Do differences in *Toxoplasma* prevalence influence global variation in secondary sex ratio? Preliminary ecological regression study

MADHUKAR S. DAMA[1], LENKA MARTINEC NOVÁKOVÁ[2,3] *and* JAROSLAV FLEGR[4]*

[1] *Institute of Wildlife Veterinary Research, KVAFSU, Doddaluvara, Kodagu 571232, India*
[2] *Department of Anthropology, Faculty of Humanities, Charles University, Prague 158 00, Czech Republic*
[3] *National Institute of Mental Health, Klecany, 250 67, Czech Republic*
[4] *Department of Biology, Faculty of Science, Charles University, Prague 128 44, Czech Republic*

(*Received 25 February 2016; accepted 12 March 2016; first published online 13 April 2016*)

SUMMARY

Sex of the fetus is genetically determined such that an equal number of sons and daughters are born in large populations. However, the ratio of female to male births across human populations varies significantly. Many factors have been implicated in this. The theory that natural selection should favour female offspring under suboptimal environmental conditions implies that pathogens may affect secondary sex ratio (ratio of male to female births). Using regression models containing 13 potential confounding factors, we have found that variation of the secondary sex ratio can be predicted by seroprevalence of *Toxoplasma* across 94 populations distributed across African, American, Asian and European continents. *Toxoplasma* seroprevalence was the third strongest predictor of secondary sex ratio, $\beta = -0{\cdot}097$, $P < 0{\cdot}01$, after son preference, $\beta = 0{\cdot}261$, $P < 0{\cdot}05$, and fertility, $\beta = -0{\cdot}145$, $P < 0{\cdot}001$. Our preliminary results suggest that *Toxoplasma gondii* infection could be one of the most important environmental factors influencing the global variation of offspring sex ratio in humans. The effect of latent toxoplasmosis on public health could be much more serious than it is usually supposed to be.

Key words: Toxoplasmosis, manipulation hypothesis, Trivers Willard effect, secondary sex ratio, ecological regression.

INTRODUCTION

Toxoplasma gondii (Coccidia, Apicomplexa) is a highly successful protozoan parasite infecting about one third of humans worldwide. Felines, especially cats, are the definitive hosts of *Toxoplasma*, while any warm-blooded animal can be an intermediate host and many vertebrates as well as invertebrates can serve as the paratenic host of this parasite (Dubey, 1998). While congenital infection can lead to an abortion or serious health consequences for infected children, postnatal infection of immunocompetent humans with a common non-virulent strain of *Toxoplasma* usually takes a subacute course that spontaneously converts into latent toxoplasmosis.

Until the beginning of the 21st century, latent toxoplasmosis, the lifelong presence of slowly dividing bradyzoites encysted in various tissues of infected hosts, had mostly been considered asymptomatic and harmless in immunocompetent subjects. However, within the past 20 years several independent studies have shown that latent toxoplasmosis could produce a plethora of consequences in humans. Most strikingly, latent toxoplasmosis is associated with an increased risk of various psychiatric and neurological disorders, such as schizophrenia, bipolar disorder, personality disorder, Parkinson disease, Alzheimer disease, obsessive-compulsive disorder, cryptic epilepsy, recurrent migraines, autism, suicides, homicides and even brain tumours, for a recent review see Flegr (2013*a*). Further, latent toxoplasmosis enhances the occurrence of chronic heart failure, myocarditis, arrhythmia (Paspalaki *et al*. 2001; Yazar *et al*. 2006), inflammatory bowel disease (Prandota, 2012), liver cirrhosis (Ustun *et al*. 2004) and diabetes mellitus types 1 and 2 (Gokce *et al*. 2008; Krause *et al*. 2009) in the infected individuals. Incidence of many diseases positively correlates with the prevalence of latent toxoplasmosis in different countries (Flegr *et al*. 2014).

A series of interesting reproductive phenotypes are seen in women with latent toxoplasmosis, like slower prenatal (Flegr *et al*. 2005; Kaňková and Flegr, 2007) and postnatal (Kaňkova *et al*. 2012) development of offspring, increased length of pregnancy with significantly higher foetal weight gain post 16th week, especially in RhD negative women (Kaňková *et al*. 2010), and most conspicuously, the duration-specific effect on offspring sex ratio at birth (SRB) (defined as the ratio of male births to female births in a population). Women with a relatively high concentration of anamnestic antibodies, i.e. the women with already latent but still relatively recent *Toxoplasma* infection, bear a higher proportion of sons while women with a lower concentration of anamnestic antibodies, i.e. women with older

* Corresponding author: Department of Biology, Faculty of Science, Charles University, Prague 128 44, Czech Republic. E-mail: flegr@cesnet.cz

 © Cambridge University Press 2016
doi:10.1017/S0031182016000597

https://doi.org/10.1017/S0031182016000597 Published online by Cambridge University Press

Toxoplasma infections, tend to produce a significantly higher proportion of daughters (Kaňková *et al*. 2007*b*). These two phenomena have been confirmed in experimentally infected mice, which delivered 59% male pups 86–120 days post-infection, which was reduced to 40% by 121–222 days (Kaňková *et al*. 2007*a*).

Prevalence of toxoplasmosis in different countries depends on local environmental factors, especially on temperature and moisture, kitchen habits and hygienic standards. *Toxoplasma* seroprevalence shows a drastic global variation, with as low as 4% in Korea to a very high of 78% in Nigeria. Even within Europe it varies from 11% in Norway to 63% in Germany (Table 1). While the global presence of *Toxoplasma* is bound to affect all the populations, the striking variation in prevalence across populations is likely to produce seroprevalence dependence in these outcomes. Building upon the results from human and laboratory mice studies, we tested whether the variation of offspring sex ratio across countries is influenced by the prevalence of *Toxoplasma* in humans. We studied the offspring sex predictive power of *Toxoplasma* seroprevalence by accounting for the known confounding factors across the world, and show that *Toxoplasma* could be an important mediator of global variation in the offspring sex ratio.

MATERIAL AND METHODS

Dependent variable

SRB for the period of 2002–2007 was taken from the United States Central Intelligence Agency (CIA) (CIA, 2011) and averaged. A ratio above or below 1 means there are more males or females, respectively, whereas a ratio of 1 indicates equality of both the sexes at birth. These datasets have been criticized for not matching census data of some countries such as Switzerland, Sweden, Norway, Ireland, India and Japan. However, these differences are minor and the CIA data are accepted and widely used by cross-cultural researchers (Navara, 2009; Dama, 2012). There were three cases of a very high (China, South Korea) and low (Grenada) sex ratio in the sample. In the former two cases, this was due to a strong son preference and female feticide. The reason for the low ratio in Grenada is not known. However, having rerun the analyses excluding these three cases yielded results similar to those reported below.

