An Update on the Italian Twin Register: Advances in Cohort Recruitment, Project Building and Network Development

Sonia Brescianini,1 Corrado Fagnani,1 Virgilia Tocaceli,1 Emanuela Medda,1 Lorenza Nisticò,1 Cristina D’Ippolito,3 Sabrina Alviti,1 Antonio Arnofi,1 Bruno Caffari,2 Davide Delfino,1 Maurizio Ferri,3 Luana Penna,1 Miriam Salemi,1 Silvia Sereni,1 Laura Serino,1,4 Rodolfo Cotichini,1,5 and Maria A. Stazi1

1Genetic Epidemiology Unit, National Centre for Epidemiology, Surveillance and Health Promotion, Istituto Superiore di Sanità, Rome, Italy
2Pharmacoepidemiology Unit, National Centre for Epidemiology, Surveillance and Health Promotion, Istituto Superiore di Sanità, Rome, Italy
3Data Management, Documentation, Library and Publishing Activities Service, Istituto Superiore di Sanità, Rome, Italy
4School of Specialization in Hygiene and Preventive Medicine, Tor Vergata University, Rome, Italy
5Institute of Clinical Physiology, National Research Council, Pisa, Italy

The Italian Twin Register has been in place for more than 10 years. Since its establishment, it has been focusing, on the one hand, on a continuous update of the existing information, and on the other hand, on new phenotypes and sample collection. Demographic data on about 140,000 twins have been updated using the municipality registries. The Italian Twin Register has been carrying out several new studies during the last few years. A birth cohort of twins, Multiple Births Cohort Study, has been started and the enrolment is ongoing. For this cohort, data on pregnancy and birth are collected, and periodical follow-ups are made. DNA is being collected for the twins and their parents. In the area of behavioral genetics, most efforts have been directed to psychological well being assessed with self-reported tools. Research on age-related traits continues with studies on arteriosclerosis development, early biomarkers in mild cognitive impairment, and the relation between lifestyle habits and mutagen sensitivity. The Italian Twin Register biobanking has grown in its size and in its know-how in terms of both technical issues and ethical procedures implementation. Furthermore, attitudes toward biobank-based research, together with willingness and motivation for donation, are being investigated. A valuable key resource for the Italian Twin Register is the possibility of linking twin data with disease registries. This approach has been yielding several important results, such as the recent study on the heritability of type 1 diabetes.

Keywords: twins register, genetic epidemiology, birth cohort, age-related traits, mental health, biobanking

Short History of the Register, Sources, and Recruitment Procedures

The Italian Twin Register (ITR) was established at the National Institute of Health (Istituto Superiore di Sanità) in 2001 when a research project for its implementation was funded by the Ministry of Health. At the beginning, the ITR started with a dataset of all the potential Italian twins born before December 31, 1996. This dataset was provided by the Ministry of Finance and contained demographic data on all persons born on the same day, in the same town with the same last name, and consisted of about 650,000 possible twin pairs. A number of these individuals were not actually twins, and for this reason the database was updated using municipalities’ records. In particular, the ITR contacted municipalities in all big cities and several medium-size towns and queried names and addresses of residents with the same last name born on the same day, in the same town, and from the same parents. In this way it was possible...
to update information on about 140,000 twins. This was an important step forward for the ITR because it allowed the enrolment of children and adolescents born after 1996, and the detection and elimination of the ‘non-twins’ from the old database. The resulting database is a list of twins with demographic information; on the basis of this list, the twins are enrolled in the ITR. The enrolment procedure, the conduction of the studies and the biobanking activities respect the national current privacy rules on sensitive data, genetic data, and biological samples treatment (legislative decree n. 196/2003 and the General Authorisation for the Processing of Genetic Data, 2011). For twins’ enrolment, three different strategies are usually followed. The first, which corresponds to a ‘population-based approach’, consists of sending, by mail, an enrolment kit to all twins of a specific age group living in a specific area. This kit contains a questionnaire, the informed consent form to be signed by the twin or the twins’ parents for minors, and an informative letter on the ITR research activities. The second, which corresponds to a ‘hospital-based approach’, relies on linking the ITR database with disease registries that could be locally or nationally based, in order to detect potential twins to be enrolled in studies of specific diseases. The third approach consists of voluntary requests for enrolment by twins who have become familiar with the ITR through the Web site (www.iss.it/gemelli). These volunteers are also sent the enrolment kit. All twins are added to the database and are re-contacted according to inclusion criteria (e.g., zygosity, targeted age groups, geographic area) or by target outcomes requested by specific studies (e.g., complex traits, specific pathologies, discordance for trait/pathology, etc.).

