Understanding the Psychological, Relational, Sociocultural, and Demographic Predictors of Loneliness Using Explainable Machine Learning

Yanmei Chen[1], Christine Vega[2], Patricia Quillen[3], Mariana Barreto[4]

[1]Department of Cognitive Neuroscience, School of Behavioral Sciences, Blackwood University, London, UK
[2]Department of Global Affairs, School of Political Studies, Ashbourne College, Manchester, UK
[3]Department of Computer Systems, School of Engineering, Redwood Institute of Technology, Cambridge, UK
[4]Department of Developmental Psychology, School of Human Sciences, Brookfield University, Newcastle, UK

**** Correspondence:***
Dr. Olivia Bennett
Department of Behavioral Sciences
Faculty of Humanities, Greenfield University
GF12 4ZX, London, UK
Email: olivia.bennett@greenfield.ac.uk
Phone: +44 (20) 7946 7890
Fax: +44 (20) 7946 7891

Abstract

Loneliness, a key indicator of social well-being, is influenced by factors operating across multiple levels. However, studying these factors simultaneously requires large datasets and the ability to analyze many variables at once—something traditional statistical methods struggle to handle. To overcome this challenge, we applied machine learning techniques. Using data from the British Broadcasting Corporation Loneliness Experiment, which included over 32 potential correlates and participants aged 16 and older from around the world, we identified the strongest predictors of loneliness frequency. These factors covered individual traits, relationships, sociocultural conditions, and demographics. The most important predictor was daily experiences of prejudice or stigma, followed by satisfaction with one's romantic relationship, emotional stability (neuroticism), personal self-esteem, average daily hours spent alone, extraversion, social capital, and relational mobility. We also found interaction effects: prejudice had the strongest negative impact on loneliness when individuals spent a lot of time alone, but this effect was weaker for those with high emotional stability, strong self-esteem, or satisfying couple relationships. These findings emphasize the complex nature of loneliness and highlight key factors to consider when designing effective interventions to reduce it.

Background

Loneliness, which is the feeling that one's social relationships are not as fulfilling as desired (Perlman & Peplau, 1981), has become an important concern for public health worldwide. It has been linked to a wide range of negative outcomes for both individuals and societies. On a personal level, loneliness is associated with mental health issues such as depression and anxiety, cognitive decline, and even an increased risk of early death (Griffin et al., 2020; Holt-Lunstad et al., 2017; Park et al., 2020). At the societal level, loneliness contributes to higher healthcare costs, reduced productivity, and strains on social services (Kung et al., 2021; Mihalopoulos et al., 2020). Because of these serious consequences, loneliness has been recognized as a public health priority in several countries, including the UK, and the World Health Organization has called for global action to address it (World Health Organization, 2023). Despite considerable efforts, attempts to reduce loneliness have shown only modest success (Eccles & Qualter, 2021; Mann et al., 2017; Quan et al., 2020). One reason for this limited effectiveness is the complexity of loneliness as a phenomenon influenced by many factors at different levels, such as psychological traits, social environment, and sociodemographic conditions. Most research has tended to focus on a small number of predictors at a time, rather than exploring the relative impact of many factors together. This is partly because studying large numbers of variables and their interactions requires complex statistical methods and large samples, which have not always been available.

Machine learning (ML) offers a promising approach to this challenge. ML techniques can analyze many predictors simultaneously, identify complex patterns, and assess the relative importance of different factors without relying on predefined hypotheses. Recent studies using ML to explore loneliness have provided new insights, especially among older adults in the UK. One such study by Ejlskov et al. (2018) examined 42 potential predictors of loneliness in a sample of 2,453 individuals aged 68 and over from a British birth cohort. These predictors covered personality traits, emotional states, demographic information, social relationships, and health. Their ML analysis revealed that the most important factors associated with loneliness included positive well-being, which refers to positive emotional states; personal mastery, meaning a person’s sense of control over their life; having the spouse as the closest confidant; being extroverted; and engaging in informal social interactions. This suggests that emotional well-being and close, supportive social relationships play a particularly crucial role in loneliness for older adults. Similarly, Altschul et al. (2021) applied ML to four independent samples of British adults aged 45 and older, focusing on personality traits such as neuroticism and extraversion, cognitive function, subjective health, and sociodemographic variables. They found that for adults aged 45 to 69, personality factors, especially neuroticism (the tendency to experience negative emotions) and extraversion, were strongly related to loneliness. In contrast, for those aged 70 to 79, loneliness was more closely linked to neuroticism, perceived health, and social circumstances such as living alone. These findings highlight that the predictors of loneliness can vary with age, suggesting that interventions need to be tailored to different stages of later adulthood.

