Title

The association of a single nucleotide polymorphism in the nuclear factor erythroid derived-2 like 2 (Nrf-2) gene with adverse drug reactions, multimorbidity and frailty in older people

Authors

[1]Greg Scutt BPharm PhD, [1]Andrew Overall BSc PhD, [1]Prijay Bakrania MPharm, [2]Eliseveta Krasteva MBBS, [3]Nikesh Parekh MBBS PhD, [3]Khalid Ali MBBS MD, [4]J Graham Davies BPharm MSc PhD, [3]Chakravarthi Rajkumar MBBS MD PhD

Affiliations

1. School of Pharmacy and Biomolecular Sciences, University of Brighton, Brighton, UK
2. Dartford and Gravesham NHS Trust, UK
3. Brighton and Sussex Medical School, Brighton, UK
4. Institute of Pharmaceutical Science, King's College London, London, UK

Corresponding author

Dr Greg Scutt
School of Pharmacy and Biomolecular Sciences
University of Brighton
Brighton, UK
Email: g.scutt@brighton.ac.uk

Word count

4,455 (excluding references)

Abstract

Susceptibility to adverse drug reactions (ADRs), multimorbidity, and frailty are associated with human ageing, yet there is wide variation in the severity and age at which individuals are afflicted. Identifying genetic markers of increased risk of this phenotype would help stratify individuals to specialist interventions. Nuclear factor erythroid derived-2 like 2 (Nrf2) regulates a cell's response to stressors, including the expression of enzymes involved in drug-metabolism. Its expression has been shown to decline in animal ageing models. In this study we tested the hypothesis that Nrf2 gene transcription/translation decline in human ageing, and that single nucleotide polymorphisms (SNPs) in the Nrf2 gene are associated with increased ADR risk, multi-morbidity, and frailty in older people. Gene expression and protein levels were measured in peripheral blood mononuclear cells (PBMCs) donated from healthy patients aged 18-80 years old. Nrf2 genotypes were determined at three loci in a sub-population of patients recruited to the PRIME study (a multicentre prospective cohort study that followed older adults for 8-weeks post-discharge to determine ADR). Both Nrf2 gene and protein expression declined significantly with age in human PBMCs. In the PRIME sub-study population, the rs35652124 Nrf2 SNP was associated with increased ADR risk, and decreased frailty and multi-morbidity scores.

Keywords

1. Introduction

Older people are the largest consumers of healthcare interventions and the recipients of the majority of prescribed medications. As a population, they are characterized by multimorbidity, frailty, and polypharmacy, and suffer more adverse drug reactions (ADRs) requiring hospitalization than younger adults (1). Recent studies have attempted to identify clinical and social predictors of this phenotype, notably multimorbidity, frailty, and an increased susceptibility to ADRs, in an attempt to build models to predict those in need of early, or specialist intervention. For example, several tools have been developed which use clinical factors to identify older patients at risk of developing an ADR (2-4). However, it is clear that clinical and social factors alone do not explain all variation in the populations studied. Uncertainty remains regarding why some individuals, with limited exposure to known risk factors, still develop marked multimorbidity, frailty and increased susceptibility to ADRs in old age.

Oxidative stress is at the core of many theories of ageing (5). Indeed, Reactive Oxygen Species (ROS) are thought to play a central role in the pathology of several age-related diseases such as Parkinson's disease (6) and cognitive impairment (5, 7) where they are thought to disrupt signaling pathways and cause cellular damage. Organisms have, however, evolved an elaborate system to protect against oxidative damage. Under conditions of acute oxidative stress, the transcription factor Nuclear Factor (Erythroid Derived 2)-Like 2 (Nrf2) becomes activated and induces the expression of a myriad of antioxidant genes. However, recent evidence from rodent models shows that the activity of Nrf2 declines with age, along with a decrease in downstream anti-oxidant protein levels (8), increasing susceptibility to age-related diseases and the signs of ageing (9-13). What is particularly interesting is that Nrf2 not only induces the expression of antioxidant proteins, but also enzymes and

transporters involved in Phase I and II drug-metabolism, such as glutathione-s-transferase, and p-glycoprotein (14). If a decline in Nrf2 activity is mirrored in human ageing, not only could this provide an explanation for increasing multi-morbidity and frailty, but also for the increased risk of ADR and altered pharmacokinetics seen in human ageing. Interestingly, there is some emerging evidence that both the nuclear concentration of Nrf2 protein, and its activity are reduced with age in human bronchial epithelial cells (15). Whether a reduction is Nrf2 levels during ageing is exhibited in other human cells types, and whether it is the cause of the ageing phenotype has yet to be elucidated.

There are several Single Nucleotide Polymorphisms (SNPs) in the promoter region of the human Nrf2 gene which affect the expression of Nrf2 *in vivo*. These SNPs have been shown to associate with specific age-related diseases, including Acute Lung Injury (16), impaired forearm vasodilator response (17), and Parkinson's Disease (18). We might hypothesize therefore, that in the context of an age-related decline in Nrf2, individuals possessing a variant allele may be more sensitive to the adverse effects of medicines, have a greater number of co-morbidities, and are frailer. This explorative study therefore has two main aims: (1) to test for an age-related decline of Nrf2 transcription and translation in peripheral blood mononuclear cells (PBMCs), and (2) to establish whether common SNPs in the promoter region of the Nrf2 gene are associated with an increased susceptibility to ADRs, multimorbidity, and frailty in older individuals.