Independent variables

Prevalence of toxoplasmosis: Data for 88 countries (Flegr *et al*. 2014) were supplemented with six other countries: Botswana (Joubert and Evans, 1997), Kenya (Kamau *et al*. 2012), Lebanon (Usta *et al*. 2006), Namibia (Joubert and Evans, 1997), Uganda (Lindstrom *et al*. 2006), and Zambia (Kistiah *et al*. 2011). The final list contains data on prevalence of toxoplasmosis (seroprevalence) in women of childbearing age published mostly between 1995 and 2008 for 94 countries, 30 European; see Table 1. It is known that prevalence of toxoplasmosis varies between different regions of the same country and also between different age cohorts (Liu *et al*. 2009; Gao *et al*. 2012). Therefore, many of the data points are most probably imprecise, which can highly increase the risk of a Type II error.

The obtained data were adjusted to a standard age of 22 years to eliminate differences in prevalence caused by a different childbearing age in the particular countries using the formula $\text{Prevalence}_{adj} = 1 - (1 - \text{Prevalence})^{(22/\text{childbearing age})}$ (Lafferty, 2006). As pointed out by an anonymous referee, this formula was based on some non-realistic assumptions, for example on an assumption of a constant infection rate, and could therefore provide biased results; however, it is widely used in toxoplasmosis research. Also, results obtained with adjusted and unadjusted data in the present study are virtually identical (see Table 2).

Toxoplasma prevalence as well as SRB are believed to be influenced by various socioeconomic and environmental factors. To deal with the confounding effect of these factors on the correlation of SRB with toxoplasmosis prevalence, we controlled for the effects of all known factors in regression modelling. These factors are fertility, maternal age, polygyny intensity, wealth, son preference, latitude, parasite stress, nutritional stress, contraceptive use and health status. The rationale behind inclusion of each control variable is presented below.

Coital rate hypothesis of James (1971), that predicts more female births from intercourse around the time of ovulation, was tested by Barber (2004) for 148 countries revealing a significant correlation of SRB with the intensity of polygyny ($r = -0{\cdot}41$) and total fertility ($r = -0{\cdot}60$). In the same analysis, SRB was found to be significantly positively correlated with wealth, the proportion of women using any form of contraception and maternal age. While maternal age had been known to influence offspring sex long ago (Lowe and McKeown, 1950), influence of parental wealth on offspring sex gained further support by recent studies (Cameron and Dalerum, 2009). Indicator of polygyny for the year of 2009 was obtained from Gender, Institutions and Development Database (OECD, 2009). Polygyny is defined as men having multiple wives simultaneously. Countries were coded as 1 = generally not accepted/polygyny is not legal in a country, 2 = accepted by part of the population/polygyny is only legal for some people, or 3 = generally accepted/polygyny is legal in a country.

https://doi.org/10.1017/S0031182016000597 Published online by Cambridge University Press

Table 1. Prevalence of latent toxoplasmosis in women of childbearing age in various countries. Third column shows prevalence (%) adjusted to a standard age of 22 years to account for variation in childbearing age across countries using the formula $Prevalence_{adj} = 1 - (1 - Prevalence)^{(22/childbearing\ age)}$ (Lafferty, 2006). Year in which the given study has been carried out is shown in the fourth column, the fifth column states the number of women in the sample and the last one gives sex ratio at birth (SRB). Data for 88 countries published in Flegr *et al*. (2014) have been supplemented with six other countries: Botswana (Joubert and Evans, 1997), Kenya (Kamau *et al*. 2012), Lebanon (Usta *et al*. 2006), Namibia (Joubert and Evans, 1997), Uganda (Lindstrom *et al*. 2006), and Zambia (Kistiah *et al*. 2011)

Country	Prevalence	Age Adj. prevalence	Period	No.	Sex ratio at birth
1. Albania	49	42	2004–2005	496	1·062
2. Argentina	60	53	2001	1007	1·047
3. Australia	23	16	2001	308	1·052
4. Austria	42	36	1997	4601	1·052
5. Bahrain	22	16	2005	3499	1·037
6. Bangladesh	38	38	1995–1996	286	1·056
7. Belgium	49	42	2004	16 541	1·048
8. Benin	54	47	1993	211	1·034
9. Botswana	11	9	1978	1063	1·030
10. Brazil	50	50	2012	2136	1·050
11. Burkina Faso	25	25	2006	336	1·035
12. Cameroon	77	70	1992	1014	1·030
13. Canada	20	17	2006	–	1·051
14. Chile	39	33	1996	7536	1·047
15. China	11	11	2006	235	1·143
16. Colombia	54	54	2006	630	1·037
17. Costa Rica	76	76	1996	1234	1·050
18. Croatia	29	24	2000	1109	1·060
19. Cuba	55	55	2004	526	1·059
20. Czech Republic	20	16	2007	1053	1·059
21. Democratic People's Republic of Korea	4	3	2008	351	1·092
22. Democratic Republic of the Congo	60	60	1990	2897	1·027
23. Denmark	28	20	1999	89 873	1·058
24. Egypt	42	36	1995	62	1·050
25. Estonia	68·6	45	1999–2000	1277	1·058
26. Ethiopia	74	66	2012	1016	1·030
27. Finland	20	17	1989	16 733	1·041
28. France	54	47	1995	13 459	1·051
29. Gabon	71	71	1997	767	1·030
30. Germany	63	50	1999	4854	1·058
31. Greece	25	21	2004	5532	1·063
32. Grenada	57	50	2006	534	1·017
33. Hungary	45	39	2000	31 759	1·062
34. Iceland	13	8	1998	440	1·049
35. India	35	35	2003	180	1·060
36. Indonesia	53	46	2006	17 735	1·050
37. Iran (Islamic Republic of)	39	33	2007	576	1·050
38. Iraq	49	42	2002	254	1·055
39. Ireland	34	25	2008	20 252	1·070
40. Israel	21	17	1989	213	1·051
41. Italy	23	16	2004	3426	1·068
42. Jamaica	57	57	1986	1604	1·050
43. Japan	10	8	2011	4466	1·052
44. Jordan	47	40	2005	280	1·057
45. Kenya	54	54	2011	–	1·030
46. Kuwait	46	53	2002–2005	225	1·037
47. Lebanon	77	70	200–2004	1227	1·060
48. Libya (State of)	45	34	2007	143	1·054
49. Lithuania	40	34	1991	–	1·055
50. Macedonia	22	18	2005	–	1·080
51. Madagascar	84	84	1992	599	1·028
52. Malaysia	49	42	2003	200	1·067
53. Mexico	49	49	2006	–	1·050
54. Montenegro	27	23	–	–	1·080
55. Morocco	51	44	2007	2456	1·054

https://doi.org/10.1017/S0031182016000597 Published online by Cambridge University Press

Table 1. (*Cont.*)

