Twin Research and Human Genetics (2022), **25**, 63–66
doi:10.1017/thg.2022.14

CAMBRIDGE UNIVERSITY PRESS

Article

The Heritability of Twinning in Seven Large Historic Pedigrees

David L. Duffy and Nicholas G. Martin
QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia

Abstract

It is widely recognized that dizygotic twinning (DZT) runs in families, but estimates of heritability from twin and family data are remarkably scarce and vary considerably. Here, we traced seven large, sometimes historical, multigeneration pedigrees from West Africans, fin de siècle French Jews, Canadians (two pedigrees), and the French royal family, in which twin births were recorded. We estimated heritability of twinning (of all types) as zygosity information was not available, diluting the true DZT heritability by a third or so. The estimates in the range 8−20% are remarkably consistent across time (8−19 generations) and ethnicities and also consistent with twin and family estimates.

Keywords: Dizygotic twins; twinning; heritability; fertility; multifactorial threshold model

(Received 30 March 2022; accepted 30 March 2022; First Published online 20 May 2022)

That twins run in families is well-known folk wisdom. Less well understood is that, with a few rare exceptions, monozygotic (MZ) twinning is largely a stochastic event, so that the familial nature of twinning is almost entirely due to dizygotic (DZ) twins running in families. There have been a few attempts to explain the inheritance of DZ twinning (DZT) in terms of a single gene with incomplete penetrance (e.g. Bulmer, 1970; Davenport 1920a, 1920b; Meulemans et al., 1996). More likely, in view of the lessons from recent advances in complex trait genetics due to the advent of microarray genotyping and genome-wide association studies (GWAS) is that the familial aspect of DZT is due to multifactorial inheritance — the combined effects of many different genes — and the first such GWAS has been completed and provides support for this hypothesis. Regardless of the truth of this, which is beyond the scope of this article, a starting point for all enquiries in genetic epidemiology is an estimate of the total heritability of the trait.

Here, we report new estimates using an entirely novel source of data — seven large historic pedigrees from Africa and Europe collected for anthropological purposes, which have twin births reported going back many generations. A weakness is that zygosity of same-sex twins is not available so any estimates of heritability we obtain will be diluted by the inclusion of MZ twins which are, as discussed, largely nonheritable.

Data

Data were obtained from www.kinsources.org (Christifoli & Garcia-Fernandez 2016), a repository for archiving, analyzing, and comparing kinship data for scientific use. We have analyzed those datasets where a date of birth or year of birth has been included, and diagnosed twin pairs as offspring with the same (or next) day of birth, or same year of birth. Obviously there may be some incorrect diagnoses of twin birth in the datasets providing only year of birth. The seven datasets used are numbered below as they are listed in Table 1.

1. The Watchi dataset is a corpus of genealogical, residential, and related data stemming from ongoing fieldwork research conducted since 2004 in the Watchi-speaking area of South-East Togo (Maritime Region, prefectures of Bas-Mono, Yoto and Vo, and partly Lacs and Golfe), with a special focus on the village of Afagnan-Gbléta in the Bas-Mono prefecture (6.5000 N, 1.6167 E). Twin ties are very frequent in this region of West Africa where twins are venerated as voodoos (https://www.kinsources.net/kidarep/dataset-275-watchi-2017.xhtml).
2. The Bwa are an ethnic group living in central Burkina Faso and south-east Mali (https://www.kinsources.net/kidarep/dataset-308-bwa-slam-1988-2009.xhtml).
3. This dataset is a collection of genealogies for the study of demographic and genetic evolution of an isolated Dogon population around Boni in Mali (https://www.kinsources.net/kidarep/dataset-186-dogon-boni.xhtml).
4. This genealogical dataset was assembled by Cyril Grange and includes members of the major Jewish financial dynasties in Paris at the end of the 19th century; it covers the period extending from 1750 to 1950. It is limited to those families mentioned in the Who's Who of the 1899 Livre d'Or des Salons (https://www.kinsources.net/kidarep/dataset-125-dynasties-juives.xhtml).
5. Genealogy of the Parish of Saint-Cathérine de Fossambault (Quebec) collected by Caroline Legrand. Based on parish registers (https://www.kinsources.net/kidarep/dataset-93-sainte-catherine.xhtml).
6. Genealogies of the inhabitants of a village in Charlevoix (Québec). Collected by Chantal Collard (2014). Genealogies of the inhabitants of a village in Charlevoix (Québec). (https://www.kinsources.net/kidarep/dataset-115-charlevoix.xhtml).

