BMJ Open

Role of personality traits in reporting the development of adverse drug reactions: a prospective cohort study of the Estonian general population

Anu Realo,1,2 Henriët van Middendorp,3 Liisi Kõöts-Ausmees,2 Jüri Allik,2,4 Andrea W M Evers3,5

ABSTRACT

Objective To examine the role of the Five Factor Model (FFM) personality traits in reporting the development of adverse drug reactions (ADRs) when controlling for sociodemographic variables and health status.

Design Prospective cohort study.

Setting The Estonian Biobank of the Estonian Genome Centre, University of Tartu.

Participants 814 women and 543 men (mean age=47.9 years; SD=15.2) who after the initial enrolment in the Estonian Biobank were re-contacted for follow-up purposes about 5.3 years after the enrolment and for whom both self- and informant-reported personality data were available.

Main outcome measure Participants who did not report having any ADRs at baseline but who reported ADRs at the follow-up about 5.3 years later versus participants who did not report any ADRs at either time point. The reports of developing ADRs were predicted from the FFM personality traits after statistically controlling for sociodemographic variables (age, gender and education), baseline indicators of health status (number of diagnoses and medicines taken, body mass index and blood pressure), and the change in health status between the two measurements.

Results The results of a hierarchical binary logistic regression analysis showed that participants who reported the development of ADRs between the two measurements had higher levels of conscientiousness, were more likely to be women, were taking more medicines at baseline and had a higher increase in the number of medicines taken during the study period than participants who did not report any ADRs at either time point (all p values <0.05). Higher neuroticism (p=0.067) and a higher number of diagnosed diseases at baseline (p=0.053) also made marginal contributions to predicting the development of ADRs.

Conclusions This study shows for the first time that higher levels of conscientiousness and neuroticism are associated with reporting the development of ADRs.

INTRODUCTION

Adverse drug reactions (ADRs) are very common and represent significant issues in healthcare, as they, for example, may negatively impact treatment adherence. ADRs have been defined as ‘an appreciably harmful or unpleasant reaction, resulting from an intervention related to the use of a medicinal product, which predicts hazard from future administration and warrants prevention or specific treatment, or alteration of the dosage regimen, or withdrawal of the product’ (p1255).1 ADRs have a major effect on public health resulting in a broad range of clinically severe conditions (including a significant numbers of deaths) as well as an increased burden on the healthcare system due to a high number of hospitalisations and outpatient visits.2-4 The prevalence of ADRs among the general population is not well established,5 but the few existing studies have shown that the 1-month prevalence of
suspected or self-reported ADRs in the general public is about 6% to 8%.6–8

All medicines can cause ADRs, the severity of which may range from minor to life-threatening. However, the propensity and likelihood of developing and reporting ADRs does not only depend on the chemical formula or the pharmacological action of the drug, but also on a multitude of other factors, as has been shown, for instance, by the prevalence of reported ADRs in studies examining the placebo and nocebo effects of drugs.9,10

Up to now, factors associated with the reporting of the development of ADRs include genetics,11 gender,12 age,13 polypharmacy,14 drug metabolism15 and several other individual and social factors.16

A less studied factor that may be relevant for reporting on the development of ADRs relates to people’s personality characteristics. Personality traits are usually understood as enduring tendencies to feel, think and behave in a characteristic way in similar life situations17 and there has been an increasing amount of literature showing that personality has an important impact on people’s health and health-related behaviour.18–20 The most prevalent personality framework is the Five-Factor Model (FFM),21 which proposes neuroticism, extraversion, openness to experience (openness), agreeableness and conscientiousness as the main factors of personality differences. Most notably, low neuroticism22,23 and high conscientiousness24–26 have been most frequently shown to predict longevity and better physical and mental health. There is also compelling evidence showing that personality traits shape people’s subjective interpretations of their health status in important ways.27 People high in neuroticism, for instance, are more likely to report different symptoms and medical problems28 and to ask for medical help29 than people low in neuroticism. Higher conscientiousness has been found to be related to a bias toward reporting disease among persons who do not meet clinical criteria for disease,30 while people with higher levels of agreeableness are more likely to get pain relief from a placebo.31

To the best of our knowledge, only a few studies have specifically examined the relationship between personality traits and the reporting or developing of ADRs. These studies—that have mostly been conducted in small (65 to 120 participants) samples of non-representative participants, either patients or healthy volunteers in phase 1 drug trials—have shown that people with higher levels of neuroticism32,33 and/or lower levels of extraversion34 are likely to report more ADRs than less neurotic and more extraverted individuals. Therefore, affirming the role of personality traits in developing and reporting suspected ADRs in the general population is of vital importance.

