Twin Research and Human Genetics (2020), **23**, 114–115
doi:10.1017/thg.2020.33

CAMBRIDGE UNIVERSITY PRESS

Article

Cohorts

Jaakko Kaprio[1,2] and Dorret I. Boomsma[3]

[1]Department of Public Health, University of Helsinki, Helsinki, Finland, [2]Institute for Molecular Medicine FIMM, University of Helsinki, Helsinki, Finland and [3]Netherlands Twin Register, Dept Biological Psychology, Vrije Universiteit, Amsterdam, the Netherlands

Abstract

Cohort studies are essential for conducting large studies of multiple exposures and outcomes in humans. Recently, the ability to combine data from multiple cohorts in, for example, meta-analyses, and the willingness in the genetics community to collaborate to enable replication studies has led to many new insights into the genetic and environmental determinants of human health and behaviors. The contribution of Professor Nicholas Martin to the development of cohort studies, particularly of twin and twin-family studies, over a period of several decades is reviewed. He has contributed to the development and use of both Australian and international resources. The contributions of Australian twin studies to genomewide association projects are multiple, and across multiple domains, from biomarkers, lifestyle and behavior to disorders and disease.

Keywords: twin; twin cohort; GWAS; complex traits

(Received 24 February 2020; accepted 25 March 2020; First Published online 26 May 2020)

Over the past five to six decades, our knowledge of the etiology of common, complex diseases has deepened enormously. Early epidemiological studies of coronary heart disease in the 1950s, such as the Framingham study and the international multisite Seven Countries study, identified major risk factors such as smoking, blood lipids and blood pressure. Other large-scale epidemiological studies of cardiovascular diseases, cancer and other common diseases followed over the next decades. In the course of these studies, the importance of family history became evident, and large-scale twin studies were established to enable the distinction between exposures and experiences shared by family members ('shared environment') and genetic factors common to family members.

While the first nationwide cohorts were established in Nordic countries (first Denmark, Sweden and then Norway and Finland), large cohorts of twins and their family members have been established in the Netherlands and Australia as well as in many other countries later on. Nick Martin established the Australian Twin Registry in 1978. He has been responsible for the development and expansion of twin and twin-family studies based at Queensland Institute of Medical Research (QIMR) Berghofer Medical Research Institute ever since. The importance of twin studies for medical and lots of other traits was emphasized in a key paper that he published in 1997 (Martin et al., 1997). As summarized in this review paper, multiple twin and twin-family studies confirmed the role of genetic factors for nearly every human trait, including those known to be risk factors for common diseases. As relatively rare conditions, twin studies of diseases from any single cohort were generally underpowered to provide reliable estimates of heritability except for the most common conditions. Combining data and analyses from individual cohorts permit more reliable estimates to be made, but also permit analysis of whether variance components differed by country or other aspects of the contributing cohorts, such as an analysis of smoking behavior from Australia, Finland and Sweden (Madden et al., 1999), which then led to the large family-based Nicotine Addiction Genetics study (Saccone et al., 2007). While such ad hoc studies had been done earlier, the GenomEUtwin study was the first large-scale effort to enable a pooled analysis of twin data of complex diseases and their risk factors.

The GenomEUtwin study undertook large-scale analyses of European twin and population cohorts to realize meta-analyses of heritability studies and to identify genes in common diseases and traits, including migraine, BMI, lipids and body height. Datasets were derived from eight twin cohorts from Istituto Superiore di Sanità, Italy, Karolinska Institutet, Sweden, University of Helsinki, Finland, Vrije Universiteit Amsterdam, Netherlands, University of Southern Denmark, Odense, Denmark, Norwegian Institute of Public Health, Norway, St Thomas Hospital, London, UK, and Queensland Institute of Medical Research, Australia as well as other centers that contributed expertise in analyses and genotyping. Though funded (13.7 million euros) by the EU through its Framework Programme 5 integrated project funding mechanism, agreements between the EU and Australia allowed Australian researchers to join. QIMR under Nick's leadership was an important component of the project. A first key set of papers were published in the October 2003 issue of *Twin Research*. The collaborative papers confirmed and extended knowledge of the genetic basis of these traits and became key cited papers. The heritability estimates were very close to each other in all these eight populations of European ancestry, despite quite divergent geographical, cultural and health system environments, for example, on migraine (Mulder et al., 2003). Many other analyses followed, and at the end of the project, large-scale genotyping of MZ pairs using the Illumina 370 chip was

Author for correspondence: Jaakko Kaprio, Email: jaakko.kaprio@helsinki.fi
Cite this article: Kaprio J and Boomsma DI. (2020) Cohorts. *Twin Research and Human Genetics* **23**: 114–115, https://doi.org/10.1017/thg.2020.33

conducted to study variability genes in lipid traits (Surakka et al., 2012).

