

Article

Variance Components Models for Analysis of Big Family Data of Health Outcomes in the Lifelines Cohort Study

Nino Demetrashvili^{1,2}, Nynke Smidt¹, Harold Snieder¹, Edwin R. van den Heuvel^{1,3} and Ernst C. Wit^{4,5}

¹Department of Epidemiology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands, ²Department of Medical Statistics, National Center for Disease Control and Public Health, Tbilisi, Georgia, ³Department of Mathematics and Computer Science, Eindhoven University of Technology, Eindhoven, The Netherlands, ⁴Bernoulli Institute, University of Groningen, Groningen, The Netherlands and ⁵Institute of Computational Science, Università della Svizzera italiana, Lugano, Switzerland

Abstract

Large multigenerational cohort studies offer powerful ways to study the hereditary effects on various health outcomes. However, accounting for complex kinship relations in big data structures can be methodologically challenging. The traditional kinship model is computationally infeasible when considering thousands of individuals. In this article, we propose a computationally efficient alternative that employs fractional relatedness of family members through a series of founding members. The primary goal of this study is to investigate whether the effect of determinants on health outcome variables differs with and without accounting for family structure. We compare a fixed-effects model without familial effects with several variance components models that account for heritability and shared environment structure. Our secondary goal is to apply the fractional relatedness model in a realistic setting. Lifelines is a three-generation cohort study investigating the biological, behavioral, and environmental determinants of healthy aging. We analyzed a sample of 89,353 participants from 32,452 reconstructed families. Our primary conclusion is that the effect of determinants on health outcome variables does not differ with and without accounting for family structure. However, accounting for family structure through fractional relatedness allows for estimating heritability in a computationally efficient way, showing some interesting differences between physical and mental quality of life heritability. We have shown through simulations that the proposed fractional relatedness model performs better than the standard kinship model, not only in terms of computational time and convenience of fitting using standard functions in R, but also in terms of bias of heritability estimates and coverage.

Keywords: Determinants and health outcome; founders and non-founders; fractional relatedness of family members; genetic and environmental factors; heritability and its confidence interval; kinship model; mental and physical health scores; mixed effects models

(Received 27 November 2018; accepted 29 November 2018)

Excessive weight, especially obesity, is a major public health concern in the western world. Around two-thirds of the adult population in the USA and at least half of the population of many European countries are currently overweight or obese (Berghöfer et al., 2008; Flegal et al., 2010; Wang et al., 2008). In the Netherlands, the prevalence of overweight is 48.3% (Volksgezondheid, Nationaal Kompas, 2012), and 12.7% of the Dutch population is classified as obese. It is known that obesity — defined as a BMI of 30 or more — is a major risk factor for many chronic diseases, such as hypertension, stroke, coronary heart disease, diabetes, arthritis and overall mortality (Flegal et al., 2013; Prospective Studies Collaboration et al., 2009). Furthermore, increased BMI has also been shown to be associated with reduced physical health-related quality of Life (HRQoL) (Ul-Haq et al., 2012, 2013). However, evidence on the relationship between BMI and mental HRQoL is inconclusive. Some studies have found that BMI is associated with poor mental health (Baumeister & Härter, 2007; Ohayon, 2007; Petry et al., 2008), whereas others did

not find such a relationship (Crisp & McGuinness, 1976; Goldney et al., 2009; Palinkas et al., 1996; Petry et al., 2008). Ul-Haq et al. (2014) also found that the association between BMI and mental health among the Scottish adult population ($N = 37,272$) was moderated by age and sex. In contrast to the above-mentioned studies, Ul-Haq et al. (2014) used the full spectrum of BMI and adjusted for potential confounders and found that only young obese women (<45 years of age with BMI >29.9 kg/m²) had significantly reduced mental health. Furthermore, being underweight was also associated with diminished mental health among women of all ages, but not men.

Background on Lifelines

Lifelines is a large, population-based cohort study and biobank investigating the biological, behavioral and environmental determinants of healthy aging among 167,729 inhabitants from the northern part of the Netherlands. The cohort profile of the Lifelines study has been extensively described in Scholtens et al. (2015). Summarizing, the participants' baseline visit took place between December 2006 and December 2013. All general

Author for correspondence: Ernst C. Wit, Email: e.c.wit@rug.nl

Cite this article: Demetrashvili N, Smidt N, Snieder H, van den Heuvel ER, and Wit EC. (2019) Variance Components Models for Analysis of Big Family Data of Health Outcomes in the Lifelines Cohort Study. *Twin Research and Human Genetics* 22: 4–13, <https://doi.org/10.1017/thg.2019.1>

© The Author(s) 2019. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

practitioners in the three northern provinces of the Netherlands were asked to invite their registered patients aged 25–49 years. All persons who consented to participate were asked to provide contact details to invite their family members (i.e., partner, parents and children), resulting in a three-generation study. In addition, participants could also register their participation via the Lifelines website. Lifelines adopted a multigenerational study design to disentangle the genetic, lifestyle and environmental contributions to the development of chronic diseases, study the between-generation similarities and identify the preclinical stages of ageing at an early age (Stolk et al., 2008). Baseline data were collected from 167,729 participants, aged from 6 months to 93 years. Follow-up is planned for at least 30 years, with questionnaires administered every 1.5 years and a physical examination scheduled every 5 years. The physical examinations, including anthropometry, lung function, blood pressure, electrocardiogram (ECG) and cognition tests, are conducted at one of the Lifelines research sites. In addition, fasting blood and 24-h urine samples are collected from all participants. A comprehensive questionnaire on history of (chronic) diseases, HRQoL, lifestyle (physical activity, alcohol use, diet and smoking status), individual socioeconomic status (income and education level), psychosocial stress, work (profession, working hours), psychosocial characteristics and medication use is completed at home.

Lifelines is a facility that is open to all researchers. Information on application and data access procedure is summarized on www.lifelines.net. An overview of the available data is presented in the online Lifelines Data Catalogue.

Motivation and Goal

Although the explicit familial structure of the data is an advantage in studying the genetic and environmental components of various health-related outcomes, it is in fact a complicating factor in studying the effect of traditional epidemiological covariates. Given that participants are related (e.g., grandparent–grandchild, parent–child, and sibling–sibling relations), genetic, shared environmental and health behavioral factors may confound the effects of BMI on physical and mental HRQoL.

The standard way to control for familial effects is through the kinship model (Almasy & Blangero, 1998), which is a variance components model whereby genetic relatedness is modeled through a kinship covariance matrix. Despite the genetic plausibility of the model, it is computationally prohibitive for thousands of individuals and hundreds of families. This means that it cannot be used for the Lifelines study.

Our first objective is to study the effect of BMI on mental and physical components of HRQoL scores with and without accounting for relatedness in a family. Another objective is to determine a computationally efficient model that incorporates genetic and shared environment to assess the contribution of epidemiological determinants on health outcomes. Given the large sample size and inclusion of extended families in the study, Lifelines allows the identification of the effect of relatedness with higher precision than other studies that are often (much) smaller and limited to particular types of family relationships (e.g., Lichtenstein et al., 2009; Noh et al., 2006; Pawitan et al., 2004; Rabe-Hesketh et al., 2008; Yip et al., 2008). The current Lifelines study includes extended families with up to 19 members.