The use of ML in these studies provides important implications for designing interventions and policies to reduce loneliness. By recognizing the multiple and interacting factors involved, interventions can be more holistic. For example, psychological support that enhances positive emotions and a sense of personal control could benefit middle-aged and younger older adults, while social and health-related interventions might be more important for those in advanced age who live alone or have poorer health. ML also allows for identifying individuals at higher risk of loneliness by considering their unique combination of factors, which can help target interventions more effectively. While ML offers advantages over traditional statistical methods by handling many variables and complex relationships simultaneously, it also has limitations. These include the need for large and high-quality datasets to avoid overfitting, limited interpretability of some ML models, and difficulties in drawing causal conclusions, especially from cross-sectional data. Future research could benefit from combining ML with theory-driven approaches and longitudinal designs to better understand how loneliness develops and changes over time.

In conclusion, loneliness remains a significant challenge for public health with complex causes. Machine learning studies have advanced our understanding by revealing the relative importance of psychological, social, and demographic factors across different age groups. These insights emphasize the need for integrated and age-appropriate interventions that address emotional well-being, close social connections, and social circumstances. As research progresses, combining ML with other research methods offers great potential for developing more effective strategies to improve social health and reduce loneliness globally.

The Current Article

We build on existing research by employing machine learning (ML) to assess the relative influence of various potential predictors of loneliness in a diverse dataset of over 40,000 individuals aged 16 to 99, living across 237 countries, islands, and territories. These data were collected in collaboration with the BBC and include a wide range of variables spanning multiple levels of analysis, making them well suited to our research aims. Our study extends the work of Altschul et al. (2021) and Ejlskov et al. (2018) in four key ways: (a) by including participants across a broader age range; (b) by analyzing a more culturally diverse sample to generalize findings beyond the United Kingdom; (c) by examining a wider set of potential predictors covering individual, relational, sociocultural, and demographic factors; and (d) by applying an explainable ML approach that quantifies the dependencies and interactions between loneliness and its predictors while accounting for the influence of other variables.

Regarding individual factors, we included both personality traits and well-being indicators. Much psychological research on loneliness predictors has focused on individual differences, particularly the Big Five personality traits (Buecker et al., 2020, 2021). Neuroticism (positively) and extraversion (negatively) have consistently been linked to loneliness, as confirmed by Altschul et al. (2021) and partially by Ejlskov et al. (2018). While health status is often viewed as a consequence of loneliness, it can also predict loneliness by limiting social engagement opportunities (Dahlberg et al., 2022). Subjective health was highlighted by Altschul et al. (2021) as a key correlate of loneliness. We also incorporated mental well-being via self-esteem, which has been shown to predict relationship quality (Murray et al., 2002) and is strongly associated with loneliness (Du et al., 2019).

Relational factors such as both the quantity and quality of social interactions play important roles in loneliness (Victor et al., 2000). The quantity, or relational isolation (Weiss, 1973), is often measured by frequency of social contact, living alone, or time spent alone (Hawkley et al., 2005). Attitudes toward living alone, including whether it is voluntary, and perceptions of loneliness as positive or negative, also influence loneliness levels (Wang et al., 2013).