Author for correspondence: Nick Martin, Email: Nick.Martin@qimrberghofer.edu.au
Cite this article: Duffy DL and Martin NG. (2022) The Heritability of Twinning in Seven Large Historic Pedigrees. *Twin Research and Human Genetics* **25**: 63–66, https://doi.org/10.1017/thg.2022.14

© The Author(s), 2022. Published by Cambridge University Press on behalf of International Society for Twin Studies

https://doi.org/10.1017/thg.2022.14 Published online by Cambridge University Press

Table 1. Heritability estimates for being a mother of twins (zygosity unknown) in seven large multigenerational pedigrees

Dataset	*N*	Gen	SS	OS	MOT	h^2 95% CI	LMM-h^2
Watchi (Togo)[1]	38,587	18	173	118	263 (1.3%)	14% (5–24)	8%
Bwa (Mali)[2]	14,883	18	155	110	220 (2.9%)	8% (3–15)	3%
Dogon (Mali)[3]	10,906	16	102	53	130 (2.4%)	16% (6–27)	4%
Juives (France)[4]	9075	19	214	23	85 (1.9%)	13% (11–25)	5%
St Catherine (Canada)[5]	9595	8	63	52	100 (2.2%)	10% (3–22)	4%
Charlevoix (Canada)[6]	3688	16	17	10	26 (1.4%)	19% (4–40)	6%
Ancien (European)[7]	4810	14	14	4	17 (0.7%)	20% (4–42)	12%

Note: *N*, sample size; Gen, generations; SS, same-sex twin pairs; OS, opposite-sex twin pairs; MOT, mother of twins (number, as percentage of all mothers); h^2, heritability; LMM, linear mixed model. Zygosity is unknown.
Numbered references refer to pedigree descriptions in the text (Methods).

7. Public domain genealogy of European high nobility of the 16th −18th century. Data are quasi-complete. The genealogy contains the complete cognatic (ascending and collateral) kin of Louis XVI and Marie Antoinette up to the 6th canonical degree (https://www.kinsources.net/kidarep/dataset-53-ancien-regime.xhtml).

Statistical Methods

The proportion of same-sex twin pairs to opposite-sex twin pairs has traditionally been used to estimate the proportion of MZ twins among all pairs (method of Weinberg [1909] — number of same-sex DZ twin pairs should equal that of the opposite-sex pairs, so number of MZ twin pairs is total less twice the number of opposite-sex twin pairs). We have not explicitly incorporated this information in the genetic analyses (in a missing data type framework), as the datasets are not large enough. Deviations in this ratio from dataset to dataset may reflect differences in population DZ twinning rate (given MZ rates are relatively constant worldwide), but also may point to data quality issues, at the very least in correct identification of the sexes for a particular sample.

We have estimated the heritability of the trait of a twin maternity by a variance components model that includes polygenic additive effects. These unobservable effects are assumed to be correlated between pedigree members proportionately to the pairwise kinship.

Estimating the heritability of binary traits from large pedigrees encounters analytic and computational difficulties. One simple approach is to analyze the data by the standard linear mixed model (LMM), coding mothers of twins as 1 and mothers of singletons as 0. As long as the trait lifetime risk is low, and heritability not too high, then this does not lead to numerical difficulties, and can be fitted using standard software. Obviously, this assumes that the effects of genotypes at each locus interact additively and linearly on the (raw) probability of expressing the trait. This has been the most commonly applied method to analyze GWAS data. A more elaborate model is used to fit a generalized linear mixed model (GLMM), using either a logistic or Gaussian link function, where the latter is equivalent to the multifactorial threshold model (MFT). This requires numerical multidimensional integration up to the size of the kinship matrix and so various approximations are used to solve this problem. The probit and logit links usually give quantitatively similar parameter estimates. These assume genotypes interact multiplicatively to increase risk, but can be flexibly fitted over the full range of trait risk and heritability. In many biological examples, a multiplicative interaction model for binary trait risk is better fitting than a linear model.

Here, we have calculated heritability by three methods: (1) the LMM using WOMBAT (Meyer, 2007); (2) the logit-normal GLMM using the R AnimalINLA (Holand, 2013); (3) probit-normal (and logit-normal) using MCMCglmm (Hadfield, 2010), which was only feasible for the smaller pedigrees. We fitted models with and without year of birth and generation as fixed effects, but in all cases, the effects of this adjustment on the heritability estimates were small. We provide the LMM Wombat results mainly because these are on the same scale as GWAS-based heritability estimates. We provide 95% confidence intervals for the GLLM heritability estimates from AnimalINLA as these are quantiles from the estimated marginal distribution of the binomial heritability hyperparameter (Gómez-Rubio, 2020; Rue et al., 2009).

We calculated the expected proportional difference between the DZT heritability and 'total' twinning heritability under the MFT by estimating maternal recurrence for each twin type for a given risk and tetrachoric correlation, then combining the resulting relative-pair fourfold tables under the assumption that risks of each twinning type are independent. We took the MZ twinning rate at 5 per 1000, the DZ twinning rate at 10 per 1000, and the heritability of MZ twinning as zero. When the observed total twinning heritability is 20%, the underlying DZT heritability is 28%.