Not only are the variance components due to specific factors of interest, but also the relative contribution of these variances in the total variance of the outcome. Related to this is the concept of

intra-class correlation coefficient (ICC). The ICC represents the heritability coefficient in a narrow sense when applied to additive genetic models. The concept of heritability originates from Fisher (1918) and Wright (1920) and was formalized by Lush (1940). An extensive review on the concept and misconceptions of heritability is given by Visscher et al. (2008). In the Methods section, we provide the definitions of shared environmental, unique environmental and hereditary ICCs in the context of our models. We also briefly introduce the beta-approach (Demetrashvili et al., 2016) to construct confidence intervals for the ICC. The beta-approach has been successfully applied to construct confidence intervals for ratios of sums of variance components in linear and nonlinear mixed-effects models (Demetrashvili & Van den Heuvel, 2015). This approach will use the first and second moments of the ICC estimate in combination with a beta distribution for accurate confidence intervals.

In the next section, we provide the background on the reconstruction of families and outcome measures. Then we describe various models and give a motivation of their use, including criteria for their selection. Following, we provide the analysis of familial confounding of BMI-related physical and mental HRQoL in the Lifelines study. We then compare our fractional relatedness model with more traditional kinship models through simulations. We conclude the article with an extensive discussion.

Methods

In this section, we explain how we reconstructed extended families from the available local kinship relationships. The outcomes of interest are mental and physical health, which were reconstructed from a RAND-36 questionnaire. Then we applied mixed-effects models beside a fixed-effect model to study the effect of BMI on mental and physical health.

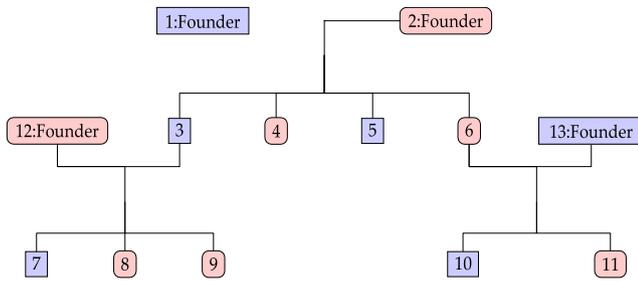
Family Reconstruction

For a large number of participants in the Lifelines study, their parents, partner and children are also included in the study. Such information is relevant for disentangling the genetic, behavioral and shared environmental variances. In biometric genetics, the *coefficient of relatedness* or *genetic correlation* for two individuals is defined as the expected proportion of genes of two individuals that are identical by descent (Sham, 1998, p. 208). A related concept in biometric genetics, which is used in this study, is that of a *founder*. Individuals without ancestors in the study are called *founders*, whereas others are called *non-founders* (Almgren et al., 2003, p. 10). Founders in our study population are assumed unrelated.

Considering the information provided by Lifelines participants, we define a *family* as a group of related individuals sharing environmental and/or genetic factors. For example, health responses of mother and child may be similar due to both genetic similarity and shared behavior and environment, whereas the health responses of partners are related only due to the latter. Within the context of the Lifelines study, we define the concept of an extended family as a connected graph of individuals either via parent–child or partner–partner relationships. An example of a family is given in Figure 1. Note that the sibling relationships in this graph are inferred from common parent relationships. Sibling information itself is not recorded within Lifelines. We do not have extensive information for all reconstructed families in the Lifelines study, as some members might not participate in the study.

Table 1. Algorithm for Construction of a Fractional Relatedness Matrix

1. For family, i determine the set of founders, that is, individuals without parents
2. Initialize a fractional relatedness matrix of size n_i by m_i as a matrix with zeros, where n_i is number of individuals and m_i is number of founders in family i
3. For each founder do the following: assign fractional relatedness equal to one to himself/herself and assign fractional relatedness equal to $(1/2)_g$ to offspring, where g is a generational distance (e.g., $g = 1$ for child, $g = 2$ for grandchild, etc.)

**Fig. 1.** Family example consisting of 13 members.

Information on children from previous marriages is, in principle, also available. However, not all familial information is complete in Lifelines, since people may not be willing, for example, to identify ex-partners. Some of the information, however, can be reconstructed from the partial information provided by the participants. For example, if a child declares both parents and the parents declare this child, but do not declare each other as partners, we make a link between such parents as a couple. For family reconstruction, we identify a set of related individuals through parent–child and/or partner–partner relationships and call this set a family. Once an individual is assigned to a particular family, no further reassessments take place. The model we propose in this study requires reconstruction of the relatedness matrix. Relatedness between founders and non-founders is fractional, as explained in the next section. In Table 1, we outline the algorithm used for construction of fractional relatedness matrix.

Outcome Measures

A sample consisting of 91,759 participants from the baseline Lifelines cohort study data release was available for analysis. HRQoL was measured using the Dutch version of the RAND-36 questionnaire (Hays & Morales, 2001; Van der Zee & Sanderman, 1993; Van der Zee et al., 1996). HRQoL refers to how health impacts on an individual's ability to function and his or her perceived well-being in physical, mental and social domains of life. The RAND-36 consists of 36 items measuring eight health concepts, that is, physical functioning, role limitations caused by physical health problems, bodily pain, general health perceptions, role limitations caused by emotional problems, social functioning, emotional well-being, and energy/fatigue. The first four reflect the physical health and the last four reflect the mental health of an individual. The scales of these eight concepts are combined into two summary measures of HRQoL, the physical component score (PCS) and the mental component score (MCS), using the scoring algorithm of Ware et al. (1994). PCS and MCS are between 0% and 100% and higher scores correspond to better quality of life.

Statistical Models

We compare four models to examine whether the effect of BMI on HRQoL differs with and without accounting for family structure. These models are: M_0 : multiple regression model; M_1 : mixed-effects model with random intercept for the family, capturing the environmental familial effect; M_2 : mixed-effects model with random slopes for founders within a family, capturing the genetic familial effect. The fourth model, M_3 , is a combination of the last two models. Thus, M_3 is a mixed-effects model with random intercept for family and random slopes for founders, capturing both environmental and genetic familial effects.

Assume that a total of n individuals and I families are included in the study. Suppose for the i th family n_i members have been observed ($i = 1, 2, \dots, I$), such that $n = \sum_{i=1}^I n_i$. The multiple regression model M_0 of the HRQoL response y_{ij} for the j th member of the i th family can be written as:

$$y_{ij} = \mathbf{x}_{ij}^T \boldsymbol{\beta} + \varepsilon_{ij}, \quad (1)$$

where \mathbf{x}_{ij} is a $p \times 1$ vector for the j th individual in the i th family measured on p covariates, including BMI and possible confounders, $\boldsymbol{\beta}$ is a $p \times 1$ vector of coefficients and residual errors ε_{ij} across all observations are assumed to be identically and independently distributed (iid) having a normal distribution with mean zero and variance σ_R^2 , $\varepsilon_{ij} \stackrel{iid}{\sim} N(0, \sigma_R^2)$. The fixed effects \mathbf{x}_{ij} can be quantitative or dummy variables to represent categorical variables, so the effective number of covariates may be less or equal to p .

