Heterozygote advantage maintains Rhesus factor blood group polymorphism: Ecological regression and case-control studies

Jaroslav Flegr¹*, Rudolf Hoffmann², Mike Dammann ¹

¹ Division of Biology, Faculty of Science, Charles University in Prague, Prague, Czech Republic
² Department of Hematology, Na Homolce Hospital, Prague, Czech Republic

* Corresponding author

JF E-mail: flegr@cesnet.cz
RH E-mail: rhoffmann@vaclavka-medical.cz
MD E-mail: rhesusnegatives@gmail.com
Abstract

Background: The existence of the Rhesus factor polymorphism has been an evolutionary enigma since the discovery of this factor in 1939. Carriers of the less common allele should be eliminated by selection pressure against Rhesus positive children born to Rhesus negative mothers unless some form of frequency-dependent selection, e.g. heterozygote advantage, stabilizes the equilibrium frequencies of two major alleles in a population. Here we report the first direct evidence of this heterozygote advantage.

Methods: Ecologic regression study performed on a set 65 countries for which RhD phenotypes frequencies data and disease burdens data were available and cross sectional internet study performed on population of 2,779 subjects.

Results: The ecologic regression study has shown that the burden associated with many diseases correlated with the frequencies of particular Rhesus genotypes in a country. In agreement with predictions of the heterozygote advantage hypothesis, the direction of the relation was nearly always opposite for the frequency of Rhesus negative homozygotes and that of Rhesus positive heterozygotes. The case-control has shown that RhD negative subjects have reported more frequent allergic, digestive, heart, hematological, immunity, mental health, and neurological problems. Incidence of various diseases also differed from RhD negative to RhD positive subjects, mostly being higher in the former.

Conclusions: On the population level, RhD-negativity-associated burden could be compensated by the heterozygote advantage. However, from a clinical medicine point of view, RhD negativity could represent a serious problem.

Key words: frequency dependent selection; mortality; morbidity; RhD; blood group; polymorphism; evolution
Background

Polymorphism in the Rhesus factor, i.e. the existence of a large deletion in the RHD gene [1] in a substantial fraction of the human population, has been an evolutionary enigma since the discovery of this factor in the Thirties of the past century [2-5]. Before the introduction of prophylactic treatment in 1968, the carriers of the rarer variant of the gene, namely Rhesus negative women in mostly Rhesus positive populations or Rhesus positive men in mostly Rhesus negative populations, had lower fitness as RhD-positive children born to pre-immunized RhD-negative mothers were at a higher risk of fetal and newborn death or health impairment from the haemolytic disease. Therefore, mutants or migrants with the rarer variant of the RHD gene could not invade the population and any already existing RhD polymorphism should be unstable. It was suggested that the polymorphism can be stabilized when the disadvantage of carriers of the locally less abundant allele is counterbalanced by, for example, higher viability of their heterozygote children or by another form of frequency-dependent selection [6]. In the past five years, several studies have demonstrated that Rhesus positive and Rhesus negative subjects differ in resistance to the adverse effects of latent toxoplasmosis [6-10] and possibly also other health-related factors, such as aging, fatigue and smoking [11]. More specifically, latent toxoplasmosis is associated with excessive weight gain [4 kg vs. 2.2 kg in the 16th week of pregnancy) in Rhesus negative women [12]. A large case-control study performed on 3,800 soldiers has shown that smoking has about 3 times stronger effects on the reported numbers of bacterial and viral diseases in Rhesus negative than in Rhesus positive subjects [11]. A study performed on 250 blood donors has further shown that the resistance to effects of toxoplasmosis in Rhesus positive heterozygotes is higher than in Rhesus positive homozygotes and substantially higher than in Rhesus negative homozogotes [6]. This is the first direct evidence for the role of selection in favour of heterozygotes in stabilization of the RHD gene polymorphism in populations. This mechanism reminds of widely known situations with polymorphism in genes associated with sickle cell anaemia in geographic regions with endemic malaria [13].

The frequencies of Rhesus negative subjects (and therefore also Rhesus positive heterozygotes) as well as the incidences of particular diseases and disorders vary between countries. If the protective effect of Rhesus positivity or Rhesus heterozygosity is strong enough, then the relationship between the frequencies of Rhesus negative homozygotes (and Rhesus positive
heterozygotes) should correlate with the incidences of specific diseases when important confounding variables are controlled. Here, we have studied the correlation of disease burden estimates compiled by the WHO with the frequencies of Rhesus negative homozygotes and Rhesus positive heterozygotes, using partial correlation with the gross domestic product (GDP), latitude, humidity, medical expenses per capita, and percentages of smokers within the population in a set of 65 countries for which the data on the frequencies of Rhesus-negative individuals are available. To independently confirm the prediction of the heterozygote advantage hypothesis, namely the prediction of a worse health status of Rhesus-negative homozygotes, we have also run a large questionnaire study in a population of healthy Czech and Slovak volunteers. Using an electronic questionnaire distributed with a Facebook-based snowball method, we have screened a population of 2,779 subjects for indices of various health problems as well as for incidences of 225 diseases and disorders.

Methods

Sources of data

The data on the disease burden were obtained from the table “Mortality and Burden of Diseases Estimates for WHO Member States in 2004” published by WHO [18, 19] and are available at: www.who.int/evidence/bod. Because of the expected effects of the frequencies of the Rhesus factor genotypes on age structures of the populations, the age-non-standardized disease burden data were used. The frequencies of Rhesus negative homozygotes in particular countries [21] were taken from the internet compilation by RhesusNegative.net available at http://www.Rhesusnegative.net/themission/bloodtypefrequencies/ from December, 22nd of 2013 and from the monograph of Mourant [18, 19]. The frequencies of Rhesus positive heterozygotes were calculated from the frequencies of Rhesus negative homozygotes using the Hardy-Weinberg equation. If the conclusions of the present study are correct, then these theoretical frequencies are influenced by a selection in middle-age and especially in high-age strata. However, they are being re-established in every generation. Geographical latitude and annual mean of relative humidity for particular countries were derived using the data available at http://data.worldbank.org/indicator/NY.GDP.PCAP.CD (accessed December, 10th of 2013) and

Subjects

The medical questionnaire prepared by two medical doctors was distributed as a Qualtrics survey. The subjects were invited to participate in the study using a Facebook-based snowball method, namely by posting an invitation to participate in “an experiment searching for associations between the blood group of a subject and his/her personality, performance, morphology and health” on the wall of the Facebook page “Guinea pigs” for Czech and Slovak nationals willing to take part in diverse evolutionary psychological experiments (www.facebook.com/pokusnikralici). The participants were informed about the aims of the study on the first page of the electronic questionnaire. They were also being provided with the following information: “The questionnaire is anonymous and obtained data will be used exclusively for scientific purposes. Your cooperation in the project is voluntary and you can terminate it at any time by closing this web page. If you can check up on what your blood group is, please do it now. We need the data from subjects with all blood groups, not only from Rh negative subjects. Therefore, please share the link to this questionnaire with your friends, for example on Facebook. Press the “continue” button if you agree to your anonymous participation in the study”. The share button was pressed by 400 participants, which resulted in obtaining data from 3,854 responders in total between 28.4. -27.11. 2014. The study, including the method of obtaining an electronic consent with a participation in the study, was approved by the IRB of the Faculty of Science, Charles University (Komise pro práci s lidmi a lidským materiálem Přírodovědecké Fakulty Univerzity Karlovy) - No. 2014/21.