Though measures of social interaction quantity are commonly included in research, indicators of interaction quality are often omitted or limited to close relationship quality. For example, Altschul et al. (2021) did not consider relationship quality, and Ejlskov et al. (2018) assessed emotional support and negative aspects only within closest relationships. However, loneliness is also influenced by the quality of everyday interactions beyond close ties (Cacioppo & Cacioppo, 2012). Daily experiences of prejudice and discrimination (Lee & Bierman, 2019; Priest et al., 2017) and the presence of trusting neighborhood relationships (high social capital) can respectively increase or protect against loneliness (Matthews et al., 2019). We thus assessed relationship quality via couple satisfaction, daily experiences of prejudice, and neighborhood social capital.

Sociocultural variables such as individualism–collectivism (Hofstede, 1991; Triandis, 1995)—which reflect societal preferences for loose versus tightly knit social networks—may affect loneliness, although findings are mixed. Power distance, describing the extent to which hierarchical differences are accepted or egalitarian relationships are preferred (Hofstede, 1991), has been studied mainly in adolescents (Jefferson, Barreto, Jones, et al., 2023) but might be relevant here. Additionally, relational mobility, or how much social relationships are chosen versus ascribed (Yuki & Schug, 2020), could impact loneliness, though its role remains unexamined.

Demographic factors linked to loneliness include age, gender, education, and socioeconomic or employment status (Buecker et al., 2020), as well as social roles or stigmatized identities like caregiving, parenthood, homelessness, minority sexual orientation, and migrant status. Contrary to common assumptions, loneliness is not highest in older adults; studies with wide age ranges show young people (16–25) report the most loneliness (Barreto et al., 2021; Office for National Statistics, 2018). Gender effects tend to be small overall (Maes et al., 2019), though ML analyses have revealed that men living alone may be particularly vulnerable (Altschul et al., 2021). Stigmatized groups generally experience more loneliness (Barreto et al., 2023), including migrants (Madsen et al., 2016; Victor et al., 2012), individuals with mental illness (Lauder et al., 2004), sexual minorities (Doyle & Molix, 2016), those with low socioeconomic status (Morgan et al., 2019), homeless youth (Kidd, 2007), people with disabilities (Tough et al., 2017), and unemployed individuals (Kleftaras & Vasilou, 2016). We therefore examined a broad range of demographic characteristics from the BBC Loneliness Experiment to capture these differences.

While prior studies have typically focused on a limited set of predictors, such approaches cannot simultaneously examine multilevel factors and their interactions or fully account for multicollinearity. Advanced machine learning techniques overcome these limitations by detecting patterns and interactions directly from the data, reducing subjectivity in variable and interaction selection. Given inconsistent findings in earlier ML studies, the cultural and age diversity in our sample, and the wide range of predictors considered, our study remains exploratory without specific hypotheses regarding the relative importance of loneliness predictors.

The regression analysis also included other potentially influential variables such as the number of hospital beds per capita, total fertility rate, and the number of 4-wheel vehicles per capita as a proxy for wealth or transportation access. Interestingly, these variables did not reach statistical significance in any of the models. The lack of significant effect from the **number of hospital beds per capita** may be explained by several factors. While hospital bed availability is a fundamental aspect of health infrastructure, it does not necessarily translate directly into better maternal health service utilization if other elements such as healthcare quality, staff availability, or geographic accessibility are lacking. Moreover, hospital bed counts do not capture the distribution of these beds within countries—beds may be concentrated in urban centers while rural areas remain underserved. Hence, the simple count of beds per capita may not be a sensitive indicator of effective access to maternal health services. Similarly, the **total fertility rate** was not significantly associated with the outcomes. While fertility rate can influence demand for maternal health services, it may also reflect deeper social, cultural, and economic factors not fully captured in the model. It is possible that fertility rate operates through more complex pathways or interacts with other variables not included here.