Results

Results are presented in Table 1. Listed are the total number of individuals in each pedigree, the number of generations recorded, the number of same-sex and opposite-sex twin pairs and percentage of mothers of twins in the pedigree as a proportion of all women who have given birth.

The ratio of same-sex to opposite-sex twins differs significantly across the different samples, but this is completely due to the 'Juives' sample, where only 10% of twin pairs are of opposite sex (including all samples, $\chi^2_6 = 82.8$, $p = 9e\text{-}16$; excluding sample 4, $\chi^2_5 = 6.3$, $p = .28$).

It is widely known that the heritability estimates from the LMM are always lower than those from a GLMM, and we see this for all these samples as well. The GLMM heritabilities all lie between 8% and 20%, and all are significantly different from zero. The 95% confidence intervals from even the largest dataset ('Watchi') extend upwards nearly two-fold above the point estimate, confirming the large sample sizes required for this type of analysis. The

https://doi.org/10.1017/thg.2022.14 Published online by Cambridge University Press

Table 2. Overall population twinning rate for each of seven samples. African twinning rates are 150–200% higher than European rates

Dataset	Twinning rate (twin births out of total maternities)	Contemporaneous country-level twinning rate
Watchi (Togo)[1]	0.01169 (363/31061)	0.0214 (Smits & Monden, 2011)
Bwa (Mali)[2]	0.02754 (323/11729)	0.0169 (Smits & Monden, 2011)
Dogon (Mali)[3]	0.01604 (165/10287)	0.0169 (Smits & Monden, 2011)
Juives (France)[4]	0.02096 (140/6679)	0.010 (Pison & D'Addato, 2006)
St Catherine (Canada)[5]	0.02287 (139/6077)	0.0091 (1974) (Millar et al., 1992)
Charlevoix (Canada)[6]	0.00990 (32/3233)	0.0091 (1974) (Millar et al., 1992)
Ancien (European)[7]	0.00582 (23/3955)	0.010 (Pison & D'Addato, 2006)

Note: Numbered references refer to pedigree descriptions in the text (Methods). Note that for datasets 4 and 7, twinning rates used are for France in 1860, as read from Figure 1 of Pison and D'Addato (2006).

heritability from the 'Juives' is the third lowest, so one would infer that the large number of putative same-sex twins may represent mislabelling of sexes (Weinberg's formula would interpret the reported numbers as implying 81% of the sample twins being MZ). Table 2 compares the twinning rates in the present samples with comparable country level data from other sources. The 'Juives' dataset rate seems higher than overall French rates for that era, as is true of the other outlier, the 'St Catherine' dataset. It is possible that either could represent a founder effect (mean inbreeding coefficient 'Juives' $F = 0.003537$; 'St Catherine' $F = 0.00013$).

Discussion

We have estimated heritabilities for giving birth to twins (of either zygosity) from an online repository containing seven unique large multigeneration pedigrees from West Africans, fin de siècle French Jews, Canadians, and the French royal family, in which twin births were recorded, though not zygosity. Our estimates in the range 8–20% are remarkably consistent across time (8–19 generations) and ethnicities. Our estimates are comparable with estimates of biometric (total genome) MFT heritability of 15–18%, based on the same 1% prevalence and recurrence risks of DZ twinning in sisters (1.67) and female cousins (1.56–1.86) as reported in Lewis et al. (1996). These estimates are lower than those derived from Bulmer (1970; his Table 6.6) where sister–sister recurrence data are compatible with a tetrachoric correlation in liability of 0.14 so a heritability of 28%. Our results are also compatible with recurrence data of DZ twinning in MZ twins (Lewis et al., 1996); in 41 Australian and Finnish MZ female pairs where at least one had DZ twins, between 0 and 2 cotwins also had a DZ pregnancy, equivalent to a liability correlation (and heritability) between 0 and 26%. Note that these samples are either historical or from the developing world, so that assisted reproductive techniques are not affecting any of these populations. Given the inclusion of MZ twins in the large pedigrees analyzed here, we estimate that our estimates of heritability for total twinning are 32–45% lower than the true DZT heritability and this upward adjustment would put them in the same range as those from the other studies cited above.

Our estimate of the heritability of DZ twinning is probably lower than of its principal antecedent of multiple ovulation. This is hard to measure in humans in samples large enough to estimate heritability (Martin et al, 1991), but in sheep, the heritability of ovulation rate, measured by counting corpora lutea at laparoscopy, is considerably higher (0.34–0.58; Dewi et al., 1996; Hanrahan, 1990) than for twinning. Mechanistically, a large number of stochastic factors can intervene to cause one or both fetuses to be lost prior to term after multiple implantation.