Note, observations y_{ij} within the same family are most likely correlated, but the multiple regression model M_0 does not account for this. Model M_1 will account for this correlation by introducing a random intercept u_i for every family:

$$y_{ij} = \mathbf{x}_{ij}^T \boldsymbol{\beta} + u_i + \varepsilon_{ij}, \quad (2)$$

where u_i is normally distributed with mean zero and variance σ_u^2 , $u_i \stackrel{iid}{\sim} N(0, \sigma_u^2)$; the definitions of \mathbf{x}_{ij} , $\boldsymbol{\beta}$ and ε_{ij} are the same as in (1).

Model M_1 does not disentangle the genetic and shared environmental variation. Since one of the goals is to estimate the variance contribution in MCS and PCS due to various factors, model M_2 will assume that the genetic correlation between family members is due to additive genetic effects of alleles (with no dominant and epistatic effects). By assuming all genetic information is in the founders, and consequently imposing the fractional relatedness effect between founders and other members of the family, model M_2 introduces the random slopes v_i for the set of founders m_i in family i and can be formulated as:

$$y_{ij} = \mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{F}_{ij}^T v_i + \varepsilon_{ij} \quad (3)$$

where definitions of \mathbf{x}_{ij} , $\boldsymbol{\beta}$ and ε_{ij} are the same as in (1) and \mathbf{F}_{ij} is the $m_i \times 1$ vector of founders for individual j in family i , where F_{ijk} is a fractional relatedness of individual j to founder k in family i with $\sum_{k=1}^{m_i} F_{ijk} = 1$; v_i is the $m_i \times 1$ vector of random slopes for founders in family i , where $v_i = (v_{i1}, \dots, v_{im})$, $v_{ik} \stackrel{iid}{\sim} N(0, \sigma_f^2)$. We assume independence between random terms.

An important computational advantage of model M_2 compared to the traditional kinship model is employment of a substantially smaller design matrix F_{ij} . Namely, in M_2 , the design matrix across all families is of size m by n , where m is the maximum number of founders among all families and n is the total number of participants in all families. In the classical kinship model, the variance component matrix would be of size n by n . Dimensionality reduction is particularly crucial when one analyzes large number of participants, such as in Lifelines.

An example of a family consisting of 13 members in the Lifelines study is shown in Figure 1: oval shapes refer to females and squares to males. The associated fractional relatedness matrix F is demonstrated below. The M_2 model assumes that the health outcome has a hereditary component. For example, the founders F1 and F2 both share half of their random hereditary effects with their child, member 3, and one-quarter with their grandchild, member 7.

$$F_i = \begin{pmatrix} 1 & F1 & F2 & F12 & F13 \\ 2 & 1 & 0 & 0 & 0 \\ 3 & 0 & 1 & 0 & 0 \\ 4 & 1/2 & 1/2 & 0 & 0 \\ 5 & 1/2 & 1/2 & 0 & 0 \\ 6 & 1/2 & 1/2 & 0 & 0 \\ 7 & 1/4 & 1/4 & 1/2 & 0 \\ 8 & 1/4 & 1/4 & 1/2 & 0 \\ 9 & 1/4 & 1/4 & 1/2 & 0 \\ 10 & 1/4 & 1/4 & 0 & 1/2 \\ 11 & 1/4 & 1/4 & 0 & 1/2 \\ 12 & 0 & 0 & 1 & 0 \\ 13 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Model M_3 is a combination of models M_1 and M_2 , and can be written:

$$y_{ij} = \mathbf{x}_{ij}^T \beta + u_i + \mathbf{F}_{ij}^T v_i + \varepsilon_{ij}, \tag{4}$$

where definitions and assumptions used in models M_1 and M_2 remain the same for this model. Variance components σ_u^2 , σ_f^2 and σ_R^2 are due to shared environmental, genetic and unique factors, respectively.

Inference and Model Selection

The overall significance of each covariate is tested using a conditional F test. In this test, as in the usual F test of covariates for regression models, the conditional estimate of the residual error variance is used. More details on the F test are given in Pinheiro and Bates (2009, §2.4.2). Confidence intervals for marginal coefficients β_l are constructed based on conditional t tests. Each fixed-effect coefficient can be tested marginally in the presence of other fixed effects in the model (Pinheiro & Bates, 2009, pp. 92–96). The approximate 100% $(1-\alpha)$ confidence limits on the β_l are computed as:

$$\hat{\beta}_l \pm t_{df_l}(1 - \alpha/2) \sqrt{\hat{var}(\hat{\beta})_{ll}}, \tag{5}$$

where $\hat{\beta}_l$ is an estimate of l th fixed effect, $t_{df_l}(q)$ denotes the q th quantile of a t distribution with df_l degrees of freedom and

$\hat{var}(\hat{\beta})_{ll}$ is an estimate of the variance of $\hat{\beta}_l$. Clearly $\hat{var}(\hat{\beta})$ is the variance-covariance matrix of the vector $\hat{\beta}$ of fixed-effects estimates. More on the determination of the degrees of freedom for our model is in the below ‘Estimation of fixed effects’ section.

To select the best model we used the Bayesian information criterion (BIC; Schwarz, 1978). The BIC for these models is defined as:

$$BIC(M) = -2l_M(\theta|y) + df(M) \ln(n), \tag{6}$$

where $l_M(\cdot)$ is the log-likelihood function for the estimated model M with θ a vector of all parameters, $df(M)$ denotes the overall number of parameters in the model, that is, the regression and variance components parameters and n is the total number of observations used to fit the model. The model with the smallest BIC is preferred.

Intraclass Correlation Coefficient

The models we defined above allow us to define the following three types of ICC: (1) the behavioral and shared environmental ICC (c^2), as the proportion of total variance due to shared environmental components; (2) the hereditary ICC (h^2), as the proportion of total variance due to the additive genetic component; and (3) the unique environmental ICC (e^2), as the proportion of total variance due to unique environmental components. We define these ICCs for model M_3 as follows:

$$c^2 = \frac{\sigma_U^2}{\sigma_U^2 + \sigma_f^2 + \sigma_R^2} \quad h^2 = \frac{\sigma_f^2}{\sigma_U^2 + \sigma_f^2 + \sigma_R^2} \quad e^2 = \frac{\sigma_R^2}{\sigma_U^2 + \sigma_f^2 + \sigma_R^2} \tag{7}$$

For M_1 and M_2 models, the c^2 , h^2 and e^2 are deduced from formula (7), by ignoring (setting to zero) those variance components that are not present in the model.

