BMJ Open

A novel measure of socioeconomic status using individual housing data to assess the association of SES with Rheumatoid Arthritis and its mortality: a population-based case-control study

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2014-006469
Article Type:	Research
Date Submitted by the Author:	28-Aug-2014
Complete List of Authors:	Ghawi, Husam; Mayo Clinic, Department of Pediatric and Adolescent Medicine Crowson, Cindy; Mayo Clinic, Rand-Weaver, Jennifer; Mayo Clinic, Community Pediatric and Adolescent Medicine Krusemark, Elizabeth; Mayo Clinic, Community Pediatrics and Adolescent Medicine Gabriel, Sherine; Mayo Clinic, Medicine and Epidemiology Juhn, Young; Mayo Clinic, Community Pediatrics and Adolescent Medicine
Primary Subject Heading:	Epidemiology
Secondary Subject Heading:	Rheumatology
Keywords:	EPIDEMIOLOGY, Rheumatology < INTERNAL MEDICINE, RHEUMATOLOGY

A novel measure of socioeconomic status using individual housing data to assess the association of SES with Rheumatoid Arthritis and its mortality: a population-based case-control study

Running title: Socioeconomic status and Rheumatoid Arthritis Outcomes

Husam Ghawi, MD, Cynthia S. Crowson, MS, Jennifer Rand-Weaver, MS, Elizabeth Krusemark, BA, Sherine E. Gabriel, MD, MSc, Young J. Juhn, MD, MPH

Send correspondence and requests for reprint to:

Young J. Juhn, MD, MPH

Department of Pediatric and Adolescent Medicine/Internal Medicine

Mayo Clinic

200 1st Street SW, Rochester, MN 55905

TEL: 507-284-9528, FAX: 507-284-9744

E-mail: juhn.young@mayo.edu

Abstract word count: 299

Manuscript word count: 3295

Keywords: Socioeconomic status, rheumatoid arthritis, mortality, disparities, and HOUSES

Author Affiliations: Husam Ghawi, MD, Department of Pediatric and Adolescent Medicine, Rochester, MN. Cynthia Crowson, MS, Department of Health Sciences Research and Department of Internal Medicine, Division of Rheumatology, Mayo Clinic, Rochester, MN. Jennifer Rand-Weaver, MS, Department of Pediatric and Adolescent Medicine, Mayo Clinic,

Rochester, MN. Elizabeth Krusemark, BA, Department of Pediatric and Adolescent Medicine, Mayo Clinic, Rochester, MN. Sherine Gabriel, MD, MSc, Department of Health Sciences Research and Department of Internal Medicine, Division of Rheumatology, Mayo Clinic, Rochester, MN. Young Juhn, MD, MPH, Department of Pediatric and Adolescent Medicine/Internal Medicine, Mayo Clinic, Rochester, MN

ABSTRACT

Objectives: To assess whether HOUSES (HOUsing-based index of SocioEconomic Status), a recently developed index of socioeconomic status (SES) measure, is associated with risk of and mortality after rheumatoid arthritis (RA).

Design: We conducted a population-based case-control study, which enrolled population-based RA cases and their controls without RA.

Setting: The study was performed in Olmsted County, Minnesota.

Participants: Study subjects were all residents of Olmsted County, Minnesota, with RA identified using 1987 American College of Rheumatology criteria for RA from January 1, 1988, to December 31, 2007, using the auspices of the Rochester Epidemiology Project (REP). For each patient with RA, one control was randomly selected from all Olmsted County residents of similar age and sex without RA.

Primary and secondary outcome measure: The primary and secondary outcome measures were the risk of RA and post-RA all-cause mortality, respectively. The associations of SES measured by HOUSES with the study outcomes were assessed using logistic regression and Cox models. HOUSES was formulated based on a summed z-score for housing value, square footage, and numbers of bedrooms and bathrooms.

Results: Among 650 RA patients; 604 (93%) were successfully geo-coded to real property data. Of these 604 subjects, 418 (69%) were female; the mean age was 56±15.6 years. HOUSES was associated with risk of developing RA (0.5±3.8 for controls vs. -0.2±3.1 for RA cases, P=0.003) adjusting for age, gender, calendar year of RA index date, smoking status and BMI. The lowest quartile of HOUSES was significantly associated with increased mortality after RA compared to higher quartiles of HOUSES (HR: 1.74; 95%CI: 1.10-2.74; P=0.017) in multivariate analysis.

Conclusions: Lower SES, as measured by HOUSES, is associated with increased risk of RA and mortality after RA. HOUSES may be a useful tool for health disparities research concerning rheumatologic outcomes when conventional SES measures are unavailable.

Abbreviations: HOUSES (HOUsing-based index of SocioEconomic Status); RA, rheumatoid arthritis; REP, Rochester Epidemiology Project; SES, socioeconomic status

ARTICLE SUMMARY

Strengths and limitations of this study

- The main strength is a population-based study design in a study setting with self-contained health care environment and availability of medical records for nearly all Olmsted County, Minnesota, residents.
- Another strength is the HOUSES index as an individual-level SES measure based on an objective measure derived from real property data instead of self-report.
- The main limitation includes an inherent limitation as a retrospective study.
- Another limitation is a modest sample size and predominantly white study subjects, which might limit generalizability of our results in other settings.

1. Article focus: We addressed the following question in this study.

- As a new socioeconomic measure termed HOUSES was recently developed and validated, does the newly developed HOUSES index predict the risk of rheumatoid arthritis (RA) and post-RA mortality?

2. Key messages

- Lower SES, as measured by HOUSES, is associated with increased risk of RA and mortality after RA.

- HOUSES may be a useful tool for health disparities research concerning rheumatologic outcomes when conventional SES measures are unavailable.

INTRODUCTION

Rheumatoid arthritis (RA) causes a significant morbidity and burden to our society affecting about 1.5 million US adults in 2005. The overall prevalence of RA increased from 0.62% in 1995 to 0.72% in 2005.[1] Although the literature has shown that both genetic and environmental factors determine the risk of RA,[2] [3] the cause for RA remains unknown.

Socioeconomic status (SES) influences various health outcomes including RA. Indeed, impacts of SES on RA have been widely reported in the United States and other countries.[4-6] For example, Pederson, et. al. reported individual SES measures, including educational level, which were shown to be inversely associated with risk of developing RA among those with the highest educational level compared with those having the lowest level of education (adjusted odds ratio = 0.43, 95% CI: 0.24 – 0.76, $P = 0.001$).[7] However, unavailability of individual-level SES measures in commonly used data sources for clinical research has been a significant impediment to advancing health disparities research concerning a broad range of health outcomes including RA.[8] Given the rising trends of utilizing large-scale administrative datasets for health outcome or service research, the absence of individual-level SES measures might deter proper interpretation and application of the results. Census (or area)-level SES measures have been used as a proxy measure for individual SES measures.[9-11] However, they often result in misclassification bias,[12 13] and area-level SES measures might not be a mere proxy measure for individual-level SES given the influence of neighborhood environment on health outcomes independent of individual SES.[14 15]

To address the unavailability of individual SES measures in the common data sources, we recently developed and validated a novel individual-level SES measure based on housing features termed HOUSES (i.e., HOUsing-based SES measure) index.[16] HOUSES is a

composite index derived from individual housing features by linking property address information in clinical dataset to enumerated real property data that is available from local government assessors' offices. Previous studies have shown HOUSES index to be significantly associated with health outcomes in children.[16-18] However, it has not been applied to research concerning health outcomes in adults.

The main aims of this study were to assess whether HOUSES is associated with risk of RA and mortality after RA. As a comparison, we assessed the relationship between educational levels as a reference SES measure and RA outcomes. To address these study aims, we conducted a population-based case-control study.

PATIENTS AND METHODS

This study was approved by the Institutional Review Boards of both the Mayo Clinic and the Olmsted County Medical Center.

Study population and settings: Olmsted County, Minnesota, is an excellent setting to conduct a population-based study. The Rochester Epidemiology Project (REP)[19-21] links the medical records from all healthcare providers in the county, including Mayo Clinic, Olmsted Medical Center and others for all Olmsted County residents. All diagnoses are electronically indexed and information from every episode of care is contained within detailed patient-based medical records.[22]

Study design: The study was designed as a population-based case-control study. The primary aim was to determine the association between HOUSES index and risk of RA, which was addressed by comparing HOUSES index between RA cases and matched controls without a

history of RA. The secondary aim was to determine the association between HOUSES index and post-RA mortality between index date of RA and the last follow-up date.

Case ascertainment: Details about the study subjects have been previously reported.[21] Briefly, all eligible incident RA cases (≥ 18 years of age) who had fulfilled the 1987 American College of Rheumatology (ACR) criteria for RA, between January 1, 1988, and December 31, 2007, were retrospectively identified and assembled using the REP medical records linkage system. RA incidence date was defined as the earliest date at which each patient fulfilled ≥ 4 ACR criteria for RA. Excluded cases are those who were non-Olmsted County, Minnesota, residents during the study time, those who denied authorization for using medical records for research per Minnesota statute, and those who could not be geocoded to address information and real property data.

The 1987 ACR criteria for RA ascertainment:*[19]* A patient shall be said to have RA if he/she has satisfied at least 4 out of the following 7 criteria (Criteria 1 through 4 must have been present for at least 6 weeks); 1-Morning stiffness in and around the joints lasting ≥ 1 hour before maximal improvement, 2- Arthritis of three or more joint areas simultaneously with soft tissue swelling or fluid, 3- Arthritis of hand joints with at least 1 area swollen in a wrist, MCP, or PIP joint, 4- Symmetrical arthritis with simultaneous involvement of the same joint areas (as defined in 2), 5- Rheumatoid nodules, 6- Rheumatoid factor present in abnormal amounts, and 7- Radiographic changes typical of RA on hand and wrist radiograph.

