BMJ Open is committed to open peer review. As part of this commitment we make the peer review history of every article we publish publicly available.

When an article is published we post the peer reviewers' comments and the authors' responses online. We also post the versions of the paper that were used during peer review. These are the versions that the peer review comments apply to.

The versions of the paper that follow are the versions that were submitted during the peer review process. They are not the versions of record or the final published versions. They should not be cited or distributed as the published version of this manuscript.

BMJ Open is an open access journal and the full, final, typeset and author-corrected version of record of the manuscript is available on our site with no access controls, subscription charges or pay-per-view fees (http://bmjopen.bmj.com).

If you have any questions on BMJ Open's open peer review process please email info.bmjopen@bmj.com

BMJ Open

Education, sex, and risk of stroke: a prospective cohort study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-024070
Article Type:	Research
Date Submitted by the Author:	08-May-2018
Complete List of Authors:	Jackson, Caroline A.; University of Edinburgh, Usher Institute of Population Health Sciences & Informatics Sudlow, Cathie; University of Edinburgh, Centre for Clinical Brain Sciences Mishra, Gita; University of Queensland, Faculty of Medicine, School of Public Health
Keywords:	SOCIAL MEDICINE, Stroke < NEUROLOGY, PUBLIC HEALTH, EPIDEMIOLOGY

Education, sex, and risk of stroke: a prospective cohort study

Caroline A Jackson (PhD)[1,2*], Cathie LM Sudlow (DPhil)[1,3], Gita D Mishra (PhD)[2]

[1]Institute of Population Health Sciences & Informatics

University of Edinburgh

Nine Bioquarter

9 Little France Road

Edinburgh

EH16 4UX

[2]School of Public Health

University of Queensland

Herston

Brisbane

QLD 4006

[3]Centre for Clinical Brain Sciences

University of Edinburgh

Chancellor's Building

49 Little France Crescent

Edinburgh

EH16 4SB

*Corresponding author details:

Email: caroline.jackson@ed.ac.uk

Tel: +44 (0) 131 651 7872

Word count: 3552

ABSTRACT

Objective: to determine whether: the association between highest educational attainment and stroke differed by sex and age; and whether potential mediators of observed associations differ by sex.

Design: prospective cohort study

Setting: population based, New South Wales, Australia

Participants: 253,657 stroke-free participants from the New South Wales 45 and Up Study.

Outcome measures: first-ever stroke events, identified through linkage to hospital and mortality records

Results: During mean follow-up of 4.7 years, 2031 and 1528 strokes occurred among men and women, respectively. Age-standardised stroke rate was inversely associated with education level, with the absolute risk difference between the lowest and highest education group greater among women than men. In relative terms, stroke risk was more pronounced in women than men when comparing low versus high education (age-adjusted HRs: 1.41, 95% CI 1.16, 1.71 and 1.25, 95% CI 1.07, 1.46, respectively), but there was no clear evidence of statistical interaction. This association persisted into older age, but attenuated. Much of the increased stroke risk was explained by modifiable lifestyle factors, in both men and women.

Conclusion: Low education is associated with increased stroke risk in men and women, and may be marginally steeper in women than men. This disadvantage attenuates but persists into older age, particularly for women. Modifiable risk factors account for much of the excess risk from low education level. Public health policy and governmental decision-making should reflect the importance of education for positive health throughout the life-course

Keywords: education; socioeconomic disparities; stroke; cohort study; data linkage

ARTICLE SUMMARY

Strengths and limitations of this study

- A strength is the large study population and large number of outcome events, allowing stratification by sex
- This is one of just a few studies on this topic to have included both men and women and both fatal and non-fatal stroke outcome events
- Breadth of data enabled adjustment of many potential confounders/mediators, allowing identification of candidate mediators of the observed association
- Limitations include the 'healthy cohort' effect which may mean findings are less generalizable to the general population
- We could not identify non-fatal strokes for which participants were not admitted to hospital

FUNDING

This work was supported by a University of Queensland Early Career Researcher Grant awarded to CAJ (grant number 2013002357).