The aim of the present study
The aim of the present study was to examine the role of the FFM personality traits in reporting the development of ADRs in a large population-based adult sample when controlling for relevant sociodemographic and health indicators. All ADRs were self-reported by the participants of the study during the computer-assisted personal interview (CAPI) conducted by a clinician.

Our study went beyond earlier research in several important aspects. (1) Differently from earlier studies, we used a relatively large population-based adult sample, which allows us to generalise our findings to the general population. (2) The prospective nature of our study (the participants were followed up for a mean period of 5.3 years) let us examine not just the associations between personality and self-reported ADRs, but also the role of the FFM personality traits in developing ADRs during the period of the study. (3) In addition to self-reports of personality, we employed informant-reports of personality by knowledgeable others to minimise the common method bias (ie, response tendencies within raters) and thereby increase both the reliability and validity of our findings.35 (4) When examining the role of the FFM personality traits in reporting the development of suspected ADRs, we controlled for the relevant sociodemographic (gender, age and education) and health indicators (number of medicines taken, number of diagnoses, body mass index and blood pressure) that may affect the reporting and development of ADRs.

METHOD
Sample
Participants for the present study came from the Estonian Biobank cohort that is a volunteer-based sample of the Estonian resident adult population of the Estonian Genome Centre at the University of Tartu (EGCUT).36 The participants were recruited randomly by general practitioners (GPs), physicians or other medical personnel in private practices and clinics or in the recruitment offices of the EGCUT. Each participant signed an informed consent form (available at www.biobank.ee) and the GPs or physicians performed a standardised health examination of the participants. Participants also donated blood samples for DNA, white blood cells and plasma tests as well as completed a CAPI together with a clinician on health-related topics such as lifestyle, diet and clinical diagnoses.30

Our sample for the current study includes 1357 people (814 women, 60.0%) who joined the Estonian Biobank cohort during the years of 2002–2010, who were followed up longitudinally by the EGCUT in the years of 2007–2014, and for whom both self- and informant-reported personality data were available. At baseline (T1), the mean age of the participants was 47.9 years (SD=15.2, ranging from 18 to 86 years). About 11% of the participants had basic vocational secondary education (n=151), 23.7% (n=322) had secondary, 33.8% (n=458) had vocational secondary31 and 31.4% of the participants (n=426) had higher education. The second health examination and the completion of the CAPI (ie, T2) took place on average 5.3 years (SD=3.2) after T1.36

Vocational secondary education means that besides vocational training the student also acquires upper secondary education.
The mean age of informants (996 women, 75.1%) was 45.2 (SD=16.1) years. On average, the informants had known the targets for 27.4 (SD=14.6) years. About 41% of the informants were spouses or partners, 24.8% were parents, 14.4% were friends and 18.7% were other relatives (eg, children and siblings) or acquaintances.

MATERIALS

Personality

The Estonian version of the NEO Personality Inventory-3 (NEO PI-3) was used to assess personality. The NEO PI-3 has 240 items that measure 30 personality facets, which are grouped into the five FFM domains—neuroticism, extraversion, openness, agreeableness and conscientiousness—such that each domain score is a composite of six facet scores. Neuroticism is a general tendency to experience negative emotions such as sadness, anger, hostility and fear. People who score high on neuroticism are prone to anxiety and psychological distress, whereas people scoring low on neuroticism are usually emotionally stable and well-adjusted. Extraversion is a tendency to experience joy and other positive emotions, to seek out stimulation, to be active and sociable. People who score high on openness to experience are usually intellectually curious, attentive to their inner feelings, open to new ideas and opportunities. People scoring low on openness tend to be conservative and conventional, preferring the ‘old way’ of doing things rather than the ‘new’. Agreeableness is mostly about interpersonal tendencies referring to people’s altruistic, sympathetic, cooperative and trustful nature. Finally, conscientiousness refers to individual differences in an active process of planning, organising, and carrying out tasks. People scoring high in conscientiousness are thorough, reliable, and task-oriented and goal-oriented, whereas people scoring low in conscientiousness are more laid-back and less anxious about setting and reaching their goals.