Country	Prevalence	Age Adj. prevalence	Period	No.	Sex ratio at birth
56. Mozambique	19	13	2008	150	1·030
57. Namibia	10	10	1978	1063	1·030
58. Nepal	55	55	1998	345	1·050
59. Netherlands	35	26	2004	7521	1·051
60. New Zealand	35	26	2004	500	1·044
61. Nigeria	78	71	1992	352	1·040
62. Norway	11	9	1993	35 940	1·054
63. Pakistan	33	28	1997	105	1·050
64. Papua New Guinea	18	15	1990	197	1·060
65. Peru	39	33	–	–	1·050
66. Poland	40	34	2003	4916	1·060
67. Portugal	24	17	2011	401	1·067
68. Qatar	35	30	2005–2008	1857	1·048
69. Romania	44	38	2008	184	1·061
70. Sao Tome and Principe	75	68	2007	499	1·030
71. Saudi Arabia	32	27	1991	921	1·043
72. Senegal	40	34	1993	353	1·031
73. Serbia	31	26	2007	765	1·080
74. Singapore	17	14	–	120	1·078
75. Slovakia	22	18	2008	656	1·050
76. Slovenia	25	21	2002	21 270	1·063
77. South Africa	10	8	2010?	779	1·030
78. Spain	32	23	2004	16 362	1·068
79. Sudan	42	36	2003	487	1·051
80. Sweden	18	13	2001	40 978	1·060
81. Switzerland	35	26	2006	–	1·051
82. Thailand	13	11	2001	1200	1·054
83. Togo	75	68	1991	620	1·026
84. Trinidad and Tobago	43	43	2008	450	1·046
85. Tunisia	43	37	1996	2231	1·065
86. Turkey	54	47	2005	1149	1·050
87. Uganda	54	54	2005	130	1·030
88. United Arab Emirates	23	19	1997	1503	1·050
89. UK	9	6	2005	1897	1·050
90. United Republic of Tanzania	35	35	1991	549	1·030
91. USA	11	9	2007	–	1·050
92. Venezuela	38	38	2006	446	1·068
93. Vietnam	11	9	2003	300	1·068
94. Zambia	7	7	1990	375	1·020

Contraceptive use, which indicates the proportion of women of reproductive age who are using (or whose partner is using) a contraceptive method for the period of 2005–2009, was obtained from the World Bank (2011). Total fertility estimates for the year of 2008 were taken from the World Bank (2011). Total fertility rate represents the number of children who would have been born to a woman if she were to live to the end of her childbearing years and bear children in accordance with the current age-specific fertility rates. The estimate includes all the children born dead or alive. Gross national income *per capita* based on purchasing power parity (GNI) was used as a measure of wealth. GNI is the sum of value added by all resident producers plus any product taxes (less subsidies) not included in the valuation of output plus net receipts of primary income (compensation of employees and property income) from abroad. GNI, calculated in national currency, is usually converted to US dollars according to official exchange rates for comparisons across economies. GNI data were taken from the World Bank (2011). Values were log transformed for normality. Mother age was calculated as a mode of the 5-year age block with the highest fertility in a country as explained by Barber (2004).

Navara proposed that latitudinal variation of climatic factors, by its ability to create cross-cultural variation in resource availability and consumer density, could influence offspring sex ratio (Navara, 2009). She examined SRB in relation to latitude and associated climatic variables across 202 countries, and showed that SRB was positively correlated with latitude such that tropical populations produce more daughters compared with temperate and subarctic populations. Latitude values, namely rounded latitude for the centroid or centre point of a *country* expressed in degrees and minutes, for nations were obtained from the CIA World Factbook (CIA, 2011) and numerical values were used irrespective of direction.

https://doi.org/10.1017/S0031182016000597 Published online by Cambridge University Press

Table 2. Side-by-side comparison of categorical regressions of sex ratio at birth on known confounding factors and toxoplasmosis seroprevalence (both adjusted and unadjusted for mother age), respectively

	Toxoplasmosis prevalence not included	Toxoplasmosis prevalence (adjusted) included	Toxoplasmosis prevalence (unadjusted) included
R^2 (adj.)	0·538 (0·515)	0·606 (0·530)	0·607 (0·531)
F	7·004***	6·457***	6·529***
ΔR^2		0·068***	0·069***
Contraceptive use	0·048	0·046	0·044
Fertility	−0·148***	−0·145***	−0·148***
Health factor	0·091***	0·085***	0·086***
Latitude	0·042	0·031	0·033
Log wealth	0·002	−0·006	−0·008
Mother age	0·003	−0·004	0·001
Nutrition stress	−0·024	−0·023	−0·022
Parasite stress	−0·093***	−0·093***	−0·095**
Polygyny intensity	−0·078[a]	−0·087**	−0·089**
Son preference	0·268*	0·261*	0·266*
Toxoplasmosis prevalence		−0·097**	−0·095**

ΔR^2 in the two models in which toxoplasmosis seroprevalence has been included refers to a change in R^2 relative to the model in which the variable has not been included. *Denotes $P < 0·05$, **$P < 0·01$ and ***$P < 0·001$.
[a] Denotes a trend at $P < 0·1$.

Testing the relation of high resource availability and male-biased sex ratios frequently found in many animals, Mathews *et al*. (2008) studied a large cohort of British Women and found that 56% of women in the highest third of preconception energy intake bore boys, compared with 45% in the lowest third. Disability-adjusted life years (DALY) lost due to protein energy malnutrition, iodine deficiency, vitamin A deficiency, and iron deficiency were used as an independent measure of nutrition stress. These variables were obtained from the World Health Organization (WHO, 2008) and log transformed for normality. DALY has also been used by sex ratio researchers before (Dama, 2011). Henceforth, DALY owing to nutrition will be referred to as nutritional stress.

Some of the cultures, especially from Asian countries, show strong preference for sons. These populations frequently terminate female pregnancies leading to a significant increase in population SRB values (Hesketh and Xing, 2006). Son preference (missing women) describes the difference between the number of women that should be alive (assuming no son preference) and the actual number of women in a country. Values for prevalence of son preference for the year of 2009 were obtained from Gender, Institutions and Development Database (OECD, 2009). As son preference results in female selective abortions (Hesketh and Xing, 2006), it is necessary to adjust for the level of son preference in the statistical analysis. The variable was treated as ordinal and coded in the following fashion: 1 = male babies generally not preferred over female ones, 2 = male babies preferred over female ones by part of the population, or 3 = male babies generally preferred over female ones.

Dama (2011) tested the relation of measures of parental condition with SRB and found that SRB shows a strong positive correlation with health indicators and a strong negative correlation with mortality rates across the world. As a measure of mortality level for each nation, health adjusted life expectancy at birth (HALE for the year of 2007), estimated by WHO (2009), was used. While life expectancy at birth summarizes the mortality pattern that prevails across all age groups, HALE adds up expectation of life for different health states and measures the average number of years that a person can expect to live in 'full health' by taking into account the years lived in less than full health due to disease and/or injury. These two measures reflect the age-standardized summary of mortality in a population; however, only HALE will be used in the present analysis as it is a more complete estimate of health than standard life-expectancy rates. Mortality rates at different stages of life were: infant mortality rate (2009), under-five mortality rate (2009), maternal mortality rate (2008), and adult mortality rate (2008), obtained from the World Bank (2011). While infant mortality rate and maternal mortality ratio are the actual numbers of deaths of infants (during the first year of life per 1000 live births in a given year) and mothers (per 100 000 live births in a given year), under 5 mortality rate and adult mortality rate are the probabilities of dying before reaching the age of five and between the age of 15–65, respectively.