We outline the beta-approach for obtaining the CI for h^2 , though this approach can be similarly applied to c^2 and e^2 . The distribution of the estimator h^2 is approximated with a beta distribution, $\hat{h}^2 \sim Beta(a, b)$ with parameters $a > 0$ and $b > 0$. If \hat{h}^2 is an estimate of the variance of the mean and $\hat{\tau}_{h^2}^2$ is an estimate of the variance of \hat{h}^2 , the method of moment estimates for a and b are given as:

$$\hat{a} = \frac{\hat{h}^2 [\hat{h}^2(1-\hat{h}^2) - \hat{\tau}_{h^2}^2]}{\hat{\tau}_{h^2}^2}, \tag{8}$$

$$\hat{b} = \frac{(1-\hat{h}^2) [\hat{h}^2(1-\hat{h}^2) - \hat{\tau}_{h^2}^2]}{\hat{\tau}_{h^2}^2}.$$

A first-order Taylor expansion is used to approximate $\hat{\tau}_{h^2}^2$, as shown in Demetrashvili et al. (2016). The approximate 100% $(1-\alpha)$ confidence interval on the h^2 in (7) is then given by the lower and upper confidence limits as:

$$LCL_{\hat{h}^2} = B_{a,b}^{-1}(1 - \alpha/2), \tag{9}$$

$$LCL_{\hat{h}^2} = B_{a,b}^{-1}(\alpha/2),$$

with $B_{a,b}^{-1}(q)$ being the q th quantile of the beta (a and b) distribution. A detailed description of the beta-approach is given in Demetrashvili et al. (2016).

Table 2. Counts of Family Sizes for the Original Set of 91,759 Participants and the Remaining 89,353 Participants after Removing Incomplete Records

Family size	1	2	3	4	5	6–7	8–12	13–15	16–19	total
Total	253	18,585	6990	3756	1389	1046	586	18	7	32,531
In analysis	883	18,446	6863	3542	1303	962	433	14	6	32,452

Lifelines Analysis Results

Analysis of the Lifelines data was conducted using R (R Core Team, 2017), version 3.4.0. Mixed-effect models were fitted by applying the `lme` function of the `nlme` package using maximum-likelihood. Unlike for the traditional kinship model, we were able to fit all models to thousands of subjects and families using the standard R function. All results below are presented with two-sided 95% confidence intervals.

A sample of the baseline Lifelines cohort was used consisting of 91,759 participants, from which we constructed 32,531 families. The distribution of family sizes is summarized in Table 2. The largest family has 19 members. There are 253 singletons, 18,585 families have two members, and so on. About 99% of all reconstructed families consist of at least two members. Among all participants, 44% were recruited via their general practitioner, 13% via self-registration and the remaining 43% were recruited as family members of the first two groups. The number of declared partners is 56,560, meaning that 62% of all individuals in Lifelines currently have a partner. The number of declared fathers is 18,342 (20%), whereas the number of declared mothers is 26,627 (29%); 34% of all individuals in Lifelines have at least one child. The maximum number of declared children is 7, which occurred once.

We omitted 2406 (2.6%) observations for both MCS and PCS analyses. These observations were incomplete with respect to PCS scores, MCS scores or BMI. There were no missing values for sex or age. Finally, 89,353 observations were included in the analysis. The distribution of both outcomes, MCS and PCS, is slightly left-skewed. The median (25th, 75th percentile) MCS and PCS were 53.4 (48.9, 56.4) and 54.6 (50.5, 56.8), respectively. The distribution of family sizes for 89,353 participants is shown in Table 2. About 97% of the 32,452 reconstructed families with complete data consisting of at least two family members. The maximum number of founders is 9.

Descriptive Statistics of Covariates

BMI is calculated by dividing a person's weight measurement (in kilograms) by the square of their height (in meters) and subsequently categorized into six categories: underweight (<18.5) kg/m², normal weight (18.5–24.9) kg/m², overweight (25.0–29.9) kg/m², class I obese (30.0–34.9) kg/m², class II obese (35.0–39.9) kg/m² and class III obese (>40) kg/m² (World Health Organization, 1995). Out of 89,353 observations, 686 (0.7%) were underweight, 39,277 (44%) were normal, 35,824 (40.1%) were overweight, 10,494 (11.7%) were obese I, 2306 (2.6%) were obese II and 766 (0.9%) were obese III.

All subjects were 18 years and older, with an average age of 45. Out of 89,353 observations, 38,841 (44%) were men. The same proportions of males and females were found when all 91,759 observations were summarized.

Model Selection, Variance Components and Heritability

In the analysis of all models, we used 89,353 observations with 32,452 constructed families. BMI is treated as a categorical variable with six levels. Age and sex are included in all models. Age is

treated as a continuous variable. Sex is a categorical variable with male being a reference category. Model M_0 is a multiple regression model with BMI, age and sex. Models M_1 , M_2 and M_3 are the variance components models. M_1 models the environmental random component of the family. M_2 models the genetic random components of the founders. M_3 models both the random components of the family and founders.

We used BIC for model selection. BIC consistently selects the true model for large sample sizes (Claeskens & Hjort, 2008) and tends to choose parsimonious models (i.e., models with few explanatory variables). Using the BIC model, M_1 provides the best fit for MCS and M_3 for PCS. Since the best model for MCS is modeling the shared environmental factors, the estimated ICC shown in Table 3 implies that approximately 12–14% variation in MCS is determined by shared environmental variation. Regarding the best model M_3 for PCS, both the shared environmental and the genetic contribution are included; approximately 12–14% variation in PCS is determined by genetic variation and 3–4% by shared environmental variation.

Estimation of Fixed Effects

Besides accounting for the familial correlation structure, the aim of the study was to estimate the effect of BMI, age and sex on the HRQoL scores, PCS and MCS. The results of the conditional F tests are presented in Table 4. The degrees of freedom for the tests of significance of slopes are 56,894, which were calculated by subtracting the number of families and the number of parameters of fixed effects from the number of observations (i.e., 89,353 – 32,452 – 7). These degrees of freedom are also used in the t test. The results for individual effects are presented in Figures 2 and 3 and in Table 5.

Figures 2 and 3 show the effect sizes of BMI (middle line) surrounded by confidence intervals (outer lines) of these effects. Plain lines are used for the selected models of MCS and PCS. Dashed lines are used for the other three models. Obviously, the effects of BMI on MCS and PCS of HRQoL match very closely across four models (lines overlap) and this is true for all categories of BMI. Confidence intervals also match very closely. This implies that the BMI effects do not change with and without accounting for relatedness in a family.

We see inverted parabolic shape of MCS across increasing BMI categories. A somewhat different shape is observed for PCS across increasing BMI categories. Interestingly, the MCS slightly increases for overweight people compared to the normal category. It also shows a dramatic drop for underweight individuals. The PCS decreases for all categories of BMI compared to the normal category, and the decreasing trend of PCS is increasingly steeper with increasing BMI. The confidence intervals of all BMI categories for PCS are beyond the confidence limits of the normal category, meaning that all BMI categories have substantially different PCSs than the normal category.