The enrolment questionnaire collects demographic information not available in the original database, such as occupation and education; it includes a set of questions on physical resemblance (Kyvik et al., 1995) from which zygosity is derived and a few other items such as current weight and height. To date, all data have been collected through paper questionnaires. The questionnaires are either sent by mail or given by hand to the twins if a face-to-face meeting is included in a specific study. The questionnaires are then scanned using a character recognition system and the data are stored in a central database and server.

During 2012, the ITR has been experimenting with online questionnaires, using an ad-hoc software developed for the specific research purposes. This is an efficient, fast, and cheap way of collecting data, and is becoming suitable for large population subgroups, given that the use of personal computer and Internet is increasingly spreading across age groups. To date, the ITR has enrolled about 25,000 twins; Table 1 shows enrolled twins by year of birth, gender, and zygosity.

The Register Web site, where information about ‘being twins’ can also be found, contains data on enrolment, ongoing research and biobanking activities, updated studies results, and publications.

Major Goals of the Register

The Register is involved in many projects: some deal with particular age groups, some with specific diseases or phenotypes, some are exclusively based on data from questionnaires, while others require a biological sample. The following sections provide the reader with a general overview of the work that has been going on at the ITR in the past years and the new projects that will be started in the years to come.

Perinatal and Pediatric Outcomes

A major impediment to fully understanding the relationship between birth/early life factors and development of adult disease (known as the ‘Barker hypothesis’; Barker, 1995) comes from a lack of longitudinal data, beginning in the prenatal period and continuing across the life course. The rate of postnatal growth and early childhood body size has been shown in some studies to be just as or even more important than birth weight in predicting many adult health outcomes, including cardiovascular disease and diabetes (Barker, 2005; Barker et al., 2005). Twin data offers a chance to explore these correlations even more deeply than singleton birth cohort studies. Besides estimating the genetic and environmental component of main pediatric outcomes, data on twins allows the estimation of the interaction between genes and environment that result in disease in adulthood. For all these reasons, the ITR is following up a birth cohort of twins (MUBICOS, Multiple Birth COhort Study) that will be a valuable source for future research, not only for pediatric outcomes, but also for adolescent and adult health.

MUBICOS has benefited from the collaboration with the Italian Society of Neonatology, through which eight hospitals representing the different geographical areas of Italy have agreed to participate in the study. The hospitals are located in the north (Turin and Trieste), center (Pisa,
Bologna, and Rome), and south (Foggia and Palermo) of Italy. These hospitals have a total of about 25,000 births per year, yielding about 700 twins (assuming a twinning rate of 1.4% of total deliveries).

All twins eligible for the study who were born in one of the eight hospitals are invited to join the study. Twin pairs are eligible if they are born with a gestational age of 32 completed weeks or more, if both twins are born alive, and if at least one of the two parents can speak Italian. In the rare and special case that one of the two twins is transferred to another hospital or dies a few days after birth, the twin pair is withdrawn from the study. Twins from assisted reproductive techniques (ART) are eligible, as well as spontaneously conceived twins. Up to now, about 300 pairs have entered the study and enrolment is ongoing.

The enrolment consists of a questionnaire filled in by the medical personnel, with medical information on pregnancy, delivery, and the first few days of life of the twins. This information is derived directly from medical records and relates to pregnancy complications and drug use, prenatal diagnosis, ART use, labor, type of delivery, gestational age, birth weight, chorionicity, Apgar score, perinatal outcomes (e.g., respiratory distress syndrome), screening results for congenital hypothyroidism, blood type, and treatment during hospital stay. Another questionnaire is filled in by the parents with information on lifestyles before and during pregnancy, occupational exposure, reproductive history, and folic acid and drug use during and before pregnancy.

At the time of enrolment the parents are asked to donate a saliva sample from themselves and the newborns. The saliva and DNA samples are stored in the ITR Biobank. Informed consent is signed by both parents after the medical personnel have explained the study and given them the informative letter.

Follow-up questionnaires are sent to the parents every 6 months in the first 2 years of life, and every year or more afterwards. Main outcomes of the study are: growth, neuropsychological development, sleeping behavior, and nutrition. This cohort study is part of a European Network of birth cohorts (www.birthcohorts.net).

All information collected by questionnaires is stored in an online database. Access to the database is provided through a platform (http://vcms.iss.it/moodle19). Selected hospital and ITR staff are given access to the database. There are different levels of access: the hospital can access and edit its own data only, while the ITR staff access and edit data from all the hospitals.