Statistical methods

5
Ecological regression study: Factor analyses and general linear model (GLM) analysis of the obtained factors were performed with Statistica v. 8.0 and all other tests, including the tests of statistical test assumptions, with IBM SPSS v. 21. Factor analysis of the specific disease burden data (principal components method, raw Varimex rotation) was used for data reduction. When the factors were extracted from the mortality rate variables, the diseases with not enough data were excluded and the missing data were substituted with means. The Disability Adjusted Life Year (DALY) data for nearly all diseases except trypanosomiasis, Chagas disease, and onchocerciasis were available for all countries in our set. All factors with eigenvalues>1.0 were extracted. Type III sum of squares and models including the intercept were used in the GLM (multivariate and multiple-multivariate) analyses; for details, see [22].

Case-control study: Before statistical analysis, suspicious data (too high or too short body height, too low or too high body mass or age, too short duration of the test etc.) were filtered out. In the test, we have also measured simple reaction times, operational, short-term and long-term memory, psychomotor performance, intelligence and personality profiles. However, here we have analysed only data concerning the health status.

SPSS v. 21. was used for all statistical tests. Ordinal and binary data were analysed by partial Kendall’s correlation test [23, 24]. This test could measure strength and significance of association between binary, ordinal and continuous data regardless of their distributions. This technique enables to control for one confounding variable, for example age of a responder. The Excel sheet for computing partial Kendall’s Tau and significance between variables A and B after the variable C is controlled based on Kendall Tau’s AB, AC and BC. It is available here: http://web.natur.cuni.cz/flegr/programy.php. Certain diseases have very different incidence in men and women. Also, some biological factors, including RhD phenotype, could have different impacts on men and women. Therefore, we have performed all analyses also separately for the male and female responders.

Results

A) Ecological regression study
Association between disease burden and frequencies of RhD genotypes – a simple model

Mortality and morbidity data for more than one hundred diseases are available in the WHO database. A search for associations between Rhesus genotype and particular diseases would necessarily result in many spurious associations. Therefore, factor analysis was used at first for data reduction. This analysis performed on 125 diseases and burden of disease categories yielded 23 factors with eigenvalues>1.0, together explaining 85 % of variability in Disability Adjusted Life Year (DALY) among 192 WHO member countries. The DALY has been defined [14] as “a health gap measure that extends the concept of potential years of life lost due to premature death and also to include equivalent years of ‘healthy’ life lost by virtue of being in a state of poor health or disability”. Then multiple multivariate analysis was performed with these factors as dependent variables and frequencies of Rhesus negative homozygotes, Rhesus positive heterozygotes and also five potential confounding variables, GDP, latitude (distance from the equator), humidity, medical care expenses per capita and frequencies of smokers in the population for 65 countries with the RhD genotype frequency data as independent variables. Effects of the frequency of RhD heterozygotes, smokers, latitude, and humidity were significant (Table 1). The model explains significant parts of variability of factors 1-3, 6-8, 11, and 22 (adjust $R^2 = 0.372$, 0.771, 0.497, 0.461, 0.197, 0.237, 0.224, and 0.296, respectively). Post hoc simple multivariate analyses have shown that the frequency of Rhesus negative homozygotes correlated with five of 23 factors, including the second most important one, that together explained 17.7 % of between-countries variability in DALY. Similarly, the frequency of RhD positive heterozygotes significantly correlated with six factors that together explained 13.9 % of between-countries variability in DALY. Expected number of false significant results for 46 statistical tests was not eleven but 2.30. In accord with the heterozygote advantage hypothesis of maintaining Rhesus factor polymorphism, the regression coefficients (B-values) for Rhesus negative homozygotes and Rhesus positive heterozygotes went into opposite directions whenever any of these correlations was significant.

Association between disease burden and frequencies of RhD genotypes – a complex model

Previously published data have shown differences in the reaction to smoking between RhD negative and RhD positive subjects [11]. Smoking has probably different effects on health for men and
women. Therefore, we have also analysed more complex model that included the frequency of male
and female smokers and four interaction terms, namely male smokers-RhD negatives, male
smokers-RhD heterozygotes, female smokers-RhD negatives, and female smokers-RhD
heterozygotes. The model explained significant and mostly large parts of variability of factors 1-4,
6-8, 11, 12, and 22 (adjust $R^2 = 0.472, 0.820, 0.629, 0.180, 0.488, 0.352, 0.264, 0.464, 0.182,$ and
0.300, respectively). This time, not only the main factors, i.e. the frequency of RhD negatives and
that of RhD heterozygotes, but also the interactions of RhD genotypes’ frequencies with the
frequencies of either male or female smokers correlated with many factors. The Supplementary
Table 1 shows 14 significant effects and interactions (6.9 false positive results were expected).

**Correlation between particular disease burdens and frequencies of RhD genotypes – post hoc
tests**

Both non-parametric (Spearmen) and parametric (Pearson) univariate correlations have shown
many strong associations between disease burdens measured with DALY or with incidences of
deaths per population of 100,000 and the frequencies of Rhesus negative subjects or Rhesus positive
heterozygotes (Fig. 1). However, specific disease burdens also strongly correlated with some of the
confounding variables, most often with latitude, smoking and humidity. Therefore, the GLM
analysis was used to search for the association between Rhesus genotypes and disease burden using
the frequency of Rhesus negative homozygotes and Rhesus positive heterozygotes and GDP,
latitude, humidity, medical care expenses per capita and frequencies of smokers within the
populations of 65 countries with the RhD genotype frequency data. The results presented in the
Supplementary Table 2 show that the frequencies of Rhesus negative subjects correlated with
DALY for 21 of 121 diseases and disease categories with available DALY data (twelve positively
and nine negatively). The frequency of Rhesus positive heterozygotes correlated with DALY for 25
of 121 diseases and disease categories (eleven positively and fourteen negatively). Similarly, the
frequencies of Rhesus negative subjects correlated with the mortality rates (incidences of deaths per
population of 100,000) for 10 of 97 diseases and disease categories with the mortality rate data
available (all positively). The frequencies of Rhesus positive heterozygotes correlated with the
mortality rates for 8 of 97 diseases and disease categories (two positively and six negatively).
Expected number of false significant results for 436 statistical tests was not 65 but 22.
Figure 1. Correlation of the frequencies of Rhesus negative subjects in 65 countries with the mortality rates for nine diseases or disease categories.