In Table 5, we present the coefficients for fixed effects with measures of uncertainty, namely standard error and lower and upper confidence limits. The results show that each additional year of age

Table 3. Estimates of variance components, ICC and its confidence interval for outcomes MCS and PCS

Outcome	Model	Variance component		ICC estimate	ICC confidence		BIC estimate
		Estimator	Estimate		$\hat{L}CL$	$\hat{U}CL$	
MCS	M ₀	$\hat{\sigma}_R^2$	64.842				626,443
	M ₁	$\hat{\sigma}_u^2$	8.438	$\hat{c}^2 = \mathbf{0.130}$	0.123	0.137	625,134
		$\hat{\sigma}_R^2$	56.551				
	M ₂	$\hat{\sigma}_f^2$	5.855				626,302
$\hat{\sigma}_R^2$		60.206					
MCS	M ₃	$\hat{\sigma}_u^2$	8.438				625,145
		$\hat{\sigma}_f^2$	2.11×10 ⁻⁶				
		$\hat{\sigma}_R^2$	56.551				
MCS	M ₀	$\hat{\sigma}_R^2$	45.546				594,882
	M ₁	$\hat{\sigma}_u^2$	3.063				594,495
		$\hat{\sigma}_R^2$	42.499				
	M ₂	$\hat{\sigma}_f^2$	8.094				594,374
$\hat{\sigma}_R^2$		39.161					
MCS	M ₃	$\hat{\sigma}_u^2$	1.535	$\hat{c}^2 = \mathbf{0.130}$	0.026	0.040	594,324
		$\hat{\sigma}_f^2$	5.981				
		$\hat{\sigma}_R^2$	39.302				
				$\hat{h}^2 = \mathbf{0.128}$			

Note: LCL and UCL stand for lower and upper confidence limits, respectively; best model for outcomes MCS and PCS is in bold; \hat{c}^2 shows the proportion of behavioral and shared environmental variance in total phenotype variance and \hat{h}^2 shows the proportion of additive genetic variance in total phenotype variance.

Table 4. Conditional F Tests for BMI, Age, and Sex of Selected Models

Model	Term	Numerator df	Denominator df	F-value	p-value
MCS: M ₁	BMI	5	56,894	72.6	<.0001
	Sex	1	56,894	1386.4	<.0001
	Age	1	56,894	1573.4	<.0001
PCS: M ₃	BMI	5	56,894	695	<.0001
	Sex	1	56,894	412	<.0001
	Age	1	56,894	2460	<.0001

is associated with a 0.078-unit increase (0.1% of maximum observed value) in MCS and a decrease by about the same amount in PCS, on average, holding BMI and sex constant. Women have lower MCS, by approximately 1.891 units (2.6% of maximum observed value), and lower PCS, by approximately 1.011 units (1.4%), on average, holding BMI and age constant.

Simulation Study: Design and Results

We conducted a simulation study to compare the commonly used variance component model with kinship matrix (Almasy & Blangero, 1998) to our suggested model containing a reduced matrix of founders. The kinship model is computationally demanding and therefore we had to simulate relatively small sample sizes. We generated data using the model (10) shown below:

$$y_{ij} = \mathbf{x}_{ij}^T \beta + u_i + k_{ij} \sigma_f^2 + \varepsilon_{ij}, \tag{10}$$

$$\varepsilon_{ij} \stackrel{iid}{\sim} N(0, \sigma_R^2),$$

where definitions of \mathbf{x}_{ij} , β , u_i , σ_u^2 , σ_f^2 and ε_{ij} are the same as in models (2), (3) and (4), σ_f^2 is the variance due to genetics and k_{ij} is the kinship effect for the j th member of the i th family. k_{ij} is generated from the multivariate normal distribution with mean zero

and variance–covariance matrix Ω . The Ω consists of coefficients of pairwise relationships formed in the following way: in the first degree of relationship (parent–child and siblings), the coefficient of relationship is 1/2, in the second degree of relationship (grandparent/grandchild, half-sibling and avuncular), the coefficient of relationship is 1/4 and similarly for other degrees, as shown in Table 1 of Almasy and Blangero (2010). Technically, Ω is a block-diagonal matrix with coefficient of relationships among family members on the diagonal blocks and zeros (the coefficients of relationship between families) on the off-diagonal blocks.

Simulation parameters were selected from the results of the best-fitted model for Lifelines PCS data (see parameters in PCS: M₃ of Table 5) and equal to: $\beta_{\text{underweight}} = 56.5$, $\beta_{\text{normal}} = 57.5$, $\beta_{\text{overweight}} = 56.7$, $\beta_{\text{obese1}} = 54.9$, $\beta_{\text{obese2}} = 53.0$, $\beta_{\text{obese3}} = 50.2$, $\beta_{\text{age}} = -0.08$, $\beta_{\text{sex}} = -1.0$, $\sigma_u^2 = 1.5$, $\sigma_f^2 = 6.0$, $\sigma_R^2 = 40.0$. The variable for age was generated from the normal distribution with a mean of 45 and a standard deviation of 14. The variable for sex was generated from Bernoulli distribution with a probability of .44 (for men). The number of observations for BMI categories ‘underweight’, ‘normal’, ‘overweight’, ‘obese 1’, ‘obese 2’, and ‘obese 3’ were generated from the multinomial distribution with probabilities of .007, .44, .401, .117, .026, and .009, respectively. These probabilities were calculated from the Lifelines data. Then the BMI was generated from a normal distribution based on the number of observations from the multinomial distribution and the following mean and standard deviation parameters: 17.7, 0.77 for ‘underweight’; 23.0, 1.5 for ‘normal’; 27.0, 1.4 for ‘overweight’; 32.0, 1.4 for ‘obese 1’; 37.0, 1.4 for ‘obese 2’; 43.0, 3.0 for ‘obese 3’. For family sizes, we set 2, 3, 4, 5 and 6 members, each size repeated 2, 4, 10, 20 and 38 times, respectively. Even though family sizes were repeated, in fact all families were different in their composition (e.g., for a family size of 2, we constructed parent–child and partner–partner families; for a size of 3, we constructed families of parent–parent–child, parent–child–grandchild, parent–child–child and parent–child–partner of child). In total, 74 families were constructed containing 384 numbers of

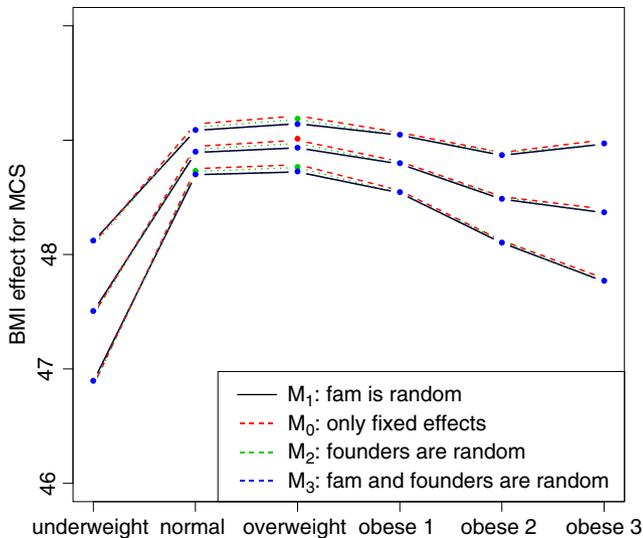


Fig. 2. BMI effects for MCS surrounded by confidence intervals.