CONFLICTS OF INTEREST

None declared.

INTRODUCTION

Socioeconomic disparities in health are well-recognised, with lower socioeconomic position (SEP) associated with greater risk of mortality and disease, including cerebrovascular disease.[1] Based on cause-specific mortality studies, absolute socioeconomic inequalities in health appear to be greater for men than women.[2-4] There is, however, accumulating evidence that the SEP-cardiovascular disease relationship might actually be steeper in women than men.[5, 6] This has been investigated far less in relation to stroke, with the few, generally small, existing studies having reported conflicting findings.[7-9] Furthermore, few studies have explored whether the potential underlying mechanisms of the SEP-stroke association might differ by sex.

Age differences in the SEP-stroke association have also been rarely studied.[7, 10] There are conflicting views as to whether SEP disparities in health persist into older age. Some believe that the inequality gap narrows with age, whilst others propose that this is an artefact of mortality selection.[11, 12] It is important to not only determine whether lower SEP poses a particular disadvantage to women's health in terms of stroke risk, but also to ascertain whether such an effect, if it exists, persists over the life-course in order to inform approaches to reducing health inequalities.

Valid examination of sex and age differences in SEP inequalities in disease relies on careful measurement of SEP in both sexes. Educational attainment is a particularly useful SEP measure since it is easy to measure, elicits a high response rate, is relevant regardless of age and employment status and performs well regardless of sex.[13] In addition, since education is generally completed in young adulthood and is strongly related to parental characteristics, it also partly captures early life SEP.[13] Importantly, this also minimises potential for reverse causation between SEP and health outcomes (and their risk factors).

Using data from the 45 and Up study, a large Australian prospective cohort study, we aimed to determine whether: the association between highest educational attainment and stroke differed by sex and age; and whether potential mediators of observed associations differ by sex.

METHODS

Study population

We included participants from the Sax Institute's 45 and Up Study, a prospective cohort recruited between 2006 and 2009 from New South Wales (NSW), Australia, general population aged 45 years or over. Recruitment methods are described in detail elsewhere.[14] Briefly, potential participants were randomly sampled from the Department of Human Services (formerly Medicare Australia) enrolment database and mailed a self-administered questionnaire and information leaflet. Participants gave informed consent, including for follow-up via linkage to routinely collected health datasets. For this study, the cohort was linked to the NSW Admitted Patient Data Collection, the Australian Capital Territory (ACT) Admitted Patient Collection and the Australian Bureau of Statistics Death Data, with linkage performed by the Centre for Health Record Linkage.[15] We excluded participants with a previous hospitalised stroke record or self-reported stroke at baseline.

The conduct of the 45 and Up Study was approved by the University of NSW Human Research Ethics Committee. Ethical approval for the present study was obtained from the NSW Population and Health Services Research Ethics Committee, the ACT Health Human Research Ethics Committee and the University of Queensland Institutional Human Research Ethics Committee.

Education

Education was self-reported, with participants provided with the following list and asked to provide their highest attained qualification: college/university degree; high school certificate; school certificate; certificate/diploma/trade or apprenticeship; or no qualifications.

Stroke

We identified incident stroke from hospital admission discharge records and mortality records and defined all strokes based on ICD-10 codes (I60, I61, I63 and I64). Ischaemic stroke was defined using I63 and I64 (since the majority of 'undetermined' strokes coded as I64 will be ischaemic).[16]

Covariates

Definitions of all included covariates are given in supplementary Table 1. We adjusted for demographic and other socioeconomic measures (including marital status and geographical remoteness). Whilst area-based deprivation was available, we did not adjust for it since in univariate analyses it was not associated with stroke risk (perhaps because area is not measured at a small enough level). We also did not adjust for income. Average household income was missing in a substantial proportion of people, particularly among women and older people. Due to extensive missing data and the issue of income being an inadequate measure of income in retired people, we did not include it in any of our models. We also adjusted for: psychological distress (as measured by the Kessler Psychological Distress scale[17]), a range of lifestyle factors (including fruit, vegetable and fish intake); history of disease/stroke risk factors (including physical comorbidity based on a modified Charlson comorbidity index[18] using hospital admission data in the five years prior to recruitment); and, among women, reproductive factors.