In a follow-up study, Dama (2012) showed that the relation of SRB with latitude and health variables is most likely to be driven by the cross-national variation of parasite stress level as all the statistical analyses pointed towards a strong negative association of parasite intensity and possibility of son births

https://doi.org/10.1017/S0031182016000597 Published online by Cambridge University Press

(Dama, 2012). Parasite stress across nations was measured by DALY lost due to parasitic diseases (WHO, 2008). One DALY owing to parasites equals one healthy year of life lost per one million people. This measure covers disability due to 28 important parasites across the world. The variable was log transformed for normality. Henceforth, DALY owing to parasites will be referred to as parasite stress.

Control variables

Survival of oocysts and therefore the effectiveness of transmission of parasites from definitive to intermediate host depend on moisture of the soil (Ruiz *et al*. 1973), consumption of meat, and number of cats. To control for these factors, we included the average relative humidity in each country (data taken from http://www.climatemps.com/, accessed 2. 4. 2013), sanitation rate, yearly *per capita* consumption of meat and a number of cats *per capita*.

Descriptive statistics of the raw variables and their inter-correlation characteristics are summarized in online Supplementary Table 1 and 2, respectively. It must be noted that the datasets used above, especially the control variables, are for different years. However, we have tried to use the data for the dependent and main predictor variable for the same period of time. This discrepancy is owing to a longer sampling frequency for most of the social variables. These differences are unlikely to affect the statistical outcomes, as changes, if any, in social factors are produced very gradually. Hence, most of the cross-cultural studies use data from the nearest sampling year when data for the desired year are not available (Ember and Ember, 2001; Mace *et al*. 2003; Barber, 2004).

Estimation of missing data

The following five variables had high counts of missing values: son preference (50·5%), polygyny prevalence (42·1%), contraceptive use (6·3%), maternal mortality rate (2·1%) and adult mortality rate (1·1%). To deal with these, we used the missForest package (Stekhoven, 2013*a*), available from the Comprehensive R Archive Network (CRAN) and run in the R (R Development Core Team, 2008). Recommended for conducting multiple imputation of mixed data (numeric and factor variables in one data frame) (Starkweather, 2014), it has been compared with other imputation methods and found to have the least imputation error for both continuous and categorical variables and the smallest prediction difference (error) (Waljee *et al*. 2013). Default settings were used (Stekhoven, 2013*b*).

Data reduction

To reduce the number of variables, principal component analysis (PCA) using the IBM SPSS (IBM Corp., 2012) categorical PCA (CATPCA) Optimal Scaling option was performed on the 5 mortality variables (Health adjusted life expectancy, Adult mortality rate, Maternal mortality ratio, Under-five mortality rate, and Infant mortality rate), after confirming that the data were suitable for reduction. The CATPCA settings involved discretizing the numeric variables by means of ranking and selecting variable principal as the normalization method. Dimensions in solution were determined upon several trials in order to obtain the most interpretable structure of loadings. The recommendation of Stevens (1992) on factor loadings with respect to sample size was followed, and loadings >0·512 were considered significant (all >0·9). The CATPCA of the five mortality variables revealed one factor with eigenvalue greater than one (4·69), accounting for 93·82% of the total variance. This factor was labelled 'health factor'. Cronbach's alpha reliability coefficient (=0·710) for the health factor indicated satisfactory internal consistency of the computed variable.

Statistical analysis

Analyses were carried out with IBM SPSS 21.0 (IBM Corp., 2012). Normality of the raw data was checked, firstly, by producing skewness and kurtosis values and their respective S.E., from which z-scores were computed and compared with the value of 1·96, as suggested by Field (2013). Secondly, we have visually examined individual histograms of all relevant variables. Finally, Shapiro–Wilk's W tests were run. Given the violation of the normality assumption in all analysed variables non-parametric tests were preferred. To regress the SRB on the above-specified variables, a categorical regression analysis was run using the SPSS Optimal Scaling (CATREG) feature. The assumptions of the test were met since the number of valid cases exceeded the number of predictor variables plus one. Scale and ordinal variables were treated as numeric and ordinal, respectively, and all were discretized by multiplying. A numerical initial configuration was selected, as recommended when no variables are treated as nominal (IBM Corp., 2011). Perfect multicollinearity (intercorrelations >0·9) did not appear a serious problem, which was further supported by reviewing the variance inflation factors (VIF), which were nowhere near the value of 10, and the average VIF was not greater than 1, as recommended by Field (2013). Moreover, parallel analyses with multiple linear regression (not reported here in detail) showed comparable results. In both the cases, we nevertheless decided to err on the side of caution and employ the ridge regression option using default settings and 0·632 bootstrap with 50 samples for resampling. Ridge regression artificially reduces correlation coefficient of each pair of variables by incorporating a ridge parameter to the diagonal of

https://doi.org/10.1017/S0031182016000597 Published online by Cambridge University Press

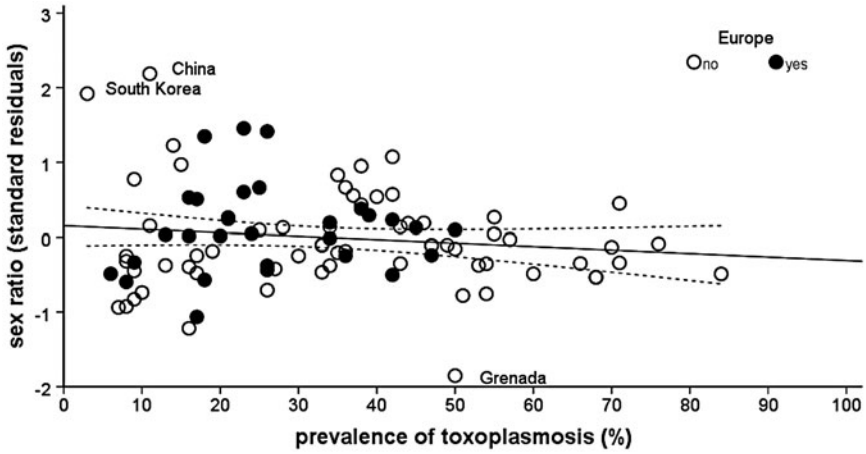

Fig. 1. Relation of toxoplasmosis seroprevalence and population secondary sex ratio across 94 populations. The standard residuals were computed for the model containing all 10 covariates but not the toxoplasmosis prevalence.

a correlation matrix of highly collinear independent variables, leading to reduced error variance of estimators. Based on this principle, ridge regression copes with the collinearity problem (Fox, 1991).