The ITR has been working on pediatric outcomes for a few years. In particular, the interest has focused on sleeping behavior in toddlers (Brescianini et al., 2011). The study was performed on more than 300 pairs for whom parents answered a sleeping behavior questionnaire when the twins were 18 months old. A few sleeping phenotypes were analyzed: number of night awakenings per week, and night and daytime sleep duration (in hours). For all these phenotypes, heritability was around 33%, while almost all the remaining proportion of variance was due to shared environment, leaving very little to unique environment. ITR focused also on infant growth during the first month of life (Brescianini et al., 2012). The study was performed on preterm twins born in the same hospital in Rome. Heritability of growth rate in the first month of life in this very homogeneous set of infants was around 90%. The ITR was also involved in a study on the association between folic acid supplementation before conception and the probability of a twin pregnancy. Data on folic acid intake were collected for cases (mother of twins) and controls (mother of singletons). DNA was also collected and extracted from saliva. The data analysis is currently underway.

**Mental Health and Psychological Well Being**

During the last few years, behavioral genetics has become a major area of investigation for the ITR. A close collaboration has been established with psychologists, psychiatrists, and mental health professionals operating in several Italian universities and research institutions, including the Mental Health Unit at the Istituto Superiore di Sanità of Rome, the psychology departments at the Universities La Sapienza of Rome and Vita-Salute San Raffaele of Milan, and the Inter-University Center for Behavioural Neurosciences (ICBN) at the Universities of Udine and Verona. So far, all studies in this area have been conducted in the general population rather than in clinical settings, but new projects involving psychiatric patients are being set up.

Most efforts have been directed toward understanding the genetic and environmental bases of adult psychological well being, which has been assessed using a broad spectrum of self-reported tools, all of them previously validated in the Italian population. Relevant measures of psychological well being include the Rosenberg Self-Esteem Scale, the Satisfaction with Life Scale, the Life Orientation Test, and the Ryff’s Scales of Psychological Well-Being (Caprara et al., 2009; Gigantesco et al., 2011). Moreover, a new well being measure — the Positivity Scale (Caprara et al., 2012) — has been validated by our group, and its genetic and environmental architecture are currently being investigated on a large twin sample. Attachment style in close relationships, measured with the Experiences in Close Relationships questionnaire, has also been explored as a putative determinant of positive mental status (Picardi et al., 2011b). Recently, interest has been focused on the biological correlates of psychological well being, and new projects are being planned on the relationship of the aforementioned measures with autoimmune, neuroendocrine, and cardiovascular parameters. Beside the ‘positive psychology’ approach, some studies in the adult population have been performed — or are in progress — on various phenotypes that have more direct clinical relevance. These phenotypes encompass psychotic and obsessive symptoms (Symptom Check List — SCL-90; Fagnani et al., 2011), psychosis proneness...
estimates were about 40% for stiffness of the arteries and only moderately influenced by genetic factors (heritability). Preliminary results showed that investigated parameters are validated questionnaires such as the Rosenberg self-esteem (Pesenti-Gritti et al., 2011) and between these and competence scales (Pesenti-Gritti et al., 2011).

Age-Correlated Traits

The ITR is involved in age-related trait research. One of the study’s aims is to explore the contribution of genetic versus environmental factors in arteriosclerosis development in a sample of adult twins selected from the ITR database and resident in Rome, Padua, and Perugia. A sample of 164 monozygotic (MZ) and 184 dizygotic (DZ) twin pairs, aged 22–74 years old (mean ± SD, 54.6 ± 12.4 years), was subjected to carotid Doppler ultrasound to measure intima media thickness (IMT) of common and internal carotid arteries and to determine the anatomic extent of atherosclerosis (Tarnoki et al., in press). A TensioMed Arteriograph was used to evaluate, non-invasively, brachial and aortic augmentation indices (Aixao and Aixbra) and aortic pulse wave velocity (PWVao), indirect measures of arterial wall stiffness and wave reflection (Tarnoki et al., 2011). In addition, twins answered a questionnaire regarding demographic data, medical history, and personal habits. In order to explore the relationship between atherosclerotic vascular condition and mental or emotional factors, psychological status of twins was also investigated in depth using validated questionnaires such as the Rosenberg self-esteem scale, Life Orientation Test, Ryff Scale, and Diener’s Satisfaction With Life Scale (Diener et al., 1985; Marsh, 1996; Ryff & Keyses, 1995; Scheier et al., 1994). Saliva samples of all twins were collected and stored in the ITR Biobank. Preliminary results showed that investigated parameters are only moderately influenced by genetic factors (heritability estimates were about 40% for stiffness of the arteries and around 30% for carotid IMT). Relevant environmental factors for these measures appeared not to be shared within the family, but related to individual experience. From 10% to 44% of individual differences in IMT or arterial stiffness could be explained by age (Medda et al., 2012). The discordant MZ approach will be also used to explore unshared environmental factors in early traits of arteriosclerosis.