Statistical analyses

We performed analyses using Stata version 12. We calculated stroke rates by sex, education level and age-group, age-standardised to the Australian sex-specific standard population. We initially categorised age as follows: 45-59; 60-69; 70-79 and ≥80 years. For the purpose of investigating interactions, we later collapsed the two lowest and two highest groups, to dichotomise age into <70 and ≥70 years.

Overall, the frequency of missing values was less than 5% for most covariates, with the exception of geographical remoteness, which was less complete.

Overall, missing values were widely spread across participants, with 33% of men and 43% of women having missing values for at least one variable. Missing data patterns indicated that data were likely to be missing at random but not missing completely at random. We therefore used multiple imputation by chained equations to impute missing values of included covariates. We imputed data for men and women separately, since we had additional sex-specific covariates for women, performing 35 imputations for men and 45 imputations for women.

We used Cox regression to obtain hazard ratios (HRs) with 95% confidence intervals (CIs) for the association between education level and stroke. We used age as the time axis, following participants from recruitment date and censoring for stroke event date, non-stroke death and end of follow-up (31st Dec 2012). We obtained sex-specific unadjusted HRs, before serially adjusting for groups of confounders and potential mediators including: other sociodemographic factors; psychological distress; lifestyle factors; and disease history. Our primary analysis included the multiple imputed dataset. In sensitivity analyses, we restricted our outcome to ischaemic strokes and also performed a complete case analysis. There was no clear violation of the proportional hazards assumption for education or any covariates in any analyses.

We investigated effect modification by sex and age by testing for multiplicative interaction (through testing for statistical interaction in the models). Since additive interaction is more important for understanding public health, we also investigated supra-additive interaction[19] between education (comparing the lowest versus highest education categories) and stroke, by calculating the relative excess risk of interaction (RERI) and synergy index with accompanying 96% CIs. A relative excess risk of interaction (RERI) of greater/lesser than 0 and a synergy index of greater/lesser than 1 indicate that the combined effects of each exposure is more or less than expected from adding the individual effects.

Results are reported in accordance with the Strengthening the Reporting of Observational Studies (STROBE) statement.[20]

Patient involvement

We did not include patient involvement in the present study.

RESULTS

Among 257,843 participants eligible for inclusion, we excluded 4186 (1.6%) with missing information on education, ultimately including 253,657 participants (Figure 1). Compared to included participants, excluded participants were older and more likely to be: female; from areas of higher deprivation; and less healthy (Supplementary Table 2).

We included 116,810 men and 136,847 women (mean ± SD age in years 63.3 ± 11.0 and 61.4 ± 10.9 respectively). Sex-specific cross-tabulations of characteristics by education level are given in Supplementary Tables 3 and 4. Distribution of education level differed by sex and age. The proportion with a college or university qualification was slightly higher in men than women (25.6% versus 22.2%) and the proportion with no qualifications slightly higher in women than men (12.4% versus 10.7%). The proportion with no qualifications or a school/intermediate certificate was higher among older age groups (and was greater in women than men), whereas the proportion with a college/university degree was higher among younger age groups (with far less disparity between women and men). Lower educational status was associated with poorer lifestyle behaviour and clinical stroke risk factors, with similar patterns observed in men and women. Similarly, most characteristics were associated with stroke risk (Supplementary Table 5).

Absolute and relative stroke risk by education, sex and age

Mean follow-up was 4.70 (± SD 0.98) years. Among men, 2031 strokes (1696 ischaemic) occurred during 545,543 person-years of follow-up and among women 1528 strokes (1225 ischaemic) occurred during 644,362 person-years.

