of consenting twins. As of today, data on more than 8,000 non-consenting, deceased twins are also available for linkage pending prior approval by the Regional Committees for Medical and Health Research Ethics and the registries being linked. Additionally, there are approximately 1,500 twins from Panel II who have not been approached for consent and, therefore, linkage is not currently an option. Figure 1 depicts the structure of the cohort data in the NTR.

Phenotypes in NTR

The NTR contains a wide range of phenotypes, obtained by questionnaires, clinical assessments, and interviews. These have been detailed for Panel III elsewhere (Harris et al., 2002, 2006) and described to some extent for Panel II (Bergem 2002). Panel II conducted two large questionnaire surveys in 1978–1982 and again in 1990–1998. Twins

who responded to a previous general zygosity questionnaire and agreed to further contact were invited to participate (Magnus et al., 1983). Two questionnaires were sent to each twin: the first (1978–1982) contained items on general health, including current weight and height plus basic demographics and education, lifestyle, tobacco and alcohol use, physical health history, and degree of contact with co-twin. For all health history questions (diseases and medical conditions), the twin was also asked to report about their first-degree relatives, thus constituting twin family data. Data on the first-degree relatives only indicates whether said disease or condition afflicts any member of the family, including the twin. For these family members neither national identity number nor name is available. The second, companion questionnaire focused on reproductive health, maternal health, and lifestyle (including alcohol and tobacco use) during pregnancy, complications during pregnancy or birth, congenital disease, health history, and medical conditions of the children from infancy onwards. These two questionnaires were repeated in the second survey (1990–1998), where those who completed the first survey were invited. The general health questionnaire was expanded to include more items on health history, seizures, diet, anxiety and

TABLE 2
Selected Health and Population Registries Available for Linkage

Registry	Year	Contents	Completeness
Cancer Registry	1952	Incidence/diagnosis/mortality Clinical registries for specific cancers.	98.8%
Medical Birth Registry	1967	Births/stillbirths, complications/newborn health, maternal health before and during pregnancy.	99%–100%
Causes of Death Registry	1925/ (1951 electronic form)	Causes of death	99%–100%
Norwegian Immunization Registry	1995	Individual vaccination status/monitor vaccination coverage in the Norwegian population.	99%
Norwegian Prescription DB	2004	Overview about dispensing of prescribed medicines to patients, doctors and institutions from pharmacies.	99%–100%
Norwegian Patient Registry	2007	Special health service/Secondary health care: somatic and psychiatric.	somatic ~98% psychiatry ~95%*
Social Security DB	1992 (whole population)	Pensions/sickness benefits/unemployment insurance/social insurance/disability insurance.	Complete*
National Education DB	1970	Primary school through university. Results and direction.	Complete*
Employment/Income	1983/1993	Place/type of work/position/hours/	Complete*

Note: *These are large and complex registries, quality and completeness, may vary depending on the variables and population group and years of interest.

depression, whereas the reproductive health questionnaire was reduced, excluding many items regarding the children's health history, but including items on menstruation and fertility. Responses were received from 14,000 of the 21,885 consenting twins in Panel II. In total, 6,800 twins (including 1,900 complete pairs) responded to both of the general health questionnaires and 2,700 twins (including 700 complete pairs) responded to the two reproduction questionnaires, thus constituting a longitudinal sample (1978–1982 and 1990–1998). The discrepancy in response rates is primarily due to the reproduction questionnaire being targeted just to women whereas the general health questionnaire was targeted to the full twin sample.

Registry Linkage Opportunities

Norway has a comparative advantage in registry-based research. The first national patient registry in the world on leprosy was established in Bergen in 1856 (Irgens, 2002). Since the 19th century, in particular with the creation of The Central Bureau of Statistics (now Statistics Norway) in 1876, Norway has built an extensive registry infrastructure covering health and general population data. A large effort is underway, funded by the Norwegian government, to modernize all major health registries, and improve quality, scope, and access (www.nhrp.no). Being a welfare state, with free healthcare and education (primary school through university) and mandatory national insurance systems for unemployment, pension, sickness, and other benefits, these registries cover the whole population, thus selection and attrition bias is virtually eliminated. Through the national identity number, which is assigned at birth or registration with the National Population Registry, linkage between various registries and NTR can be conducted. The data available on each individual is potentially very extensive and detailed, both cross-sectionally and longitudinally. Table 2 provides an overview of some of the principal health and population

registries available for linkage to the NTR (Dahl et al., 2009; Statistics Norway, 2011).