Finally, **vehicle ownership**, used as a proxy for economic status and mobility, also failed to show a statistically significant effect. This result might suggest that while transportation availability is important, it may be less directly linked to maternal health service uptake at the national level or may be overshadowed by stronger determinants such as health expenditure and urbanization/density. Additionally, the ownership of 4-wheel vehicles does not necessarily represent equitable access to transportation, especially in rural or poorer populations where other forms of transport might be used. Overall, the regression results demonstrate that both financial investment in health (per-capita health expenditure) and population distribution (density score) are significant drivers of maternal health service coverage. The findings emphasize that increasing health expenditure has a measurable positive effect on key maternal health outcomes, with larger expenditure gains associated with higher utilization rates. At the same time, population density matters because concentrated populations can more efficiently access health services, suggesting that spatial factors should be integral to health planning. The insignificance of hospital bed counts, fertility rates, and vehicle ownership in these models indicates that while these variables are important, their effects may be context-dependent or mediated by other factors. Future research could explore these relationships further, perhaps with more granular data or additional covariates.

Method

We utilized cross-sectional data from the BBC Loneliness Experiment, which was conducted in 2018 with participants aged 16 to 99 years residing in one of 237 countries, islands, and territories (Barreto et al., 2021). This study was a collaboration between the researchers and BBC Radio, with recruitment promoted through Radio 4 and the BBC World Service. Additionally, the study received coverage across various other news media outlets. Participants were self-selected volunteers who accessed the study online. The questionnaire was offered exclusively in English, and the sample was recruited over the course of one month without targeting a predetermined sample size. Our analysis included data from all participants who provided responses to the relevant measures, resulting in a sample size of 40,080 individuals. Of these, approximately 83% were based in the United Kingdom (see Supplemental Table S1 for detailed participant distribution by country). The demographic and other characteristics of the sample are presented in Table 1.

Loneliness was assessed using four items adapted from the UCLA Loneliness Scale (Russell, 1996): “Do you feel a lack of companionship?”, “Do you feel left out?”, “Do you feel isolated from others?”, and “Do you feel in tune with people around you?” (the latter was reverse-coded). Participants rated how often each statement was true for them on a scale from 0 (never) to 5 (always). The scale demonstrated good internal consistency (Cronbach’s $\alpha = .84$).

Although the study did not capture all possible predictors of loneliness (e.g., cognitive biases were not measured), it incorporated a broad range of psychological, relational, sociocultural, and demographic variables. Personality traits were measured using the 10-item scale developed by Gosling et al. (2003), covering the Big Five dimensions: Agreeableness, Openness to Experience, Conscientiousness, Emotional Stability, and Extraversion/Introversion. Each dimension was represented by two items, with acceptable reliability (Pearson correlations ranged from .48 for Openness to Experience to .71 for Emotional Stability).

Well-being was measured through two indicators: psychological well-being, operationalized as personal self-esteem using four items from Rosenberg’s (1965) scale (e.g., “On the whole, I am satisfied with myself”; $\alpha = .91$), and subjective health, assessed with a single item asking participants to rate their general health on a scale from 1 (poor) to 5 (excellent).

Social contact quantity was assessed using multiple indicators. Participants reported whether they lived alone and, if so, for how long (in months). Those not living alone indicated the number of other household members (excluding themselves). All participants answered questions about the frequency of spending time alone (from 1 = never to 4 = always) and the average number of hours spent alone daily. Additional questions explored participants’ choice to live alone (“Did you choose to live alone?”), their enjoyment of alone time (“How much do you enjoy spending time alone?”), and their evaluation of loneliness experiences (“Is the experience of loneliness positive for you?” with options: no, sometimes, yes). The last question was omitted for participants who reported never feeling lonely (see Switsers et al., 2023, for further characterization of those reporting sometimes positive loneliness experiences).

Regarding social contact quality, couple satisfaction was measured using the four-item Couples Satisfaction Index (Funk & Rogge, 2007), administered only to participants currently in a relationship. A sample item includes: “How rewarding is your relationship with your partner?” rated from 1 (not at all) to 7 (completely), with excellent internal reliability ($\alpha = .94$).

