

# Establishment of the Avera Twin Register in the Midwest USA

Julie Kittelsrud,<sup>1</sup> Erik A. Ehli,<sup>1</sup> Vikki Petersen,<sup>1</sup> Tammy Jung,<sup>1</sup> Gonneke Willemsen,<sup>2</sup> Dorret Boomsma,<sup>1,2</sup> and Gareth E. Davies<sup>1,2</sup>

<sup>1</sup>Avera Institute for Human Genetics, Sioux Falls, South Dakota, USA

<sup>2</sup>Department of Biological Psychology, Vrije Universiteit, Amsterdam, the Netherlands

The Avera Twin Register (ATR) aims to study environmental and genetic influences on health and disease using a longitudinal repository of biological specimens, survey data, and health information provided by multiples and their family members. The ATR is located in Sioux Falls, South Dakota, which is a rural and frontier area in the Midwestern United States with a density of four people per square kilometer. The target area of the ATR is South Dakota and the four surrounding states: Minnesota, Iowa, North Dakota, and Nebraska. Enrollment of twins and higher-order multiples of all ages and their family members started on May 18, 2016. A description of the first 13 months of enrollment in this longitudinal register will be provided. The ATR will collect longitudinal data on lifestyle, including diet and activity levels, aging, complex traits, and diseases. Upon registration, all participants are genotyped on the Illumina Global Screening Array (GSA) and twins and higher order multiples receive information on their zygosity. The ATR aims to contribute to large international GWAS consortia and collaborates closely with the Netherlands Twin Register, allowing for the comparison of collected data and analyses of results. In addition, the ATR will address twin-specific questions.

■ **Keywords:** twins, multiples, twin register, lifestyle, genotyping, Global Screening Array, GSA

The Avera Twin Register (ATR) began enrolling twins and multiples of all ages and their family members on May 18, 2016 by launching a media campaign and a website, <http://www.avera.org/innovation-research/institute-for-human-genetics/twin-register>. The media campaign announced both the ATR and an official partnership with the Netherlands Twin Register (NTR), Vrije Universiteit (VU), Amsterdam. The intention of the ATR is to establish a longitudinal repository of genetic samples and data of twins and their families to understand genetic influences of health and disease and to contribute valuable information for our region of twins and multiples as well as to the wider consortium of twin research. As a health system, Avera and its researchers' interests include the understanding of traits, accurate diagnosis, treatment, and disease prevention.

Previous collaborations between the NTR and Avera have focused on the incorporation of molecular genetic information into twin research, including copy number variation in monozygotic twins with attention deficit disorder (Abdellaoui et al., 2015; Ehli et al., 2012; Scheet et al., 2012), telomere repeat mass comparisons (TRM) between buccal and blood-derived DNAs (Finnicum et al., 2017), as well as studies to quantify optimizing genotyping assays using the