A second study focuses on reliable early biomarkers useful for identifying asymptomatic subjects or those with early cognitive impairment. A sample of about 120 MZ and DZ twin pairs aged over 62, living near Rome, was enrolled and evaluated by a wide neuropsychological test battery: Mini Mental state examination, Depression Scale CES-D, Purdue Pegboard test, Stroop test (Desrosiers et al., 1995; Folstein et al., 1975; Radloff, 1977; Vreeling et al., 1993). Blood samples were collected at the time of medical check between 2002 and 2003 and stored in the ITR Biobank, and several functional and biochemical markers were measured. Data analysis is now ongoing according to a multidisciplinary approach, which includes epidemiological and basic research aspects: parameters predictive of cognitive impairment and decline will be explored, and the role of the genetic component in cognitive and physical skills and in some biomarkers will be examined.

The third area of interest focuses on the relation between lifestyle habits and mutagen sensitivity. MZ and DZ twin pairs with available information on lifestyle (diet, job, smoking), psychological wellness (Life Orientation Test, Rosenberg self-esteem scale, Ryff scale), anthropometric, physiological, and hematological data (e.g., body mass index [BMI], blood pressure, cholesterol) are being evaluated for lymphocytes DNA damage and DNA damage response to in vitro γ-irradiation treatment.

Biological Sample Collection and Storage: The ITR Biobank

Up to now the main source of DNA has come from twin donors’ saliva (about 2,000). Saliva specimens have been generally self-collected during visits and interviews for specific research projects (Alessandri et al., 2010; Caprara et al., 2009; Fagnani et al., 2011; Nisticò et al., 2012; Picardi et al., 2011; Tarnoki et al., 2011; Tarnoki et al., in press; Tocaceli et al., 2009a), and also during ‘twin-days’ organized by our unit with the collaboration of the Italian blood donors association (AVIS), or in other gatherings set up by twins associations (Tocaceli & Nisticò, 2006; Tocaceli et al., 2009b). Approximately one-third of donations have instead been obtained by a saliva collection kit mailed to twins, with twins mailing samples back to the ITR. Furthermore, saliva swabs have been collected from 320 baby twins unable to spit. Whole blood, lymphocytes, buffy coat, serum, and plasma aliquots of about 500 adult twins have also been collected during various studies and stored in the ITR Biobank in liquid/vapor nitrogen tanks or −80°C freezers with nitrogen back-up.
Donors signed an informed consent that authorizes a 20-year storage period of their biological material. For under 18-year-old subjects, storage is allowed up to full age, when consent must be newly requested. Moreover, a new consent has to be required when new studies on donors’ DNA are planned.

DNA extraction, quantification, quality check, back-up cryopreservation, and zygosity assessment have been entrusted to an external service provider. Zygosity has been determined using a set of nine tetranucleotide markers, for over 500 pairs of twins and for few triplets of the same gender. This company has no access to identification data.

Ethics and the ITR Biobanking Research

With the increasing importance of the genetic epidemiological investigations in the field of common diseases and the need for huge amounts of biological specimens from large groups in healthy populations, the establishment of the ITR Biobank at the National Center for Epidemiology, Surveillance and Health Promotion has been a great challenge in the past years. Besides managerial and technical efforts, it has required professional skills and careful attention to the ethical and legal issues involved. National ethical guidelines on topics such as suitability of informed consent when biological material use and reuse is at stake were quite recently issued (Presidenza del Consiglio dei Ministri, 2009), but in our opinion, they did not contribute to providing a clear ethical framework to conduct such research. Moreover, in the background there is a widespread international debate on the ethical tools for biobanking (Murphy et al., 2009; Wendler, 2006). We thought it worthwhile to start an ITR line of research on the ethical aspects involved in biobanking activities for epidemiological research.