The x and y axes show the frequency of Rhesus negative homozygotes in the population of a country and mortality (numbers of deaths per population of 100,000), respectively. The figures represent the results of the Pearson correlation analysis, namely the level of significance (p) and the coefficient of determination, i.e. the fraction of a country’s variability of the specific mortality rate that can be explained by the differences within the frequencies of Rhesus negative subjects.

B) Case-control study

Descriptive statistics of data

Among 3,856 Czech and Slovak participants, 2,779 subjects (766 RhD positive men, 281 RhD negative men, 1,184 RhD positive women and 548 RhD negative women) provided information about their RhD phenotype. RhD negative subjects, especially women, have had a higher motivation to care about and to remember their RhD phenotype. Therefore, the frequency of RhD negative subjects (29.8 %) differed from the 16 % general frequencies within the Czech and Slovak populations and also between men (26.8 %) and women (31.6 %). The mean age of RhD positive men (37.5, SD 13.4) was approximately the same as that of RhD negative men (37.5, S.D. 12.8), t(1045) = 0.07, p = 0.942. RhD positive women were younger (33.5, S.D. 11.6) than RhD negative women (35.4, S.D. 12.7), t(1730) = 3.04, p = 0.002. The numbers of men and women in the particular 10-age strata were comparable, except for the 20-30 age stratum, which consisted of 537 men and 880 women (Supplementary Fig. 1).

Correlation of RhD phenotype with self-reported health problems (ordinal variables)

Twenty-one dependent variables (mostly ratings of particular health problems on a scale from 1-6, 1: “no problems at all”, 6: “frequent or serious”) were ordinal and had a highly skewed distribution. Therefore, the nonparametric partial Kendall’s correlation test (which enables to control one confounding variable, here the age) was used to search for the association between the RhD phenotype and intensity of fifteen categories of health problems (allergic, cancer, digestive, fertility, genitourinary, heart, hematological, immunity, mental health, metabolic including...
endocrine, musculoskeletal, neurological, respiratory organs, sense organs, and sexual life
problems) and also another six health-related variables, namely the numbers of drugs prescribed by
doctors the subject currently takes per day, numbers of “different herbs, food supplements,
multivitamins, superfoods etc.” the subject currently takes per day, how many times the subject has
maximally used antibiotics in the same calendar year in the past years, how many times the subject
had to seek acute medical care for a serious illness (not injury) that lasted more than 3 days in the
past 5 years, how many specialised medical doctors the subject had to regularly attend (not for
prevention) at least once in the past two years, how often the subject felt tired (not after exertion,
e.g. sports) and how often the subject has experienced a headache. The results have shown that the
RhD negative subjects had more serious health problems in 9 of 22 analysed variables than the RhD
positive subjects (Tab. 2).

Association between RhD phenotype and incidence of particular diseases (binary variables)

After rating each particular category of health problems, the subjects were asked to check which
specific disorders they suffered from on the lists of 225 disorders. They were also asked to check
which specialised medical doctors they had to visit regularly (not for prevention) at least once in the
past two years from the list of 10 types of specialists. The associations were analysed with the
partial Kendall’s Tau correlation test with age being a covariate. One hundred fifty one (151) of 225
diseases/disorders were reported by at least 10 subjects. In this subset, there were 35 expressed
significant associations with RhD negativity (26 positive and 9 negative) in all subjects. In male
subjects, the number of significant association was 27 (16 positive and 11 negative) and in female
subjects, the number of significant associations was 32 (21 positive and 11 negative). Expected
number of false significant results for 546 statistical tests was not 94 but 22. For example, RhD
negative men more often reported certain mental health disorders including unipolar depressive
disorders, antisocial personality disorders and attention deficits, ticks, thyroiditis, immunity
disorders, allergies, especially skin allergies, infectious diseases and liver disease and they less
often reported gall bladder attacks, warts and some types of cancers, especially prostate
hypertrophy. RhD negative women reported more frequently psoriasis, constipation and diarrheas,
ischemic diseases, type 2 diabetes, thyroiditis, some types of cancers, tonsil stones, too high sex
appetency, both precocious and delayed puberty, urinary tract infections and they less often reported
accented blood clotting, hearing loss, sense of smell problems, muscle twitch /fasciculation and
warts. RhD negative subjects had to visit more often otolaryngology (p = 0.019), psychiatry (p =
0.005), gynecology (p=0.001), and dermatology (p = 0.001) specialists (theoretical number of false
positive results was 0.5). Supplementary Table 3 shows the associations between RhD negativity
and disease incidences and Table 3 shows the associations between RhD negativity and visiting
specialised doctors.

Discussion

The results of the ecological regression analysis have shown that both, the frequencies of Rhesus
negative homozygotes and those of the Rhesus positive heterozygotes correlated (mostly in the
opposite direction) with specific disease burdens. The general pattern was that the countries with a
high frequency of Rhesus negative homozygotes had lower congenital-anomalies-associated burden
and neuropsychiatric condition-associated burden (except Alzheimer´s and Parkinson´s Disease
burden) and higher cardiovascular and especially malignant neoplasm-associated burden.
Noticeably, the only form of cancer expressing a clear opposite trend was cervix uteri cancer, i.e.
cancer of viral origin. In the case-control study, the RhD negative subjects expressed many indices
of a worse health status. Men, women or both sexes reported more frequent allergic, digestive,
heart, hematological, immunity, mental health and neurological problems. They also reported the
usage of more drugs prescribed by doctors per day, attended more specialised doctors, namely,
dermatologists, gynecologists, internal medicine doctors, neurologists, and psychiatrists (men) in
the past two years, a higher frequency of headaches and being tired more often than RhD positive
subjects. Incidence of various diseases and disorders also differed between RhD negative and RhD
positive subjects, mostly being higher in the former.

The results of the ecological regression and case-control studies are difficult to compare. For
example, in an ecological regression study, the effect of RhD negativity on disease burden is
seemingly decreased by the opposite effect of RhD heterozygosity on the same disease burden
while in the case-control study, the effect of RhD negativity on the health status is seemingly
increased by an opposite effect of RhD heterozygosity on RhD positive controls. Moreover, a
positive correlation between the focal disease burden and the frequency of a particular RhD
genotype observed in an ecological regression study could be either due to an increased sensitivity
of Rhesus negative individuals towards the focal disease or a relatively higher resistance or
tolerance of these subjects towards other diseases. For example, the protective effect of RhD
negativity against many neuropsychiatric disorders observed in the ecological study could be caused
by the fact that RhD negative subjects usually die at an earlier age due to their higher susceptibility
to cardiovascular diseases.