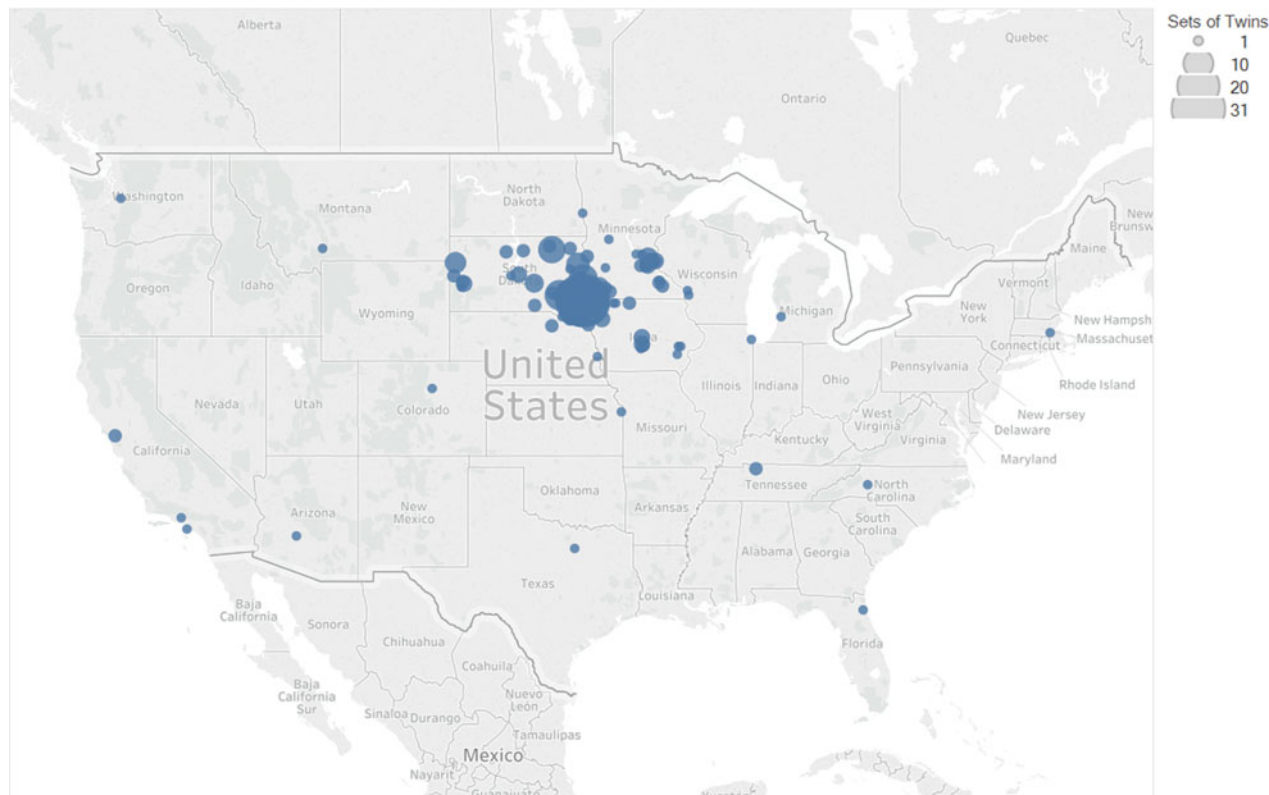
Axiom-NL array for risk profiling, new variant identification, and ancestry testing (Ehli et al., 2017).

Avera Health is based in Sioux Falls, with a population of approximately 178,000 people, and is the largest city in South Dakota. It is a regional, vertically integrated healthcare delivery network with a mission to make a positive impact on the lives and health of persons and communities. Avera Health is sponsored by the Sisters of St. Benedict of Sacred Heart Monastery of Yankton, South Dakota, and the Sisters of the Presentation of the Blessed Virgin Mary of Aberdeen, South Dakota Avera. The health system includes 330 locations in 100 communities, and the service area spans 71,550 square miles (185,313 kilometers) over five states (South Dakota, Minnesota, Iowa, Nebraska, and North Dakota) and includes a service population of 986,120 people. The health system has established the Avera

RECEIVED 14 July 2017; ACCEPTED 17 July 2017. First published online 22 August 2017.

ADDRESS FOR CORRESPONDENCE: Julie Kittelsrud, Avera Institute for Human Genetics, 3720 W. 69th Street, Sioux Falls, SD 57108, USA. E-mail: [Julie.Kittelsrud@Avera.org](mailto:Julie.Kittelsrud@Avera.org)

## National Twin Registrants - Excludes Alaska

**Figure 1**

(Colour online) A map depicting the geographic location of participants enrolled in the Avera Twin Register based on zip codes reported during enrollment. The size of the circles indicate the number of twins from the respective zip code. One participant in Alaska was excluded from the map.

Institute for Human Genetics (AIHG), a molecular genetics laboratory, with the facility to complete both small-scale pilot studies and large-scale high-throughput genome-wide association studies (GWAS). The AIHG provides a complete personalized pharmacogenomic service, which includes pain, psychotropic, and cardiovascular medication management. As participants enroll, the AIGH will process genetic materials and provide genotype information for all ATR participants.

Recruitment is through the Avera Health system and through word of mouth in the communities it services. The ATR is enrolling twin families throughout the United States with a focus on South Dakota and the bordering states of North Dakota, Iowa, Nebraska, and Minnesota (Figure 1). Future ATR studies will address questions about comparisons between the Netherlands twin population and the twin populations in the Midwestern United States. South Dakota spans a primarily rural and frontier area with a population of 865,545 individuals and a density of four people per square kilometer. North Dakota and Nebraska have similar population densities, 4/km<sup>2</sup> and 9/km<sup>2</sup> respectively, while Iowa (21/km<sup>2</sup>) and Nebraska (26/km<sup>2</sup>) have more people per square kilometer (United States Census