Many more disease specific health registries also exist, and several others are under creation; for example, the Registry for Cardiovascular Diseases, which will be a national health registry. The registry creation and infrastructure are parts of a national strategy for improving healthcare, medical research, and disease surveillance. If linkages are not specifically mentioned in the consent, the Regional Committees for Medical and Health Research Ethics decide on a study-to-study basis whether the twins have to be individually notified or provide consent about the linkage study. However, through the NTR web pages and newsletters all twins are informed about current research, as required by law.

Access to NTR

Access to data from the NTR is applied for through an application (www.fhi.no/english) and granted by a steering committee that reviews applications according to a set of NIPH guidelines. These guidelines are to ensure that projects have a sound scientific basis, fall within the scope of the NTR mandate, have the pre-requisite permits and meet legal obligations pertaining to the registry and project host country. Within the European Economic Area (EEA), this is regulated through the EU Data Protection Directive 95/46/EC. For countries outside EEA, compliance with 95/46/EC is required. NTR also has responsibility to evaluate ethical aspects of all projects (even though they have been approved by the applicant's ethical review board) to ensure that they follow the proper consent procedures. Access to biological samples is more restricted, being a limited resource, and results from analyses must be returned to NTR for general use. Access fees are charged to cover administrative and data management costs incurred by the project.

TABLE 3
Overview of New Main Research Projects Analyzing and/or Collecting Additional Twin Data

Study	Sample/Data collection	Description purpose	Participants (PI)
Epigenetic surveying in immune-mediated diseases. A twin study. 2007 (Gervin et al., 2012)	Cohorts 1915–1960, 1967–1979 (140 cases, 220 controls) Data collection: Blood drawn from study participants/Questionnaire	Epigenetic factors involvement in disease. Pathogenesis of common immune-mediated diseases like asthma and psoriasis and IBD.	NIPH/(University of Oslo)
Genetics Consortium for Anorexia Nervosa (GCAN) Start 2009	From Mental Health Interview Study (1999–2003) DNA from 142 cases/controls	GWAS study, link between DNA sequence variations and Anorexia Nervosa.	International Consortium (University of North Carolina, Chapel Hill)
Nordic Twin Cancer (NorTwinCan) 2010	1895–1960, 1967–1979, linkage with Norwegian Cancer Registry/Norwegian Causes of Death Registry +30,000 twins	Aetiology of cancers, incl. breast, lung, colorectal, stomach and ovarian cancer. Heritability of onset of disease, cancer survival.	Nordic Consortium (Harvard School of Public Health)
Consequences of common mental disorders and personality disorders 2009	Mental Health Interview Study 1999–2003 and Questionnaire (cohort 1967–1979), linkage with Socio-Economic registries at Statistics Norway/National Educational DB	Consequences of personality disorders on several socio-economic indicators.	(NIPH)/Virginia Institute for Psychiatric and Behavioral Genetics
Axis I & Axis II Psychiatric Disorders in Norwegian Twins: A Follow Up Study 2010	Mental Health Interview data (cohort 1967–1979) Data collection: New interview and questionnaire, selected SID-P IV and CIDI disorders. 2,758 twins invited.	To study the course of selected psychiatric disorders, including personality disorders in early adulthood and how genetic and environmental factors influence their development over time.	(NIPH)/Virginia Institute for Psychiatric and Behavioral Genetics, University of Minnesota
Pain sensibility 2010	Twins 1915–1945. Data collection: NEO Personality Inventory Revised, Giessen Symptom Checklist, Satisfaction With Life Scale, assessment of chronic pain, SCL-5, Corah's Dental Anxiety Scale	Heritability of somatic complaints, genetic and environmental sources of covariance between personality traits and pain/complaints indicators. (esp. chronic pain)	NIPH, (University of Oslo)
Socioeconomic determinants of cancer (twin, family, and adoptee study) 2009	Cohorts 1915–1960, 1967–1979, linkage Norwegian Cancer Registry, Socio-Economic registries at Statistics Norway	To which extent the environmental factors driving cancer risk can be ascribed to socio-economic background.	University of Amsterdam, ENSAE Paris, (Statistics Norway)

NTR Research Projects

Table 3 provides an overview of main ongoing NTR-based projects. As seen from this table, new data are continually added to the registry (data obtained by linkage cannot be transferred to NTR) from new projects collecting study-specific data. During the project period, exclusive rights to the data are given to the research project, and after the project ends the data are released to the NTR and made available for research use through application.