RESULTS

Regression of SRB on the above-specified independent variables revealed that SRB was negatively predicted by *Toxoplasma* prevalence, fertility, parasite stress, and polygyny intensity, and positively by the health factor, and son preference. This means that a greater number of male offspring was produced in populations with lower *Toxoplasma* prevalence, fertility, less intense parasite stress and polygyny, better health status, and a tendency to prefer sons (Table 2; Fig. 1). Importantly, side-by-side comparisons of regression models including *Toxoplasma* prevalence (adjusted and unadjusted) and not, respectively, revealed a significant change in R^2 following the inclusion of the variable in the model (Table 2 and online Supplementary Table 4 and 5), showing it to be an important predictor of SRB.

When countries were divided into European and non-European ones, the results were different. While in the European countries the model was not significant, outside Europe SRB was predicted in the same direction by the same variables as in the whole set of countries excluding polygyny intensity but including contraception, with countries in which more males are born exhibiting greater contraceptive use (online Supplementary Table 3 and 4).

Four additional regression models were fitted to investigate the influence of control variables on the prediction of SRB by *Toxoplasma* prevalence (Table 3). Results show that all the models were significant at least at $P < 0{\cdot}01$ and the βs of *Toxoplasma* prevalence were not much affected by the sequential addition of humidity, sanitation rate, cat ownership, and meat consumption. This further strengthens the hypothesis that *Toxoplasma* is likely to influence the offspring sex ratio in humans.

Furthermore, as suggested by an anonymous referee, SRB may explain a large part of son preference, and thus the latter may not be suitable for inclusion in the regression models. Although this was clearly not the case in the present study (Kendall's Tau = 0·226; $P < 0{\cdot}01$), we ran the same regression models as those reported in Table 2 with the only difference that the son preference variable has not been included. These analyses yielded results virtually identical to those reported in Table 2 (see online Supplementary Table 5 for comparison).

DISCUSSION

Our analysis showed that prevalence of toxoplasmosis is negatively correlated with offspring sex ratio in 94 countries. The correlation can be detected in the whole data set and in the subset of non-European countries, but could not be detected in European countries.

The interpretation of results of ecological regression studies is 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 (Guthrie and Sheppard, 2001; Wakefield and Salway, 2001). Therefore, the existence of strong effects of prevalence of toxoplasmosis on SRB, suggested by results of an ecological study, should be confirmed by independent case controls or cohort studies. The results of a cohort study performed on a population of European women as well as a study performed on experimentally infected female mice suggest that *Toxoplasma* infection increases SRB in the first phases of latent toxoplasmosis while decreases the offspring sex ratio in the latter phases. Women with high titres of specific anti-

https://doi.org/10.1017/S0031182016000597 Published online by Cambridge University Press

Table 3. Sequential addition of control variables did not change the significant prediction of sex ratio at birth by toxoplasmosis prevalence in the categorical regression models

Predictor variables	1	2	3	4	5
Toxoplasmosis prevalence	−0·098**	−0·093**	−0·092*	−0·137**	−0·137**
Humidity		−0·057†	−0·002	−0·010	−0·010
Sanitation rate			−0·007	−0·021	−0·020
Cat ownership				−0·009	0·009
Meat consumption					−0·012
Standard independent variables included#	Yes	Yes	Yes	Yes	Yes
R^2	0·607***	0·615***	0·648***	0·651**	0·652**
Adjusted R^2	0·532	0·535	0·500	0·481	0·469
N	94	94	62	59	59

*Denotes a model/predictor significant $P < 0·05$, **<0·01, and ***<0·001 # Standard independent variables included are Contraceptive use, Health factor, Latitude, log Wealth, Mother age, Nutrition stress, Parasite stress, Polygyny intensity, Son preference, Total fertility and Toxoplasmosis prevalence. The predictor-related values represent βs.

Toxoplasma IgG antibodies (i.e. women probably infected with *Toxoplasma* less than 2 years before pregnancy) had a SRB of 0·72, while women with low titres of specific anamnestic antibodies had a SRB of 0·45 (Kaňková *et al*. 2007*b*). Similarly, female mice that delivered 86–120 days after *Toxoplasma* infection had a SRB of 0·59 while those that delivered at 121–222 days had a SRB of 0·40 (Kaňková *et al*. 2007*a*). The increased sex ratio is speculated to be caused either by a higher probability of survival of more immunogenic male embryos (Kaňková and Flegr, 2007) or by an increased level of testosterone around the time of fertilization (James, 2010). The decreased sex ratio in women and female mice infected with *Toxoplasma* for a long time is considered to be a manifestation of the Trivers–Willard effect, i.e. the increased probability of birth of female offspring in females in poor physical conditions (Flegr, 2010). It is not important in the present context whether the decreased sex ratio is just a side-effect of the impaired health of infected females or if it is the evolutionary adaptation that results in increased or, rather, not-so-impaired fitness of the infected females. On the population level, the direction of *Toxoplasma* infection-associated effect on sex ratio will depend on the fraction of recently infected women in a birth giving age in a particular population. In the high prevalence countries, most women are probably infected by contact with contaminated food and water in the early childhood, i.e. long before their birth giving age. In such countries, prevalence of toxoplasmosis should correlate negatively with sex ratio, which is the pattern observed in non-European countries. The situation will be more complicated in the low prevalence, low fertility countries. For example in the Czech Republic, an increased rate of incidence of toxoplasmosis occurs in women in a birth giving age; in men and women the seroprevalence of toxoplasmosis increases between the ages of 19–39 from 25·5 to 27 and 31·1 to 46·3%, respectively (Kodym *et al*. 2000). Such an increase is hypothesized to occur either due to manipulation with undercooked/raw meat when the young women start cooking in their own households or by transmission during unprotected intercourse with infected men (Dass *et al*. 2011; Flegr, 2013*a*; Holub *et al*. 2013). Therefore, a higher fraction of *Toxoplasma*-infected women in low prevalence countries could acquire their infection a short time before pregnancy and increased offspring sex ratio of recently-infected women could neutralize or even reverse the negative correlation between prevalence of toxoplasmosis and the offspring sex ratio. Such an effect in recently infected women could be especially strong in low-fertility countries because the probable duration of *Toxoplasma* infection positively correlates with the age of women and therefore also with parity. Therefore, a stronger positive effect of toxoplasmosis on sex ratio can be expected to occur rather in primiparous women than in multiparous women in low prevalence countries.