Bureau, 2010; United States Census Bureau, 2016). These states are all characterized by a rural lifestyle and an average life expectancy for males of 77.1 years and for females, 82.0 years of age (Institute for Health Metrics and Evaluation, 2017; Segal et al., 2016). The total population for this Midwestern region is 9,011,510 (United States Census Bureau, 2010). In contrast, the Netherlands is a small densely populated country with 17,032,845 inhabitants and has a population density of 406 per square kilometer (United Nations, 2017). Life expectancy in 2015 in the Netherlands is 80.0 years for males and 83.6 years for females. Interestingly, South Dakota, North Dakota, Minnesota, and Iowa have been shown, based on self-reported ancestry and verified via common genetic variation, to have genetic ancestry within two to nine generations from Western European countries, including the Netherlands, (Bryc et al., 2015). Descendants in southeastern South Dakota, southern Minnesota, and northern Iowa also have heritage from Italy, Ireland, and Scotland (Bryc et al., 2015). Comparisons between the Dutch population participating in the NTR and the ATR populations may allow us to gain insights into environmental and genetic influences as families in these registries come from genetically similar ancestry.

## Aims of the Avera Twin Register (ATR)

There are three objectives for the establishment of the ATR. The first objective is to formulate connections with the community to develop and enroll a prospective longitudinal repository of biological specimens, survey data, and health information from all ages of twins or multiples, and their non-twin family members throughout the Avera Health System and the Midwestern United States. Second, the ATR seeks to study environmental and genetic influences on health and disease in the ATR population in collaboration with the NTR, resulting in American–European comparisons of data and results and contributions to international consortia to study genome-wide associations. Third, the ATR seeks to serve the needs of twins and twin families by addressing questions through the ‘twins for twins’ approach, which focuses on questions that are exclusive to twins, such as those related to their zygosity, language development, separation in schools, and long-term health outcomes (Segal, 2017).

## Materials and Methods

The study was approved by the Avera Institutional Review Board and the Avera Department of Human Subject’s protection. The first announcement of the start of the ATR was made via local media, which initiated recruitment by snowball sampling. All parents and family members of twin and multiples born within the Avera system are invited to participate in the Register. Participants who live in the vicinity of Sioux Falls, South Dakota, are invited to come to the AIHG office to meet with researchers who conduct face-to-face informed consent and coordinate the collection of samples and questionnaire data. When participants live too far away to come to the AIHG office, the process is completed at a distance. In that case, participants are mailed a consent form and schedule a time to review the document and answer their questions via phone. Once participants have mailed back a completed informed consent, they are sent an enrollment kit. The enrollment kit includes a copy of their signed informed consent, the baseline questionnaire, and a buccal swab collection kit with instructions. The buccal cell samples collected at registration are analyzed via Global Screening Array (GSA) for zygosity, with same-sex twins or higher order multiples receiving results. Extra DNA from their samples are stored for future studies.

During the informed consent process, participants are given information about the longitudinal nature of the register and are informed that their participation may include yearly questionnaires plus the option of volunteering in specific short-term studies or cross-sectional studies, which will be conducted as substudies of the register. As part of this process, participants are informed that requests for other biological samples may be sought to be used for specific cross-sectional research projects. The initial questionnaire includes several topics such as health information and

**TABLE 1**  
Background Information on ATR Participants from the First Survey

Information	Data
Demographics	Contact information, race, living conditions
Family information	Relationship to twin, eye color, handedness.
Twin or multiple characteristics	Zygosity, birth information, physical characteristics of twins
Physical health	General health, height, weight, sleep, health status
Education and occupation	Educational background, type of employment, SES
Religion	If religious and what religion
Sports and Physical Activity	Activity level, sports participation
NEO-PI	Personality inventory
For women	Questions related to childbirth and fertility

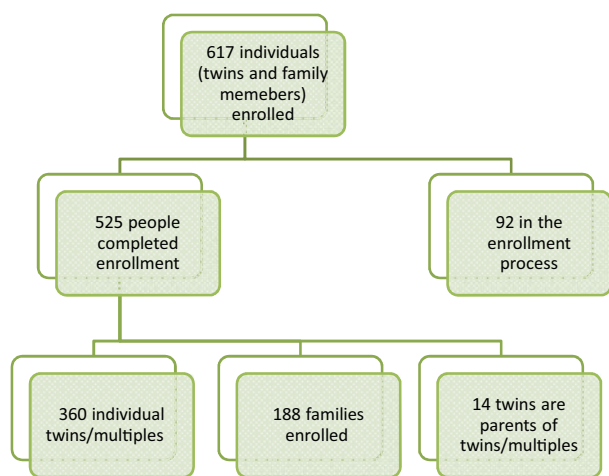
a personality assessment, the Revised NEO Personality Inventory (Costa et al., 1995). For an overview of the data collected in the first survey, refer to Table 1. Participants are asked if they would like to receive periodic updates about the registry, such as newsletters and emails. Those who respond affirmatively provide contact information and those participants who decline further contact from the ATR will participate in only cross-sectional study designs.