WHO data used in the ecological regression study refer to all age groups of a population
while the study population of present case control study consisted of mainly younger and early-
middle age subjects. It is highly probable that the effects of the RhD phenotype will be different
depending on the age of the population under study.

The frequency of RhD negative homozygotes is positively but non-linearly related to that of
RhD positive heterozygotes. Therefore, the same case control study probably gives different results
for different countries, dependent on the frequencies of the allele for RhD negativity and resulting
differences in the homozygotes vs. heterozygotes ratio. For example, in countries with high
frequencies of RhD negative subjects, larger fractions of RhD positive subjects are heterozygotes
and therefore case-control studies will probably show larger differences in health status between
RhD negative and RhD positive subjects in comparison with countries with low frequencies of the
RhD- allele.

Both studies suggested that RhD negative subjects have increased risks of certain heart
diseases, respiratory diseases and some immunity and autoimmunity related diseases, for example
rheumatoid arthritis. The case-control study suggests that RhD negative subjects could have
problems with autoimmunity, could be more resistant to infections of viral origin and less resistant
to infections of bacterial origin.

The mechanism of the effect of the RhD phenotype on the human health status is not clear.
RhD protein together with strongly homologous RhCE protein and with also homologous RhAG
glycoprotein is a component of a membrane complex of which the function is not quite clear. It is
most probably involved in NH$_3$ transport and possibly also in CO$_2$ transport [15, 16]. This complex
is associated with spectrin-based cytoskeleton and therefore plays an important role in maintaining
the typical shape (biconcave discoid) of human erythrocytes [17]. The biological functions of the
RhD protein containing complexes are unknown, however, they might be involved in NH₃/NH₄⁺
detoxification of organs. Ammonia, the product of protein catabolism is extremely toxic, especially
for brain cells and must be quickly removed from the sensitive organs. It was observed that the
concentration of ammonium is three times higher in red cells than in plasma [17] and it was
suggested that the RhD containing complex plays a key role in its capturing and its transport to the
kidneys and the liver [17]. It was also suggested that the complex might participate in the
intracellular pH regulation [17] and consequently also in the regulation of local oxygen tension. It
was suggested that RhD-negativity-associated anoxia in certain parts of the nervous system could be
responsible for physiological (and also behavioral) effects of the RhD phenotype [17]. The variation
of the oxygen tension in various organs and tissues could, of course, influence also other biological
functions, including the functions of the immune system. This could explain why RhD negativity
seems to be associated with neurological, mental health and immunological disorders. The probable
roles of the RhD-containing complex in keeping the normal morphology and adhesiveness of red
cells (for review see [17]) could be responsible for the observed associations of RhD negativity with
some haematological and inflammation-related diseases, including arthritis.

The “European” RhD- variant of the RHD gene carries a deletion covering the whole
protein-coding part of the gene [1]. Therefore, no product of this allele is synthetized in the cells of
RhD negative homozygotes and is most probably substituted in the corresponding molecular
complex by the related protein RhCE. Therefore, erythrocytes of RhD- and RhD+ homozygotes
differ in molecular complexes on their cell membranes and probably also in their biological
activities. An important difference was also observed between erythrocytes of RhD positive
homozygotes and heterozygotes. About 33,560 and 17,720 D antigen sites were detected on the
surfaces of an erythrocyte in RhD homozygotes and heterozygotes, respectively [17]. This suggests
that the susceptibility of RhD positive homozygotes and RhD positive heterozygotes (and even
more so RhD negative homozygotes) to various aberrant conditions, including various diseases,
could dramatically differ. Due to the general trade-off principle, heterozygotes could be more
resistant to one disease and more prone to another disease and the opposite could be true for
homozygotes. Such trade-offs could explain the heterozygote advantage hypothesis and all other
observed phenomena.
Limitations of present study: The interpretations of ecological regression studies are sometimes complicated, especially if aggregated data are used for the estimation of the strength and direction of the influence of particular factors within a population [18, 19]. Therefore, one must be very careful with the interpretation of the observed associations between particular disease burden and the frequency of Rhesus negative subjects within a population. For example, a positive correlation could be due to an increased sensitivity of Rhesus negative individuals to the focal disease or just to higher resistance or tolerance of these subjects to other diseases. Therefore, future studies of mechanisms of effects of the Rhesus phenotype on risks of particular diseases must be grounded within individual based (case-control and cohort) studies. However, the objective of the present ecological study was to test the heterozygote advantage hypothesis of maintaining the genetic polymorphism in the RHD gene, resulting in polymorphism of the Rhesus factor phenotypes. Based on this hypothesis, we suggested that the frequencies of Rhesus negative homozygotes as well as Rhesus positive heterozygotes should correlate with certain disease burden and the direction of such correlation will be the opposite for the two phenotypes. The results of this ecological regression study have confirmed these two predictions.

In the case-control study, we have obtained data from a large number of subjects. However, most of them were relatively young people (mean age was 37.7). Most of the diseases and disorders with a largest public health impact (but possibly not the largest economic impact) start at a higher age in developed European countries such as the Czech and Slovak Republics. This can largely distort the whole picture of the RhD negativity impacts on public health. Future studies (which could by easily done in countries with available national databases of medical records of all citizens) should aim to middle age subjects and to seniors. Our study compared the health status of RhD negative subjects (16 % in general population of Czech and Slovak Republics) with RhD positive subjects, i.e. with the health status of mixed population of RhD positive homozygotes (36 % of the general populations within the Czech and Slovak Republics) and heterozygotes (48 % the general populations in the Czech and Slovak Republics). The results of the published case-control study on the effects of the RhD genotype on psychomotor performance [18, 19], as well as the heterozygote advantage hypothesis, however, suggest that the health status of RhD positive homozygotes and heterozygotes differs. In further studies concentrated on particular disorders, smaller populations of subjects should be RhD genotyped using molecular biology techniques and
then the health status of all three RhD genotypes have to be compared. In the present study, the health status data were collected using a questionnaire. This approach enabled a check-up of the results of the first ecological regression study very quickly and also the study of the effects of the RhD phenotypes on not so common disorders in a large population sample. Of course, more precise and more detailed data could be obtained from medical records. Primarily, we have run the study to confirm or disprove the alarming results of a previous ecological regression study. However, due to large differences between character of data collected with the ecological regression study and case-control study, the present study had a more or less explorative character. Therefore, we have reported the results of post hoc tests without formal correction to multiple tests. It should be noted, however, that, for example, we have obtained 41% positive results for the ordinal health status variables and 20% positive results for binary health status variables. Theoretically, only 5% of false positive results should be expected in multiple tests.