2. Methods

2.1 Ethical approval

Ethical approval was granted by the University of Brighton Tier 2 Research Ethics committee (CRECLHPS-16-063) and by the South Central – Berkshire B Research Ethics Committee (16/SC/0240).

2.2 Changes to the expression of Nrf2 with age

2.2.1 Participants

A convenience sample of healthy individuals (volunteers), $\geq$ 18-years-old were recruited over an 8-month period from a population of University students and staff. Participant's age, gender, details of regular medicines used, details of chronic medical conditions, and smoking, alcohol and recreational drug taking behaviour were recorded. Each participant donated 6 mL blood. To detect a medium to large effect size (Cohen's *f*) of age on Nrf2 expression, with 80% power at a significance level of 0.05 we calculated the required sample size to be in the region of 80 participants.

2.2.2 Isolation of peripheral blood mononuclear cells (PBMCs)

PBMCs were isolated from donated whole blood by density gradient centrifugation (Ficoll-Paque Plus (GE Healthcare Life Sciences, UK)) within 4 hours of donation according to established protocols. Isolated PBMCs were split into two microcentrifuge tubes for determination of Nrf2 gene and protein expression.

2.2.3 Extraction of messenger RNA

Messenger RNA (mRNA) was extracted from PBMC pellets within 2 hours of isolation using a RNeasy Mini Kit (Qiagen, Hilden, Germany) in accordance with the manufacturer's

instructions. Extracted mRNA was immediately converted to cDNA using QuantiTect Reverse Transcriptase (Qiagen, Germany). cDNA samples were then stored at -20°C until later analysis.

2.2.4 *Quantitative real time polymerase chain reaction (q-rtPCR)*

Quantitative real time polymerase chain reaction (q-rtPCR) was carried out using a Rotor-Gene Q 5plex PCR machine (Qiagen, Germany). Separate cDNA samples from each participant were mixed with SYBR green MasterMix (Qiagen, Germany) along with forward and reverse primers (Eurofins Genomics, Germany) for either Nrf2 (gene of interest), glyceraldehyde-3-phosphate dehydrogenase (GAPDH) or 18s ribosomal subunits (both housekeeping genes). Each participant sample was run in duplicate. Primer sequences are listed in Table 1. The conditions for q-rtPCR were: initial denaturation 95°C for 15 mins, followed by 40 cycles at 94°C for 15 s, 55°C for 30 s, and 72°C for 30 s.

The log abundance, and log fold change of Nrf2 was determined using the Matz analytical method (MCMC.qpcr package) (19) in the statistical software program R (20).

2.2.5 *Determination of Nrf2 protein levels in PBMC samples*

Expression of Nrf2 protein was carried out using enzyme-linked immunosorbent assay (ab207223; Abcam, UK). Isolated PBMCs were suspended in lysis buffer, and then prepared for the assay according to the manufacturer's instructions. Absorbance was measured using an Elisa Plate Reader at a wavelength of 450 nm with a reference set at 665 nm.

2.3   Association of Nrf2 genotype / haplotype with adverse drug reactions, multimorbidity and frailty

2.3.1 Participants and recruitment

A subset of older patients who had recently participated in the PRIME study (4, 21) were invited to send a cheek-swab of genetic material to establish their genotype at three Nrf2 SNP loci (rs35652124, rs6706649 and rs6721961) and to provide consent for study co-ordinators to access participant data collected during the original PRIME study . The PRIME study was a large multi-centre study investigating the incidence and cost-implications for medication-related harm (including ADR) in older people 8 weeks post-hospital discharge. Inclusion criteria were: (1) prior participation in the PRIME study. Exclusion criteria were: (1) PRIME participants who measured <7/10 on the abbreviated mental test score (AMTS) during the PRIME study, or those who during a follow-up phone call were deemed to have insufficient capacity and (2) deceased following participation in PRIME study (as determined by examination of hospital electronic records).

2.3.2 PRIME study participants and dataset

The PRIME study recruited 1280 older adults at hospital discharge from 5 hospitals in England between 2013 to 2015 (21). Patients were assessed for the presence of an ADR over an 8-week period by a senior pharmacist using the Narajo algorithm (22). A frailty index was developed and internally validated (23) based on the Rockwood approach (24). This index included 55 age-related deficits from multiple domains (e.g. morbidity, cognition, mood, strength and mobility, nutrition, daily function). Each patient's disease burden was calculated by (1) a summation of comorbidities and (2) calculating a Charlson Index score using established criteria (25).

2.3.2 Sample size

In a study investigating the association of the -617 (rs6721961) single nucleotide polymorphism (SNP) with Acute Lung Injury, Marzec et al. (16) demonstrated the SNP to increase the odds of disease by approximately 6-fold. If we consider the odds ratio of Nrf2 SNPs on the ageing phenotype to be similar to this value, and if we assume the population ADR rate to be around 10%, then this equates to an effect size of w=0.280. By setting the significance level to 0.05 and power to 0.8, we therefore estimated a suitable sample size to be 100 participants.

2.3.3 Genotyping of genetic material

Genomic DNA was extracted from donated cheek swabs (Isohelix SK-1S; Isohelix, UK) using DNeasy Blood and Tissue Kits (Qiagen, Germany). Isolated gDNA was checked for purity using a Nanodrop® Lite Spectrophotometer (Sigma, UK), before storage at -20°C prior further analysis. DNA deemed to be of insufficient quality or quantity was discarded (e.g., yield < 1 ng/microL and A260/A280 < 1.4).