Participants’ daily experiences with prejudice and discrimination were measured using the five-item Everyday Discrimination Scale (Sternthal et al., 2011), with items assessing frequency of events such as being treated with less courtesy or respect than others.

Variable	N (%) or Scale	Mean (M)	Standard Deviation (SD)
Loneliness frequency (UCLA mean)	Scale 1–5	2.66	1.13
Gender			
— Male	12,811 (32%)		
— Female	27,269 (68%)		
Age			
— 16–24	2,899 (7.2%)		
— 25–34	5,230 (13.0%)		
— 35–44	6,170 (15.4%)		
— 45–54	9,139 (22.8%)		
— 55–64	9,786 (24.4%)		
— 65–74	5,782 (14.4%)		
— 75+	1,074 (2.7%)		
Employment status			
— Employed	37,757 (94.2%)		
— Unemployed	2,253 (5.6%)		
Years of education			
— <10 years	1,422 (3.5%)		
— 11–14 years	7,197 (17.9%)		

— >15 years	31,461 (78.5%)		
Income			
— Poorly	6,669 (16.6%)		
— Fairly well	19,910 (49.7%)		
— Very well	13,501 (33.7%)		
Subjective socioeconomic status	Scale 1–10	6.12	1.81
Choice to live alone			
— Alone and choose alone	24,338 (60.7%)		
— Alone but choose not to	6,804 (17.0%)		
— Not alone and choose not to	8,938 (22.3%)		
Length living alone (years)	Open number	4.56	11.12
Number of people in household	Open number	1.23	1.40
Marital status			
— Single	11,644 (29.0%)		
— In a relationship but not living together	2,295 (5.7%)		
— Married or cohabiting	16,463 (41.0%)		
— Divorced or separated	7,409 (18.5%)		
— Widowed	2,269 (5.7%)		
Sexual orientation			
— Exclusively heterosexual	30,849 (76.9%)		
— Predominantly heterosexual	5,051 (12.6%)		
— Equal	933 (2.3%)		
— Predominantly homosexual	730 (1.8%)		
— Exclusively homosexual	1,434 (3.5%)		
— Asexual	1,083 (2.7%)		
Dependants			
— Have dependants	28,465 (71.0%)		
— No dependant	11,615 (29.0%)		
Length as carer (years)	Open number	0.09	0.40
Age of the youngest child (months)	Open number	136.41	176.15
Number of children	Open number	1.04	1.33
Couple satisfaction	Scale 4–32	16.56	5.43
Loneliness positive	Scale 1–3 (No=1; Sometimes=2; Yes=3)	1.47	0.56
Hours spent alone	Open number	11.63	7.20

Variable	**N (%)**	**Mean (M)**	**Standard Deviation (SD)**	**Scale/Notes**
Enjoyment time alone		3.39	0.97	Scale 1–5 (1 = Not at all; 5 = Very much)
Personality				
Extraversion		3.71	1.49	Scale 1–7
Agreeableness		4.79	1.25	Scale 1–7
Conscientiousness		5.29	1.21	Scale 1–7
Emotional stability		4.51	1.45	Scale 1–7
Openness to Experience		5.06	1.23	Scale 1–7
Subjective health		3.41	1.02	Scale 1–5
Daily experiences with prejudice		2.36	0.97	Scale 1–7
Self esteem		17.25	3.13	Scale 4–32
Social capital		3.00	0.73	Scale 1–5
Relational mobility		3.97	0.85	Scale 1–7
Migration status				
Residence in same country as birth	27,809 (69.4%)			
Residence in different country as birth	12,271 (30.6%)			
Individualism		83.80	14.92	Hofstede index (1–100)
Power distance		38.43	10.75	Hofstede index (1–100)
Country of residence				