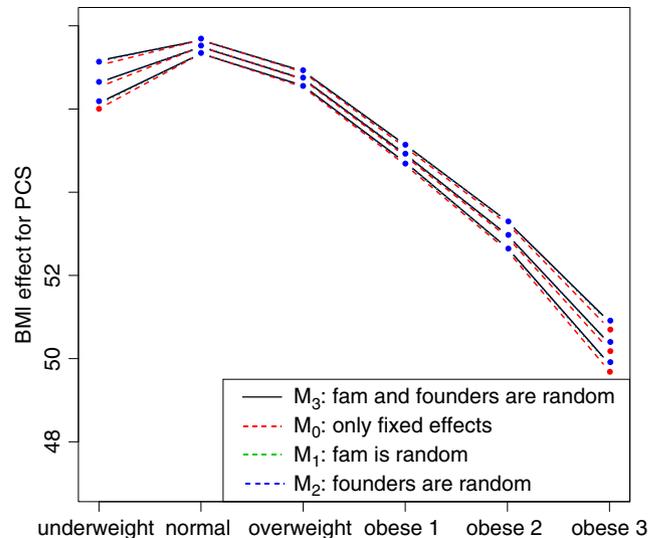


Fig. 3. BMI effects for PCS surrounded by confidence intervals.

observations. The PCS outcomes were generated using model (10). Afterwards, we fitted both models (kinship and M_3) and compared the biases of heritability estimates and coverage probabilities of 95% confidence intervals of heritabilities. The heritability in model (10) is equivalent to hereditary ICC (h^2) and computed as shown in (7). Confidence intervals for heritability were calculated using the beta-approach, as described in Section 2.5. We conducted 100 simulations. The kinship model was fitted using the `lmekin` function of the `coxme` (Therneau, 2018) package and model M_3 using the `lme` function of the `nlme` (Pinheiro et al., 2017) package in R. Results of the simulation studies are summarized in Table 6.

Results from setting 3 are shown in Figure 4. The kinship model results in larger bias (0.18) in comparison with the one (0.09) in M_3 model. The M_3 model results in larger variation than the kinship model on average, as shown in Figure 4. Subsequently, the kinship model demonstrates substantial undercoverage (0.86) while the M_3 model shows some overcoverage (0.98) for a two-sided 95% confidence interval.

We compared the computational time needed for fitting the kinship and M_3 models. We varied the number of families and, correspondingly, the total number of observations. We examined three settings, each with 100 simulations. Settings and results are shown in Table 6. M_3 is 8 times faster than the kinship model (setting 1). As the number of families quadruples (from setting 1 to 3), the computational time increases 2.5 times for the M_3 model (`lme` function) and 15.2 times for the kinship model (`lmekin` function). Thus, there is an exponential increase in computational time for the kinship model as the number of observations increases while there is much slower increase (roughly linear) in computational time for the M_3 model.

Discussion

In this study, the primary goal was to answer whether relatedness in a family must be accounted for when estimating the effects of risk factors of interest in large family-based cohort studies. The answer to this question is of particular importance for researchers analyzing the Lifelines data and data from other large cohort studies of families. From our study, it is clear that the effects of BMI on MCS and PCS of HRQoL scores do not change when accounting

for family structure. This conclusion confirms theoretical considerations within longitudinal data analysis given by Diggle et al. (2002, chapter 1). The authors state that when the focus of the study is on modeling the dependence between the response and explanatory variable, then the nature of correlation among responses is unimportant if there is a large number of families relative to the number of individuals per family. Our Lifelines data clearly satisfy this criterion.

McArdle et al. (2007) conducted a simulation study with the objective to compare the performance of association analysis of family-based designs that account for and ignore family structure in assessment of the phenotype–genotype association. They concluded that effect size estimates and power are not significantly affected by ignoring family structure, although type 1 error rates increase when family structure is ignored, and the magnitude of the increase depends on trait heritability and pedigree configuration. Induced type 1 error is directly related to diminished standard errors (and narrow confidence intervals), leading to liberal inference about regression coefficients, that is, falsely claiming significance when there is none. In our analysis of both PCS and MCS, we saw that ignoring the correlation (or family structure) led instead to larger standard errors of the regression coefficients (although the increase was very small), thereby leading to conservative inference about the covariate effects. Therefore, the standard errors of regression parameters can be larger or smaller when ignoring family structure, and therefore may lead not only to liberal inference and inflated type I errors, but also to conservative inference and deflated type I errors. Increase or decrease of the standard errors depends on (1) the relationship between the family structure and the covariates of interest and (2) the family structure effect size on the outcome.

We compared the Lifelines analysis results on association between BMI and HRQoL with similar results in the literature. Increased BMI has been shown to be associated with reduced physical HRQoL (Ul-Haq et al., 2012, 2013); however, evidence on the relationship between BMI and mental HRQoL was antagonistic. In the Lifelines study, we see reduced mental HRQoL for all categories of BMI in comparison with the overweight category. This conclusion matches with the conclusion of Ul-Haq et al. (2013) from meta-analysis study. Furthermore, an inverted

Table 5. Estimates of Coefficients for BMI, Age and Sex of Selected Models

Model Parameter	Estimate	SE	\hat{LCL}	\hat{UCL}
MCS: M_1	47.502	0.313	46.889	48.114
BMI (underweight)				
BMI (normal)	48.890	0.099	48.696	49.083
BMI (overweight)	48.931	0.108	48.720	49.142
BMI (obese 1)	48.790	0.129	48.538	49.042
BMI (obese 2)	48.482	0.195	48.100	48.865
BMI (obese 3)	48.364	0.307	47.763	48.965
Sex (female)	-1.891	0.053	-1.995	-1.788
Age	0.078	0.002	0.074	0.081
PCS: M_3	56.653	0.259	56.145	57.161
BMI (underweight)				
BMI (normal)	57.491	0.082	57.331	57.650
BMI (overweight)	56.731	0.090	56.556	56.906
BMI (obese 1)	54.927	0.107	54.716	55.137
BMI (obese 2)	52.978	0.164	52.656	53.299
BMI (obese 3)	50.373	0.258	49.868	50.879
Sex (female)	-1.011	0.046	-1.100	-0.921
Age	-0.081	0.002	-0.084	-0.078

U-shape of mental HRQoL across increasing BMI categories is seen in both our Lifelines study (Figure 2) and that of Scottish study conducted by Ul-Haq et al. (2012). Similarly to Ul-Haq et al. (2013), we see that increasing BMI is associated with impaired HRQoL in Lifelines. In addition, our study reveals a shape of association between BMI and physical HRQoL. There is an inverted J-shape negative association in PCS across increasing BMI categories (Figure 3).