Selection of controls: For each patient with RA, one control was randomly selected from all Olmsted County residents of similar age and sex without RA using the REP medical records linkage system (i.e., same source population). Each subject without RA was assigned an index date corresponding to the RA incidence date of the patient with RA.

Mortality after RA: For this secondary outcome, we limited our analysis to only RA cases. All subjects were followed until death or 12/31/2008. All death certificates for Olmsted County residents are obtained every year from the county office. In addition, the Mayo Clinic registration office monitors the notice of death in the local newspapers to update the record. Finally, electronic files of death certificates are obtained from the State of Minnesota Department of Vital and Health Statistics.[20]

Socioeconomic indicators and HOUSES index: Demographic questionnaire was used to ascertain self-reported individual-level education status (i.e., years of school completed). Educational years were categorized into four groups: less than 12 years, 12 years, 13-15 years, and 16 years or longer. Details about HOUSES index have been previously reported.[16] Briefly, HOUSES is a composite index that is derived from individual housing features combined with neighborhood (census tract level) socioeconomic characteristics ascertained by linking address information at the time of interest (index date of RA or mortality) to enumerated real property data available at most local government assessors' offices. A factor (construct) that is composed of the number of bedrooms, number of bathrooms, square footage of housing unit, and estimated value of housing unit was extracted. HOUSES index was formulated by summing a z-score for the number of bedrooms, number of bathrooms, square footage of housing unit, and estimated value of housing unit and used as continuous and categorical variables (quartiles): the higher the HOUSES z-score, the higher the SES. It has been recently developed and tested by conducting studies in both Olmsted County, Minnesota, and Jackson County, Missouri. Results for these studies have been reported in previous publications.[16-18]

Other variables: All factors known to be associated with mortality in RA based on previous work were included in multivariate models for adjustment.[21] These factors were ascertained

through the review of medical charts and REP by trained nurse abstractors blinded to the original study hypothesis and they include the following: smoking status, rheumatoid factor positivity, history of alcoholism, obesity, cardiovascular disease (including hospitalized or silent myocardial infarction, heart failure, revascularization, angina, or a physician diagnosis of coronary artery disease), renal disease, liver disease, cancer, metastases, dementia, severe extra-articular RA manifestations (i.e. pleuritis, pericarditis, Felty's syndrome, RA vasculitis, scleritis, neuropathy, or glomerulonephritis), as well as use of glucocorticoids and other RA therapies (i.e. methotrexate, hydroxychloroquine, other disease modifying antirheumatic drugs and biologics). The Charlson comorbidity index was assessed using an electronic adaptation developed by Deyo.[23 24]

Statistical analysis: We first assessed sociodemographic characteristics of subjects with and without missing HOUSES index. Spearman correlation methods were used to assess the association between HOUSES index and educational levels. For the association between HOUSES and risk of RA, we compared sociodemographic and pertinent clinical characteristics of RA cases and their corresponding controls. HOUSES index was categorized into quartiles using 604 patients with non-missing data [1st quartile (lowest SES) < -2.0215; 2nd quartile -2.0215 – 0.265; 3rd quartile 0.265 – 1.81; and 4th quartile (highest SES) ≥ 1.81)]. Logistic regression models were used to determine the association of HOUSES as both continuous and categorical (quartiles) variables with risk of RA to adjust for pertinent covariates and confounders. Linearity (or non-linearity) for the association between HOUSES and risk of RA was assessed using smoothing splines.

For the association between HOUSES and mortality after RA as a secondary outcome, overall survival rates after RA among subjects with HOUSES in quartiles were estimated using

Cox proportional hazards models to allow adjustment for age, sex and calendar year of RA incidence. Estimated survival rates were directly adjusted for the age, sex, and calendar year of RA incidence of the RA cohort. The adjusted survival rates were plotted against time since RA incidence is in accordance with quartiles of HOUSES index. Associations of SES with time to death were evaluated using Cox proportional hazards regression models and summarized with hazard ratios and 95% confidence intervals (CIs). Similar to the relationship between HOUSES and risk of RA, non-linearity of the associations between HOUSES and post-RA mortality was examined using smoothing splines. Statistical analyses were performed using the SAS software package version 9.3 (SAS Institute, Cary, NC) and R software version 3.0.2 (www.r-project.org). All tests were two-sided and P <0.05 were considered statistically significant.

RESULTS

Subject characteristics: The study cohort consisted of 650 RA cases between January 1, 1988, and December 31, 2007. Compared to RA patients with HOUSES index, RA patients with missing HOUSES index had relatively older age (mean±SD: 60.0±16.5 years vs 55.5±15.6 years, P=0.042) and earlier year of index dates (mean±SD: 1993.9±5.8 vs 1999.0±5.4, P<0.001). Otherwise, cases included and those excluded from the study were similar in regards to gender, educational level, and marital status. Baseline characteristics of cases and their matched controls are summarized in Table 1. Address information was successfully geocoded and HOUSES index was formulated for 604 RA cases (93%) and 564 matched controls (87%). The association between HOUSES and educational levels was measured by Spearman correlation coefficient (r=0.40, P<0.001 for controls and r=0.28, P<0.001 for RA cases).

HOUSES and risk of RA: The HOUSES index was significantly lower among RA cases compared to their matched controls (*P*=0.003; Table1). Age, gender and calendar year of index date were all significantly associated with HOUSES index (data not shown). HOUSES was inversely associated with the odds of RA (adjusted OR: 1.06 per 1 unit *decrease* in HOUSES index; 95% CI: 1.02-1.09; *P*=0.022), adjusting for age, gender, and calendar year of RA incidence/index date. This association persisted after additional adjustment for smoking status and body mass index (adjusted OR: 1.06 per 1 unit *decrease* in HOUSES index; 95% CI: 1.02-1.09; *P*=0.003). Non-linearity for the association between HOUSES and risk of RA was examined using smoothing splines. Non-linearity component was not statistically significant (*P*=0.38). When we examined the association of HOUSES in quartiles with odds of RA, we observed similar findings. Controlling for the same variables as above, HOUSES index in quartiles was inversely associated with the risk of RA (*P*-value for trend =0.02, Table 1). As a reference SES measure, when we examined the association between educational levels and odds of RA, we found that educational levels as a discrete variable was similarly associated with odds of RA, adjusting for age, gender, calendar year of RA, smoking status, and body mass index (*P*=0.002). Controlling for the same variables, educational levels in categorical variable were associated with the risk of RA (Table 1).

HOUSES and mortality after RA: During the follow-up period, 109 of the RA patients with an available HOUSES index had died. The cumulative mortality rates at 20 years following RA were 50% (95%CI: (32 – 69), 44% (25 – 62), 39% (20 – 57), and 40% (22 – 59) for patients in the first (lowest SES), second, third, and fourth quartiles (highest SES) of the HOUSES index, respectively adjusted for age, sex and calendar year. The results are summarized in Figure 1. Patients in the lowest quartile of HOUSES index (≤ -2.1) were associated with an increased risk

of mortality (HR: 1.58; 95% CI: 1.05, 2.36; *P*=0.027) compared to patients with higher HOUSES index values adjusted for age, sex, and calendar year of RA incidence. This association remained significant following additional adjustment for all factors (i.e., comorbid conditions, RA therapies, smoking history, and BMI) associated with post-RA mortality listed in the Method section (HR; 1.74; 95% CI; 1.10-2.74; *P*=0.017). None of the risk factors, such as RA therapy or extra-articular manifestations of RA, accounted for the association of HOUSES index with mortality after RA. While HOUSES index, as a continuous variable, was not linearly associated with the outcome of mortality in patients with RA (HR:1.03, per 1 unit *decrease* of HOUSES index, 95%CI: 0.96 -1.11, *P*=0.37) following adjustment to age, gender, and calendar year of RA incidence/index date, a significant non-linear (inverse) association between HOUSES and post-RA mortality was found after adjusting for the same variables (*P*=0.010). This finding is consistent with the significant findings by quartiles. No significant association between educational level and mortality after RA was found (*P*=0.98, Table 2).

DISCUSSION

Our study results showed that lower SES, as measured by HOUSES, was independently associated with both increased risk of RA and mortality after RA. However, educational levels as a reference SES measure were only associated with the risk of RA.

As we previously reported, HOUSES and educational levels of subjects showed a moderate correlation (rho=0.28-0.40), and this finding is consistent with the literature (r=0.33 for education and income, r=0.40 for occupation and income, and r=0.61 for occupation and education).[8 16] These results not only confirm our previous study findings but also suggest little redundancy in measuring one's SES by other SES measures. Despite the moderate

correlation with educational levels, HOUSES showed a significant inverse association with risk of RA both as a continuous variable and as a categorical variable (quartiles). This association persisted after adjustment for pertinent covariates and confounders. Educational levels were similarly associated with risk of RA. There was a linear inverse relationship between HOUSES and risk of RA based on linearity testing using smoothing spline, suggesting no specific cut-offs or threshold for the impact of HOUSES on risk of RA.

For the association between HOUSES and post-RA mortality during the follow-up period, there was no significant linear relationship, but a significant non-linear (inverse) relationship between HOUSES and post-RA mortality was found. When we examined HOUSES as a categorical variable, we found a significant association of HOUSES with post-RA mortality after adjusting all potential factors associated with outcomes of RA (HR; 1.74 for the lowest quartile of HOUSES; 95% CI; 1.10-2.74; *P*=0.017). However, educational levels were not associated with post-RA mortality. Therefore, despite the moderate concordance between HOUSES and educational levels, HOUSES index is a more suitable SES measure in predicting post-RA mortality than educational levels. Also, HOUSES index was originally developed from a sample of young families with children in Olmsted County, Minnesota, and Jackson County, Missouri. Given the mean age of study subjects, the present study findings suggest that HOUSES is potentially generalizable to adult populations highlighting external validity in terms of age of study subjects.