Isolated gDNA samples were amplified using HotStar Taq Mastermix Kit (Qiagen, Germany) with forward and reverse primers for the Nrf2 gene (Eurofins Genomics, Germany; Table 1). Polymerase Chain Reaction conditions were as follows: 95°C for 4 mins (initial denaturation), then 35 cycles of 95°C for 45 s, 56°C for 45s and 72°C for 45 s, followed by a final elongation step at 72°C for 10 mins. PCR products were sequenced by Sanger sequencing. Genotype was determined by visual analysis of the electropherogram at the rs35652124, rs6706649 and rs6721961 loci.

2.4   Data and statistical analysis

When testing for age-related changes to Nrf2 gene expression, participants were stratified in age categories (18-29, 30-39, 40-49, 50-59, 60-69, >70). Log abundance and log-fold change of Nrf2 as a function of age category were then analysed using the Markov Chain Monte Carlo (MCMC) analytical method described by Matz *et al.* (MCMC.qpcr R statistical package, 2013) (19). The MCMC.qpcr package uses a generalised linear mixed model based on lognormal-Poisson error distribution, that is fitted using MCMC. Data from the qPCR experiments (gene of interest, and house-keeping genes) were organised according to participant and age-category. Results were tested for significance using a One-way ANOVA with Dunnet's post-hoc correction for multiple comparisons. The Changes in Nrf2 protein expression were compared between participants aged <30 and >30 years old using an unpaired Student's t-test.

Genotype and haplotype association studies across all ages were conducted using Haploview software (Broad Institute, www.broadinstitute.org/haploview). The software uses a two marker expectation maximisation (EM) algorithm to estimate the maximum-likelihood values of the four gamete frequencies (26), given as Chi^2 with significance at 5%. Odds ratios and 95% confidence intervals were determined from 2x2 contingency tables for single gene or haplotype of interest in cases and controls. Single gene association analyses were conducted on the whole cohort, and then on a sub-population of participants aged 65-79 years old. Mann-Whitney tests were used to test for significance in single gene association analyses as the data followed a non-Gaussian distribution.

Data handling was conducted in Microsoft® Excel. Descriptive statistics and statistical testing (Chi^2, analysis of variance (ANOVA), Student's t-test, Mann-Whitney U test and

post-hoc analyses) were carried out using Graphpad Prism for Mac Version 6.00 (Graphpad Software, La Jolla California, USA). Sample and power analysis, and Monte Carlo Markov Chain q-rtPCR analysis were conducted in R (R Core team, 2014 (20)), using the 'pwr' and 'MCMC.qpcr' packages respectively. Statistical significance was accepted if $p < 0.05$. Data are presented as mean $\pm$ standard deviation unless indicated otherwise.

Results

2.5   Changes to the expression of Nrf-2 with age

2.5.1   Participant demographics

A total of 55 healthy volunteers provided consent and were enrolled into this part of the study. The median age was 30 years (range 18-75), with 25 male participants (45%). Twenty-two participants (40%) self-reported at least 1 chronic medical condition, with the same number taking at least 1 regular medication. Seven participants reported smoking at least 1 cigarette regularly each week, and thirty-nine participants (71%) reported consuming alcohol regularly, with a median of 4.0 units consumed each week (range 0-16). As might be expected, we found a significant relationship between age category and both the number of reported chronic medical conditions, and the number of regular medications taken by participants (Chi^2 test, $p<0.001$ and $p<0.01$ respectively). A breakdown of the sample characteristics according to age category is provided in Table 2.

2.5.2   Nrf-2 gene expression changes with age

Using the Markov Chain Monte Carlo analytical method described by Matz *et al.* (2013) (19), we found that the Log abundance of Nrf2 is significantly lower in those aged >70 compared to all younger age groups ($p<0.0001$, One-way ANOVA with Dunnet's post-hoc correction for multiple comparisons; Figure 1A). Similarly, we found that the negative Log-fold change in Nrf2 is significantly greater in the >70-year-old age-group compared to all other ages ($p<0.0001$, One-way ANOVA with Dunnet's post-hoc correction for multiple comparisons; Figure 1B). Despite recruiting fewer participants than our target sample size, post-hoc analysis revealed a larger than anticipated effect-size of f=0.64, and power (β-1)=0.95.

2.5.3   Nrf-2 protein levels changes with age

To test whether translation of the Nrf2 gene is reduced with age, we measured Nrf2 protein levels in PMBCs extracted from volunteers. The reason for looking at both Nrf2 mRNA and protein levels is because the basal level of Nrf2 protein is determined by a combination of both proteasomal degradation of Nrf2 protein, and Nrf2 gene transcription / translation. A decrease in Nrf2 protein levels coupled with a decrease in Nrf2 gene expression will tell us that the mechanism is primarily transcription related, whereas a reduction in protein levels without a decrease in gene-expression will suggest that the mechanism is translation / degradation related.

Unfortunately, several protein samples were used up from our original sample of 55 which meant that we were unable to compare protein levels across our originally determined age-categories (as several categories contained only single measurements). Instead, we compared Nrf2 protein expression between volunteers aged <30 years old (n=9) to those aged >30 years old (n=9) in two equally sized groups. We found that Nrf2 protein expression was significantly lower in the >30-year-old age-group compared to the younger group ($0.065 \pm$ SEM 0.013 vs. $0.033 \pm$ SEM 0.006 AU in <30-year-old and >30-year-old respectively; $p < 0.05$, unpaired Student's t-test; Figure 1C), in line with changes observed in gene transcription.