In this work, we studied real data and did not assume any a priori family inheritance structure. The main strength of our study is the use of Lifelines data which makes it possible to estimate the relatedness in more complex families than other studies can. Consequently, our results are practically more relevant. We learned that with or without accounting for family structure, the effect of determinants on health outcomes does not significantly change. Nevertheless, the ability to incorporate a family structure into our model in a computationally efficient way allows one to disentangle the genetic, shared behavioral and environmental variances. Furthermore, our proposed model allows for estimating hereditary, behavioral and shared environmental, and unique environmental ICCs of both HRQoL outcomes in computationally an efficient way through fractional relatedness of founders and non-founders.

The kinship model as implemented in the SOLAR (Almasy & Blangero, 1998) software is unable to handle the analysis of the Lifelines data with over 89,000 individuals, although we did not study exhaustively all methods, such as generalized estimating equations (Liang & Zeger, 1986) that could have been used to implement the kinship model. We overcame the computational infeasibility with the kinship model by introducing and fitting a variance component model with fractional relatedness using standard functions in R.

We have made the assumption in our study that founders are unrelated, but Lifelines may not have complete information. For example, siblings' information itself is not recorded within

Table 6. Comparison of Heritability Parameters and Computational Time between Kinship (Imekin Function) and M_3 (Ime Function) Models

Setting	1	2	3
Number of families	19	37	74
Total number of observation	96	192	384
Time per simulation in M_3 model (seconds)	0.11	0.12	0.27
Time per simulation in kinship model (seconds)	0.92	2.69	13.99
True heritability	0.13	0.13	0.13
Bias of heritability ^a in M_3 model	0.09	0.08	0.09
Bias of heritability ^a in kinship model	0.19	0.18	0.18
Coverage of 95% confidence intervals of heritability in M_3 model	0.78	0.91	0.98
Coverage of 95% confidence intervals of heritability in kinship model	0.66	0.88	0.86

^aHeritability is equivalent to hereditary ICC (h^2) and computed using formula (7).

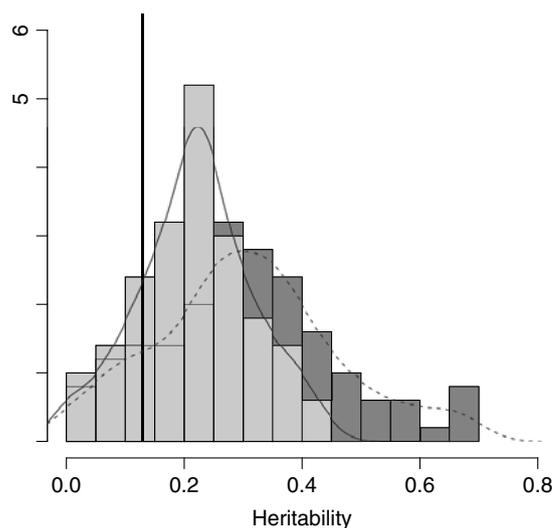


Fig. 4. Overlapping histograms for comparison of estimated heritabilities across M_3 and kinship models in setting 3: vertical bar (0.13) shows true heritability, light grey histogram (left) and smoothed plain line show distribution of estimated heritabilities in M_3 model and dark grey histogram (right) and smoothed dashed line show distribution of estimated heritabilities in kinship model.

Lifelines, and therefore siblings might be modeled as two founders while they do have a common ancestor and are related. It would be interesting to examine whether modeling just a subset of founders would impact the variance component parameter estimates.

Furthermore, MCS and PCS may be genetically correlated, meaning that there is genetic overlap between these two traits (i.e., the same set of genes may regulate these traits). If interest lies in separation of genetic and environmental contributions simultaneously in MCS and PCS, then bivariate models could be used, similarly to the way that others (Lichtenstein et al., 2009; Yip et al., 2008) modeled schizophrenia and bipolar disorder using multivariate generalized linear mixed models.

In summary, the proposed model offers to solve the computational issues involved in modeling family structure when thousands of families are analyzed, and subsequently to fit the model accurately using standard functions of R.

Acknowledgments. The authors wish to acknowledge the services of the Lifelines cohort study, the contributing research centers delivering data to Lifelines and all the study participants. The authors declare no conflicts of interest.

Financial support. During the first six months, this work was supported by the University of Groningen, University Medical Center Groningen as part of PhD thesis and later, it was supported by a NWO STAR travel grant from the University of Groningen. For the remaining work, this research received no specific grant from any funding agency, commercial or not-for-profit sectors.