Previous studies have shown the similar inverse association of socioeconomic status (SES) with risk of developing RA and post-RA health outcomes. For example, Bengtsson, et. al. reported high SES, defined as individuals with high education and less manual work, has shown to be associated with decreased risk of developing RA.[25] In addition, Maiden, et. al. reported

in a prospective study, conducted in West of Scotland to evaluate 200 patients with RA, that 12-year mortality percentage was higher in deprived areas compared to more affluent areas, according to level of male unemployment, overcrowding, car ownership, and distribution of social class within the area's population (61% vs. 36% death percentage).[26]

The mechanisms underlying the association of HOUSES index to risk of and mortality after RA have not been fully elucidated and need to be studied in the future, using the suggested framework for health disparities (genetic model, fundamental, pathway model, and gene-environmental interaction model).[27] To reduce the gap in differential mortality after RA among individuals with different SES, clinicians or health care systems need to consider their preventive and therapeutic strategies in the patients' socioeconomic status. In understanding and addressing heterogeneity of outcomes among patients with RA, in addition to immunogenetic factors, one's SES may provide additional and important information; in this respect, HOUSES index may be helpful in identifying patients with a mismatch between medical needs (greater medical needs) and access to resources (decreased availability of resources). Also, research efforts should be made to determine the extent to which SES modifies the effect of RA therapies on its outcomes.

Our study has several limitations. Although, the sample size was relatively modest in this study, we were able to address the study aims. Another limitation is that the study sample subjects were predominantly white, which might limit generalizability of our results in other settings. Moreover, despite excluding RA cases with missing HOUSES index, due to the retrospective nature of this study, comparison of excluded cases to those with non-missing HOUSES index has shown they have similar baseline characteristics. The study population was predominantly Caucasians, which might limit generalizability to populations with different

ethnic composition. Generalizability of HOUSES to other countries needs to be determined in future studies.

The main strength of this study is the population-based study design. Another important strength is the epidemiologic advantages of the study setting including self-contained health care environment with availability of comprehensive medical records of nearly all Olmsted County, Minnesota, residents under the auspices of REP. Also, the HOUSES index has the advantage of being an individual-level SES measure that consists of available objective measures based on real property data instead of self-reported subjective measure. In addition, housing data are publicly available and are maintained and updated because they are the basis of real property assessment and taxation. Also, given that the median duration of residence in the US was only 4.7 years and 1.9 years for people aged 25-34,[28] HOUSES can capture changes in individual SES over time.[29]

In conclusion, while educational levels were only associated with the risk of RA, lower SES, as measured by HOUSES, is associated with both increased risk of RA and mortality after RA. HOUSES may be a useful tool for epidemiological research concerning rheumatologic disease outcomes among adults when conventional SES measures are unavailable in commonly used datasets.

FUNDING

The study was supported by the Office of Health Disparities Research Award from Mayo Foundation. Research reported in this publication was supported by the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS) of the National Institute of Health under Award Number R01AR046849 and the National Institute on Aging of the National Institutes of Health under Award Number R01AG034676. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

ACKNOWLEDGEMENT

We thank Dr. Chung Wi for his comments and suggestions and Denise Chase for administrative assistance.

COMPETING INTERESTS STATEMENT

The study investigators have nothing to disclose that poses a conflict of interest.

CONTRIBUTORSHIP STATEMENT

Husam Ghawi: participated in the study design, interpretation of the results, drafted the manuscript, and approved the manuscript.

Cynthia Crowson: participated in the study design, performed data analysis, interpreted the results, and reviewed and approved the manuscript.

Jennifer Rand-Weaver: participated in the study design, data collection, interpreted the results, and, reviewed and approved the manuscript.

Elizabeth Krusemark: participated in the study design, data collection, and reviewed and approved the manuscript.

Sherine E. Gabriel: participated in the study design, interpretation of the results, and reviewed and approved the manuscript.

Young J. Juhn: participated in the study design, interpretation of the results, and reviewed and approved the manuscript.

REFERENCES

1. Myasoedova E, Crowson CS, Kremers HM, et al. Is the incidence of rheumatoid arthritis rising?: Results from olmsted county, minnesota, 1955-2007. *Arthritis Rheum* 2010;62:1576-82. doi: 10.1002/art.27425published Online First: Epub Date.

2. MacGregor AJ, Snieder H, Rigby AS, et al. Characterizing the quantitative genetic contribution to rheumatoid arthritis using data from twins. *Arthritis Rheum* 2000;43:30-7. doi: 10.1002/1529-0131(200001)43:1<30::AID-ANR5>3.0.CO;2-Bpublished Online First: Epub Date.

3. Hoovestol RA, Mikuls TR. Environmental exposures and rheumatoid arthritis risk. *Curr Rheumatol Rep* 2011;13:431-9. doi: 10.1007/s11926-011-0203-9published Online First: Epub Date.

4. Massardo L, Pons-Estel BA, Wojdyla D, et al. Early rheumatoid arthritis in latin america: Low socioeconomic status related to high disease activity at baseline. *Arthritis Care Res (Hoboken)* 2012;64:1135-43. doi: 10.1002/acr.21680published Online First: Epub Date.

5. Parks CG, D'Aloisio AA, DeRoo LA, et al. Childhood socioeconomic factors and perinatal characteristics influence development of rheumatoid arthritis in adulthood. *Ann Rheum Dis* 2013;72:350-6. doi: 10.1136/annrheumdis-2011-201083published Online First: Epub Date.

6. Marra CA, Lynd LD, Esdaile JM, et al. The impact of low family income on self-reported health outcomes in patients with rheumatoid arthritis within a publicly funded health-care environment. *Rheumatology (Oxford)* 2004;43:1390-7. doi: 10.1093/rheumatology/keh334published Online First: Epub Date.

7. Pedersen M, Jacobsen S, Klarlund M, et al. Socioeconomic status and risk of rheumatoid arthritis: A danish case-control study. *J Rheumatol* 2006;33:1069-74.

8. Liberatos P, Link BG, Kelsey JL. The measurement of social class in epidemiology. *Epidemiol Rev* 1988;10:87-121.

9. Simon AE, Lukacs SL, Mendola P. Emergency department laboratory evaluations of fever without source in children aged 3 to 36 months. *Pediatrics* 2011;128:e1368-e75.

10. Foster SB, Paul ME, Kozinetz CA, et al. Prevalence of asthma in children and young adults with hiv infection. *Journal of Allergy and Clinical Immunology* 2007;119:750-52. doi: http://dx.doi.org/10.1016/j.jaci.2007.01.002published Online First: Epub Date.

11. Lang DM, Polansky M, Sherman MS. Hospitalizations for asthma in an urban population: 1995-1999. *Annals of Allergy, Asthma and Immunology* 2009 103:128-33(6).

12. Pardo-Crespo MR, Narla NP, Williams AR, et al. Comparison of individual-level versus area-level socioeconomic measures in assessing health outcomes of children in olmsted county, minnesota. *J Epidemiol Community Health* 2013;67:305-10. doi: 10.1136/jech-2012-201742published Online First: Epub Date.

13. Marra CA, Lynd LD, Harvard SS, et al. Agreement between aggregate and individual-level measures of income and education: A comparison across three patient groups. *BMC Health Serv Res* 2011;11:69. doi: 10.1186/1472-6963-11-69published Online First: Epub Date.

14. Juhn YJ, Sauver JS, Katusic S, et al. The influence of neighborhood environment on the incidence of childhood asthma: A multilevel approach. *Soc Sci Med* 2005;60:2453-64.

15. Kim D, Diez Roux AV, Kiefe CI, et al. Do neighborhood socioeconomic deprivation and low social cohesion predict coronary calcification?: The cardia study. *Am. J. Epidemiol.* 2010:kwq098. doi: 10.1093/aje/kwq098published Online First: Epub Date.

16. Juhn YJ, Beebe TJ, Finnie DM, et al. Development and initial testing of a new socioeconomic status measure based on housing data. *Journal of urban health : bulletin of the New York Academy of Medicine* 2011;88:933-44. doi: 10.1007/s11524-011-9572-7published Online First: Epub Date.

17. Butterfield MC, Williams AR, Beebe T, et al. A two-county comparison of the houses index on predicting self-rated health. *J Epidemiol Community Health* 2011;65:254-9. doi: 10.1136/jech.2008.084723published Online First: Epub Date.

18. Johnson MD, Urm SH, Jung JA, et al. Housing data-based socioeconomic index and risk of invasive pneumococcal disease: An exploratory study. *Epidemiol Infect* 2013;141:880-7. doi: 10.1017/S0950268812001252published Online First: Epub Date.

19. Arnett FC, Edworthy SM, Bloch DA, et al. The american rheumatism association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum* 1988;31:315-24.

20. Melton LJ, 3rd. History of the rochester epidemiology project. *Mayo Clin Proc* 1996;71:266-74. doi: 10.1016/S0025-6196(11)63966-9published Online First: Epub Date.

21. Myasoedova E, Crowson CS, Nicola PJ, et al. The influence of rheumatoid arthritis disease characteristics on heart failure. *J Rheumatol* 2011;38:1601-6. doi: 10.3899/jrheum.100979published Online First: Epub Date.

22. Melton L. History of the rochester epidemiology project. *Mayo Clinic Proc* 1996;71:266-74.

23. Sykes A, Macintyre J, Edwards MR, et al. Rhinovirus-induced interferon production is not deficient in well controlled asthma. *Thorax* 2013 doi: 10.1136/thoraxjnl-2012-202909published Online First: Epub Date.

24. Trompette A, Gollwitzer ES, Yadava K, et al. Gut microbiota metabolism of dietary fiber influences allergic airway disease and hematopoiesis. *Nature medicine* 2014;20:159-66. doi: 10.1038/nm.3444

http://www.nature.com/nm/journal/v20/n2/abs/nm.3444.html#supplementary-informationpublished Online First: Epub Date.