Conclusions

Some of the associations observed in both studies were relatively strong. For example, the slopes (B) values of 1,812 and -463 for cardiovascular diseases mean that a 1% increase within the frequencies of Rhesus negative homozygotes and Rhesus positive heterozygotes would result in 1,812 more and 463 less cardiovascular disease associated deaths per 100,000 inhabitants, respectively. A positive nonlinear correlation between the frequencies of Rhesus negative homozygotes and Rhesus positive heterozygotes exists in the population equilibrium and therefore an increase of the frequency of Rhesus negative homozygotes is always accompanied by an increase of the frequency of heterozygotes within a population. The advantage or disadvantage of an increased frequency of Rhesus negative homozygotes in a population is therefore usually at least partly compensated by an increased frequency of heterozygotes. However, from the point of view of human medicine and especially that of an individual, the increased risk of a particular disease associated with one genotype is not compensated by the decreased risk of a disease in individuals with another genotype.

From the point of basic science, the most important merit of this study is its strong and two-independent-ways’ support of the heterozygote advantage hypothesis. The results suggest that the Rhesus factor polymorphism is maintained in human populations due to a higher resistance or
tolerance of heterozygotes against specific diseases. It could be speculated to what extent the highly uneven distributions of RHD minus alleles in world populations might be the result of a founder event and a gene flow [20] and to what extent it is also modulated by specific selection pressures caused by differences in the geographical distribution of some disease or diseases. From the practical point of view, its potential merit for prevention of certain diseases and for public health policy in general are more important.

**Competing interests**
The authors declare that they have no competing interests.

**Authors' contributions**
JF designed both studies and performed all data analyses, RH prepared the questionnaire for the cross sectional study, MD collected RhD frequency data for the ecologic study and collected the data for the cross sectional study. All authors participated on writing the manuscript.

**Acknowledgements**
We would like to thank Marek Maly from Department of Statistics, National Health Institute, Prague for his help with the statistics and Zdeněk Hodný Ins. Mol. Genetics, Czech Academy of Science, for his help with preparing the medical questionnaire.

**References**


Legends

Figure 1. Correlation of the frequencies of Rhesus negative subjects in 65 countries with the mortality rates for nine diseases or disease categories. The x and y axes show the frequency of Rhesus negative homozygotes in the population of a country and mortality (numbers of deaths per population of 100,000), respectively. The figures represent the results of the Pearson correlation analysis, namely the level of significance (p) and the coefficient of determination, i.e. the fraction of a country’s variability of the specific mortality rate that can be explained by the differences within the frequencies of Rhesus negative subjects.

Table 1. Multiple correlations between 23 factors explaining 85% of variability in disease burden (DALY) between 192 WHO member countries and the frequencies of Rhesus genotypes and five confounding variables. The first column shows the percentages of variability between countries in DALY explained by a particular factor, other columns show partial correlation (Beta) and statistical significance (p) for the effect of a particular factor, including the frequencies of Rhesus factor genotypes in particular
countries. Significant results (p < 0.05) and trends (p < 0.10) are printed in bold. Asterisks indicate results significant in two-sided tests. p values < 0.0005 are coded as 0.000.

**Table 2.** Difference in various health status related variables between RhD negative and RhD positive subjects

N, Tau and p shows number of responders, partial Kendall’s Tau and statistical significance, respectively. Mostly significant effect of age on health status was controlled in this non parametric test. Positivity of Tau indicates that RhD negative subjects have higher values of particular health related variables, i.e. a worse health status. Significant results (p < 0.05) and trends (p < 0.10) are printed in bold. Asterisks indicate results significant in two-sided tests. p values < 0.0005 are coded as 0.000. Asterisks indicate results significant in two-sided tests. p values < 0.0005 are coded as 0.000.

**Table 3.** Differences between RhD positive and RhD negative participants in specialised medical doctors the subject had to regularly visit at least once in the past two years

Columns 2-5 show numbers of RhD- or RhD+ subjects that had to (V+) and had not to (V-) visit a doctor of particular specialization within the past 2 years. Columns 6 and 7 show Tau and p computed with partial Kendall’s correlation between two binary variables, i.e. the RhD phenotype and the Visiting doctor, controlled for the confounding variable age of a subject. Positivity of Tau indicates that RhD negative subjects have had to more frequently visit a doctor of particular specialization. Significant results (p < 0.05) and trends (p < 0.10) are printed in bold. Asterisks indicate results significant in two-sided tests. p values < 0.0005 are coded as 0.000. Asterisks indicate results significant in two-sided tests.

**Supplementary Figure 1.** Age distribution for male and female participants of a case-control study
**Supplementary Table 1.** Multiple correlation between 23 factors explaining 85 % of variability in disease burden (DALY) between 192 WHO member countries and the frequencies of Rhesus genotypes and five confounding variables – Large model

The first column shows the percentages of variability between countries in DALY explained by a particular factor, other columns show partial correlation (Beta) and statistical significance (p) for the effect of a particular factor, including the frequencies of Rhesus factor genotypes and the interactions between these frequencies and prevalences of male and female smokers in particular countries. Significant results (p < 0.05) and trends (p < 0.10) are printed in bold. Asterisks indicate results significant in two-sided tests. p values < 0.0005 are coded as 0.000.

**Supplementary Table 2.** Correlation between the frequencies of Rhesus factor genotypes and specific disease burden

The correlations were estimated with the General Linear Model with GDP per capita, latitude, humidity, medical care expenses and frequencies of smokers in the population as covariates. Positive B corresponds to a positive correlation and negative B to a negative correlation between particular Rhesus factor genotype frequency and the specific disease burden. Significant results (p < 0.05) and trends (p < 0.10) are printed in bold. Asterisks indicate results significant in two-sided tests. p values < 0.0005 are coded as 0.000. The effect size is shown as Eta\(^2\).

**Supplementary Table 3.** Differences in incidences of particular disorders between RhD positive and RhD negative subjects

Numbers of RhD negative subjects without particular disorders, RhD negative subjects with particular disorders, RhD positive subjects without particular disorders, RhD positive subjects with particular disorders, partial Kendall’s Tau and statistical significance, respectively, are shown in six columns of each section. The effect of age on health status was controlled in partial Kendall’s correlation (non-parametric) test. Positive Tau corresponds to a positive association and negative B to a negative association of RhD negativity with incidence of particular disorder. Significant results (p < 0.05) and trends (p < 0.10) are printed in bold. Asterisks indicate results significant in two-sided tests. p values < 0.0005 are coded as 0.000. The effect size is shown as Tau.
Table 1. Multiple correlations between 23 factors explaining 85 % of variability in disease burden (DALY) between 192 WHO member countries and the frequencies of Rhesus genotypes and five confounding variables.