Of course, the absence of a significant effect of prevalence of toxoplasmosis on SRB in the European countries subset could also have another explanation. It has been observed, for example, that Rhesus factor (Rh) phenotype modifies the effects of toxoplasmosis on the human organism (Kaňková *et al*. 2010) and positive heterozygotes are strongly protected against these effects (Novotna *et al*. 2008; Flegr, 2016). The frequency of Rh heterozygotes is much higher in the Caucasian population of European countries than populations in other parts of the globe. It is, however, also possible that the observed association between prevalence of toxoplasmosis and SRB is caused by an unknown factor that is not present in Europe and that independently correlates both with prevalence of toxoplasmosis and SRB.

In our opinion, it is probable that each set of the countries used in the previous nine ecological regression studies dealing with toxoplasmosis published by other authors in the past 10 years would give a slightly different result of the analysis. To maximally

https://doi.org/10.1017/S0031182016000597 Published online by Cambridge University Press

decrease the risk of subjectivity in the selection of countries, we included data for all the countries ($n = 94$) available as of 15th August 2013; which represented the largest ever data set analysed for biological influence of *Toxoplasma*.

It is also highly probable that at least some of the *Toxoplasma* prevalence data are incorrect. As far as we know, national surveys for latent toxoplasmosis are not systematically performed in any country, or at least, no results of such studies have been published. Our experience with such a survey performed in a relatively small and rather ethnically and sociologically homogeneous Czech Republic showed that the prevalence varies drastically among different parts of a country (Kodym *et al*. 2000). Therefore, it is difficult to estimate the average prevalence of toxoplasmosis in women of birth giving age in a particular country in a single study. Hence, we always searched for all available data concerning the time period of 1990–2008 and, whenever possible, we also took into consideration the results on the prevalence published earlier. It should be noted that a lack of precision in the prevalence data increases the risk of false negative but not of false positive results of statistical tests. Lack of precision (stochastic noise) in the prevalence data could bias the estimate of regression parameters (e.g. betas) and increase the risk of false negative results of statistical tests. Stochastic noise alone cannot increase the risk of false positive results of studies. However, it is possible that the precision of the prevalence data (or the SRB data) is somehow related to the prevalence of toxoplasmosis. It is, for example, known that toxoplasmosis influences some personality traits of infected humans, including conscientiousness (Lindová *et al*. 2012) and neuroticism (Flegr *et al*. 2013), and such effects explain a statistically significant portion of the variance in aggregate neuroticism among populations (Lafferty, 2006). It is therefore possible that such a behavioural effect of toxoplasmosis could result in a systematic error in the available prevalence (or other) data, which could result in false positive results of an ecological study. Therefore, even the positive results of our study should be approached with caution and should only be considered preliminary until confirmed in other cross-sectional or longitudinal studies. For example, our study can be repeated on regional data from a large country, e.g. the USA, Mexico, or France, for which the sex ratio as well as toxoplasmosis prevalence data are probably available. It is also necessary to confirm the results of an earlier cohort study (Kaňková *et al*. 2007*b*), showing decreased probability of birth of male offspring in women with 'old' *Toxoplasma* infections, on another population. This study has been supported by the results of the latter mice-infection study (Kaňková *et al*. 2007*a*); however, it has not been repeated on any other human population.

Another possible source of error are temporal changes in the incidence of toxoplasmosis. In many parts of world, the prevalence of toxoplasmosis is changing, mostly having decreased in the past two decades. It is highly probable that the offspring sex ratio reflects the situation in an (unknown) past, rather than the current prevalence of toxoplasmosis. Again, the existence of such a delay could increase the risk of false negative, rather than false positive results of studies.

The observed correlation can be most parsimoniously explained by the effect of latent toxoplasmosis on secondary sex ratio, i.e. by the effect demonstrated in earlier studies including the experimental infection study (Kaňková *et al*. 2007*a*). The opposite causality, i.e. influence of the offspring sex ratio on prevalence of toxoplasmosis, is rather improbable. However, we also cannot exclude a possibility that some unknown factor can correlate both with the offspring sex ratio and the prevalence of toxoplasmosis. Such a factor could be either a particular disease, particular epidemiological factor for toxoplasmosis, e.g. rural *vs* urban style of life, or even some cultural habit. The difficulty with studying causality relationships on the *Toxoplasma*–human model are discussed in detail in Flegr (2013*b*).

The increased probability of producing a baby girl by a *Toxoplasma*-infected woman could be the result of impaired health. The incidence of many diseases, including cardiovascular diseases and certain forms of cancer, positively correlates with the prevalence of toxoplasmosis in different countries (Flegr *et al*. 2014). The possible physiological basis for the correlation of SRB and *Toxoplasma* prevalence could be derived from the sex differences in the costs of offspring production. Male foetus grows faster (Marsal *et al*. 1996) and requires significantly higher parental investment during gestation (Tamimi *et al*. 2003), which means that, to produce sons, women should be in superior body condition to meet higher physiological costs required. Indeed, male foetuses are more often aborted spontaneously than a female fetus due to various stressors (Mace *et al*. 2003; Boklage, 2005). It is, therefore, possible that various factors that influence maternal investment ability (health impairment caused by toxoplasmosis in the present analysis) are more likely than genetic and geographical factors to form the basis for the striking cross-cultural variation in SRB.

Of course, this explanation of the phenomenon is not the only explanation available. Alternatively, the decreased offspring sex ratio can be the result of a decreased level of testosterone (James, 2010, 2012) observed in infected women (Flegr *et al*. 2008*a*, *b*) and mice (Kaňková *et al*. 2011). However, the Trivers–Willard effect-related hypothesis of the observed effect should be preferentially tested because if it is correct, then the effect of latent toxoplasmosis

https://doi.org/10.1017/S0031182016000597 Published online by Cambridge University Press

on public health could be much more serious than it is usually supposed to be. *Toxoplasma gondii* infects most species of homoeothermic animals. This suggests that abundance of this parasite could be one of the most important ecological factors influencing the global variation of offspring sex ratio not only in humans, but possibly also in many other animal species.

SUPPLEMENTARY MATERIAL

The supplementary material for this article can be found at http://dx.doi.org/10.1017/S0031182016000597.

ACKNOWLEDGEMENT

We thank L. Lanchava for improving the manuscript.

FINANCIAL SUPPORT

J.F. and L.M.N.'s work was supported by the Charles University Research Centre (UNCE 204004). J.F. was further supported by the Grant Agency of the Czech Republic (Grant No. P303/16/20958). L.M.N.'s work was further funded by the project 'National Institute of Mental Health (NIMH-CZ)', under grant number ED2.1.00/03.0078, and the European Regional Development Fund, and the Ministry of Education, Youth and Sports – Institutional Support for Longterm Development of Research Organizations – Charles University, Faculty of Humanities (project PRVOUK P20).

REFERENCES

Barber, N. (2004). Sex ratio at birth, polygyny, and fertility: a cross-national study. *Social Biology* **51**, 71–77.

Boklage, C. E. (2005). The epigenetic environment: secondary sex ratio depends on differential survival in embryogenesis. *Human Reproduction* **20**, 583–587.