25. Bengtsson C, Nordmark B, Klareskog L, et al. Socioeconomic status and the risk of developing rheumatoid arthritis: Results from the swedish eira study. *Ann Rheum Dis* 2005;64:1588-94. doi: 10.1136/ard.2004.031666published Online First: Epub Date.

26. Maiden N, Capell HA, Madhok R, et al. Does social disadvantage contribute to the excess mortality in rheumatoid arthritis patients? *Ann Rheum Dis* 1999;58:525-9.

27. Diez Roux AV. Conceptual approaches to the study of health disparities. *Annu Rev Public Health* 2012;33:41-58. doi: 10.1146/annurev-publhealth-031811-124534published Online First: Epub Date.

28. Schachter JP, Kuenzi JJ. Seasonality of moves and the duration and tenure of residence: 1996: The US Bureau of Census, 1996.

29. Dunn JR. Housing and health inequalities: Review and prospects for research. *Housing Studies* 2000;15:341-66.

Table 1. Characteristics of RA cases and their matched controls and the association of HOUSES index and each variable with risk of RA

	Controls (N=650)	RA cases (N=650)	Adjusted Odds Ratios*(95% CI)	*P*-Value
Age at index date of RA, years, Mean (SD)	55.8 (15.7)	55.8 (15.7)	1.00 (0.93-1.07) per 10 year increase	-
Gender (female):	448 (69%)	448 (69%)	1.00 (0.79-1.27)	-
Year of Index Date, Mean (SD)	1998.6 (5.6)	1998.6 (5.6)	1.00 (0.82-1.21) per 10 year increase	0.97
Race: Caucasian	618 (96%)	604 (94%)	0.59 (0.36-0.99)	0.046
HOUSES Index:				0.002
N	564	604	1.06 (1.02-1.09) per 1 unit decrease	
Mean (SD)	0.5 (3.8)	-0.2 (3.1)		
Median (Q1, Q3)	0.0 (-1.9, 2.4)	-0.5 (-2.1, 1.5)		
HOUSES in quartiles				
Quartile1 (lowest SES)	130 (23)	162 (27)	1.04 (0.75-1.44)	0.02
Quartile2	128 (23)	164 (27)	1.42 (1.02-1.98)	
Quartile3	151 (27)	141 (23)	1.37 (0.98-1.91)	
Quartile4 (Highest SES)	155 (27)	137 (23)	Referent	
Educational Level:				<0.001
Missing	13	22		
<High School	60 (9%)	57 (9%)	1.92 (1.15-3.22)	
High School	192 (30%)	201 (32%)	2.13 (1.41-3.20)	
Technical-School/College	292 (46%)	324 (52%)	2.28 (1.54-3.37)	
Graduate School	93 (15%)	46 (7%)	1 (reference)	
Married in Lifetime	581 (90%)	590 (91%)	1.20 (0.82-1.74)	0.35
Body mass Index:				0.46
Underweight <18.5 kg/m^2	7 (1%)	13 (2%)	1.92 (0.75-4.95)	
Normal ≥18.5 – 24.9 kg/m^2	213 (33%)	208 (32%)	1 (reference)	
Overweight ≥25 - 29.9 kg/m^2	225 (35%)	212 (33%)	0.96 (0.73-1.26)	
Obesity ≥30 kg/m^2	204 (31%)	216 (33%)	1.09 (0.82-1.43)	
Smoking status:				0.015
Current	104 (16%)	120 (18%)	1.36 (1.00-1.84)	
Former	194 (30%)	229 (35%)	1.41 (1.09-1.81)	
Never	352 (54%)	301 (46%)	1 (reference)	
Pertinent Comorbidities:				
Diabetes Mellitus	56 (9%)	61 (9%)	1.10 (0.75-1.62)	0.62
Congestive heart failure	19 (3%)	17 (3%)	0.89 (0.45-1.75)	0.73
Alcohol abuse	42 (6%)	43 (7%)	1.03 (0.66-1.60)	0.91
Hypertension	337 (52%)	412 (63%)	1.82 (1.42-2.34)	<0.001
Hyperlipidemia	348 (54%)	373 (57%)	1.20 (0.95-1.51)	0.14
Charlson comorbidity Index:**				<0.001
0	429 (66%)	259 (40%)	1 (reference)	
1-2	150 (23%)	281 (43%)	3.20 (2.46-4.17)	
≥3	71 (11%)	110 (17%)	2.61 (1.79-3.18)	

*Adjusted for age, sex and calendar year of RA incidence/index date; **Excluding rheumatologic disorders for comparability

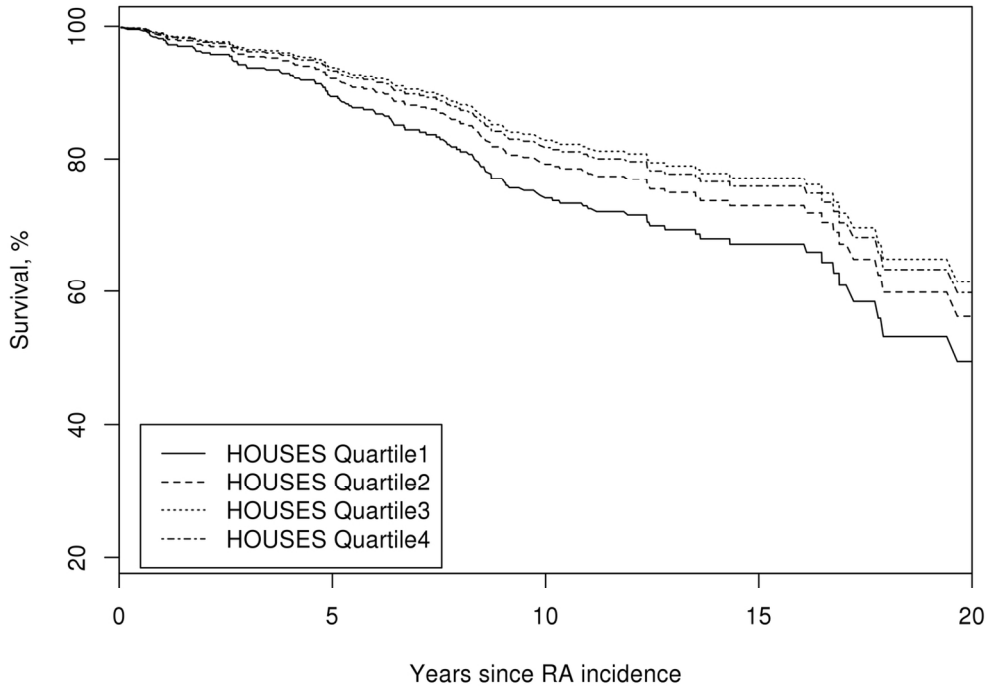
Table 2: Association of HOUSES and other variables with post-RA mortality during the study period

	Adjusted hazard ratios* (95% CI)	*P*-Value
Age (years) at RA incidence date	2.76 (2.35- 3.22) per 10 year increase	<0.001
Gender: male	1.13 (0.78-1.64)	0.50
Year of RA incidence date	0.90 (0.57-1.43)	0.67
Race: Caucasian	0.64 (0.28-1.49)	0.30
HOUSES index as a continuous variable	1.03 (0.96-1.11)	0.37
HOUSES in quartiles		
Quartile 1 (lowest SES)	1.73 (0.91- 3.30)	0.09
Quartile 2	1.21 (0.63-2.35)	
Quartile 3	0.92 (0.44-1.91)	
Quartile 4 (highest SES)	1 (reference)	
Educational Level:		
<High School	1.25 (0.47-3.32)	0.98
High School	1.21 (0.48-3.07)	
Technical-School/College	1.21 (0.48-3.06)	
Graduate School	1 (reference)	
Married in Lifetime (vs. never married)	0.67 (0.34-1.36)	0.27
Body mass Index:		0.020
Underweight <18.5 kg/m^2	2.84 (1.30-6.20)	
Normal ≥18.5 – 24.9 kg/m^2	1 (reference)	
Overweight ≥25 - 29.9 kg/m^2	0.89 (0.59-1.34)	
Obesity ≥30 kg/m^2	0.78 (0.48-1.27)	
Smoking status:		0.001
Current	2.48 (1.51-4.05)	
Former	1.42 (0.94-2162)	
Never	1 (reference)	
Charlson comorbidity Index:**		<0.001
0	1 (reference)	
1-2	1.65 (1.02-2.68)	
≥3	2.89 (1.66-5.02)	

*Adjusted for age, sex and calendar year of RA incidence/index date; **Excluding rheumatologic disorders

FIGURE LEGEND

Figure 1: Estimated survival curves adjusted for age, sex and calendar year of RA incidence according to quartiles of HOUSES index (the numbers in parenthesis are the range of HOUSES index for each group).