<table>
<thead>
<tr>
<th>Factor 1</th>
<th>RhD negat.</th>
<th>RhD heteroz.</th>
<th>HDP</th>
<th>Latitude</th>
<th>Humidity</th>
<th>Medic. care</th>
<th>Smokers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beta</td>
<td>-0.385</td>
<td>0.697</td>
<td>-0.527</td>
<td>-0.335</td>
<td>0.261</td>
<td>0.336</td>
<td>-0.224</td>
</tr>
<tr>
<td>(29.5 %)</td>
<td>p 0.337</td>
<td><strong>0.058</strong></td>
<td><strong>0.053</strong></td>
<td><strong>0.081</strong></td>
<td><strong>0.028</strong></td>
<td>0.187</td>
<td>0.157</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Factor 2</th>
<th>Beta</th>
<th>-0.576</th>
<th>0.138</th>
<th>-0.063</th>
<th>-0.069</th>
<th>-0.139</th>
<th>-0.329</th>
<th>-0.168</th>
</tr>
</thead>
<tbody>
<tr>
<td>(10.1 %)</td>
<td>p <strong>0.020</strong></td>
<td>0.528</td>
<td>0.695</td>
<td>0.545</td>
<td><strong>0.053</strong></td>
<td><strong>0.035</strong></td>
<td><strong>0.080</strong></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Factor 3</th>
<th>Beta</th>
<th>-0.363</th>
<th>0.388</th>
<th>-0.465</th>
<th>0.439</th>
<th>0.033</th>
<th>-0.326</th>
<th>0.377</th>
</tr>
</thead>
<tbody>
<tr>
<td>(7.5 %)</td>
<td>p 0.313</td>
<td>0.233</td>
<td><strong>0.056</strong></td>
<td><strong>0.012</strong></td>
<td>0.750</td>
<td>0.154</td>
<td><strong>0.009</strong></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Factor 4</th>
<th>Beta</th>
<th>-0.595</th>
<th>0.474</th>
<th>-0.265</th>
<th>0.331</th>
<th>0.138</th>
<th>0.050</th>
<th>0.063</th>
</tr>
</thead>
<tbody>
<tr>
<td>(5.4 %)</td>
<td>p 0.245</td>
<td>0.304</td>
<td>0.436</td>
<td>0.172</td>
<td>0.352</td>
<td>0.875</td>
<td>0.752</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Factor 5</th>
<th>Beta</th>
<th>0.240</th>
<th>-0.229</th>
<th>-0.199</th>
<th>0.286</th>
<th>-0.096</th>
<th>-0.074</th>
<th>-0.258</th>
</tr>
</thead>
<tbody>
<tr>
<td>(3.9 %)</td>
<td>p 0.645</td>
<td>0.626</td>
<td>0.567</td>
<td>0.248</td>
<td>0.525</td>
<td>0.821</td>
<td>0.210</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Factor 6</th>
<th>Beta</th>
<th>-0.122</th>
<th>0.996</th>
<th>0.463</th>
<th>-0.725</th>
<th>0.219</th>
<th>-0.233</th>
<th>0.162</th>
</tr>
</thead>
<tbody>
<tr>
<td>(3.8 %)</td>
<td>p 0.742</td>
<td><strong>0.004</strong></td>
<td><strong>0.065</strong></td>
<td><strong>0.000</strong></td>
<td><strong>0.046</strong></td>
<td>0.322</td>
<td>0.269</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Factor 7</th>
<th>Beta</th>
<th>1.593</th>
<th>-1.524</th>
<th>-0.035</th>
<th>0.159</th>
<th>-0.038</th>
<th>0.057</th>
<th>-0.268</th>
</tr>
</thead>
<tbody>
<tr>
<td>(2.9 %)</td>
<td>p <strong>0.001</strong></td>
<td><strong>0.000</strong></td>
<td>0.908</td>
<td>0.458</td>
<td>0.773</td>
<td>0.840</td>
<td>0.134</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Factor 8</th>
<th>Beta</th>
<th>0.841</th>
<th>-0.985</th>
<th>-0.216</th>
<th>0.005</th>
<th>0.036</th>
<th>-0.225</th>
<th>-0.259</th>
</tr>
</thead>
<tbody>
<tr>
<td>(2.6 %)</td>
<td>p <strong>0.061</strong></td>
<td><strong>0.016</strong></td>
<td>0.464</td>
<td>0.982</td>
<td>0.778</td>
<td>0.420</td>
<td>0.138</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Factor 9</th>
<th>Beta</th>
<th>1.148</th>
<th>-0.962</th>
<th>-0.335</th>
<th>-0.449</th>
<th>-0.014</th>
<th>0.301</th>
<th>-0.092</th>
</tr>
</thead>
<tbody>
<tr>
<td>(2.2 %)</td>
<td>p <strong>0.019</strong></td>
<td><strong>0.029</strong></td>
<td>0.294</td>
<td><strong>0.050</strong></td>
<td>0.919</td>
<td>0.319</td>
<td>0.622</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Factor 10</th>
<th>Beta</th>
<th>-0.159</th>
<th><strong>0.093</strong></th>
<th>0.148</th>
<th>0.017</th>
<th>0.206</th>
<th>-0.135</th>
<th>-0.183</th>
</tr>
</thead>
<tbody>
<tr>
<td>(2.1 %)</td>
<td>p</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Factor</td>
<td>Beta</td>
<td>p</td>
<td>Factor</td>
<td>Beta</td>
<td>p</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>------</td>
<td>------</td>
<td>--------</td>
<td>------</td>
<td>------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>0.658</td>
<td>0.143</td>
<td>0.005*</td>
<td>0.658</td>
<td>0.143</td>
<td>0.005*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>0.062</td>
<td>0.898</td>
<td>0.087</td>
<td>0.062</td>
<td>0.898</td>
<td>0.087</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>-1.008</td>
<td>0.341</td>
<td>0.039*</td>
<td>-1.008</td>
<td>0.341</td>
<td>0.039*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>0.634</td>
<td>0.201</td>
<td>0.014*</td>
<td>0.634</td>
<td>0.201</td>
<td>0.014*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0.457</td>
<td>0.341</td>
<td>0.033*</td>
<td>0.457</td>
<td>0.341</td>
<td>0.033*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>-0.812</td>
<td>0.107</td>
<td>0.014*</td>
<td>-0.812</td>
<td>0.107</td>
<td>0.014*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>1.040</td>
<td>0.758</td>
<td>0.016*</td>
<td>1.040</td>
<td>0.758</td>
<td>0.016*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>-0.159</td>
<td>0.758</td>
<td>0.017*</td>
<td>-0.159</td>
<td>0.758</td>
<td>0.017*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>-0.543</td>
<td>0.848</td>
<td>0.045*</td>
<td>-0.543</td>
<td>0.848</td>
<td>0.045*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>0.651</td>
<td>0.204</td>
<td>0.142</td>
<td>0.651</td>
<td>0.204</td>
<td>0.142</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>0.032</td>
<td>0.948</td>
<td>0.045*</td>
<td>0.032</td>
<td>0.948</td>
<td>0.045*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The first column shows the percentages of variability between countries in DALY explained by a particular factor, other columns show partial correlation (Beta) and statistical significance (p) for the effect of a particular factor, including the frequencies of Rhesus factor genotypes in particular countries. Significant results (p < 0.05) and trends (p < 0.10) are printed in bold. Asterisks indicate results significant in two-sided tests. p values < 0.0005 are coded as 0.000.
Table 2. Difference in various health status related variables between RhD negative and RhD positive subjects