The pattern of absolute stroke risk was not entirely linear with education level. Absolute stroke risk was lowest in the group with highest educational attainment vice versa. The stroke risk of the three middle categories of educational attainment were very similar for both sexes (Figure 2a and Table 1). Among those with no qualifications, stroke risk was slightly greater in women than men. The absolute risk difference between lowest and highest education categories was greater in women than men (Table 1). After stratifying by age group, a similar pattern was observed for those aged under 70 years (Figure 2b and supplementary Figure 1) and, among women, for those aged 70 years or over. Among men, stroke risk in those with no qualifications was actually lower than in those with a trade/apprenticeship or school certificate (Figure 2c and Supplementary Table 6).

In relative terms, low education level was significantly associated with increased stroke risk in men and women, but was slightly more pronounced in women than men (age-adjusted HRs 1.41, 95% CI 1.16 to 1.71 and 1.25, 95% CI 1.07 to 1.46, respectively; Table 2). This pattern persisted when stratifying by age-group. The association between education and stroke among those aged 45-69 was slightly stronger in women than men (HR lowest versus highest education level: 1.73, 95% CI 1.20 to 2.49 and 1.58, 95% CI 1.19 to 2.10, respectively). The association between education and stroke was somewhat weaker in men aged 70 years or over, but persisted among women in this age group, with having no qualifications associated with a 21% increased risk of stroke (HR 1.21, 95% CI 0.96 to 1.52), although this was not statistically significant (Supplementary Table 6). However, having only a school certificate was statistically significantly associated with increased stroke risk in men aged 70 or over (HR 1.21, 95% CI 1.02, 1.44).

Similar associations were found when we restricted our analyses to ischaemic strokes only.

Sex and age interactions with education

Although the effect of education on stroke risk appeared to be slightly greater among women than men there was no evidence of statistical interaction on either the multiplicative or additive scale. Although the RERI and synergy index were less than 0 and 1, respectively, confidence intervals did not exclude the possibility of no interaction (Supplementary Table 7). Similarly, there was no clear evidence of statistical interaction between age and education for men or women. (Supplementary Table 7).

Fully-adjusted education-stroke associations

Adjusting for confounding by marital status and geographical remoteness had little effect on the associations (Table 2). Following additional adjustment for behavioural factors and disease history, which could confound and/or possibly mediate a potentially causal association between education level and stroke risk, the magnitude of effect was attenuated and, for almost all education categories, was no longer statistically significant (fully-adjusted HRs for lowest versus highest education level in men and women: 1.10, 95% CI 0.94 to 1.30 and 1.21, 95% CI 0.97 to 1.51, respectively; Table 2). Whilst the same pattern of association was observed in the complete-case analysis, when compared to the analysis of the imputed dataset, the magnitude of effect was slightly greater for all education categories (Supplementary Table 8).

DISCUSSION

We found an inverse association between education level and stroke risk in both men and women, which was slightly more pronounced in women than men. The association weakened somewhat in older ages for men and, to a lesser extent, for women. Modifiable lifestyle factors explained much of the increased stroke risk in both sexes.

Our study benefits from key strengths. We included a very large stroke-free study population, among which a large number of strokes occurred, providing sufficient power to stratify by sex. The breadth of data collected allowed us to adjust for a wide range of potential confounders/mediators, including those less commonly adjusted for in previous studies, to identify candidate mediators through which education level might affect stroke risk. There are however, some limitations. First, the participation rate in the 45 and Up Study is about 18% and given the 'healthy cohort effect', it is unlikely to be representative of the general NSW population aged 45 and over.[14] However, the cohort is heterogeneous across collected variables. Thus, whilst people with a low education level may well be under-represented in this cohort, given the cohort size and heterogeneity, this is unlikely to have had an effect on internal comparisons of exposure and outcome.[14] Second, it is important to recognise that the meaning of education differs by birth cohort. The consequences of having no high school education may differ among those born in the 1950s versus the 1920s for example. However, we did take this into account by examining the association between education and stroke risk within different age-groups. Third, we were unable to identify non-fatal strokes for which people were not admitted to hospital (estimated to be about 15% of all strokes, but likely higher among older people[21]) or strokes that occurred outside NSW and the ACT (which will have been few in number). Finally, misclassification of stroke diagnosis may have occurred within hospital and mortality records. However, a recent systematic review of the accuracy of hospital and mortality records suggests that the use of appropriately selected, stroke-specific codes (rather than broad cerebrovascular disease codes) yields positive predictive values of greater than 70% in most studies and greater than 90% in some studies.[16]