## Whole Genome Single Nucleotide Polymerase (SNP) Typing and Zygosity Assessment

All ATR participants who give consent for DNA collection and genotyping upon registration are genotyped by AIHG. Genotyping and zygosity assessments are completed utilizing buccal cell-derived DNA. The AIHG has joined the Infinium GSA Consortium initiated by Illumina, Inc. in 2016, in order to gain access to a customized array containing highly predictive markers for translational research applications. In conjunction with collaborators at the NTR, the Avera-NTR GSA array was designed to include a core imputation backbone (approximately 660,000 markers) and thousands of other custom add-on markers of interest concerning Alzheimer’s disease, drug metabolism, fertility, twinning, and schizophrenia (approximately 40,000 markers). Initial validation of the array content for imputation was assessed by checking concordance, coverage, and imputation quality using the Genome of the Netherlands reference sequence (Ehli et al., 2017). A database of SNP genotypes for each ATR participant will be included in potential GWAS.

## Results

During the first 13 months of enrollment, May 18, 2016 to June 19, 2017, no active advertising was conducted outside of the initial announcement made via local media coverage. Participants were referred to the ATR through a number of avenues. There were 360 people who completed an



**Figure 2**

(Colour online) A diagram depicting the recruitment of twins, multiples, and families into the Avera Twin Register (ATR). In total there are 617 participants including twins/multiples and their family members. This number includes 360 twins and multiples, 188 families, 14 twins who are also parents of twins, and 92 participants in the process of enrollment.

online interest form, located on the Avera website. Interested parties were then contacted either with a follow-up phone call or emailed to schedule an appointment or send an informed consent to start the enrollment process. Thirty-two forms were not fully completed and no contact information was provided. A cohort of mothers contacted the ATR because they had heard about the ATR through a twin specific mother's group. Other ways in which participants learned about the ATR was through employees of the Avera Health system, through reading the ATR twin brochures that are provided in obstetric offices and hospital units, and additional local media coverage of the experiences of MZ twins who had enrolled in the study.

The ATR has enrolled a total of 617 participants, including 360 individuals identifying as a twin ( $n = 180$  twin pairs). This comprises multiples between the ages of 6 months to 88 years old ( $M = 28.33$ ,  $SD = 24.25$ ). One hundred and seventy-six twin participants are monozygotic (56.6%), while  $n = 74$  are dizygotic (23.8%), and the remaining ( $n = 61$ ) are in the process of completing the enrollment procedures for the register. There are three sets of triplets enrolled ( $n = 9$ ), with 14 individuals identifying themselves as a twin or a parent of a twin or multiple (Figure 2). Two hundred and forty-two family members (59.9% female) have completed enrollment, including the first questionnaire and donation of a buccal sample for DNA isolation and zygosity testing. Family members have identified themselves as parents (60.7%), siblings (24.8%), children of twins (8.3%), grandparents and aunts (3.7%), and adoptive or foster siblings (2.5%). Thus far, 100% of enrolled participants have requested to be contacted for future studies.

## Future Directions

The ATR continues to include new participants while also asking for completion of ongoing questionnaires from those participants already registered. Within half a year from publication of this article, ATR participants will be completing their next annual surveys, establishing this register as a prospective study with repeated assessments of twin families. The next survey will include information on activity levels and dietary preferences, and updates to personal information. Historically, immigration from European countries in the United States shows that the states of South Dakota, Iowa, Minnesota, and Nebraska have ancestry with roots from the Northern Europe. This provides an interesting genetic and environmental comparison between our region and the Netherlands population today, as there are some cultural practices that are very different between the two countries. For example, most transportation is done on a bicycle or walking in the Netherlands, while in South Dakota and much of the Midwestern United States, most transportation occurs by car. The Dutch also have different diet preferences than those in South Dakota. Such dissimilarities may be related to differences in body mass index (BMI). South Dakota and the surrounding area has an obesity rate of about 30% for both males and females (Institute for Health Metrics and Evaluation, 2017), while the Netherlands had obesity rates of 12% in 2009 (World Population Review, 2017) with a predicted decrease in 2030 to 8.5% according to the World Health Organization (WHO). Identifying the differences in activity levels and eating habits, while also having genotypic information, allows for a thorough investigation of environmental and genetic influences in health and disease.