<table>
<thead>
<tr>
<th>Condition</th>
<th>All</th>
<th></th>
<th>All</th>
<th></th>
<th>All</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>Tau</td>
<td>p</td>
<td></td>
<td>N</td>
<td>Tau</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Allergic problems</td>
<td>2507</td>
<td>0.031</td>
<td>0.020*</td>
<td></td>
<td>924</td>
<td>0.024</td>
</tr>
<tr>
<td>Cancer problems</td>
<td>2423</td>
<td>-0.008</td>
<td>0.548</td>
<td></td>
<td>901</td>
<td>-0.052</td>
</tr>
<tr>
<td>Digestive problems</td>
<td>2449</td>
<td>0.040</td>
<td>0.003*</td>
<td></td>
<td>913</td>
<td>-0.013</td>
</tr>
<tr>
<td>Fertility problems</td>
<td>2396</td>
<td>-0.002</td>
<td>0.855</td>
<td></td>
<td>891</td>
<td>0.060</td>
</tr>
<tr>
<td>Genitourinary problems</td>
<td>2312</td>
<td>-0.007</td>
<td>0.615</td>
<td></td>
<td>880</td>
<td>-0.027</td>
</tr>
<tr>
<td>Heart &amp; vascular problems</td>
<td>2398</td>
<td>0.044</td>
<td>0.001*</td>
<td></td>
<td>884</td>
<td>0.011</td>
</tr>
<tr>
<td>Hematological problems</td>
<td>2376</td>
<td>0.034</td>
<td>0.014*</td>
<td></td>
<td>889</td>
<td>-0.027</td>
</tr>
<tr>
<td>Immunity problems</td>
<td>2465</td>
<td>0.041</td>
<td>0.002*</td>
<td></td>
<td>906</td>
<td>0.040</td>
</tr>
<tr>
<td>Metabolic problems</td>
<td>2328</td>
<td>0.008</td>
<td>0.558</td>
<td></td>
<td>876</td>
<td>-0.009</td>
</tr>
<tr>
<td>Musculoskeletal problems</td>
<td>2302</td>
<td>0.011</td>
<td>0.413</td>
<td></td>
<td>866</td>
<td>-0.024</td>
</tr>
<tr>
<td>Mental health problems</td>
<td>2336</td>
<td>0.026</td>
<td>0.063</td>
<td></td>
<td>872</td>
<td>0.073</td>
</tr>
<tr>
<td>Neurological problems</td>
<td>2374</td>
<td>0.016</td>
<td>0.245</td>
<td></td>
<td>885</td>
<td>0.062</td>
</tr>
<tr>
<td>Respiratory org. problems</td>
<td>2357</td>
<td>0.009</td>
<td>0.516</td>
<td></td>
<td>883</td>
<td>0.004</td>
</tr>
<tr>
<td>Sense organs problems</td>
<td>2336</td>
<td>0.002</td>
<td>0.859</td>
<td></td>
<td>874</td>
<td>-0.006</td>
</tr>
<tr>
<td>Sexual life problems</td>
<td>2359</td>
<td>-0.017</td>
<td>0.222</td>
<td></td>
<td>887</td>
<td>-0.015</td>
</tr>
<tr>
<td>Medicine/day</td>
<td>2521</td>
<td>0.050</td>
<td>0.000*</td>
<td></td>
<td>929</td>
<td>0.031</td>
</tr>
<tr>
<td>Herbs/day</td>
<td>2510</td>
<td>-0.007</td>
<td>0.612</td>
<td></td>
<td>923</td>
<td>0.056</td>
</tr>
<tr>
<td>Antibiotics/year</td>
<td>2556</td>
<td>0.018</td>
<td>0.176</td>
<td></td>
<td>940</td>
<td>-0.021</td>
</tr>
<tr>
<td>Acute care/5 years</td>
<td>2564</td>
<td>0.007</td>
<td>0.597</td>
<td></td>
<td>942</td>
<td>0.010</td>
</tr>
<tr>
<td>Doctors/2 years</td>
<td>2381</td>
<td>0.027</td>
<td>0.047*</td>
<td></td>
<td>888</td>
<td>-0.009</td>
</tr>
</tbody>
</table>
Tired (frequency) | 2371 | 0.035 | 0.012* | 891 | 0.026 | 0.250 | 1480 | 0.025 | 0.148
Headache (frequency) | 2363 | 0.028 | 0.043* | 890 | 0.055 | 0.014* | 1473 | -0.013 | 0.446

N, Tau and p shows number of responders, partial Kendall’s Tau and statistical significance, respectively. Mostly significant effect of age on health status was controlled in this non parametric test. Positivity of Tau indicates that RhD negative subjects have higher values of particular health related variables, i.e. a worse health status. Significant results (p < 0.05) and trends (p < 0.10) are printed in bold. Asterisks indicate results significant in two-sided tests. p values < 0.0005 are coded as 0.000. Asterisks indicate results significant in two-sided tests. p values < 0.0005 are coded as 0.000.
Table 3. Differences between RhD positive and RhD negative participants in specialised medical doctors the subject had to regularly visit at least once in the past two years