Cameron, E. Z. and Dalerum, F. (2009). A Trivers-Willard effect in contemporary humans: male-biased sex ratios among billionaires. *PLoS ONE* **4**, e4195.

CIA (2011). *The World Factbook*. Central Intelligence Agency, Washington, DC.

Dama, M. S. (2011). Sex ratio at birth and mortality rates are negatively related in humans. *PLoS ONE* **6**, e23792.

Dama, M. S. (2012). Parasite stress predicts offspring sex ratio. *PLoS ONE* **7**(9), e46169.

Dass, S. A. H., Vasudevan, A., Dutta, D., Soh, L. J. T., Sapolsky, R. M. and Vyas, A. (2011). Protozoan parasite *Toxoplasma gondii* manipulates mate choice in rats by enhancing attractiveness of males. *PLoS ONE* **6**, e27229S.

Dubey, J. P. (1998). Advances in the life cycle of *Toxoplasma gondii*. *International Journal for Parasitology* **28**, 1019–1024.

Ember, C. R. and Ember, M. (2001). *Cross-cultural Research Methods*, AltaMira Press, Lanham, MD.

Field, A. (2013). *Discovering Statistics Using IBM SPSS Statistics*, 4th Edn. Sage, London.

Flegr, J. (2010). Influence of latent toxoplasmosis on the phenotype of intermediate hosts. *Folia Parasitologica (Praha)* **57**, 81–87.

Flegr, J. (2013*a*). How and why *Toxoplasma* makes us crazy. *Trends in Parasitology* **29**, 156–163.

Flegr, J. (2013*b*). Influence of latent *Toxoplasma* infection on human personality, physiology and morphology: pros and cons of the *Toxoplasma*-human model in studying the manipulation hypothesis. *Journal of Experimental Biology* **216**, 127–133.

Flegr, J. (2016). Heterozygote advantage probably maintains rhesus factor blood group polymorphism: Ecological regression study. *PLoS ONE* **11**, e0147955.

Flegr, J., Hrdá, Š. and Kodym, P. (2005). Influence of latent 'asymptomatic' toxoplasmosis on body weight of pregnant women. *Folia Parasitologica (Praha)* **52**, 199–204.

Flegr, J., Lindová, J. and Kodym, P. (2008*a*). Sex-dependent toxoplasmosis-associated differences in testosterone concentration in humans. *Parasitology* **135**, 427–431.

Flegr, J., Lindová, J., Pivoňková, V. and Havlíček, J. (2008*b*). Brief Communication: latent toxoplasmosis and salivary testosterone concentration-important confounding factors in second to fourth digit ratio studies. *American Journal of Physical Anthropology* **137**, 479–484.

Flegr, J., Preiss, M. and Klose, J. (2013). Toxoplasmosis-associated difference in intelligence and personality in men depends on their Rhesus blood group but not ABO blood group. *PLoS ONE* **8**(4), e61272.

Flegr, J., Prandota, J., Sovickova, M. and Israili, Z. H. (2014). Toxoplasmosis – a global threat. Correlation of latent toxoplasmosis with specific disease burden in a set of 88 countries. *PLoS ONE* **9**(3), e90203.

Fox, J. (1991). *Regression Diagnostics*. Sage Publications, Newbury Park.

Gao, X. J., Zhao, Z. J., He, Z. H., Wang, T., Yang, T. B., Chen, X. G., Shen, J. L., Wang, Y., Lv, F. L., Hide, G. and Lun, Z. R. (2012). *Toxoplasma gondii* infection in pregnant women in China. *Parasitology* **139**, 139–147.

Gokce, C., Yazar, S., Bayram, F. and Gundogan, K. (2008). *Toxoplasma gondii* antibodies in type 1 Diabetes mellitus. *Turkiye Klinikleri Tip Bilimleri Dergisi* **28**, 619–622.

Guthrie, K. A. and Sheppard, L. (2001). Overcoming biases and misconceptions in ecological studies. *Journal of the Royal Statistical Society Series a – Statistics in Society* **164**, 141–154.

Hesketh, T. and Xing, Z. W. (2006). Abnormal sex ratios in human populations: causes and consequences. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 13271–13275.

Holub, D., Flegr, J., Dragomirecka, E., Rodriguez, M., Preiss, M., Novak, T., Cermak, J., Horacek, J., Kodym, P., Libiger, J., Höschl, C. and Motlova, L. B. (2013). Differences in onset of disease and severity of psychopathology between toxoplasmosis-related and toxoplasmosis-unrelated schizophrenia. *Acta Psychiatrica Scandinavica* **127**, 227–238.

IBM Corp (2011). *Categorical Regression Options*. IBM Corp., Armonk, NY.

IBM Corp (2012). *IBM SPSS Statistics for Windows*. IBM Corp., Armonk, NY.

James, W. H. (1971). Cycle day of insemination, coital rate, and sex ratio. *Lancet* **1**, 112–114.

James, W. H. (2010). Potential solutions to problems posed by the offspring sex ratios of people with parasitic and viral infections. *Folia Parasitologica (Praha)* **57**, 114–120.

James, W. H. (2012). The relevance of the epidemiology of human sex ratios at birth to some medical problems. *Paediatric and Perinatal Epidemiology* **26**, 181–189.

Joubert, J. J. and Evans, A. C. (1997). Current status of food-borne parasitic zoonoses in South Africa and Namibia. *Southeast Asian Journal of Tropical Medicine and Public Health* **28** (Suppl. 1), 7–10.

Kamau, P., Jaoko, W. and Gontier, C. (2012). Seroepidemiolgy of *Toxoplasma gondii* in ante-natal women attending Kenyatta National Hospital, Kenya. *International Journal of Infectious Diseases* **16**, E162–E162.

Kaňková, Š. and Flegr, J. (2007). Longer pregnancy and slower fetal development in women with latent 'asymptomatic' toxoplasmosis. *BMC Infectious Diseases* **7**, 114.

Kaňková, Š., Kodym, P., Frynta, D., Vavřinová, R., Kuběna, A. and Flegr, J. (2007*a*). Influence of latent toxoplasmosis on the secondary sex ratio in mice. *Parasitology* **134**, 1709–1717.

Kaňková, Š., Šulc, J., Nouzová, K., Fajfrlik, K., Frynta, D. and Flegr, J. (2007*b*). Women infected with parasite *Toxoplasma* have more sons. *Naturwissenschaften* **94**, 122–127.

Kaňková, Š., Šulc, J. and Flegr, J. (2010). Increased pregnancy weight gain in women with latent toxoplasmosis and RhD-positivity protection against this effect. *Parasitology* **137**, 1773–1779.

Kaňková, Š., Kodym, P. and Flegr, J. (2011). Direct evidence of *Toxoplasma*-induced changes in serum testosterone in mice. *Experimental Parasitology* **128**, 181–183.