Participants completed the self-report form and informants the observer-report form of the Estonian NEO PI-3. The items were answered on a 5-point scale (0 = false/strongly disagree … 4 = true/strongly agree). Most of the participants (88%) completed the personality inventory in the same year as T2 or during the year after T2 (10%). Nearly all informants (96.5%) completed the NEO PI-3 in the same year as the participants. The descriptive statistics of the scales, including Cronbach alphas, are shown in online supplementary table S1. Self- and informant-reports of the NEO PI-3 personality traits correlated with each other in the expected magnitude. Pearson rs were 0.55 for neuroticism, 0.65 for extraversion, 0.61 for openness, 0.46 for agreeableness and 0.50 for conscientiousness (all p values <0.001). For all subsequent analyses, the mean score of self- and informant-ratings across the five domain scales was used to minimise the common method bias due to individual response biases.

Adverse drug reactions

The participants were asked whether there are any medicines that have caused them ADRs during their whole life and if yes, what were the specific suspected ADRs. All ADRs were self-reported by the participants of the study during the CAPI conducted by a clinician. The medicine that the respondent reported as causing their ADR was coded according to the Anatomical Therapeutic Chemical (ATC) classification system (http://www.whocc.no/atc_ddd_index). The suspected ADRs caused by medicines were coded using the 10th revision of the International Statistical Classification of Diseases and Related Health Problems (ICD-10) classification system.

Health indicators

All measures of current health status described below were retrieved from the EGCUT, which contains data gathered from the databases of healthcare institutions and registries, as well as from the information provided by the participant, using the above-mentioned CAPI. Anthropometric measures (ie, weight and height) and blood pressure were measured by EGCUT recruiters at the end of the CAPI. The means of the health indicators at T1 and T2, and the difference and correlations between the measurements are shown in the online supplementary table S2.

Clinical diagnoses. For each participant, the number of diagnoses recorded in the Estonian Health Insurance Fund (EHIF) for the year of data collection (both at T1 and T2) was used. The EHIF covers the costs of health services required by eligible persons in the case of illness and is the only organisation in Estonia dealing with compulsory health insurance (https://www.haigekassa.ee/en).

Use of medicines. First, participants were asked during the CAPI which diseases they had been diagnosed with and which medicines they had used during the previous 2 months in connection with these illnesses (‘Which medicines have you used during the last 2 months in connection with diagnosed diseases?’). Later, the participants were also asked ‘Which medicines do you use regularly (which have not been discussed earlier)’? The total number of taken medicines, either regularly or in connection with specific diseases during the last 2 months, was used in all later analyses.

Blood pressure. Systolic and diastolic blood pressure (BP) were measured in a sitting position at the end of the CAPI interview both at T1 and T2.

Body mass index (BMI). BMI was calculated on the basis of objectively measured weight and height during the CAPI at T1 and T2 as weight/height$^2$ (kg/m$^2$).

Statistical analyses

Pearson product–moment correlation coefficients were computed to examine correlations between self- and informant-reports of the NEO PI-3 personality traits. Cronbach alphas were calculated to examine internal consistency of the NEO PI-3 five domain scales. A one-way
analysis of variance was used to examine the mean differences in the continuous variables between the participants who reported developing ADRs between T1 and T2 and who did not report any medicines causing ADRs at either time point. Gender and education differences between the participants who reported developing ADRs between T1 and T2 and who did not report any medicines causing ADRs at either time point were tested using the χ² test. A hierarchical binary logistic regression analysis was performed to examine the role of the FFM personality traits in reporting the development of ADRs between T1 and T2 when controlling for sociodemographic variables, baseline measurements of health status and the change in health status between the two measurements. In logistic regression analysis, higher education was defined as the reference or baseline category (1) for the education variable. All analyses were conducted using IBM SPSS Statistics software, V.24.

Patient and public involvement
Patients and public were not involved in the development of the research question and the selection of outcome measures.

RESULTS
The reported frequency of ADRs
At baseline (T1), there were 162 participants (11.9% of the total sample) who reported taking medicines that had caused them ADRs. Among those who reported taking medicines that had caused them ADRs at T1, the average number of ADRs per participant was 1.5 (SD=1.4), with about 70% of participants reporting one ADR, 17% two ADRs and the remaining three or more ADRs.