Estimated survival curves adjusted for age, sex and calendar year of RA incidence according to quartiles of HOUSES index (the numbers in parenthesis are the range of HOUSES index for each group).
127x90mm (300 x 300 DPI)

STROBE 2007 (v4) Statement—Checklist of items that should be included in reports of *case-control studies*

Section/Topic	Item #	Recommendation	Reported on page #
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	3
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	3
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	5
Objectives	3	State specific objectives, including any prespecified hypotheses	6
Methods			
Study design	4	Present key elements of study design early in the paper	6
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	6-8
Participants	6	*(*a) Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls	7
		(*b*) For matched studies, give matching criteria and the number of controls per case	7
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	8-9
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	*6-9*
Bias	9	Describe any efforts to address potential sources of bias	6-9
Study size	10	Explain how the study size was arrived at	7
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	9
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	9
		(*b*) Describe any methods used to examine subgroups and interactions	9
		(*c*) Explain how missing data were addressed	9
		(*d*) If applicable, explain how matching of cases and controls was addressed	9
		(*e*) Describe any sensitivity analyses	9-10
Results			

Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	10
		(b) Give reasons for non-participation at each stage	10
		(c) Consider use of a flow diagram	N/A
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	10
		(b) Indicate number of participants with missing data for each variable of interest	10
Outcome data	15*	Report numbers in each exposure category, or summary measures of exposure	11-12
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	11-12
		(*b*) Report category boundaries when continuous variables were categorized	11-12
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	N/A
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	12
Discussion			
Key results	18	Summarise key results with reference to study objectives	13
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	15
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	15
Generalisability	21	Discuss the generalisability (external validity) of the study results	15
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	16

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

BMJ Open

A novel measure of socioeconomic status using individual housing data to assess the association of SES with Rheumatoid Arthritis and its mortality: a population-based case-control study

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2014-006469.R1
Article Type:	Research
Date Submitted by the Author:	09-Mar-2015
Complete List of Authors:	Ghawi, Husam; Mayo Clinic, Department of Pediatric and Adolescent Medicine Crowson, Cynthia; Mayo Clinic, Rand-Weaver, Jennifer; Mayo Clinic, Community Pediatric and Adolescent Medicine Krusemark, Elizabeth; Mayo Clinic, Community Pediatrics and Adolescent Medicine Gabriel, Sherine; Mayo Clinic, Medicine and Epidemiology Juhn, Young; Mayo Clinic, Community Pediatrics and Adolescent Medicine
Primary Subject Heading:	Epidemiology
Secondary Subject Heading:	Rheumatology
Keywords:	EPIDEMIOLOGY, Rheumatology < INTERNAL MEDICINE, RHEUMATOLOGY

A novel measure of socioeconomic status using individual housing data to assess the association of SES with Rheumatoid Arthritis and its mortality: a population-based case-control study

Running title: Socioeconomic status and Rheumatoid Arthritis Outcomes

Husam Ghawi, MD, Cynthia S. Crowson, MS, Jennifer Rand-Weaver, MS, Elizabeth Krusemark, Sherine E. Gabriel, MD, MSc, Young J. Juhn, MD, MPH

Send correspondence and requests for reprint to:

Young J. Juhn, MD, MPH

Department of Pediatric and Adolescent Medicine/Internal Medicine

Mayo Clinic

200 1st Street SW, Rochester, MN 55905

TEL: 507-284-9528, FAX: 507-284-9744

E-mail: juhn.young@mayo.edu

Abstract word count: 300

Manuscript word count: 3279

Keywords: Socioeconomic status, rheumatoid arthritis, mortality, disparities, and HOUSES

Author Affiliations: Husam Ghawi, MD, Department of Pediatric and Adolescent Medicine, Mayo Clinic, Rochester, MN. Cynthia Crowson, MS, Department of Health Sciences Research and Department of Internal Medicine, Division of Rheumatology, Mayo Clinic, Rochester, MN. Jennifer Rand-Weaver, MS, Department of Pediatric and Adolescent Medicine, Mayo Clinic,

Rochester, MN. Elizabeth Krusemark, Department of Pediatric and Adolescent Medicine, Mayo Clinic, Rochester, MN. Sherine Gabriel, MD, MSc, Department of Health Sciences Research and Department of Internal Medicine, Division of Rheumatology, Mayo Clinic, Rochester, MN. Young Juhn, MD, MPH, Department of Pediatric and Adolescent Medicine and Internal Medicine, Mayo Clinic, Rochester, MN

ABSTRACT

Objectives: To assess whether HOUSES (HOUsing-based index of SocioEconomic Status (SES)) is associated with risk of and mortality after rheumatoid arthritis (RA).

Design: We conducted a population-based case-control study, which enrolled population-based RA cases and their controls without RA.

Setting: The study was performed in Olmsted County, Minnesota.

Participants: Study subjects were all residents of Olmsted County, Minnesota, with RA identified using 1987 American College of Rheumatology criteria for RA from January 1, 1988, to December 31, 2007, using the auspices of the Rochester Epidemiology Project. For each patient with RA, one control was randomly selected from Olmsted County residents of similar age and gender without RA.

Primary and secondary outcome measure: The disease status was RA cases and their matched controls in relation to HOUSES as an exposure. As a secondary aim, post-RA mortality among only RA cases was an outcome event. The associations of SES measured by HOUSES with the study outcomes were assessed using logistic regression and Cox models. HOUSES, as a composite index, was formulated based on a summed z-score for housing value, square footage, and number of bedrooms and bathrooms.

Results: Of the eligible 604 subjects, 418 (69%) were female; the mean age was 56±15.6 years. Lower SES, as measured by HOUSES, was associated with risk of developing RA (0.5±3.8 for controls vs. -0.2±3.1 for RA cases, P=0.003) adjusting for age, gender, calendar year of RA index date, smoking status, and BMI. The lowest quartile of HOUSES was significantly associated with increased mortality after RA compared to higher of HOUSES (HR: 1.74; 95%CI: 1.10-2.74; P=0.017) in multivariate analysis.

Conclusions: Lower SES, as measured by HOUSES, is associated with increased risk of RA and mortality after RA. HOUSES may be a useful tool for health disparities research concerning rheumatologic outcomes when conventional SES measures are unavailable.

Abbreviations: HOUSES (HOUsing-based index of SocioEconomic Status); RA, rheumatoid arthritis; REP, Rochester Epidemiology Project; SES, socioeconomic status

ARTICLE SUMMARY

Strengths and limitations of this study

- The main strength is a population-based study design in a study setting with a self-contained health care environment and the availability of medical records for nearly all residents of Olmsted County, Minnesota.
- Another strength is the HOUSES index as an individual-level SES measure based on an objective measure derived from real property data rather than self-report.
- The main limitation includes an inherent limitation as a retrospective study.
- Another limitation is a modest sample size and predominantly Caucasian study subjects, which might limit generalizability of our results in other settings.

1. Article focus: We addressed the following question in this study.

- As a new socioeconomic measure termed HOUSES was recently developed and validated, does the newly developed HOUSES index predict the risk of rheumatoid arthritis (RA) and post-RA mortality?

2. Key messages

- Lower SES, as measured by HOUSES, is associated with increased risk of RA and mortality after RA.

- HOUSES may be a useful tool for health disparities research concerning rheumatologic outcomes when conventional SES measures are unavailable.

INTRODUCTION

Rheumatoid arthritis (RA) causes a significant morbidity and burden to our society affecting about 1.5 million US adults in 2005. The overall prevalence of RA increased from 0.62% in 1995 to 0.72% in 2005.[1] Although the literature has shown that both genetic and environmental factors determine the risk of RA,[2] [3] the cause for RA remains unknown.

Socioeconomic status (SES) is correlated with various health outcomes including RA. Indeed, the impact of SES on RA has been widely reported in the United States and other countries.[4-6] For example, Pederson, et. al. reported individual SES measures, including educational level, were shown to be inversely associated with risk of developing RA among those with the highest educational level compared with those having the lowest level of education (adjusted odds ratio = 0.43, 95% CI: 0.24 – 0.76, $P = 0.001$).[7] However, unavailability of individual-level SES measures in commonly used data sources for clinical research has been a significant impediment to advancing health disparities research concerning a broad range of health outcomes including RA.[8] Given the rising trends of utilizing large-scale administrative datasets for health outcome or service research, the absence of individual-level SES measures might deter proper interpretation and application of the results. Census (or area)-level SES measures have been used as a proxy measure for individual SES measures.[9-11] However, they often result in misclassification bias,[12, 13] and area-level SES measures might not be a proxy measure for individual-level SES given the influence of neighborhood environment on health outcomes independent of individual SES.[14, 15]

To address the unavailability of individual SES measures in the common data sources, we recently developed and validated a novel individual-level SES measure based on housing features termed HOUSES (i.e., HOUsing-based SES measure) index.[16] HOUSES is a

composite index derived from individual housing features by linking property address information in a clinical dataset to enumerated real property data that is available from local government assessors' offices. Previous studies have shown HOUSES index to be significantly associated with health outcomes in children.[16-18] However, it has not been applied to research concerning health outcomes in adults.

The main aims of this study were to assess whether HOUSES is associated with risk of RA and mortality after RA. As a comparison, we assessed the relationship between educational levels as a reference SES measure and RA outcomes. To address these study aims, we conducted a population-based case-control study.

PATIENTS AND METHODS

This study was approved by the Institutional Review Boards of both the Mayo Clinic and the Olmsted County Medical Center.

Study population and settings: Olmsted County, Minnesota, is an excellent setting to conduct a population-based study. The Rochester Epidemiology Project (REP)[19-21] links the medical records from all healthcare providers in the county, including Mayo Clinic, Olmsted Medical Center and others for all Olmsted County residents. All diagnoses are electronically indexed and information from every episode of care is contained within detailed patient-based medical records.[22]

Study design: The study was designed as a population-based case-control study. The primary aim was to determine the association between HOUSES index and risk of RA, which was addressed by comparing HOUSES index between RA cases and matched controls without a

history of RA. The secondary aim was to determine the association between HOUSES index and post-RA mortality between index date of RA and the last follow-up date.

Case ascertainment: Details about the study subjects have been previously reported.[21] Briefly, all eligible incident RA cases (≥ 18 years of age) who had fulfilled the 1987 American College of Rheumatology (ACR) criteria for RA, between January 1, 1988, and December 31, 2007, were retrospectively identified and assembled using the REP medical records linkage system *[19]*. RA incidence date was defined as the earliest date at which each patient fulfilled ≥ 4 ACR criteria for RA. Excluded cases are those who were non-Olmsted County, Minnesota, residents during the study time, those who denied authorization for using medical records for research per Minnesota statute, and those who could not be geocoded to address information and real property data.