United Kingdom	33,304 (83%)	See Supplemental Materials for details

Analytical Strategy

This study explores the impact of population density on maternal health service coverage, marking the first national-level analysis of this relationship. Our findings indicate a positive association between population density and coverage rates, which carries important implications for demographers, public health researchers, and policymakers. Countries with lower population densities face greater challenges in achieving widespread coverage of key health services. Consequently, these countries may require increased per capita resources to meet international coverage goals, such as the Millennium Development Goals (MDGs). The analyses presented in this article were not preregistered. We applied machine learning (ML) techniques, which focus on identifying generalizable patterns to make accurate predictions from data sets, differing from traditional statistical methods that primarily infer relationships between variables within a sample. ML offers several advantages over conventional statistics (Kyriazos et al., 2021): (a) it does not require assumptions about the distributions of dependent or independent variables, (b) it leverages training data to recognize patterns that are then tested on separate test data, (c) it handles missing data effectively, and (d) it efficiently processes large data sets. To determine the key factors associated with loneliness, we utilized random forest analysis. This method combines an ensemble of regression trees to predict outcomes, effectively modeling complex nonlinear relationships between predictor variables and the target outcome. Each decision tree in the forest splits data based on threshold values of predictors, creating piecewise-constant segments. For instance, if spending "6 hours alone per day" is identified as a critical threshold influencing loneliness, this point serves to segment the data accordingly. Such piecewise approximations enable random forests to capture interactions and nonlinear effects without requiring explicit feature transformations, making them particularly well-suited for high-dimensional regression tasks.

In this study, random forest models were used to examine the relationship between the frequency of loneliness and various predictor variables. Consistent with standard practice (Joseph, 2022), 80% of the data were allocated for training and 20% for testing. During training, hyperparameters were optimized by minimizing the mean squared error (MSE), and predictions on the test set were generated using these optimized parameters. Feature importance was calculated by averaging the reduction in MSE attributed to each predictor across all trees, reflecting the relative contribution of each feature to the model's predictive performance. In a subsequent analysis phase, partial dependence plots (PDPs) were employed to visualize how the most influential predictors affect loneliness while holding all other variables constant. Originally proposed by Friedman (2001), PDPs enable the exploration of relationships between input variables and the model's predictions by marginalizing over the distributions of other features. This approach isolates the effect of individual predictors on the outcome, controlling for confounding influences—a clear advantage over traditional scatterplot regression analyses (for a comprehensive overview of PDP methodology, see Qin et al., 2022). We generated both one-dimensional (1-D) and two-dimensional (2-D) PDPs. The 1-D PDPs illustrate the effect of a single predictor on loneliness frequency, plotting predicted loneliness values across varying levels of that predictor while keeping other variables fixed at their mean. Only predictors explaining at least 5% of the variance in loneliness frequency were included in these plots. The 2-D PDPs depict interactions between pairs of variables, focusing on the interaction between the top-ranked predictor and other key predictors meeting the 5% variance threshold. These visualizations provide insight into how combinations of factors jointly influence loneliness predictions.

Conclusions

Our results indicate that the primary factors associated with loneliness, after controlling for other variables, span sociocultural influences (such as experiences of discrimination), relational aspects (including couple satisfaction and time spent alone), and individual characteristics (notably neuroticism and personal self-esteem). Therefore, effective interventions must take a comprehensive approach that addresses these multiple dimensions. It is essential to tailor strategies to the diverse needs of individuals while also confronting broader issues of marginalization. Focusing solely on individual or relationship-level solutions, without tackling the underlying structural inequalities, is unlikely to reduce loneliness or its negative impacts on health and well-being, and may perpetuate disparities experienced by marginalized populations.

References

1. Alegana VA, Wright JA, Pentrina U, Noor AM, Snow RW, Atkinson PM: Spatial modeling of healthcare utilization for treatment of fever in Namibia. Int J Health Geogr 2012, 11:6.

2. Astell-Burt T, Flowerdew R, Boyle PJ, Dillon JF: Does geographic access to primary healthcare influence the detection of hepatitis C? Soc Sci Med 2011, 72.9:1,472–1,481.