2.6   Association of Nrf-2 genetic polymorphisms with characteristics of ageing

2.6.1   Participant demographics

In this part of the study we tested the hypothesis that in the context of an age-related decline in Nrf2 expression, individuals possessing a variant Nrf2 allele would show increased susceptibility to ADR, and greater multimorbidity and frailty. One hundred and thirty-one older patients from a sub-population of the PRIME study provided consent and donated genetic material. Four of the 131 patients were deemed to have insufficient capacity during a follow-up telephone call (i.e. they were unable to recall details of the study upon questioning) and were removed from the study. Genetic material that was supplied by a further 15 patients was of unsatisfactory quality to allow further processing and so these individuals were also excluded from the study. A total of 112/131 patients had their Nrf2 gene successfully sequenced and genotyped at all 3 loci. Table 3 provides a description of the demographics of these 112 patients. One SNP was found to be in Hardy-Weinberg equilibrium (rs6721961)1, but two, the rs35652124 and the rs6706649, deviated significantly from equilibrium (p<0.0001; Chi^2 test). The frequencies of the alleles for the three loci studied in the Nrf2 promotor region are shown in Table 4.

2.6.2   Adverse drug reactions

3.2.4.1 Single gene associations

The ADR rate across our study population (when classified as either definite, probable or possible on a modified Naranjo algorithm) was found to be 34% (38/112). We found that there was a significantly higher frequency of ADRs in individuals who were homozygous for **G** at the rs35652124 locus, compared to those who were either heterozygous or homozygous for **A** (p<0.05, Chi^2 test, n=112). Tables 5 provides details of the numbers of the sub-population of patients who developed an ADR during the original PRIME study according to

genotype. When a stricter definition is used to classify ADR (i.e. only definite/probable on the modified Naranjo algorithm) we found the association between ADR and genotype was lost (Table 5). The ADR rate in the total population when this definition was used was found to be 19.6% (22/112). No associations were found between ADR and genotype at the other two loci (rs6721961 or rs6706649).

4.2.4.2 Haplotype associations

Due to their close physical proximity on the genome, and the potential for these loci being in linkage disequilibrium, we thought that it would be of value to explore whether certain combinations of Nrf2 alleles were associated with ADR. Using our stricter definition of ADR (probable/definite according to the Naranjo algorithm), we found that the CGA haplotype (rs6721961, rs35652124 and rs6706649 respectively) showed increased odds of ADR compared with other haplotypes (OR 12.9, CI 1.01-166.0; p=0.01, Chi^2 test). The wide confidence intervals reflect the fact that the haplotype concerned was rare (n=3) in our population sample.

2.6.3   Multimorbidity

2.6.3.1 Single allele associations

We found a significant relationship between genotype at the rs35652124 locus and multimorbidity. Patients that were homozygous for the **A** allele typically had more co-morbidities than AG/GG individuals (3 [interquartile range 2-5] vs. 3 [interquartile range 1-4], AA vs. AG/GG respectively; p<0.05, Mann-Whitney U test; Figure 1A). However, we found no significant difference between median Charlson Index scores when comparing both genotypes across the entire age-range (1 [interquartile range 1-3] vs. 1 [interquartile range 0-2], AA vs AG/GG respectively; p>0.05 Mann-Whitney U test; Figure 1A).

To probe the effect of age further, we split our cohort into two groups: those aged 65-79, and those aged >80. The value of this approach is that it allows us to look at the >80-year-old group separately, as these individuals are a poorly studied age-group, typically referred to as the oldest-old, and who may have a different underpinning physiology to the 'younger' old. In doing so, we found that there was no difference in multi-morbidity between genotypes in the >80-year-old age group. Interestingly however, individuals aged between 65-79 who had the AG/GG genotype had significantly fewer co-morbidities (3 [interquartile range 2-5] vs. 2 [interquartile range 1-4] AA vs. AG/GG respectively; p<0.01, Mann-Whitney U test; Figure 2A), and lower Charlson Index scores when compared to AA individuals (1.50 [interquartile range 1.00-2.75] vs. 1.00 [interquartile range 0.00-1.00] AA vs. AG/GG respectively; p<0.01, Mann-Whitney U test; Figure 2B).

In addition to comparing the number of co-morbidities and Charlson Index scores across genotypes, we also performed association analyses comparing the numbers of individuals across each genotype with a score of ≥3 on the Charlson Index (a score of ≥3 corresponded to being in the 90th percentile of the sample population). We found that in patients aged 65-79, those with the rs35652124 **A** allele showed increased odds of having a Charlson index score of ≥3 (OR 9.03 95%CI 1.16-70.2, p=0.0127). This corresponds to a positive predictive value of 0.59 conditional on carrying an **A** allele. This is compared to *a priori* probability of having ≥3 co-morbidities in the 65-79 age group (without knowledge of genotype) of 0.13. We found no relationship between the number of co-morbidities, or the Charlson Index score for either the rs6706649 or rs6721961 SNP.

2.6.3.2 Haplotype associations

When we examined the 65-79 age-group, we found that the CGG haplotype is associated with lower odds of multimorbidity (a score of ≥ 3 on the Charlson Index) compared to other haplotypes (OR 0.11, CI 0.01-0.87; p=0.01, Chi^2 test). No association was found in the ≥ 80-year-old age group.