Selection of controls: For each patient with RA, one control was randomly selected from all Olmsted County residents of similar age and gender without RA using the REP medical records linkage system (i.e., same source population). Each subject without RA was assigned an index date corresponding to the RA incidence date of the patient with RA.

Mortality after RA: For this secondary outcome, we limited our analysis to only RA cases. All subjects were followed until death or 12/31/2008. All death certificates for Olmsted County residents are obtained every year from the county office. In addition, the Mayo Clinic registration office monitors the notice of death in the local newspapers to update the record. Finally, electronic files of death certificates are obtained from the State of Minnesota Department of Vital and Health Statistics.[20]

Socioeconomic indicators and HOUSES index: Demographic questionnaire was used to ascertain self-reported individual-level education status (i.e., years of school completed).

Educational years were categorized into four groups: less than 12 years, 12 years, 13-15 years, and 16 years or longer. Details about HOUSES index have been previously reported.[16] Briefly, HOUSES is a composite index that is derived from individual housing features by linking address information at the time of interest (index date of RA or mortality) to enumerated real property data available at most local government assessors' offices. A factor (construct) that is composed of the number of bedrooms, number of bathrooms, square footage of housing unit, and estimated value of housing unit was extracted. HOUSES index was formulated by summing a z-score for the number of bedrooms, number of bathrooms, square footage of housing unit, and estimated value of housing unit and used as continuous and categorical variables (quartiles): the higher the HOUSES z-score, the higher the SES. It has been recently developed and tested by conducting studies in both Olmsted County, Minnesota, and Jackson County, Missouri. Results for these studies have been reported in previous publications.[16-18]

Other variables: All factors known to be associated with mortality in RA based on previous work were included in multivariate models for adjustment.[21] These factors were ascertained through the review of medical charts and REP by trained nurse abstractors blinded to the original study hypothesis and they include the following: smoking status, rheumatoid factor positivity, history of alcoholism, obesity, cardiovascular disease (including hospitalized or silent myocardial infarction, heart failure, revascularization, angina, or a physician diagnosis of coronary artery disease), renal disease, liver disease, cancer, metastases, dementia, severe extra-articular RA manifestations (i.e. pleuritis, pericarditis, Felty's syndrome, RA vasculitis, scleritis, neuropathy, or glomerulonephritis), as well as use of glucocorticoids and other RA therapies (i.e. methotrexate, hydroxychloroquine, other disease-modifying antirheumatic drugs and biologics).

The Charlson comorbidity index was assessed using an electronic adaptation developed by Deyo.[23, 24]

Statistical analysis: We first assessed sociodemographic characteristics of subjects with and without missing HOUSES index. Spearman correlation methods were used to assess the association between HOUSES index and educational levels. For the association between HOUSES and risk of RA, we compared sociodemographic and pertinent clinical characteristics of RA cases and their corresponding controls. HOUSES index was categorized into quartiles using 604 patients with non-missing data [1st quartile (lowest SES) < -2.0215; 2nd quartile -2.0215 – 0.265; 3rd quartile 0.265 – 1.81; and 4th quartile (highest SES) ≥ 1.81)]. Logistic regression models were used to determine the association of HOUSES as both continuous and categorical (quartiles) variables with risk of RA to adjust for pertinent covariates and confounders. Linearity (or non-linearity) for the association between HOUSES and risk of RA was assessed using smoothing splines.

For the association between HOUSES and mortality after RA as a secondary outcome, overall survival rates after RA among subjects with HOUSES in quartiles were estimated using Cox proportional hazards models to allow adjustment for age, gender and calendar year of RA incidence. Estimated survival rates were directly adjusted for the age, gender, and calendar year of RA incidence of the RA cohort. The adjusted survival rates were plotted against time since RA incidence is in accordance with quartiles of HOUSES index. Association of SES with time to death was evaluated using Cox proportional hazards regression models and summarized with hazard ratios and 95% confidence intervals (CIs). Similar to the relationship between HOUSES and risk of RA, non-linearity of the associations between HOUSES and post-RA mortality was examined using smoothing splines. Statistical analyses were performed using the SAS software

package version 9.3 (SAS Institute, Cary, NC) and R software version 3.0.2 (www.r-project.org). All tests were two-sided and P <0.05 were considered statistically significant.

RESULTS

Subject characteristics: The study cohort consisted of 650 RA cases between January 1, 1988, and December 31, 2007. Compared to RA patients with HOUSES index, RA patients with missing HOUSES index had relatively older age (mean±SD: 60.0±16.5 years vs 55.5±15.6 years, P=0.042) and an earlier year of index date (mean±SD: 1993.9±5.8 vs 1999.0±5.4, P<0.001). Otherwise, cases included and those excluded from the study were similar in regards to gender, educational level, and marital status. Baseline characteristics of cases and their matched controls are summarized in Table 1. Address information was successfully geocoded and HOUSES index was formulated for 604 RA cases (93%) and 564 matched controls (87%). The association between HOUSES and educational levels was measured by Spearman correlation coefficient (r=0.40, P<0.001 for controls and r=0.28, P<0.001 for RA cases).

HOUSES and risk of RA: The HOUSES index was significantly lower among RA cases compared to their matched controls (P=0.003; Table1). Age, gender, and calendar year of index date were all significantly associated with HOUSES index (data not shown). HOUSES was inversely associated with the odds of RA (adjusted OR: 1.06 per 1 unit *decrease* in HOUSES index; 95% CI: 1.02-1.09; P=0.022), adjusting for age, gender, and calendar year of RA incidence/index date. This association persisted after additional adjustment for smoking status and body mass index (adjusted OR: 1.06 per 1 unit *decrease* in HOUSES index; 95% CI: 1.02-1.09; P=0.003) (data now shown). Non-linearity for the association between HOUSES and risk of RA was examined using smoothing splines. Non-linearity component was not statistically

significant (*P*=0.38). When we examined the association of HOUSES in quartiles with odds of RA, we observed similar findings. Controlling for the same variables as above, HOUSES index in quartiles was inversely associated with the risk of RA (*P*-value for trend =0.02, Table 1). As a reference SES measure, when we examined the association between educational levels and odds of RA, we found that educational levels as a discrete variable was similarly associated with odds of RA, adjusting for age, gender, calendar year of RA, smoking status, and body mass index (*P*=0.002) (data not shown). Controlling for the same variables, educational levels in categorical variable were associated with the risk of RA (Table 1).

HOUSES and mortality after RA: During the follow-up period, 109 of the RA patients with an available HOUSES index had died. The cumulative mortality rates at 20 years following RA were 50% (95%CI: (32 – 69), 44% (25 – 62), 39% (20 – 57), and 40% (22 – 59) for patients in the first (lowest SES), second, third, and fourth quartiles (highest SES) of the HOUSES index, respectively adjusted for age, gender, and calendar year. The results are summarized in Figure 1. Overall comparisons of the 4 quartiles of HOUSES summarized in Table 2 only approached statistical significance. However, patients in the lowest quartile of HOUSES index (≤ -2.1) compared to patients with higher HOUSES index values (the other 3 quartiles combined), there was a significant association for the lowest quartile of HOUSES with an increased risk of mortality (HR: 1.58; 95% CI: 1.05, 2.36; *P*=0.027) adjusted for age, gender, and calendar year of RA incidence. This association remained significant following additional adjustment for all factors (i.e., comorbid conditions, RA therapies, smoking history, and BMI) associated with post-RA mortality listed in the Method section (HR; 1.74; 95% CI; 1.10-2.74; *P*=0.017). None of the risk factors, such as RA therapy or extra-articular manifestations of RA, accounted for the association of HOUSES index with mortality after RA. HOUSES index, as a continuous

variable, was not linearly associated with the outcome of mortality in patients with RA (HR:1.03, per 1 unit *decrease* of HOUSES index, 95%CI: 0.96 -1.11, *P*=0.37) following adjustment to age, gender, and calendar year of RA incidence/index date. A significant non-linear (inverse) association between HOUSES and post-RA mortality was found after adjusting for the same variables (*P*=0.010) (data not shown). This finding is consistent with the significant findings by quartiles. No significant association between educational level and mortality after RA was found (*P*=0.98, Table 2).

DISCUSSION

Our study results showed that lower SES, as measured by HOUSES, was independently associated with both increased risk of RA and mortality after RA. However, educational levels as a reference SES measure were only associated with the risk of RA.

As we previously reported, HOUSES and educational levels of subjects showed a moderate correlation (rho=0.28-0.40), and this finding is consistent with the literature (r=0.33 for education and income, r=0.40 for occupation and income, and r=0.61 for occupation and education).[8, 16] These results not only confirm our previous study findings but also suggest that HOUSES might measure a different construct of SES, which is not captured by other SES measures such as education levels. Despite the moderate correlation with educational levels, HOUSES showed a significant inverse association with risk of RA both as a continuous variable and as a categorical variable (quartiles). This association persisted after adjustment for pertinent covariates and confounders. Educational levels were similarly associated with risk of RA. There was a linear inverse relationship between HOUSES and risk of RA based on linearity testing

using smoothing spline, suggesting no specific cut-offs or threshold for the impact of HOUSES on risk of RA.

For the association between HOUSES and post-RA mortality during the follow-up period, there was no significant linear relationship, but a significant non-linear relationship between HOUSES and post-RA mortality suggesting an inverse association was found. Although overall comparisons among HOUSES in quartiles only approached statistical significance, when we compared post-RA mortality between the lowest quartile and the rest, we found a significant association of HOUSES with post-RA mortality after adjusting all potential factors associated with outcomes of RA (HR; 1.74 for the lowest quartile of HOUSES; 95% CI; 1.10-2.74; *P*=0.017). However, educational levels were not associated with post-RA mortality. Therefore, despite the moderate concordance between HOUSES and educational levels, HOUSES index is a more suitable SES measure in predicting post-RA mortality than educational levels. Also, HOUSES index was originally developed from a sample of young families with children in Olmsted County, Minnesota, and Jackson County, Missouri. Given the mean age of study subjects, the present study findings suggest that HOUSES is potentially generalizable to adult populations highlighting external validity in terms of age of study subjects.