Biological Samples

Biological samples have been collected in several studies, primarily from twins in Panel III, and are, therefore, limited to cohorts 1967–1979. DNA was collected by mail-out buccal smear kits sent to all questionnaire participants. Four buccal samples were collected from each twin using small brushes (cytobrush, cell collector) with soft bristles used for swabbing the inside of the cheek. The twins were instructed to collect cheek swab samples at least 1 hour after their last meal and with at least 8 hours between the first two and last two swabs. Altogether, 4,800 individual samples were collected. In addition, blood samples were collected and DNA and plasma were extracted from 1,850 of these twins,

who also participated in the later Genetics and Personality Study. These samples are stored at NIPH biobank at -20°C (plasma -80°C). Biological samples collected by NTR cover only a fraction of twins in the registry, but other national sources of biological samples are or will become available. Millions of blood samples, genetic samples, and other biological materials from the Norwegian populace have been collected over many years. Norway's many biobanks are now to be reorganized into a single national research infrastructure. One aim is to develop databases and improve information and management systems that make it easier to link the large amounts of existing health data in Norwegian registries to the biobanks' biological materials. This endeavor, called Biobank Norway, is a national hub aiming to harmonize with Europe's large-scale biobank infrastructure, the Biobanking and Biomolecular Resources Research Infrastructure (BBMRI). The NTR will be integrated into Biobank Norway.

Future Plans

Of primary importance is recruitment of new twins from birth cohorts not already included in the NTR, starting from 1980 and onwards. This will encompass all twins aged 18 or older at the time of recruitment and would

encompass a target population of approximately 10,000 new twin pairs to NTR. This is vital, not only from the perspective of statistical power, but maybe even more so from linkage opportunities offered. As mentioned above, Norway has one of the most comprehensive population-based registry systems in the world; many of these registries have been established in the last decades, and older registries are becoming increasingly informative as detailed information is collected. Thus, recruiting new twins will provide important new linkage opportunities, especially as older cohorts experience natural attrition.

Acknowledgments

Development and use of the NTR has been supported from research funds granted through the European Union's Seventh Framework Programme (FP7/2007-2013), EN-GAGE Consortium, grant agreement HEALTH-F4-2007-201413; BioSHaRE-EU, grant agreement HEALTH-F4-2010-261433, the Ellison Foundation, USA, and Norwegian Research Council funds through Biobank Norway (NFR 197443/F50) and grants Axis I and Axis II Psychiatric Disorders in Norwegian Twins: A Follow-Up Study (NFR 196148) and Consequences of Common Mental Disorders and Personality Disorders (NFR 193615).

References

- Berg, K. (1973). Studies of polymorphic traits for the characterization of populations. The populations of Scandinavia. *Israel Journal of Medical Sciences*, 9(9), 1147–1155.
- Bergem, A. L. (2002). Norwegian twin registers and Norwegian twin studies — An overview. *Twin Research*, 5(5), 407–414.
- Dahl, C., Stoltenberg, C., Magnus, T., Høyer, A., Skjesol, P., Vassenden, A., Skjesol, P. O., Vassenden, A., Skau, P., Berntsen, G., Kapstad, K., Hagen, E., Opdal, V., Viste, M., & Vollset, S. E. (2009). *Gode Helseregistre — Bedre Helse*. Oslo: Sekretariatet for Nasjonalt helseregisterprosjekt, Folkehelseinstituttet.
- Gervin, K., Vigeland, M. D., Mattingsdal, M., Hammerø, Nygård, H., Olsen, A. O., Brandt, I., Harris, J. R., Undlien, D. E., & Lyle, R. (2012). DNA methylation and gene expression changes in monozygotic twins discordant for psoriasis: Identification of epigenetically dysregulated genes involved in immune response. *PLoS Genetics*, 8(1), e1002454.
- Harris, J. R., Magnus, P., & Tambs, K. (2002). The Norwegian Institute of Public Health Twin Panel: A description of the sample and program of research. *Twin Research*, 5(5), 415–423.
- Harris, J. R., Magnus, P., & Tambs, K. (2006). The Norwegian Institute of Public Health twin program of research: An update. *Twin Research and Human Genetics*, 9(6), 858–864.
- Irgens, L. M. (2000). The Medical Birth Registry of Norway. Epidemiological research and surveillance throughout 30 years. *Acta Obstetrica et Gynecologica Scandinavica*, 79(6), 435–439.
- Irgens, L. M. (2002). The discovery of the leprosy bacillus [published in Norwegian]. *Tidsskr Nor Laegeforen* 122(7), 708–709.
- Iversen, T., Tretli, S., & Kringlen, E. (2001). An epidemiological study of cancer in adult twins born in Norway 1905–1945. *British Journal of Cancer*, 84(11), 1463–1465.
- Magnus, P., Berg, K., & Nance, W. E. (1983). Predicting zygosity in Norwegian twin pairs born 1915–1960. *Clinical Genetics*, 24(2), 103–112.
- 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(5), 352–357.
- Statistics Norway. (2011). *Statistics Norway*. Retrieved from www.ssb.no