Kaňkova, S., Šulc, J., Křivohlavá, R., Kuběna, A. and Flegr, J. (2012). Slower postnatal motor development in infants of mothers with latent toxoplasmosis during the first 18 months of life. *Early Human Development* **88**, 879–884.

Kistiah, K., Barragan, A., Winiecka-Krusnell, J., Karstaedt, A. and Frean, J. (2011). Seroprevalence of *Toxoplasma gondii* infection in HIV-

https://doi.org/10.1017/S0031182016000597 Published online by Cambridge University Press

positive and HIV-negative subjects in Gauteng, South Africa. *South African Journal of Epidemiology and Infection* **26**, 225–228.

Kodym, P., Malý, M., Švandová, E., Lekatková, H., Badoutová, M., Vlková, J., Beneš, C. and Zástěra, M. (2000). *Toxoplasma* in the Czech Republic 1923–1999: first case to widespread outbreak. *International Journal for Parasitology* **30**, 11–18.

Krause, I., Anaya, J. M., Fraser, A., Barzilai, O., Ram, M., Abad, V., Arango, A., Garcia, J. and Shoenfeld, Y. (2009). Anti-infectious antibodies and autoimmune-associated autoantibodies in patients with type I diabetes mellitus and their close family members. *Annals of the New York Academy of Sciences* **1173**, 633–639.

Lafferty, K. D. (2006). Can the common brain parasite, *Toxoplasma gondii*, influence human culture? *Proceedings of the Royal Society B – Biological Sciences* **273**, 2749–2755.

Lindová, J., Příplatová, L. and Flegr, J. (2012). Higher extraversion and lower conscientiousness in humans infected with *Toxoplasma*. *European Journal of Personality* **26**, 285–291.

Lindstrom, I., Kaddu-Mulindwa, D. H., Kironde, F. and Lindh, J. (2006). Prevalence of latent and reactivated *Toxoplasma gondii* parasites in HIV-patients from Uganda. *Acta Tropica* **100**, 218–222.

Liu, Q., Wei, F., Gao, S., Jiang, L., Lian, H., Yuan, B., Yuan, Z., Xia, Z., Liu, B. and Xu, X. (2009). *Toxoplasma gondii* infection in pregnant women in China. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **103**, 162–166.

Lowe, C. R. and McKeown, T. (1950). The sex ratio of human births related to maternal age. *British Journal of Social Medicine* **4**, 75–85.

Mace, R., Jordan, F. and Holden, C. (2003). Testing evolutionary hypotheses about human biological adaptation using cross-cultural comparison. *Comparative Biochemistry and Physiology A* **136**, 85–94.

Marsal, K., Persson, P. H., Larsen, T., Lilja, H., Selbing, A. and Sultan, B. (1996). Intrauterine growth curves based on ultrasonically estimated foetal weights. *Acta Paediatrica* **85**, 843–848.

Mathews, F., Johnson, P. J. and Neil, A. (2008). You are what your mother eats: evidence for maternal preconception diet influencing foetal sex in humans. *Proceedings of the Royal Society B – Biological Sciences* **275**, 1661–1668.

Navara, K. J. (2009). Humans at tropical latitudes produce more females. *Biological Letters* **5**, 524–527.

Novotna, M., Havlicek, J., Smith, A. P., Kolbekova, P., Skallova, A., Klose, J., Gasova, Z., Pisacka, M., Sechovska, M. and Flegr, J. (2008). *Toxoplasma* and reaction time: role of toxoplasmosis in the origin, preservation and geographical distribution of Rh blood group polymorphism. *Parasitology* **135**, 1253–1261.

OECD (2009). *Gender, Institutions and Development Database 2009 (GID-DB)*. Organisation for Economic Co-operation and Development, Paris, France.

Paspalaki, P. K., Mihailidou, E. P., Bitsori, M., Tsagkaraki, D. and Mantzouranis, E. (2001). Polyomyositis and myocarditis associated with acquired toxoplasmosis in an immunocompetent girl. *BMC Musculoskeletal Disorders* **2**, 8.

Prandota, J. (2012). Gastrointestinal tract abnormalities in autism, inflammatory bowel disease and many other clinical entities may be due to *T. gondii* infection. *Open Access Scientific Reports* **1**, 256.

R Development Core Team (2008). *A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Ruiz, A., Frenkel, J. K. and Cerdas, L. (1973). Isolation of *Toxoplasma* from soil. *Journal of Parasitology* **59**, 204–206.

Starkweather, J. (2014). *A New Recommended Way of Dealing with Multiple Missing Values: Using missforest for all Your Imputation Needs*. University of North Texas, Denton, TX.

Stekhoven, D. J. (2013*a*). *missForest: Nonparametric Missing Value Imputation using Random Forest*. R Foundation for Statistical Computing, Vienna, Austria.

Stekhoven, D. J. (2013*b*). *Package 'missForest': Nonparametric Missing Value Imputation using Random Forest*. Swiss Federal Institute of Technology, Zürich, Switzerland.

Stevens, J. (1992). *Applied Multivariate Statistics for the Social Sciences*. Lawrence Erlbaum, Hillsdale, NJ.

Tamimi, R. M., Lagiou, P., Mucci, L. A., Hsieh, C. C., Adami, H. O. and Trichopoulos, D. (2003). Average energy intake among pregnant women carrying a boy compared with a girl. *BMJ (Clinical Research Ed.)* **326**, 1245–1246.

Usta, I. M., Seoud, M. A. F., Maarouf, H. H., Hobeika, E. M. and Nassar, A. H. (2006). Seroprevalence of *Toxoplasma* infection among pregnant women in Lebanon. *Obstetrics and Gynecology* **107**, 42S–43S.

Ustun, S., Aksoy, U., Dagci, H. and Ersoz, G. (2004). Incidence of toxoplasmosis in patients with cirrhosis. *World Journal of Gastroenterology* **10**, 452–454.

Wakefield, J. and Salway, R. (2001). A statistical framework for ecological and aggregate studies. *Journal of the Royal Statistical Society Series a – Statistics in Society* **164**, 119–137.

Waljee, A. K., Mukherjee, A., Singal, A. G., Zhang, Y., Warren, J., Balis, U., Marrero, J., Zhu, J. and Higgins, P. D. R. (2013). Comparison of imputation methods for missing laboratory data in medicine. *BMJ Open* **3**, e002847.

WHO (2008). *The Global Burden of Disease: 2004 Update*. World Health Organization, Switzerland.

WHO (2009). *Global Health Risks: Mortality and Burden of Disease Attributable to Selected Major Risks*. World Health Organization, France.

World Bank (2011). *World Development Indicators 2010*. World Bank, Washington, DC.

Yazar, S., Gur, M., Ozdogru, I., Yaman, O., Oguzhan, A. and Sahin, I. (2006). Anti-*Toxoplasma gondii* antibodies in patients with chronic heart failure. *Journal of Medical Microbiology* **55**, 89–92.

https://doi.org/10.1017/S0031182016000597 Published online by Cambridge University Press