References

- Almasy, L., & Blangero, J. (1998). Multipoint quantitative-trait linkage analysis in general pedigrees. *The American Journal of Human Genetics*, 62, 1198–1211.
- Almasy, L., & Blangero, J. (2010). Variance component methods for analysis of complex phenotypes. *Cold Spring Harbor Protocols*, 2010, pdb-top77.
- Almgren, P., Bendahl, P., Bengtsson, H., Hössjer, O., & Perfekt, R. (2003). *Statistics in genetics (Lecture notes)*. Lund, Sweden: Lund University.
- Baumeister, H., & Härter, M. (2007). Prevalence of mental disorders based on general population surveys. *Social Psychiatry and Psychiatric Epidemiology*, 42, 537–546.
- Berghöfer, A., Pischon, T., Reinhold, T., Apovian, C. M., Sharma, A. M., & Willich, S. N. (2008). Obesity prevalence from a European perspective: A systematic review. *BMC Public Health*, 8, 200.
- Claeskens, G., & Hjort, N. L. (2008). *Model selection and model averaging*. Cambridge, UK: Cambridge University Press.
- Crisp, A. H., & McGuinness, B. (1976). Jolly fat — Relation between obesity and psychoneurosis in general population. *BMJ*, 1, 7–9.
- Demetrashvili, N., & Van den Heuvel, E. R. (2015). Confidence intervals for intraclass correlation coefficients in a nonlinear dose-response meta-analysis. *Biometrics*, 71, 548–555.
- Demetrashvili, N., Wit, E. C., & Van den Heuvel, E. R. (2016). Confidence intervals for intraclass correlation coefficients in variance components models. *Statistical Methods in Medical Research*, 25, 2359–2376.
- Diggle, P. J., Heagerty, P. J., Liang, K.-Y., & Zeger, S. L. (2002). *Analysis of longitudinal data*. New York, NY: Oxford University Press.
- Fisher, R. A. (1918). The correlation between relatives on the supposition of Mendelian inheritance. *Transactions of the Royal Society of Edinburgh*, 52, 399–433.
- Flegal, K. M., Carroll, M. D., Ogden, C. L., & Curtin, L. R. (2010). Prevalence and trends in obesity among US adults, 1999–2008. *JAMA*, 303, 235–241.
- Flegal, K. M., Kit, B. K., Orpana, H., & Graubard, B. I. (2013). Association of all-cause mortality with overweight and obesity using standard body mass index categories: A systematic review and meta-analysis. *JAMA*, 309, 71–82.
- Goldney, R. D., Dunn, K. I., Air, T. M., Dal Grande, E., & Taylor, A. W. (2009). Relationships between body mass index, mental health, and suicidal ideation: Population perspective using two methods. *Australian and New Zealand Journal of Psychiatry*, 43, 652–658.
- Hays, R. D., & Morales, L. S. (2001). The Rand-36 measure of health-related quality of life. *Annals of Medicine*, 33, 350–357.
- Liang, K., & Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, 73, 13–22.
- Lichtenstein, P., Yip, B. H., Björk, C., Pawitan, Y., Cannon, T. D., Sullivan, P. F., & Hultman, C. M. (2009). Common genetic determinants of schizophrenia and bipolar disorder in Swedish families: A population-based study. *The Lancet*, 373, 234–239.
- Lush, J. L. (1940). Intra-size correlations or regressions of offspring on dam as a method of estimating heritability of characteristics. *Journal of Animal Science*, 33, 293–301.
- McArdle, P. F., O'Connell, J. R., Pollin, T. I., Baumgarten, M., Shuldiner, A. R., Peyser, P. A., & Mitchell, B. D. (2007). Accounting for relatedness in family based genetic association studies. *Human Heredity*, 64, 234–242.
- Noh, M., Yip, B., Lee, Y., & Pawitan, Y. (2006). Multicomponent variance estimation for binary traits in family-based studies. *Genetic Epidemiology*, 30, 37–47.
- Ohayon, M. M. (2007). Epidemiology of depression and its treatment in the general population. *Journal of Psychiatric Research*, 41, 207–213.
- Palinkas, L. A., Wingard, D. L., & Barrett-Connor, E. (1996). Depressive symptoms in overweight and obese older adults: A test of the 'jolly fat' hypothesis. *Journal of Psychosomatic Research*, 40, 59–66.
- Pawitan, Y., Reilly, M., Nilsson, E., Cnattingius, S., & Lichtenstein, P. (2004). Estimation of genetic and environmental factors for binary traits using family data. *Statistics in Medicine*, 23, 449–465.
- Petry, N. M., Barry, D., Pietrzak, R. H., & Wagner, J. A. (2008). Overweight and obesity are associated with psychiatric disorders: Results from the national epidemiologic survey on alcohol and related conditions. *Psychosomatic Medicine*, 70, 288–297.
- Pinheiro, J. C., & Bates, D. M. (2009). *Mixed-effects models in S and S-PLUS*. New York, NY: Springer Science & Business Media.
- Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & R Core Team. (2017). nlme: Linear and nonlinear mixed effects models. Retrieved from <https://CRAN.R-project.org/package=nlme>
- Prospective Studies Collaboration; Whitlock, G., Lewington, S., Sherliker, P., Clarke, R., Emberson, J., . . . Peto, R. (2009). Body-mass index and cause-specific mortality in 900 000 adults: Collaborative analyses of 57 prospective studies. *The Lancet*, 373, 1083–1096.
- Rabe-Hesketh, S., Skrondal, A., & Gjessing, H. K. (2008). Biometrical modeling of twin and family data using standard mixed model software. *Biometrics*, 64, 280–288.
- R Core Team (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Scholten, S., Smidt, N., Swertz, M. A., Bakker, S. J. L., Dotinga, A., Vonk, J. M., . . . Stolk, R. P. (2015). Cohort profile: LifeLines, a three-generation cohort study and biobank. *International Journal of Epidemiology*, 44, 1172–1180.
- Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6, 461–464.
- Sham, P. (1998). *Statistics in human genetics*. West Sussex, UK: John Wiley & Sons.
- Stolk, R. P., Rosmalen, J. G. M., Postma, D. S., de Boer, R. A., Navis, G., Slaets, J. P. J., . . . Wolffenbuttel, B. H. R. (2008). Universal risk factors for multifactorial diseases. *European Journal of Epidemiology*, 23, 67–74.
- Therneau, T. M. (2018). *coxme: Mixed effects cox models*. R package version 2.2-10. <https://CRAN.R-project.org/package=coxme>.
- Ul-Haq, Z., Mackay, D. F., Fenwick, E., & Pell, J. P. (2012). Impact of metabolic comorbidity on the association between body mass index and health-related quality of life: A Scotland-wide cross-sectional study of 5,608 participants. *BMC Public Health*, 12, 143.
- Ul-Haq, Z., Mackay, D. F., Fenwick, E., & Pell, J. P. (2013). Meta-analysis of the association between body mass index and health-related quality of life among adults, assessed by the sf-36. *Obesity*, 21, E322–E327.
- Ul-Haq, Z., Mackay, D. F., Fenwick, E., & Pell, J. P. (2014). Association between body mass index and mental health among Scottish adult population: A cross-sectional study of 37272 participants. *Psychological Medicine*, 44, 2231–2240.
- Van der Zee, K., & Sanderman, R. (1993). *Rand-36*. Groningen, the Netherlands: Northern Centre for Health Care Research, University of Groningen.
- Van der Zee, K. I., Sanderman, R., Heyink, J. W., & de Haes, H. (1996). Psychometric qualities of the rand 36-item health survey 1.0: A multidimensional measure of general health status. *International Journal of Behavioral Medicine*, 3, 104–122.
- Visser, P. M., Hill, W. G., & Wray, N. R. (2008). Heritability in the genomics era concepts and misconceptions. *Nature Reviews Genetics*, 9, 255–266.
- Volkgezondheid, Nationaal Kompas. (2012). *Gezondheidsmonitor ggd'en, cbs en rivm*. The Hague, NL: Centraal Bureau voor de Statistiek.
- Wang, Y., Beydoun, M. A., Liang, L., Caballero, B., & Kumanyika, S. K. (2008). Will all Americans become overweight or obese? Estimating the progression and cost of the US obesity epidemic. *Obesity*, 16, 2323–2330.

- Ware, J. E., Kosinski, M., & Keller, S. D.** (1994). *SF-36 physical and mental summary scales: A user's manual*. Boston, MA: The Health Institute.
- World Health Organization (WHO).** (1995). *Physical status: The use of and interpretation of anthropometry: Report of a WHO expert committee*. Technical Report Series, 854, 1–452. Geneva, CH: World Health Organization.
- Wright, S.** (1920). The relative importance of heredity and environment in determining the piebald pattern of guinea-pigs. *Proceedings of the National Academy of Sciences of the United States of America*, 6, 320–332.
- Yip, B. H., Björk, C., Lichtenstein, P., Hultman, C. M., & Pawitan, Y.** (2008). Covariance component models for multivariate binary traits in family data analysis. *Statistics in Medicine*, 27, 1086–1105.