At T2, there were roughly twice as many participants (n=329, 24.1% of the total sample) than at T1 who reported taking medicines that had caused them ADRs; difference between time points: χ²(1)=201.9, p<0.001. On average, 1.6 ADRs (SD=1.2) were reported per participant at T2, with 66% of participants reporting one, 22% two and the remaining participants three or more ADRs. The frequency of specific ADRs (as coded by ICD-10) both at T1 and T2 is shown in the online supplementary table S3.

There were 217 participants who did not report ADRs caused by medicines at T1, but did so at T2. In our further analyses, this group of participants (n=217) was compared against those participants (n=978) who did not report any medicines causing ADRs at either time point. A χ² test showed that respondents who developed ADRs between T1 and T2 were predominantly women and had less likely higher education compared with people who did not report any ADRs either at T1 or T2 (see table 1). The results of the analyses of variance showed that people who developed ADRs between T1 and T2 had higher levels of neuroticism and conscientiousness, were older, had been diagnosed with more diseases (both at T1 and T2), were taking more medicines (both at T1 and T2) and had higher BMI and systolic BP (only at T2) than those who did not report ADRs at either time point (all differences significant at p<0.05). Finally, people who developed ADRs by T2 had also a bigger increase in the number of medicines taken and in the level of BMI from T1 to T2 as compared with those who never reported any ADRs (all differences significant at p<0.05).

Next, we conducted a hierarchical binary logistic regression analysis in order to predict the reporting on the development of ADRs between T1 and T2 from the FFM personality traits when controlling for sociodemographic variables, baseline measurements of health status and the change in health status between the two measurements. Only those variables were included in the model as predictors on which there were significant differences between people who reported developing ADRs between T1 and T2 versus participants who did not report any ADRs at either time point.

The following nine variables entered in three blocks were included in the binary logistic hierarchical regression model in order to predict the reporting of the development of ADRs: (1) baseline age, gender and education; (2) number of diagnoses at T1, number of medicines taken at T1, change in the number of medicines taken from T1 to T2 and change in BMI from T1 to T2; and (3) neuroticism and conscientiousness at T2.

The nine variables together explained 11.6% of the variation (as indicated by the Nagelkerke pseudo R²) in the dependent variable (ie, developing ADRs between T1 and T2). Being a woman, taking more medicines at T1, an increase in the number of medicines taken from T1 to T2 and higher...
levels of conscientiousness were all associated with increased odds of developing suspected ADRs from medicines between T1 and T2 (p values<0.05). A higher number of diagnosed diseases (p=0.053) and higher neuroticism (p=0.067) also made marginal contribution to predicting the reporting of the development of ADRs (see table 2).

**DISCUSSION**

In the current study, we examined the role of the FFM personality traits in reporting the development of suspected ADRs in a large-scale adult cohort of >1300 individuals, when controlling for sociodemographic variables and health status. The findings of our analysis
Table 2  Results of the hierarchical binary logistic regression analysis examining the role of the FFM personality traits in reporting the development of ADRs between T1 and T2 while controlling for sociodemographic and health indicators at T1 and the change in health status from T1 to T2

<table>
<thead>
<tr>
<th>Block 1</th>
<th>Block 2</th>
<th>Block 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>SE</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>SE</td>
</tr>
<tr>
<td>Sociodemographic variables (T1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.02</td>
<td>0.01</td>
</tr>
<tr>
<td>Gender</td>
<td>−0.89</td>
<td>0.17</td>
</tr>
<tr>
<td>Education</td>
<td>7.97</td>
<td>0.047</td>
</tr>
</tbody>
</table>

Nagelkerke $R^2$=0.058

$χ^2$(5)=50.37, p<0.0001

Health indicators

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>SE</th>
<th>Wald</th>
<th>p values</th>
<th>Exp(B)</th>
<th>B</th>
<th>SE</th>
<th>Wald</th>
<th>p values</th>
<th>Exp(B)</th>
<th>B</th>
<th>SE</th>
<th>Wald</th>
<th>p values</th>
<th>Exp(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of EHIF diagnoses (T1)</td>
<td>0.04</td>
<td>0.02</td>
<td>4.37</td>
<td>0.037</td>
<td>1.04</td>
<td>0.04</td>
<td>0.02</td>
<td>3.74</td>
<td>0.053</td>
<td>1.04</td>
<td>0.10</td>
<td>0.03</td>
<td>10.07</td>
<td>0.002</td>
<td>1.11</td>
</tr>
</tbody>
</table>

Nagelkerke $R^2$=0.018

$χ^2$(2)=31.01, p<0.0001

NEO PI-3 domain scales (T2)