Previous studies have shown the similar inverse association of socioeconomic status (SES) with risk of developing RA and post-RA health outcomes. For example, Bengtsson, et. al. reported that high SES, defined as individuals with high education and less manual work, was associated with a decreased risk of developing RA.[25] In addition, Maiden, et. al. reported in a prospective study, conducted in West of Scotland which evaluated 200 patients with RA, that 12-year mortality percentage was higher in deprived areas compared to more affluent areas,

according to level of male unemployment, overcrowding, car ownership, and distribution of social class within the area's population (61% vs. 36% death percentage).[26]

The mechanisms underlying the association of HOUSES index to risk of and mortality after RA have not been fully elucidated and need to be studied in the future, using the suggested framework for health disparities (genetic model, fundamental, pathway model, and gene-environmental interaction model).[27] To reduce the gap in differential mortality after RA among individuals with different SES, clinicians or health care systems need to consider their preventive and therapeutic strategies in the patients' socioeconomic status. In understanding and addressing heterogeneity of outcomes among patients with RA, in addition to immunogenetic factors, one's SES may provide additional and important information; in this respect, HOUSES index may be helpful in clinical research, which identifies patients with a mismatch between medical needs (greater medical needs) and access to resources (decreased availability of resources). Also, research efforts should be made to determine the extent to which SES modifies the effect of RA therapies on its outcomes.

Baldassari et al recently reported home ownership is associated with a lower risk of significant clinical activity of RA.[28] In our study, home ownership was not categorized into the same factor as housing features resulting in HOUSES index in our original study. Because of a high home ownership rate in Olmsted County (~75% or so), it might not necessarily measure the same construct underlying the current HOUSES index (primarily housing size and features).

Our study has several limitations. Although, the sample size was relatively modest in this study, we were able to address the study aims. Another limitation is that the study sample subjects were predominantly Caucasian, which might limit generalizability of our results in other settings. Moreover, despite excluding RA cases with missing HOUSES index, due to the

retrospective nature of this study, comparison of excluded cases to those with non-missing HOUSES index has shown they have similar baseline characteristics.

The main strength of this study is the population-based study design. Another important strength is the epidemiologic advantages of the study setting including a self-contained health care environment with availability of comprehensive medical records of nearly all residents of Olmsted County, Minnesota, under the auspices of REP. Also, the HOUSES index has the advantage of being an individual-level SES measure that consists of available objective measures based on real property data rather than self-reported subjective measures. In addition, housing data are publicly available and are maintained and updated because they are the basis of real property assessment and taxation. Also, given that the median duration of residence in the US was only 4.7 years and 1.9 years for people aged 25-34,[29] HOUSES can capture changes in individual SES over time.[30]

In conclusion, while educational levels were only associated with the risk of RA, lower SES, as measured by HOUSES, is associated with both increased risk of RA and mortality after RA. HOUSES may be a useful tool for epidemiological research concerning rheumatologic disease outcomes among adults when conventional SES measures are unavailable in commonly used datasets.

FUNDING

The study was supported by the Office of Health Disparities Research Award from Mayo Foundation. Research reported in this publication was supported by the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS) of the National Institute of Health under Award Number R01AR046849 and the National Institute on Aging of the National Institutes of Health under Award Number R01AG034676. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

ACKNOWLEDGEMENT

We thank Dr. Chung Wi for his comments and suggestions and Denise Chase for administrative assistance.

COMPETING INTERESTS STATEMENT

The study investigators have nothing to disclose that poses a conflict of interest.

CONTRIBUTORSHIP STATEMENT

Husam Ghawi: participated in the study design, interpretation of the results, drafted the manuscript, and approved the manuscript.

Cynthia Crowson: participated in the study design, performed data analysis, interpreted the results, and reviewed and approved the manuscript.

Jennifer Rand-Weaver: participated in the study design, data collection, interpreted the results, and, reviewed and approved the manuscript.

Elizabeth Krusemark: participated in the study design, data collection, and reviewed and approved the manuscript.

Sherine E. Gabriel: participated in the study design, interpretation of the results, and reviewed and approved the manuscript.

Young J. Juhn: participated in the study design, interpretation of the results, and reviewed and approved the manuscript.

REFERENCES

1. Myasoedova, E., et al., *Is the incidence of rheumatoid arthritis rising?: results from Olmsted County, Minnesota, 1955-2007.* Arthritis and rheumatism, 2010. **62**(6): p. 1576-82.
2. MacGregor, A.J., et al., *Characterizing the quantitative genetic contribution to rheumatoid arthritis using data from twins.* Arthritis and rheumatism, 2000. **43**(1): p. 30-7.
3. Hoovestol, R.A. and T.R. Mikuls, *Environmental exposures and rheumatoid arthritis risk.* Current rheumatology reports, 2011. **13**(5): p. 431-9.
4. Massardo, L., et al., *Early rheumatoid arthritis in Latin America: low socioeconomic status related to high disease activity at baseline.* Arthritis care & research, 2012. **64**(8): p. 1135-43.
5. Parks, C.G., et al., *Childhood socioeconomic factors and perinatal characteristics influence development of rheumatoid arthritis in adulthood.* Annals of the Rheumatic Diseases, 2013. **72**(3): p. 350-6.
6. Marra, C.A., et al., *The impact of low family income on self-reported health outcomes in patients with rheumatoid arthritis within a publicly funded health-care environment.* Rheumatology, 2004. **43**(11): p. 1390-7.
7. Pedersen, M., et al., *Socioeconomic status and risk of rheumatoid arthritis: a Danish case-control study.* The Journal of rheumatology, 2006. **33**(6): p. 1069-74.
8. Liberatos, P., B.G. Link, and J.L. Kelsey, *The measurement of social class in epidemiology.* Epidemiologic reviews, 1988. **10**: p. 87-121.

9. Simon, A.E., S.L. Lukacs, and P. Mendola, *Emergency Department Laboratory Evaluations of Fever Without Source in Children Aged 3 to 36 Months.* Pediatrics, 2011. **128**(6): p. e1368-e1375.

10. Foster, S.B., et al., *Prevalence of asthma in children and young adults with HIV infection.* Journal of Allergy and Clinical Immunology, 2007. **119**(3): p. 750-752.

11. Lang, D.M., M. Polansky, and M.S. Sherman, *Hospitalizations for asthma in an urban population: 1995-1999.* Annals of Allergy, Asthma and Immunology, 2009 **103**(2): p. 128-133(6).

12. Pardo-Crespo, M.R., et al., *Comparison of individual-level versus area-level socioeconomic measures in assessing health outcomes of children in Olmsted County, Minnesota.* Journal of epidemiology and community health, 2013. **67**(4): p. 305-10.

13. Marra, C.A., et al., *Agreement between aggregate and individual-level measures of income and education: a comparison across three patient groups.* BMC health services research, 2011. **11**: p. 69.

14. Juhn, Y.J., et al., *The influence of neighborhood environment on the incidence of childhood asthma: a multilevel approach.* Soc Sci Med, 2005. **60**(11): p. 2453-64.

15. Kim, D., et al., *Do Neighborhood Socioeconomic Deprivation and Low Social Cohesion Predict Coronary Calcification?: The CARDIA Study.* Am. J. Epidemiol., 2010: p. kwq098.

16. Juhn, Y.J., et al., *Development and initial testing of a new socioeconomic status measure based on housing data.* Journal of urban health : bulletin of the New York Academy of Medicine, 2011. **88**(5): p. 933-44.

17. Butterfield, M.C., et al., *A two-county comparison of the HOUSES index on predicting self-rated health.* Journal of epidemiology and community health, 2011. **65**(3): p. 254-9.

18. Johnson, M.D., et al., *Housing data-based socioeconomic index and risk of invasive pneumococcal disease: an exploratory study.* Epidemiology and infection, 2013. **141**(4): p. 880-7.

19. Arnett, F.C., et al., *The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis.* Arthritis and rheumatism, 1988. **31**(3): p. 315-24.

20. Melton, L.J., 3rd, *History of the Rochester Epidemiology Project.* Mayo Clinic proceedings. Mayo Clinic, 1996. **71**(3): p. 266-74.

21. Myasoedova, E., et al., *The influence of rheumatoid arthritis disease characteristics on heart failure.* The Journal of rheumatology, 2011. **38**(8): p. 1601-6.

22. Melton, L., *History of the Rochester Epidemiology Project.* Mayo Clinic Proc, 1996. **71**: p. 266-274.

23. Charlson, M.E., et al., *A new method of classifying prognostic comorbidity in longitudinal studies: development and validation.* Journal of chronic diseases, 1987. **40**(5): p. 373-83.

24. Deyo, R.A., D.C. Cherkin, and M.A. Ciol, *Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases.* Journal of Clinical Epidemiology, 1992. **45**(6): p. 613-9.

25. Bengtsson, C., et al., *Socioeconomic status and the risk of developing rheumatoid arthritis: results from the Swedish EIRA study.* Annals of the Rheumatic Diseases, 2005. **64**(11): p. 1588-94.

26. Maiden, N., et al., *Does social disadvantage contribute to the excess mortality in rheumatoid arthritis patients?* Annals of the Rheumatic Diseases, 1999. **58**(9): p. 525-9.

27. Diez Roux, A.V., *Conceptual approaches to the study of health disparities.* Annual review of public health, 2012. **33**: p. 41-58.

28. Baldassari, A.R., et al., *Socioeconomic Disparities in the Health of African Americans With Rheumatoid Arthritis From the Southeastern United States.* Arthritis Care & Research, 2014. **66**(12): p. 1808-1817.