GenomEUtwin laid the foundation for twin cohorts to contribute to the genomewide association studies (GWAS) era of studies. A major and early international effort in this area was the European Network for Genetic and Genomic Epidemiology (ENGAGE) project. Starting in 2008, the five year's project (www.euengage.org) was an EU-funded large-scale project to enable meta-analyses of GWAS and develop their methodology. QIMR and Nick Martin have played a major role in developing methods in genetic epidemiology, and he and his group made important contributions to ENGAGE (Aulchenko et al, 2009). At that time, multiple other consortia for conducting meta-analyses of GWAS data were formed, and cohorts that were contributing to ENGAGE were also in many of these other projects. Lessons on how to share summary statistics from individual GWAS analyses, efficiently meta-analyses and interpret results were learnt during these years and have enabled the stunning success of GWAS studies.

ENGAGE built and expanded on the experience of and accumulated trust in collaborative research. Large-scale collaborations in human genetics have been needed to identify the myriad contributing genes of small effect size in complex disease. The multifactorial, polygenic nature of common traits and diseases was predicted by quantitative genetic theory and empirically seen in twin and family studies. After GenomEUtwin and ENGAGE, Nick has continued to be very active in multiple molecular genetic consortia, several of which he initiated. Examples include the Twinning Genetics Consortium (TGC; http://www.twinningconsortium.org/), the Psychiatric Genomics Consortium (PGC; https://www.med.unc.edu/pgc/), Enhancing NeuroImaging Genetics through Meta-Analysis (ENIGMA; http://enigma.ini.usc.edu/) and Aggression in Children: Unraveling gene-environment interplay to inform Treatment and InterventiON strategies (ACTION; http://www.action-euproject.eu/).

Major human groups of European, African and Asian ancestry have distinct genetic differences, which contribute to differences in genetic risk to common diseases and traits between populations. As large twin studies have been established in many countries, comparisons of heritability estimates and variance components between regions of the world can provide insights into the interplay of genetics and environment. The COllaborative project of Development of Anthropometrical measures in Twins (CODATwins) project (Silventoinen et al., 2019) is a collaborative effort of 54 twin projects from 24 countries, including Australian data from QIMR Berghofer. Data on weight and height as well as relevant covariates are available on 489,981 twin individuals from both twins from 228,635 twin pairs. Though this is the largest single analysis of twin data to date, Australian data have importantly contributed to many other projects such as the effort to identify genes underlying human twinning — a topic that is dear to Nick.

In addition to these individual projects, Nick has been highly influential in the International Society for Twin Studies, notably as the editor in chief of *Twin Research and Human Genetics*. The journal has published theme issues on twin cohorts and registries in 2002, 2006, 2013 and 2019. These document the scope and impact of twin research and research about and of twins and other multiples, to which Nick has made important contributions.

References

Aulchenko, Y. S., Ripatti, S., Lindqvist, I., Boomsma, D., Heid, I. M., Pramstaller, P. P.,... Peltonen, L. (2009). ENGAGE Consortium. Loci influencing lipid levels and coronary heart disease risk in 16 European population cohorts. *Nature Genetics*, *41*, 47–55.

Madden, P. A., Heath, A. C., Pedersen, N. L., Kaprio, J., Koskenvuo, M. J., & Martin, N. G. (1999). The genetics of smoking persistence in men and women: A multicultural study. *Behavior Genetics*, *29*, 423–431.

Martin, N., Boomsma, D., & Machin, G. (1997). A twin-pronged attack on complex traits. *Nature Genetics*, *17*, 387–392.

Mulder, E. J., Van Baal, C., Gaist, D., Kallela, M., Kaprio, J., Svensson, D. A., ... Palotie, A. (2003). Genetic and environmental influences on migraine: A twin study across six countries. *Twin Research*, *6*, 422–423.

Saccone, S. F., Pergadia, M. L., Loukola, A., Broms, U., Montgomery, G. W., Wang, J. C, ... Madden, P. A. (2007). Genetic linkage to chromosome 22q12 for a heavy-smoking quantitative trait in two independent samples. *American Journal of Human Genetics*, *80*, 856–866.

Silventoinen, K., Jelenkovic, A., Yokoyama, Y., Sund, R., Sugawara, M., Tanaka, M., ... Kaprio, J. (2019). The CODATwins Project: the current status and recent findings of Collaborative Project of Development of Anthropometrical Measures in Twins. *Twin Research and Human Genetics*, *22*, 800–808.

Surakka, I., Whitfield, J. B., Perola, M., Visscher, P. M., Montgomery, G. W., Falchi, M. ... GenomEUtwin Project. (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.