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>SE</th>
<th>Wald</th>
<th>p values</th>
<th>Exp(B)</th>
<th>B</th>
<th>SE</th>
<th>Wald</th>
<th>p values</th>
<th>Exp(B)</th>
<th>B</th>
<th>SE</th>
<th>Wald</th>
<th>p values</th>
<th>Exp(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neuroticism</td>
<td>0.01</td>
<td>0.00</td>
<td>3.37</td>
<td>0.067</td>
<td>1.01</td>
<td>0.01</td>
<td>0.01</td>
<td>5.48</td>
<td>0.019</td>
<td>1.01</td>
<td>0.01</td>
<td>0.01</td>
<td>5.48</td>
<td>0.019</td>
<td>1.01</td>
</tr>
</tbody>
</table>

Nagelkerke $R^2$=0.116

$χ^2$(2)=6.37, p<0.041

Constant | −2.51 | 0.31  | 68.00 | <0.001  | 0.08   | −2.37 | 0.33  | 51.44 | <0.001  | 0.09   | −4.45 | 0.90  | 24.46 | <0.001 | 0.01 |

Number of EHIF diagnoses is the number of diagnoses recorded in the Estonian Health Insurance Fund for the year of data collection.
Number of medicines taken is the number of medicines taken during the previous 2 months, either regularly or for treating specific diseases.
Gender: 0=male, 1=female; Education: higher education was defined as the reference or baseline category (1).
ADR, adverse drug reaction; BMI, body mass index; EHIF, Estonian Health Insurance Fund; FFM, Five-Factor Model; NEO PI-3, NEO Personality Inventory-3; T1, first measurement; T2, second measurement (on average 5.3 years after T1).
showed that participants who reported developing ADRs during the study had significantly higher levels of conscientiousness (p=0.019) as well as marginally higher levels of neuroticism (p=0.067), even when the relevant sociodemographic and health variables were taken into account.

Extensive research has revealed that conscientiousness is a strong lifelong predictor of longevity and good health.25 42 There are multiple pathways by which conscientiousness is associated with health, but most likely conscientiousness promotes health by engaging in healthier behaviours and by choosing healthier situations and environments, thereby minimising different small health risks.36 Among other things, conscientiousness has been found to be related to more frequent visits to the general practitioner,30 meaning that conscientious individuals are more vigilant about their health and pay greater attention to their health problems. This may explain why conscientious individuals in our study were more likely to develop and report ADRs, as they may perceive any bodily reactions as an important health risk factor.

Another personality trait that was significantly associated and, after controlling for relevant sociodemographic and health-related variables, marginally predicted the development and reporting of ADRs in our study was neuroticism. Similarly to low conscientiousness, high neuroticism is a strong and robust predictor of different mental and physical health problems as well as of the frequency of health service use.22 Again, there are probably different pathways through which neuroticism influences health, but it is possible that the development of ADRs among people high in neuroticism is related to their enhanced physiological responses that result from chronic over-activation of the autonomic nervous system18 or from greater sympathetic and hypothalamic–pituitary–adrenal reactivity.22 It has been also found that people high in neuroticism respond more strongly to sensations of pain and discomfort and, as a result, they are more likely to notice and attend to bodily sensations and minor discomforts, including unfounded symptoms without a physiological basis.28 30 Also, conscientiousness is related to a bias toward reporting disease among individuals who actually do not meet clinical criteria for disease, meaning that highly conscientious individuals ‘may use lower criteria to establish illness because they are more cautious about their health and eager to report illness so they may obtain an early diagnosis and treatment’ (p. 576).30 Thus, it is possible that individuals high in conscientiousness and high in neuroticism are over-reporting ADRs due to heightened levels of self-awareness, pain sensitivity or attentional focus, which results in heightened perception, over-reaction and vigilance about their bodily sensations and symptoms.30 41 High levels of neuroticism, accompanied by high conscientiousness—a phenomenon called ‘healthy neuroticism’—is however also seen as a good thing and is often associated with health benefits, as persons showing this combination are vigilant about their health and in case of any health problems take the necessary action.44 45

Despite several strengths of this study, including the relatively large population-based adult cohort, the prospective nature of our research, the use of both self-ratings and observer ratings of personality and the well-validated personality inventory, our research also had some limitations. One of the main limitations of our study was that we were not able to account for the severity of the reported ADRs, which has been found to be one of the main motivations for consumers to report ADRs.46 However, large-scale studies in Denmark and the UK have shown that the general public and physicians tend to report similar proportions of serious or severe ADRs and, therefore, we have no reason to believe that taking into account the severity of ADRs would have impacted our results in any substantial way.