2.6.4 Frailty

3.2.3.1 Single gene associations

We found that individuals carrying the **G** allele (AG/GG) at the rs35652124 locus had a significantly lower frailty score compared to homozygous **A** patients across the whole cohort (0.109 [interquartile range 0.073 to 0.161] vs. 0.091 [interquartile range 0.055 to 0.127] AA vs. AG/GG respectively; p<0.05, Mann-Whitney U test; Figure 2C). As with the comparisons of multi-morbidity, the effect of this genotype appears to be enhanced when comparing patients aged between 65-79 years old (0.107 [interquartile range 0.072 vs. 0.163] vs. 0.074 [interquartile range 0.036 to 0.108] AA vs. AG/GG; p<0.05, Mann-Whitney U test; Figure 2C). No significant relationship was found in the ≥ 80-year-old group between frailty score and rs35652124 genotype. No significant relationship was found between either the rs66706649 or rs6721961 genotype and frailty score.

3. Discussion

To our knowledge, this is the first human s yestudy to show an association between a single nucleotide polymorphism in the Nrf2 gene and characteristics of the ageing phenotype. It is also the first to show an age-related decrease in both Nrf2 gene, and protein expression in PBMCs of human volunteers. Nrf2 gene expression appears to decline late in life, whereas a decline in Nrf2 protein expression potentially precedes this. Further work is necessary to build on these observations to understand the mechanism and precise chronology behind these events. Together, these data may be useful in the design of tools that enable risk stratification of patients as they enter old age. This personalised approach is in line with a common agenda of using genetic and bio-psychosocial determinants to individualise treatment options to older people across Europe (27). However, due to the relatively small sample size of this study, and the limited diversity of its participants, further work in this area should be conducted to demonstrate the generalisability of these findings.

Our interest in investigating the role that this gene plays in multi-morbidity, frailty and susceptibility to ADRs stemmed from its important regulatory effect on a number of key anti-oxidant and xenobiotic response genes. There is a substantial amount of evidence that oxidative stress is a contributory factor in the ageing process of certain tissues / organ systems (5). Specifically, recent work shows that oxidative stress (induced by mutations to superoxide dismutase) is strongly correlated with frailty in mice (28). Furthermore, there is evidence in humans that frailty is associated with lower expression of stress response genes including Nrf2, and superoxide dismutase-2 (29). Interestingly, it is known that oxidative stress can trigger the release of a cocktail of pro-inflammatory cytokines and chemokines which, on a chronic time-scale, may lead to inflammageing – a condition that is associated with frailty and cardiovascular disease (30). A reduction in Nrf2 expression with age,

coupled with a polymorphism which decreases the gene's expression may therefore lead to signs and symptoms of ageing (e.g. frailty) and an increased risk of developing an ADR due to decreased downstream phase I and phase II enzyme expression.

The three SNPs that we investigated in this study are all found in the promoter region of the Nrf2 gene (16, 31). Both the rs6721961 (-617) and rs6706649 (-651) minor alleles (C→A, and G→A respectively) have consistently been shown *in vitro* to reduce the expression of Nrf2 (16, 32); we might therefore expect to see the strongest phenotypic associations with these particular variants. However, no significant associations were found between the minor alleles at these two loci and any of our outcome measures. One possible reason for this may be the low frequency of minor alleles observed in our cohort at these two loci, which limits the variation in our sample population. It is interesting to note that the minor allele frequency (MAF) of the rs6721961 was substantially lower than those observed in 3 recent studies (4.0% in our study vs. 31.2, 24.4, and 12.9%, by Song, Yu and Ran et al. (32-34)). The MAF at rs6706649 locus was however in line with those observed previously at approximately 5.4%. It should also be noted that the observed genotype frequencies at both the rs6706649 and the rs35652124 loci deviated significantly from Hardy-Weinberg equilibrium. There are a number of possible explanations for this, however, the most likely is that our sample population consisted of a number of smaller sub-populations. Indeed, it appears that the deviation of our observed genotype frequencies from the expected frequencies reaches significance at the rs6706649 locus because of one extra individual with the rare AA genotype. Similarly, at the rs35652124 locus, significant deviation from Hardy-Weinberg equilibrium appears to have been achieved by an extra 3 heterozygotes (who were expected to have been AA, and 2 x GG).

In contrast to the rs6271961 and rs6706649 SNPs, the rs35652124 locus did show an association with our various outcome measures. For example, we observed an association between the minor **G** allele at this particular locus, and increased ADR risk, which is consistent with our hypothesis. However, the presence of the same **G** allele in an individual's genotype was shown to provide some protection from frailty and multi-morbidity, which appears to be at odds with the previous finding.

In light of these data, we might therefore want to ask: what effect does the presence of the rs35652124 **G** allele have on Nrf2 gene expression and functionality. In fact, there is conflicting data regarding the influence of the rs35652124 SNP on Nrf2 expression. Yu et al. (2012) compared luciferase activity and Nrf2 mRNA expression at all 3 loci in the promotor region and showed a significant reduction in Nrf2 gene expression for the rs35652124 **G** allele. Marczak et al. (2012) showed that under conditions of oxidative stress, the **G** allele showed decreased Nrf2 expression compared to the **A** allele (17). Conversely, Song et al. (2016), demonstrated significantly increased luciferase activity with the minor **G** allele insert compared to the major **A** allele (33). Both Marzec et al. (2007) and Ran et al. (2016) showed a minimal effect of the -653 SNP on luciferase activity and mRNA expression respectively (17, 34) under basal conditions. It is unclear why such differences in the effect of the rs35652124 SNP exist – both Yu et al. and Song et al. used the same cell type (human embryonic kidney 293 T cells) for their expression assays. However, it may be the case that cell culture conditions, and exposure to stressors play an important role *in vitro*. Furthermore, *in vivo* (patients), it may be the case that the rs35652124 SNP behaves differently in different tissue types according to which transcription factors are expressed. A number of studies have reported that the Nrf2 promoter region is a target for several transcription factors (TFs), including Nrf2 itself. Some of these TFs may bind preferentially

to the **A** allele variant, and others (expressed in different tissues) to the **G** allele, resulting in different expression patterns throughout the human body.