In this context, motivation and attitude toward biobank-based research, as well as willingness to donate, awareness of donation, and the level of understanding of the aims and methods of the study participants are involved in, have so far been considered relevant issues to be investigated from both scientific and ethical perspectives (Gross et al., 2011; Johnsson et al., 2010; Laurie, 2008). On these dimensions, in 2007–2008 we conducted a cross-sectional pilot survey (Tocaceli et al., 2009b) among a group of healthy twins enrolled in our register who had donated to the Biobank and participated in a genetic epidemiological study on the genetic regulation of the end-stage clotting process (FP6 Specific Targeted Research Project EUROCLOT: Genetic Regulation of the End-Stage Clotting Process That Leads to Thrombotic Stroke). The results suggested that understanding seemed to be related to the education level, and was positively influenced by being a volunteer instead of already being enrolled in the Register and formally asked to enter the study. Awareness of the biobank (the recall of donation after 6 months to 1 year, as a proxy) was highly present among the twins, and willingness to participate and donate seemed to be supported by an ‘utilitarian/pragmatic’ attitude to contribute to research. Finally, pair-wise twin concordance for this latter trait, estimated in MZ and DZ pairs, suggested genetic influences on this attitude.

The rationale of a larger study has been recently defined, for both trying to validate former results regarding the ‘utilitarian/pragmatic’ attitude to donation, and also to assess the level of knowledge of research biobanks among all the ITR enrolled twins who have not yet donated to the Biobank, their specific concerns regarding hypothetical use of the biological samples and possible future re-use, their willingness to receive communication of individual results, their preference for a study-specific or a broadened informed consent, and their ideas about personal data protection.

Furthermore, the study will investigate possible correlates between a ‘positive attitude’ toward research donation and the overall well being of the individual (i.e., perceived health and psychological well being; Gigantesco et al., 2011); as well as possible correlates with the capacity to recognize feelings that are being experienced by another individual (i.e., empathy; Baron-Cohen & Wheelwright, 2004).

In February–March 2012, a self-administered structured questionnaire was sent by mail to 4,869 twins (2,433 pairs + one set of triplets) aged 18–65 years already enrolled in the ITR. The response rate by the end of July 2012 was around 33%. Data analysis is underway and results pertaining to the ethical aspects of the research will also have a practical impact on the outline of policies which are going to be adopted by the ITR Biobank.

Linkage of the ITR with Disease Registers

The great potential of the ITR has always been the ability to perform record linkage between data from clinical and disease registries (given the patient consent to be involved in medical research) and personal twin data (the 140,000 verified records from municipalities or the 600,000 unverified records from the Ministry of Finance held by the ITR itself). Within the framework of a multi-center clinical study, this procedure enables the identification of twins with a certain disease or medical condition, in order to estimate its heritability in a classical twin study design. This procedure, successfully carried out for the first time on celiac disease (Greco et al., 2002; Nisticò et al., 2006), has been applied in a recently published study on type 1 diabetes (Nisticò et al., 2012). The overall ITR database was linked to almost 15,000 type 1 diabetes records from 36 Italian pediatric diabetes care centers, and 173 diabetic twins were identified. Among them we were able to re-contact 104 diabetic twins with a mean age at diagnosis of 8.1 years, and their co-twins from 88 pairs (34 MZ, 54 DZ), and one triplet. Given proband-wise concordances of 45.5% in MZ and 16.4% in DZ pairs, we estimated a genetic contribution to type 1 diabetes susceptibility of 40%, an individual-specific environmental component of 9% and for the first time a
substantial shared environmental effect (51%); this suggests that in quite recently recruited cohorts, as opposed to earlier type 1 diabetes twin studies, exposures in fetal or early postnatal life may have played a major role on the occurrence of the disease and may have contributed to the increased incidence of these diseases observed in almost all of the Western world.

Although the classical twin study design still maintains its importance, the design based on detecting MZ disease-discordant pairs through the merge of the ITR databases and with disease registries is becoming more appealing. Beside the research on autoimmune diseases on which the ITR has historically focused, other clinicians, especially those interested in mental health and neuroscience, are very keen on using this resource, and various research projects have been recently submitted to our ethical committee.

Indeed, tracing and studying phenotypic discordant MZ twins is a powerful approach in the investigation of epigenetic variation for complex traits. In this regard, and according to national and international initiatives, the ITR is collecting information on MZ twins discordant for BMI in various age groups.

Conclusion

The very near future will see the ITR challenged on the enlargement of the total number of enrolled twins all over Italy. Given our sizeable population as compared to many other countries holding twin registries, we envisage we will be performing an increased amount of effective record linkage with disease databases and powerful cross-sectional surveys on health and social outcomes.

References