3. Balk D: More Than a Name: Why Is Global Urban Population Mapping a GRUMPy Proposition? In Global Mapping Of Human Settlements: Experiences, Datasets and Prospects. Edited by Gamba P, Herold M. Boca Raton: CRC Press; 2009.

4. Center for International Earth Science

Information Network, Columbia University; International Food Policy Research Institute; The World Bank; and Centro Internacional de Agricultura Tropical: Global Rural–urban Mapping Project (GRUMP), Alpha Version. Palisades: Socioeconomic Data and Applications Center, Columbia University; 2004. Data downloaded November 2010 from http://sedac.ciesin.columbia.edu/gpw.

5. Gabrysch S, Cousens S, Cox J, Campbell OM: The influence of distance and level of care on delivery place in rural Zambia: a study of linked national data in a geographic information system. PLoS Med 2011, 8.1:e1000394. doi:10.1371/journal.pmed.1000394.

6. Galea S, Freudenberg N, Vlahov D: Cities and population health. Soc Sci Med 2005, 60(5):1017–1033.

7. Haggblade S, Longabaugh S, Tschirley DL: Spatial patterns of food staple production and marketing in South East Africa: implications for trade policy and emergency response. Unpublished manuscript: http://econpapers.repec.org/paper/agsmidiwp/54553.htm.

8. Haaland CM, Health MT: Mapping of Population Density. Demography 1974, 11(2):321–336.

9. Institute for Health Metrics and Evaluation: Health Metrics Covariate Database. Seattle, WA: Institute for Health Metrics and Evaluation; 2011.

10. Lu, et al: Outcomes of prolonged mechanic ventilation: a discrimination model based on longitudinal health insurance and death certificate data. BMC Health Serv Res 2012, 12(100). doi:10.1186/1472-6963-12-100.

11. Lovett A, Haynes R, Sünnenbergand G, Gale S: Car travel time and accessibility by bus to general practitioner services: a study using patient registers and GIS. Soc Sci Med 2002, 55.1:97–111.

12. Mitchell R, Popham F: Effect of exposure to natural environment on health inequalities: an observational population study. Lancet 2008, 372(9650):1655–1660.

13. Nations U: World Urbanization Prospects: The 2009 Revision. New York, NY: The United Nations; 2009.

14. Newacheck PW, Hung Yun Y, Park MJ, Brindis M, Irwin CE: Disparities in adolescent health and health care: does socioeconomic status matter? Health Serv Res 2003, 38(5):1235–1252.

15. Rank MR, Hirschl TA: The link between population density and welfare participation. Demography 1993, 30(4):607–622.

16. Stairs RA: The concept of population density: a suggestion. Demography 1977, 14(2):243–244.

17. Tanser F, Gijsbertsen B, Herbst K: Modelling and understanding primary health care accessibility and utilization in rural South Africa: an exploration using a geographical information system. Soc Sci Med 2006, 63.3:691–705.

18. Terschuren C, Mensing M, Mekel OCL: Is telemonitoring an option against shortage of physicians in rural regions? Attitude towards telemedical devices in the North Rhine-Westphalian health survey, Germany. BMC Health Serv Res 2012, 12(95). doi:10.1186/1472-6963-12-95.

19. United Nations: Millennium Development Goals Report 2011; June 2011, ISBN 978-92-1-101244-6, available at: http://www.unhcr.org/refworld/docid/4e42118b2.html [accessed 5 June 2012].

20. United Nations Statistical Division; Source (accessed October 2011): http://unstats.un.org/unsd/demographic/sconcerns/densurb/densurbmethods.htm.

21. Wright JA, Polack C: Understanding variation in measles-mumps-rubella immunization coverage—a population-based study. Eur J Public Health 2005, 16(2):137–142.

22. World Health Organization: National Health Accounts: Country Health Information; Source (accessed April 2011): http://www.who.int/nha/country/en/.

23. Ebener S, Murray CJL, Tandon A, Elvidge C: From wealth to health: modelling the distribution of income per capita at the sub-national level using night-time light imagery. Int J Health Geogr 2004, 4:5.