If, however, we take the view that the rs35652124 **G** allele is associated with decreased Nrf2 expression, one other possible explanation for the increased risk of multi-morbidity and frailty associated with the **AA** genotype, is that elevated Nrf2 levels are pathogenic for some diseases. Indeed, Nrf2 knockout mice show reduced atherosclerotic plaque area compared to controls (35). Furthermore, in humans the rs35652124 **AA** genotype is associated with significantly increased risk of high blood pressure and cardiovascular mortality (36). It could therefore be the case in our study that the higher multi-morbidity and frailty scores are driven by cardiovascular pathology.

In conclusion, we have shown that rs35652124 genotype is a contributory factor to frailty, multi-morbidity and the risk of ADR in later life. Whilst the **G** variant is a marker of ADR risk, it is paradoxically associated with a lower risk of multi-morbidity. Although our sample population was not necessarily reflective of the entire PRIME study cohort, it does offer an important insight in to the role Nrf2 plays in the ageing phenotype in humans. However, further work needs to be conducted to determine the generalisability of these results. Additionally, experiments should be conducted to establish whether Nrf2 expression and downstream phase I and phase II enzyme expression are altered in carriers of the rs35652124 **A** allele, and also whether genotype can improve current models that aim to predict ADR, frailty and multi-morbidity risk.

Acknowledgements

Conflicts of interest

We can confirm that the authors have no actual or potential conflicts of interest.

4. References

1.   Budnitz DS, Pollock DA, Weidenbach KN, Mendelsohn AB, Schroeder TJ, Annest JL. National surveillance of emergency department visits for outpatient adverse drug events. JAMA. 2006;296(15):1858-66. DOI:10.1001/jama.296.15.1858

2.   Tangiisuran B, Scutt G, Stevenson J, Wright J, Onder G, Petrovic M, et al. Development and validation of a risk model for predicting adverse drug reactions in older people during hospital stay: Brighton Adverse Drug Reactions Risk (BADRI) model. PLoS One. 2014;9(10):e111254. DOI:10.1371/journal.pone.0111254

3.   Tangiisuran B, Davies JG, Wright JE, Rajkumar C. Adverse drug reactions in a population of hospitalized very elderly patients. Drugs Aging. 2012;29(8):669-79. DOI:10.2165/11632630-000000000-00000

4.   Parekh N, Ali K, Stevenson JM, Davies JG, Schiff R, Van der Cammen T, et al. Incidence and cost of medication harm in older adults following hospital discharge: a multicentre prospective study in the UK. Br J Clin Pharmacol. 2018;84(8):1789-97. DOI:10.1111/bcp.13613

5.   Yeoman M, Scutt G, Faragher R. Insights into CNS ageing from animal models of senescence. Nature reviews Neuroscience. 2012;13(6):435-45. DOI:10.1038/nrn3230

6.   Blesa J, Trigo-Damas I, Quiroga-Varela A, Jackson-Lewis VR. Oxidative stress and Parkinson's disease. Front Neuroanat. 2015;9:91. DOI:10.3389/fnana.2015.00091

7.   Revel F, Gilbert T, Roche S, Drai J, Blond E, Ecochard R, et al. Influence of oxidative stress biomarkers on cognitive decline. J Alzheimers Dis. 2015;45(2):553-60. DOI:10.3233/JAD-141797

8.   Suh JH, Shenvi SV, Dixon BM, Liu H, Jaiswal AK, Liu RM, et al. Decline in transcriptional activity of Nrf2 causes age-related loss of glutathione synthesis, which is

reversible with lipoic acid. Proc Natl Acad Sci U S A. 2004;101(10):3381-6.

DOI:10.1073/pnas.0400282101

9.   Shanmugam G, Narasimhan M, Conley RL, Sairam T, Kumar A, Mason RP, et al. Chronic Endurance Exercise Impairs Cardiac Structure and Function in Middle-Aged Mice with Impaired Nrf2 Signaling. Front Physiol. 2017;8:268. DOI:10.3389/fphys.2017.00268

10.   Ahn B, Pharaoh G, Premkumar P, Huseman K, Ranjit R, Kinter M, et al. Nrf2 deficiency exacerbates age-related contractile dysfunction and loss of skeletal muscle mass. Redox Biol. 2018;17:47-58. DOI:10.1016/j.redox.2018.04.004

11.   Volonte D, Liu Z, Musille PM, Stoppani E, Wakabayashi N, Di YP, et al. Inhibition of nuclear factor-erythroid 2-related factor (Nrf2) by caveolin-1 promotes stress-induced premature senescence. Mol Biol Cell. 2013;24(12):1852-62. DOI:10.1091/mbc.E12-09-0666

12.   Collins AR, Gupte AA, Ji R, Ramirez MR, Minze LJ, Liu JZ, et al. Myeloid deletion of nuclear factor erythroid 2-related factor 2 increases atherosclerosis and liver injury. Arterioscler Thromb Vasc Biol. 2012;32(12):2839-46. DOI:10.1161/ATVBAHA.112.300345

13.   Duarte TL, Caldas C, Santos AG, Silva-Gomes S, Santos-Gonçalves A, Martins MJ, et al. Genetic disruption of NRF2 promotes the development of necroinflammation and liver fibrosis in a mouse model of HFE-hereditary hemochromatosis. Redox Biol. 2017;11:157-69.