Since all ATR participants are genotyped upon registration and twin zygosity is determined, one of the first GWA projects will be participation in the 'genetics of twinning' consortium (<http://www.twinningconsortium.org/>) which continues its work on the first meta-analysis of 'being a mother of DZ twins' and will extend its GWA efforts to also include 'being a mother of MZ twins' and 'being an MZ or DZ twin' (Mbarek et al., 2016).

## Acknowledgments

The ATR is supported by Avera Health, Avera McKennan Hospital, and AIHG. The collaboration between the Netherlands and the ATR arose through NIH Grant: 1RC2MH089995-01: Genomics of Developmental Trajectories in Twins.

## References

- Abdellaoui, A., Ehli, E. A., Hottenga, J. J., Weber, Z., Mbarek, H., Willemsen, G., ... Boomsma, D. I. (2015). CNV Concordance in 1,097 MZ Twin Pairs. *Twin Research and Human Genetics*, 18, 1–12.

- Bryc, K., Durand, E. Y., Macpherson, J. M., Reich, D., & Mountain, J. L. (2015). The genetic ancestry of African Americans, Latinos, and European Americans across the United States. *American Journal of Human Genetics*, *96*, 37–53.
- Costa, P. T. Jr., McCrae, R. R., & Kay, G. G. (1995). Persons, places and personality: Career assessment using the Revised NEO Personality Inventory. *Journal of Career Assessment*, *3*, 123–139.
- Ehli, E. A., Abdellaoui, A., Fedko, I. O., Grieser, C., Nohzadeh-Malakshah, S., Willemsen, G., ... Hottenga, J. J. (2017). A method to customize population-specific arrays for genome-wide association testing. *European Journal of Human Genetics*, *25*, 267–270.
- Ehli, E. A., Abdellaoui, A., Hu, Y., Hottenga, J. J., Kattenberg, M., van Beijsterveldt, T., ... Davies, G. E. (2012). De novo and inherited CNVs in MZ twin pairs selected for discordance and concordance on Attention Problems. *European Journal of Human Genetics*, *20*, 1037–1043.
- Finnicum, C. T., Dolan, C. V., Willemsen, G., Weber, Z. M., Petersen, J. L., Beck, J. J., ... Ehli, E. A. (2017). Relative telomere repeat mass in buccal and leukocyte-derived DNA. *PLoS One*, *12*, e0170765.
- Institute for Health Metrics. (2017, May 3). US data for download. Retrieved from <http://www.healthdata.org/us-health/data-download>.
- Mbarek, H., Steinberg, S., Nyholt, D. R., Gordon, S. D., Miller, M. B., McRae, A. F., ... Boomsma, D. I. (2016). Identification of common genetic variants influencing spontaneous dizygotic twinning and female fertility. *American Journal of Human Genetics*, *98*, 898–908.
- Scheet, P., Ehli, E. A., Xiao, X., van Beijsterveldt, C. E., Abdellaoui, A., Althoff, R. R., ... Boomsma, D. I. (2012). Twins, tissue, and time: An assessment of SNPs and CNVs. *Twin Res Hum Genet*, *15*, 737–745.
- Segal, L. M., Martin, A., & Rayburn, J. (2016, June 22). The state of obesity: Better policies for a healthier America. *Trust for Americas Health*. Retrieved from <http://stateofobesity.org/files/stateofobesity2016.pdf>.
- Segal, N. (2017). *Twin mythconceptions: False beliefs, fables, and facts about twins*. London, UK: Academic Press.
- United Nations DESA/Population Division. (2017, June 21). World population prospects 2017. Retrieved from <https://esa.un.org/unpd/wpp/Download/Standard/Fertility/>.
- United States Census Bureau. (2010, March 31). 2010 Census summary files. Retrieved from <https://www.census.gov/data.html>.
- United States Census Bureau. (2016, May 3). Quick facts. Retrieved from <http://www.census.gov/quickfacts/table/PST045216/46>.
- World Population Review. (2017, June 21). Netherlands population. Retrieved from <http://worldpopulationreview.com/countries/netherlands-population/>.
-