<table>
<thead>
<tr>
<th></th>
<th>RhD-V-</th>
<th>RhD-V+</th>
<th>RhD+V-</th>
<th>RhD+V+</th>
<th>tau</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>All</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Internal medicine</td>
<td>617</td>
<td>90</td>
<td>1487</td>
<td>187</td>
<td>0.018</td>
<td>0.192</td>
</tr>
<tr>
<td>Otolaryngology</td>
<td>629</td>
<td>78</td>
<td>1523</td>
<td>151</td>
<td>0.032</td>
<td><strong>0.019</strong>*</td>
</tr>
<tr>
<td>Neurology</td>
<td>647</td>
<td>60</td>
<td>1557</td>
<td>117</td>
<td>0.025</td>
<td><strong>0.067</strong></td>
</tr>
<tr>
<td>Psychiatry</td>
<td>653</td>
<td>54</td>
<td>1580</td>
<td>94</td>
<td>0.038</td>
<td><strong>0.005</strong>*</td>
</tr>
<tr>
<td>Gynecology</td>
<td>563</td>
<td>144</td>
<td>1392</td>
<td>282</td>
<td>0.045</td>
<td><strong>0.001</strong>*</td>
</tr>
<tr>
<td>Surgery</td>
<td>665</td>
<td>42</td>
<td>1561</td>
<td>113</td>
<td>-0.015</td>
<td>0.286</td>
</tr>
<tr>
<td>Infectology</td>
<td>696</td>
<td>11</td>
<td>1651</td>
<td>23</td>
<td>0.007</td>
<td>0.598</td>
</tr>
<tr>
<td>Orthopedy</td>
<td>615</td>
<td>92</td>
<td>1479</td>
<td>195</td>
<td>0.019</td>
<td>0.161</td>
</tr>
<tr>
<td>Dermatology</td>
<td>586</td>
<td>121</td>
<td>1444</td>
<td>230</td>
<td>0.046</td>
<td><strong>0.001</strong>*</td>
</tr>
<tr>
<td>Other Doctors</td>
<td>538</td>
<td>169</td>
<td>1242</td>
<td>432</td>
<td>-0.020</td>
<td>0.149</td>
</tr>
<tr>
<td><strong>Men</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Internal medicine</td>
<td>212</td>
<td>29</td>
<td>563</td>
<td>84</td>
<td>-0.014</td>
<td>0.530</td>
</tr>
<tr>
<td>Otolaryngology</td>
<td>216</td>
<td>25</td>
<td>591</td>
<td>56</td>
<td>0.027</td>
<td>0.237</td>
</tr>
<tr>
<td>Neurology</td>
<td>223</td>
<td>18</td>
<td>613</td>
<td>34</td>
<td>0.042</td>
<td><strong>0.064</strong></td>
</tr>
<tr>
<td>Psychiatry</td>
<td>228</td>
<td>13</td>
<td>627</td>
<td>20</td>
<td>0.054</td>
<td><strong>0.015</strong>*</td>
</tr>
<tr>
<td>Gynecology</td>
<td>241</td>
<td>0</td>
<td>645</td>
<td>2</td>
<td>-0.029</td>
<td>0.199</td>
</tr>
<tr>
<td>Surgery</td>
<td>222</td>
<td>19</td>
<td>590</td>
<td>57</td>
<td>-0.015</td>
<td>0.511</td>
</tr>
<tr>
<td>Infectology</td>
<td>237</td>
<td>4</td>
<td>631</td>
<td>16</td>
<td>-0.024</td>
<td>0.279</td>
</tr>
<tr>
<td>Orthopedy</td>
<td>215</td>
<td>26</td>
<td>579</td>
<td>68</td>
<td>0.004</td>
<td>0.870</td>
</tr>
<tr>
<td>Dermatology</td>
<td>212</td>
<td>29</td>
<td>577</td>
<td>70</td>
<td>0.017</td>
<td>0.435</td>
</tr>
<tr>
<td>Other Doctors</td>
<td>200</td>
<td>41</td>
<td>514</td>
<td>133</td>
<td>-0.040</td>
<td><strong>0.072</strong></td>
</tr>
<tr>
<td><strong>Women</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Internal medicine</td>
<td>405</td>
<td>61</td>
<td>924</td>
<td>103</td>
<td>0.039</td>
<td><strong>0.025</strong></td>
</tr>
<tr>
<td>Otolaryngology</td>
<td>413</td>
<td>53</td>
<td>932</td>
<td>95</td>
<td>0.035</td>
<td><strong>0.041</strong>*</td>
</tr>
<tr>
<td>Neurology</td>
<td>424</td>
<td>42</td>
<td>944</td>
<td>83</td>
<td>0.014</td>
<td>0.421</td>
</tr>
<tr>
<td>Psychiatry</td>
<td>425</td>
<td>41</td>
<td>953</td>
<td>74</td>
<td>0.025</td>
<td>0.142</td>
</tr>
<tr>
<td>Gynecology</td>
<td>322</td>
<td>144</td>
<td>747</td>
<td>280</td>
<td>0.039</td>
<td><strong>0.023</strong>*</td>
</tr>
<tr>
<td>Surgery</td>
<td>443</td>
<td>23</td>
<td>971</td>
<td>56</td>
<td>-0.009</td>
<td>0.620</td>
</tr>
<tr>
<td>Infectology</td>
<td>459</td>
<td>7</td>
<td>1020</td>
<td>7</td>
<td>0.041</td>
<td><strong>0.019</strong>*</td>
</tr>
<tr>
<td>Orthopedy</td>
<td>400</td>
<td>66</td>
<td>900</td>
<td>127</td>
<td>0.026</td>
<td>0.132</td>
</tr>
<tr>
<td>Dermatology</td>
<td>374</td>
<td>92</td>
<td>867</td>
<td>160</td>
<td>0.055</td>
<td><strong>0.001</strong>*</td>
</tr>
<tr>
<td>Other Doctors</td>
<td>338</td>
<td>128</td>
<td>728</td>
<td>299</td>
<td>-0.016</td>
<td>0.363</td>
</tr>
</tbody>
</table>
Columns 2-5 show numbers of RhD- or RhD+ subjects that had to (V+) and had not to (V-) visit a
doctor of particular specialization within the past 2 years. Columns 6 and 7 show Tau and p
computed with partial Kendall’s correlation between two binary variables, i.e. the RhD phenotype
and the Visiting doctor, controlled for the confounding variable age of a subject. Positivity of Tau
indicates that RhD negative subjects have had to more frequently visit a doctor of particular
specialization. Significant results (p < 0.05) and trends (p < 0.10) are printed in bold. Asterisks
indicate results significant in two-sided tests. p values < 0.0005 are coded as 0.000. Asterisks
indicate results significant in two-sided tests.
Additional files provided with this submission:

Additional file 1: Supplementary Figure 1.docx, 365K
http://www.biomedcentral.com/imedia/1060547758164159/supp1.docx
Additional file 2: Supplementary Table 1.doc, 96K
http://www.biomedcentral.com/imedia/1234166341641595/supp2.doc
Additional file 3: Supplementary Table 2.doc, 229K
http://www.biomedcentral.com/imedia/3991157021641595/supp3.doc
Additional file 4: Supplementary Table 3.doc, 310K
http://www.biomedcentral.com/imedia/1769154781164159/supp4.doc