29. Schachter, J.P. and J.J. Kuenzi, *Seasonality of Moves and The Duration and Tenure of Residence: 1996*. 1996, The US Bureau of Census.

30. Dunn, J.R., *Housing and Health Inequalities: Review and Prospects for Research.* Housing Studies, 2000. **15**(3): p. 341-366.

Table 1. Characteristics of RA cases and their matched controls and the association of HOUSES index and each variable with risk of RA

	Controls (N=650)	RA cases (N=650)	Adjusted Odds Ratios*(95% CI)	*P*-Value
Age at index date of RA, years, Mean (SD)	55.8 (15.7)	55.8 (15.7)	1.00 (0.93-1.07) per 10 year increase	-
Gender (female):	448 (69%)	448 (69%)	1.00 (0.79-1.27)	-
Year of Index Date, Mean (SD)	1998.6 (5.6)	1998.6 (5.6)	1.00 (0.82-1.21) per 10 year increase	0.97
Race: Caucasian	618 (96%)	604 (94%)	0.59 (0.36-0.99)	0.046
HOUSES Index:				0.002
N	564	604	1.06 (1.02-1.09) per 1 unit decrease	
Mean (SD)	0.5 (3.8)	-0.2 (3.1)		
Median (Q1, Q3)	0.0 (-1.9, 2.4)	-0.5 (-2.1, 1.5)		
HOUSES in quartiles				
Quartile1 (lowest SES)	130 (23)	162 (27)	1.04 (0.75-1.44)	0.02
Quartile2	128 (23)	164 (27)	1.42 (1.02-1.98)	
Quartile3	151 (27)	141 (23)	1.37 (0.98-1.91)	
Quartile4 (Highest SES)	155 (27)	137 (23)	Referent	
Educational Level:				<0.001
Missing	13	22		
<High School	60 (9%)	57 (9%)	1.92 (1.15-3.22)	
High School	192 (30%)	201 (32%)	2.13 (1.41-3.20)	
Technical-School/College	292 (46%)	324 (52%)	2.28 (1.54-3.37)	
Graduate School	93 (15%)	46 (7%)	1 (reference)	
Married in Lifetime	581 (90%)	590 (91%)	1.20 (0.82-1.74)	0.35
Body mass Index:				0.46
Underweight <18.5 kg/m^2	7 (1%)	13 (2%)	1.92 (0.75-4.95)	
Normal ≥18.5 – 24.9 kg/m^2	213 (33%)	208 (32%)	1 (reference)	
Overweight ≥25 - 29.9 kg/m^2	225 (35%)	212 (33%)	0.96 (0.73-1.26)	
Obesity ≥30 kg/m^2	204 (31%)	216 (33%)	1.09 (0.82-1.43)	
Smoking status:				0.015
Current	104 (16%)	120 (18%)	1.36 (1.00-1.84)	
Former	194 (30%)	229 (35%)	1.41 (1.09-1.81)	
Never	352 (54%)	301 (46%)	1 (reference)	
Pertinent Comorbidities:				
Diabetes Mellitus	56 (9%)	61 (9%)	1.10 (0.75-1.62)	0.62
Congestive heart failure	19 (3%)	17 (3%)	0.89 (0.45-1.75)	0.73
Alcohol abuse	42 (6%)	43 (7%)	1.03 (0.66-1.60)	0.91
Hypertension	337 (52%)	412 (63%)	1.82 (1.42-2.34)	<0.001
Hyperlipidemia	348 (54%)	373 (57%)	1.20 (0.95-1.51)	0.14
Charlson comorbidity Index:**				<0.001
0	429 (66%)	259 (40%)	1 (reference)	
1-2	150 (23%)	281 (43%)	3.20 (2.46-4.17)	
≥3	71 (11%)	110 (17%)	2.61 (1.79-3.18)	

*Adjusted for age, gender and calendar year of RA incidence/index date; **Excluding rheumatologic disorders for comparability

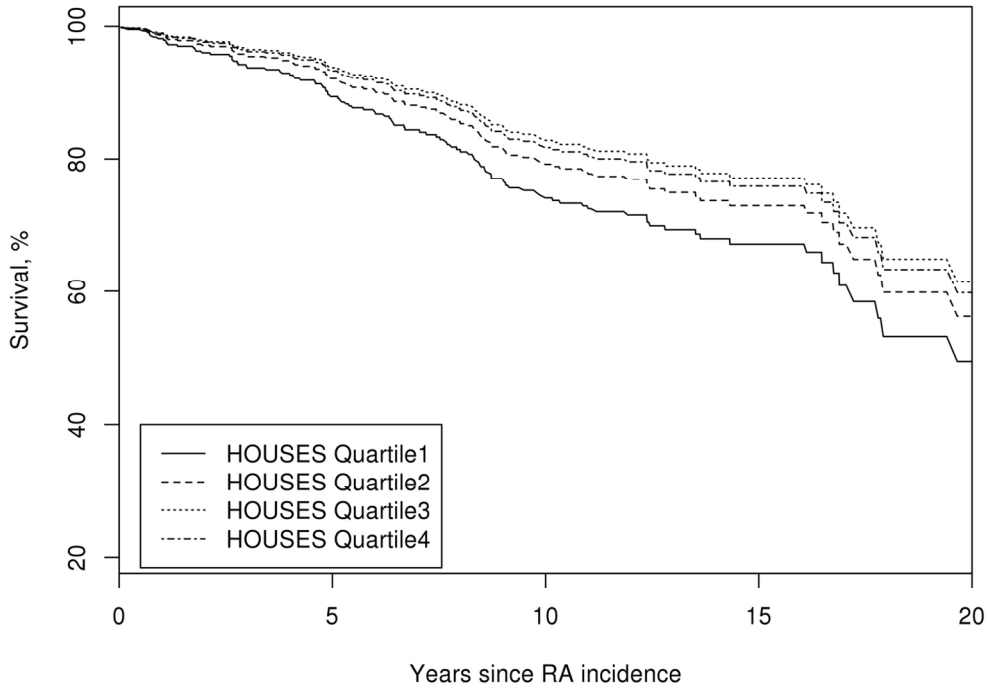
Table 2: Association of HOUSES and other variables with post-RA mortality during the study period

	Adjusted hazard ratios* (95% CI)	*P*-Value
Age (years) at RA incidence date	2.76 (2.35- 3.22) per 10 year increase	<0.001
Gender: male	1.13 (0.78-1.64)	0.50
Year of RA incidence date	0.90 (0.57-1.43)	0.67
Race: Caucasian	0.64 (0.28-1.49)	0.30
HOUSES index as a continuous variable	1.03 (0.96-1.11)	0.37
HOUSES in quartiles		
Quartile 1 (lowest SES)	1.73 (0.91- 3.30)	0.09
Quartile 2	1.21 (0.63-2.35)	
Quartile 3	0.92 (0.44-1.91)	
Quartile 4 (highest SES)	1 (reference)	
Educational Level:		
<High School	1.25 (0.47-3.32)	0.98
High School	1.21 (0.48-3.07)	
Technical-School/College	1.21 (0.48-3.06)	
Graduate School	1 (reference)	
Married in Lifetime (vs. never married)	0.67 (0.34-1.36)	0.27
Body mass Index:		0.020
Underweight <18.5 kg/m^2	2.84 (1.30-6.20)	
Normal ≥18.5 – 24.9 kg/m^2	1 (reference)	
Overweight ≥25 - 29.9 kg/m^2	0.89 (0.59-1.34)	
Obesity ≥30 kg/m^2	0.78 (0.48-1.27)	
Smoking status:		0.001
Current	2.48 (1.51-4.05)	
Former	1.42 (0.94-2162)	
Never	1 (reference)	
Charlson comorbidity Index:**		<0.001
0	1 (reference)	
1-2	1.65 (1.02-2.68)	
≥3	2.89 (1.66-5.02)	

*Adjusted for age, gender and calendar year of RA incidence/index date; **Excluding rheumatologic disorders

FIGURE LEGEND

Figure 1: Estimated survival curves adjusted for age, gender and calendar year of RA incidence according to quartiles of HOUSES index (the numbers in parenthesis are the range of HOUSES index for each group).



Estimated survival curves adjusted for age, sex and calendar year of RA incidence according to quartiles of HOUSES index (the numbers in parenthesis are the range of HOUSES index for each group).
127x90mm (300 x 300 DPI)

STROBE 2007 (v4) Statement—Checklist of items that should be included in reports of *case-control studies*

Section/Topic	Item #	Recommendation	Reported on page #
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	3
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	3
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	5
Objectives	3	State specific objectives, including any prespecified hypotheses	6
Methods			
Study design	4	Present key elements of study design early in the paper	6
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	6-8
Participants	6	*(*a) Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls	7
		(*b*) For matched studies, give matching criteria and the number of controls per case	7
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	8-9
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	*6-9*
Bias	9	Describe any efforts to address potential sources of bias	6-9
Study size	10	Explain how the study size was arrived at	7
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	9
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	9
		(*b*) Describe any methods used to examine subgroups and interactions	9
		(*c*) Explain how missing data were addressed	9
		(*d*) If applicable, explain how matching of cases and controls was addressed	9
		(*e*) Describe any sensitivity analyses	9-10
Results			

Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	10
		(b) Give reasons for non-participation at each stage	10
		(c) Consider use of a flow diagram	N/A
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	10
		(b) Indicate number of participants with missing data for each variable of interest	10
Outcome data	15*	Report numbers in each exposure category, or summary measures of exposure	11-12
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	11-12
		(*b*) Report category boundaries when continuous variables were categorized	11-12
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	N/A
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	12
Discussion			
Key results	18	Summarise key results with reference to study objectives	13
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	15
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	15
Generalisability	21	Discuss the generalisability (external validity) of the study results	15
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	16

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.