DOI:10.1016/j.redox.2016.11.013

14.   Shen G, Kong AN. Nrf2 plays an important role in coordinated regulation of Phase II drug metabolism enzymes and Phase III drug transporters. Biopharm Drug Dispos. 2009;30(7):345-55. DOI:10.1002/bdd.680

15.   Zhou L, Zhang H, Davies KJA, Forman HJ. Aging-related decline in the induction of Nrf2-regulated antioxidant genes in human bronchial epithelial cells. Redox Biol. 2018;14:35-40. DOI:10.1016/j.redox.2017.08.014

16. Marzec JM, Christie JD, Reddy SP, Jedlicka AE, Vuong H, Lanken PN, et al. Functional polymorphisms in the transcription factor NRF2 in humans increase the risk of acute lung injury. FASEB J. 2007;21(9):2237-46. DOI:10.1096/fj.06-7759com

17. Marczak ED, Marzec J, Zeldin DC, Kleeberger SR, Brown NJ, Pretorius M, et al. Polymorphisms in the transcription factor NRF2 and forearm vasodilator responses in humans. Pharmacogenet Genomics. 2012;22(8):620-8. DOI:10.1097/FPC.0b013e32835516e5

18. von Otter M, Bergström P, Quattrone A, De Marco EV, Annesi G, Söderkvist P, et al. Genetic associations of Nrf2-encoding NFE2L2 variants with Parkinson's disease - a multicenter study. BMC Med Genet. 2014;15:131. DOI:10.1186/s12881-014-0131-4

19. Matz MV, Wright RM, Scott JG. No control genes required: Bayesian analysis of qRT-PCR data. PLoS One. 2013;8(8):e71448. DOI:10.1371/journal.pone.0071448

20. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2014. p. http://www.R-project.org/.

21. Stevenson J, Parekh N, Ali K, Timeyin J, Bremner S, Van Der Cammen T, et al. Protocol for a Prospective (P) study to develop a model to stratify the risk (RI) of medication (M) related harm in hospitalized elderly (E) patients in the UK (The PRIME study). BMC Geriatr. 2016;16:22. DOI:10.1186/s12877-016-0191-8

22. Naranjo CA, Busto U, Sellers EM, Sandor P, Ruiz I, Roberts EA, et al. A method for estimating the probability of adverse drug reactions. Clin Pharmacol Ther. 1981;30(2):239-45.

23. Nikesh Parekh KA, Jennifer Stevenson, Graham Davies, Rebekah Schiff, Jatinder Harchowal, Tischa Van der Cammen, Chakravarthi Rajkumar. Frailty predicts medication-related harm requiring healthcare: A UK multicentre prospective cohort study. European Geriatric Medicine Journal. 2018;9.

24. Searle SD, Mitnitski A, Gahbauer EA, Gill TM, Rockwood K. A standard procedure for creating a frailty index. BMC Geriatr. 2008;8:24. DOI:10.1186/1471-2318-8-24

25. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis. 1987;40(5):373-83.

26. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics. 2005;21(2):263-5. DOI:10.1093/bioinformatics/bth457

27. Nimmesgern E, Norstedt I, Draghia-Akli R. Enabling personalized medicine in Europe by the European Commission's funding activities. Per Med. 2017;14(4):355-65. DOI:10.2217/pme-2017-0003

28. Viña J, Borras C, Gomez-Cabrera MC. A free radical theory of frailty. Free Radic Biol Med. 2018;124:358-63. DOI:10.1016/j.freeradbiomed.2018.06.028

29. El Assar M, Angulo J, Carnicero JA, Walter S, García-García FJ, López-Hernández E, et al. Frailty Is Associated With Lower Expression of Genes Involved in Cellular Response to Stress: Results From the Toledo Study for Healthy Aging. J Am Med Dir Assoc. 2017;18(8):734.e1-.e7. DOI:10.1016/j.jamda.2017.04.019

30. Ferrucci L, Fabbri E. Inflammageing: chronic inflammation in ageing, cardiovascular disease, and frailty. Nat Rev Cardiol. 2018;15(9):505-22. DOI:10.1038/s41569-018-0064-2

31. Yamamoto T, Yoh K, Kobayashi A, Ishii Y, Kure S, Koyama A, et al. Identification of polymorphisms in the promoter region of the human NRF2 gene. Biochem Biophys Res Commun. 2004;321(1):72-9. DOI:10.1016/j.bbrc.2004.06.112

32. Yu B, Lin H, Yang L, Chen K, Luo H, Liu J, et al. Genetic variation in the Nrf2 promoter associates with defective spermatogenesis in humans. J Mol Med (Berl). 2012;90(11):1333-42. DOI:10.1007/s00109-012-0914-z

33.   Song P, Li K, Liu L, Wang X, Jian Z, Zhang W, et al. Genetic polymorphism of the Nrf2 promoter region is associated with vitiligo risk in Han Chinese populations. J Cell Mol Med. 2016;20(10):1840-50. DOI:10.1111/jcmm.12874

34.   Ran C, Wirdefeldt K, Brodin L, Ramezani M, Westerlund M, Xiang F, et al. Genetic Variations and mRNA Expression of NRF2 in Parkinson's Disease. Parkinsons Dis. 2017;2017:4020198. DOI:10.1155/2017/4020198

35.   Sussan TE, Jun J, Thimmulappa R, Bedja D, Antero M, Gabrielson KL, et al. Disruption of Nrf2, a key inducer of antioxidant defenses, attenuates ApoE-mediated atherosclerosis in mice. PLoS One. 2008;3(11):e3791. DOI:10.1371/journal.pone.0003791