Our study makes an important contribution to the small body of existing literature on SEP-sex interactions on stroke risk, and indeed circulatory disease in general. A recent systematic review and meta-analysis concluded that there are sex differences in the relationship between education and coronary heart disease, but not stroke.[5] However, the authors included studies of stroke mortality and not just stroke incidence. Results of the latter studies tended to suggest greater education-stroke inequalities in men than women. They also included studies from low and middle-income settings, where the SEP-stroke gradient itself might differ from high-income settings.

Findings from the very few studies that have explored interactions by age and sex on stroke risk are mixed. One reported no difference by sex[22], whereas a second found an association between low education and increased stroke risk among men but not women.[23] Consistent with our findings, two studies found the education-stroke association was weaker[8] or absent[9] in men.

There is biological plausibility for a different SEP-stroke association among women compared to men. Sex differences in the association between household income and atherosclerotic processes in adolescence were observed in one study, which were partially explained by a stronger income-adiposity association in females.[24] Similarly, the associations between SEP and carotid atherosclerosis in mid-age have been reported to be greater among men than women.[25] Furthermore, based on the theory of resource substitution (which states that resources can substitute for one another), education may be more important to women's health compared to men's because they have fewer socioeconomic resources of other kinds (such as income, power, authority and wealth) to draw upon.[26] In our study, income was lower among women than men across educational categories, but there was substantial missing data on income, especially among women, which precluded more detailed analyses. Whilst we were able to adjust for marital status, we were unable to adjust for spousal educational status, which is also thought to play a role in individuals' health.[27]

We also demonstrated that the adverse effect of low education on stroke risk attenuated but did persist in older age groups. The slightly weaker association between education and stroke among elderly men may reflect mortality selection[12], since educational disparities in mortality are greater among men than women. Our findings therefore support the view that socioeconomic inequalities in health persist into older age.[11] Two previous studies on education and stroke risk stratified by age group, but they did not stratify by sex. In one, low education was associated with increased stroke risk in those aged 65-74 years, but a decreased risk in those aged 75 years or over.[7] In contrast, the other study found no association among those aged 65-74 years, but a trend towards reduced risk in those with high versus low education aged 50-64 years.[10] Two studies of women only, with mean age less than 50 years, reported a magnitude of effect comparable to that among women aged less than 70 years in our study.[28, 29]

Our results suggest that lifestyle factors predominantly account for the observed increased stroke risk but that psychological distress and disease history are also important, perhaps more so for women than men. It is difficult to compare our sex-specific findings to similar studies of education and stroke, given that so few studies stratified by sex. A formal mediation study, using a path analysis approach, suggested that behavioural factors account for much of the education-cardiovascular disease gradient[30], which concurs with our findings. There is evidence that, when stratifying by sex, lifestyle behaviours may account for all of the education-cardiovascular association in men, but not women.[31] Our findings support this to some extent in that the slight excess stroke risk seen in women does persist after full adjustment for confounders/potential mediators. In addition, previous studies have posited a possible role for psychosocial factors as partial mediators of the association between education and stroke[32] and cardiovascular disease[33] in women. In line with this, adjustment for psychological distress in our study notably attenuated the effect estimates for women, but not men.