Another limitation of our study relates to the fact the design of our study did not allow for causality assessment of ADRs, which is the approved method used for estimating the strength of relationship between drug(s) exposure and occurrence of ADRs.49 Instead, similarly to several earlier studies,7 8 our study examined people’s perceptions of an unwanted effect that they attributed to the use of a medicinal product. However, there is substantial evidence that the quality of patient reports is similar to that of health professional reports.50 It has been also suggested that since only a small percentage of the ADRs that occur are reported by health professionals, it is difficult to establish the true occurrence and extent of ADRs in the general population, and therefore, more information on the prevalence of experienced ADRs and on how patients themselves perceive ADRs, as was done in the present study, is needed.7 Therefore, we believe that our study makes a valuable contribution to this important stream of research, which has a strong potential to increase our knowledge about the possible harm of medicines.50 51

The last limitation is that FFM personality traits were measured only once, and not at baseline, but at the follow-up about 5.3 years after the enrolment in the study. Although personality traits are not set in stone and people change in terms of personality traits across the life course,52 there is little reason to believe that the scores of neuroticism and conscientiousness would have dramatically changed during our study or that having experienced a difficult life event such as serious illness or injury, for instance, would have had strong effects on the scores of the FFM personality traits.53 54 On the contrary, it is personality traits that consistently predict the occurrence of different life events,36 7 including the onset of diseases.57

To sum up, the findings from the current study, examining personality traits in addition to sociodemographic and health indicators as predictors of reporting the development of ADRs in a large population-based adult sample, suggest that people with high levels of conscientiousness and neuroticism are more likely to report on the development of ADRs, even after controlling for age, gender, medicine use and the number of diagnoses. As people with high levels of neuroticism are less likely to
volunteer to phase I drug trials than people with low levels of neuroticism, our results suggest that certain ADRs may remain undetected in phase I drug trials or the number of ADRs may be smaller than expected from the general population. Even more importantly, reporting of ADRs is no longer reserved for drug trials or healthcare professionals, but also available to the general public who have been invited as reporters of drugs and ADRs in many European and world countries. Thus, our findings also suggest that people high in conscientiousness and neuroticism may be more likely to report ADRs using spontaneous reporting systems such as the Yellow Card Scheme in the UK, for instance. However, we should note that it is not entirely clear whether people high in conscientiousness and neuroticism (so-called ‘healthy neurotics’) over-report ADRs (including physiologically unfounded symptoms) or whether ‘healthy neurotics’ are actually adequate in reporting of ADRs and it is those low in conscientiousness and neuroticism who significantly under-report potential ADRs. To that aim, future longitudinal studies need to be performed, in which participants high and low on these personality traits are followed in their health course for a longer period of time. Nevertheless, our findings are among the first to demonstrate that personality traits such as conscientiousness and neuroticism are not merely associated with reporting ADRs, but may also serve as risk factors for the development of ADRs on top of different sociodemographic and health indicators.

Acknowledgements The authors are grateful to the Estonian Genome Centre of the University of Tartu and its director, Professor Andres Metspalu, for help in collecting the data and the kind permission to use the data in the current study.

Contributors AR: conceived the study, prepared the data, and contributed to the analysis of data and writing of the manuscript. LH-A: helped with data preparation. AWME, Hvm and JA contributed to the interpretation of data and critically reviewed and commented on the manuscript. All authors reviewed and approved the final version of the manuscript.

Funding Preparation of this manuscript was supported by institutional research funding (IUT2-13) from the Estonian Ministry of Education and Science to JA and by the European Research Council Consolidator Grant to AWME (ERC-2013-CoG-617700).

Disclaimer The authors were completely independent from funders in conducting this study and writing this manuscript.

Competing interests None declared.

Patient consent Not required.

Ethics approval This research was approved by the Research Ethics Committee of the University of Tartu (approvals: 236/M-29, 14 May 2014; 206/T-4, 22 August 2011; 170/T-38, 28 April 2008; 166/T-21, 17 December 2007).

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No additional data available. The data used for this research are available for scrutiny at the Estonian Genome Centre of the University of Tartu (EGCUT), but cannot be released because of the licensing conditions to which we are obliged to adhere.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

REFERENCES