Nevertheless, our estimate is sufficiently high to support hopes of finding the particular genes responsible by GWAS. One such has already been done, and using fewer than 2000 mothers of DZ twin 'cases', two significant genes were found — *FSHB* and *SMAD3* (Mbarek et al., 2016). This has given the impetus for an expanded study with much larger sample size in the hope of further unravelling the causes of DZ twinning, albeit initially only in Europeans. The estimates for heritability in three large African pedigrees analyzed here are entirely comparable with what we see in the four European ones and challenge us to undertake GWAS in African populations, hopefully to shed light on the much higher DZ twinning rate in that continent: 20/1000 births compared with 2/1000 in east Asians and 8/1000 in Europeans.

References

Bulmer, M. G. (1970). *The biology of twinning in man.* Clarendon Press.

Christofoli, P. K., & Garcia-Fernandez, A. (2016). Kinsources.net: Archiver et diffuser les données de la récherche en SHS. *La Lettre de L'INSHS*, *43*, 17–20.

Davenport, C. B. (1920a). Heredity of twin births. *Proceedings of the Society for Experimental Biology and Medicine, 17*, 75–77.

Davenport, C. B. (1920b). Influence of the male in the production of human twins. *The American Naturalist*, *54,* 122–129.

Dewi, I. A., Ap Dewi, I., Owen, J. B., El-Sheikh, A., Axford, R. F. E., & Beigi-Nassiri, M. (1996). Variation in ovulation rate and litter size of Cambridge sheep. *Animal Science*, *62,* 489–495.

Gómez-Rubio, V. (2020). *Bayesian inference with INLA.* Chapman & Hall/CRC Press.

Hadfield, J. D. (2010). MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *Journal of Statistical Software*, *33*, 1–22.

Hanrahan, J. P. (1990). Effect of selection for increased litter size on ovulation rate and embryo survival in sheep. *Proceedings of the British Society of Animal Production (1972),* 1990, 32–32.

Holand, A. M., Steinsland Martino, S., & Jensen, H. (2013) Animal models and integrated nested Laplace approximations. *G3 (Bethseda)*, *3*, 1241–1251.

Lewis, C. M., Healey, S. C., & Martin, N. G. (1996). Genetic contribution to DZ twinning. *American Journal of Medical Genetics*, *61*, 237–246.

Martin, N. G., Shanley, S., Butt, K., Osborne, J., & O'Brien, G. (1991). Excessive follicular recruitment and growth in mothers of spontaneous dizygotic twins. *Acta Geneticae Medicae et Gemellologiae*, *40*, 291–301.

Mbarek, H., Steinberg, S., Nyholt, D. R., Gordon, S. D., Miller, M. B., McRae, A. F., Hottenga, J. J., Day, F. R., Willemsen, G., de Geus, E. J., Davies, G. E., Martin, H. C., Penninx, B. W., Jansen, R., McAloney, K., Vink, J. M.,

https://doi.org/10.1017/thg.2022.14 Published online by Cambridge University Press

Kaprio, J., Plomin, R., Spector, T. D., ... **Boomsma, D. I.** (2016). Identification of common genetic variants influencing spontaneous dizygotic twinning and female fertility. *American Journal of Human Genetics*, *98*, 898–908.

Meulemans, W. J., Lewis, C. M., Boomsma, D. I., Derom, C. A., Van den Berghe, H., Orlebeke, J. F., Vlietinck, R. F., & Derom, R. M. (1996). Genetic modelling of dizygotic twinning in pedigrees of spontaneous dizygotic twins. *American Journal of Medical Genetics*, *61*, 258–263.

Meyer, K. (2007). WOMBAT: A tool for mixed model analyses in quantitative genetics by restricted maximum likelihood (REML). *Journal of Zhejiang University Science B: Biomedicine & Biotechnology*, *8,* 815–821.

Millar, W. J., Wadhera, S, & Nimrod C. (1992). Multiple births: trends and patterns in Canada, 1974–1990. *Health Reports*, *4,* 223–250.

Pison, G., & D'Addato, A. V. (2006). Frequency of twin births in developed countries. *Twin Research and Human Genetics*, *9*, 250–259.

Rue, H., Martino, S., & Chopin, N. (2009) Approximate Bayesian inference for latent Gaussian models using Integrated Nested Laplace Approximations. *JRSS-series B (with discussion)*, *71,* 319–392.

Smits, J., & Monden, C. (2011). Twinning across the developing world. *PloS One*, *28*, e25239.

Weinberg, W. (1909). Zur Bedeutung der Mehrlingsgeburten für die Frage der Bestimmung des Geschlechts. *Arch Rass Gesellschaftsbiol*, *6,* 28–39.

https://doi.org/10.1017/thg.2022.14 Published online by Cambridge University Press