36.   Shimoyama Y, Mitsuda Y, Tsuruta Y, Hamajima N, Niwa T. Polymorphism of Nrf2, an antioxidative gene, is associated with blood pressure and cardiovascular mortality in hemodialysis patients. Int J Med Sci. 2014;11(7):726-31. DOI:10.7150/ijms.8590

Gene	Direction	Reaction type	Sequence
Nrf2	Forward Reverse	q-rtPCR	5′-GAGAGCCCAGTCTTCATTGC-3′ 5′-TGCTCAATGTCCTGTTGCAT-3′
GAPDH*	Forward Reverse	q-rtPCR	5′-ACCCACTCCTCCACCTTTGAC-3′ 5′-TCCACCACCCTGTTGCTGTAG-3′
18s ribosomal subunit*	Forward Reverse	q-rtPCR	5′-GTAACCCGTTGAACCCCATT-3′ 5′-CCATCCAATCGGTAGTAGCG-3′
Nrf2	Forward Reverse	PCR / Sanger sequencing	5′ CTTTTATCTCACTTTACCGCCC 3′

Table 1. Primer sequences for q-rtPCR, and PCR / Sanger sequencing of Nrf2 gene.

*Housekeeping genes

Age category (years)	Proportion female (n)	Chronic medical condition (n)	Consume regular medication (n)	Regular smoker (n)	Median units of EtOH consumed weekly	Median age (range)
18-29	15/27	4/27	6/27	3/27	3 (0-16)	22(18-28)
30-39	3/8	4/8	5/8	2/8	0 (0-15)	33.5 (30-38)
40-49	6/7	4/7	2/7	1/7	4 (0-10)	42 (41-47)
50-59	5/9	6/9	5/9	1/9	6 (0-10)	53 (50-57)
60-69	1/2	2/2	2/2	0/2	8 (6-10)	62 (60-64)
≥70	0/2	2/2	2/2	0/2	8.75 (7-10.5)	73.5 (72-75)
Sig.	p=0.2955	p=0.0045**	p=0.0292*	p=0.8802	p=0.2024	

Table 2. Demographics of volunteers for according to age-category. Differences between age-groups were tested using Kruskal-Wallis test for continuous, non-parametric data, or the Chi^2 test for categorical data.

	Median (range)
Age	75.5 (65-92)
Female (%)	62 (55.4%)
Number of co-morbidities	3 (0-9)
Charlson index	1 (0-5)
Number of regular medicines	7.5 (1-19)
Frailty index	0.091 (0-0.29)

Table 3. Demographics of patients enrolled from sub-population of the PRIME study.

		rs6721961 (-617)		rs35652124 (-653)		rs6706649 (-651)
Frequency of genotypes and alleles (%)	AA	0 (0%)	AA	45 (40%)	AA	2 (2%)
	AC	9 (8%)	AG	61 (55%)	AG	8 (7%)
	CC	103 (92%)	GG	6 (5%)	GG	102 (91%)
	A	9 (4%)	A	151 (67%)	A	12 (5%)
	C	215 (96%)	G	73 (33%)	G	212 (95%)
Hardy-Weinberg		p>0.05		p<0.05*		p<0.05*

Table 4. Genotype and allelic frequencies for the three single nucleotide polymorphisms studied. Deviation from Hardy-Weinberg equilibrium was considered if the observed frequencies of genotypes differed significantly ($p < 0.05$ by Fisher's exact test) from the expected frequencies, based on the frequencies of alleles observed in the sample population.

Naranjo algorithm: ADR=definite/probable/possible				
	ADR	**No ADR**	**Frequency ADR**	**Sig. (Chi^2)**
rs35652124 (-653)				
AA	14	31	0.31	p=0.031*
AG	19	42	0.31	
GG	5	1	0.83	
rs6721961 (-617)				
CC	35	68	0.34	p=0.969
CA	3	6	0.33	
rs6706649 (-651)				
GG	35	67	0.34	p=0.770
GA	2	6	0.25	
AA	1	1	0.50	
Naranjo algorithm: ADR=definite/probable				
	ADR	**No ADR**	**Frequency ADR**	**Sig. (Chi^2)**
rs35652124 (-653)				
AA	7	38	0.15	p=0.137
AG	12	49	0.20	
GG	3	3	0.50	
rs6721961 (-617)				
CC	21	103	0.20	p=0.502
CA	1	8	0.11	
rs6706649 (-651)				
GG	19	83	0.19	p=0.502
GA	2	6	0.25	
AA	1	1	0.50	

Table 5. ADR frequency according to genotype. An ADR has been classified if it scored as definite/probable/possible (top half of table), or definite/probable (bottom half of table) according to the Naranjo algorithm.

Figure legends

Figure 1. Changes to Nrf2 gene expression and protein levels with age in healthy volunteers. **A** Log Nrf2 mRNA abundance vs. age (n=55) and **B** Log-fold Nrf2 mRNA change compared to 18-29 year-olds (n=55). **B** Changes to Nrf2 protein levels with age in heathy volunteers (n=30). $*p<0.05$ (Student's t-test), $****P<0.0001$ (One-way ANOVA).

Figure 2. Association of the rs35652124 SNP with Co-morbidities (**A**), Charlson Index (**B**), and Frailty (**C**) across the whole PRIME sub-population, and those aged between 65-79 years old. $*p<0.05$, $**p<0.01$, Mann-Whitney U-test; data are presented as median, 25-75 percentiles and range.