Since this is one of the few studies to date to examine the association between education and stroke risk by sex, within the same study population, further research is needed to confirm or refute the possibility of steeper socioeconomic-stroke disparities in women compared to men. Further investigation is also needed to shed more detailed understanding on the underlying mechanisms of the relationship between socioeconomic status and stroke risk and determine where these differ between men and women. Mediation analyses to unpick these mechanisms should stratify by sex and include non-conventional risk factors (such as psychosocial factors) as well as traditional cardiovascular risk factors. Such studies will inform tailored prevention strategies aimed at reducing health inequalities. Meanwhile, ongoing public health investment is needed to facilitate healthy lifestyle behaviour and reduce uptake of poor health behaviours, particularly in vulnerable socioeconomic groups.

Conclusion

Our study suggests that the education-stroke relationship is present in both men and women, but may be marginally stronger for women than men, with low education a disadvantage to women throughout the life-course. The critical importance of education, of both men and women, for positive health throughout the life-course should be reflected in public health and educational policy and governmental decision-making.

FIGURE LEGENDS

Figure 1 Flow diagram of included participants from the 45 and Up study

Figure 2 Absolute stroke risk by education level among men and women, showing age-standardised stroke incidence (per 1000 person-years) for (a) all ages; (b) age 45-69 years and (c) age 70 years or over (at baseline)

AUTHOR CONTRIBUTIONS

CAJ designed the study, CLMS and GDM contributed to the analytical strategy and interpretation of the results. CAJ drafted the manuscript and CLMS and GDM commented on and approved the final draft.

ACKNOWLEDGEMENTS

This research was completed using data collected through the 45 and Up Study (www.saxinstitute.org.au). The 45 and Up Study is managed by the Sax Institute in collaboration with major partner Cancer Council NSW; and partners: the National Heart Foundation of Australia (NSW Division); NSW Ministry of Health; NSW Government Family & Community Services – Ageing, Carers and the Disability Council NSW; and the Australian Red Cross Blood Service. We thank the many thousands of people participating in the 45 and Up Study.

DATA SHARING STATEMENT

The 45 And Up study data is accessible to any bona fide who has a scientifically sound and feasible research proposal, has ethics approval for the study, and can cover the licence fees and costs incurred by accessing the data through the Secure Unified Research Environment provided by the New South Wales Sax Institute (https://www.saxinstitute.org.au/our-work/45-up-study/for-researchers/). Researchers accessing the data do not retain copies of the dataset and are not able to share these data with anyone else.

REFERENCES

1. Marshall IJ, Wang Y, Crichton S, et al. The effects of socioeconomic status on stroke risk and outcomes *Lancet Neurol.* 2015;**14**:1206-1218. doi: 1210.1016/S1474-4422(1215)00200-00208.

2. Huisman M, Kunst AE, Bopp M, et al. Educational inequalities in cause-specific mortality in middle-aged and older men and women in eight Western European populations *Lancet.* 2005;**365**:493-500.

3. Mackenbach JP, Hu Y, Artnik B, et al. Trends in inequalities in mortality amenable to health care in 17 European countries *Health Aff (Millwood).* 2017;**36**:1110-1118. doi: 1110.1377/hlthaff.2016.1674.

4. Mustard CA, Etches J. Gender differences in socioeconomic inequality in mortality *J Epidemiol Community Health.* 2003;**57**:974-980.

5. Backholer K, Peters SAE, Bots SH, et al. Sex differences in the relationship between socioeconomic status and cardiovascular disease: A systematic review and meta-analysis *J Epidemiol Community Health.* 2017;**71**:550-557. doi: 510.1136/jech-2016-207890. Epub 202016 Dec 207814.

6. Ernstsen L, Bjerkeset O, Krokstad S. Educational inequalities in ischaemic heart disease mortality in 44,000 Norwegian women and men: The influence of psychosocial and behavioural factors. The Hunt study *Scand J Public Health.* 2010;**38**:678-685. doi: 610.1177/1403494810380300. Epub 1403494810382010 Aug 1403494810380310.

7. Avendano M, Kawachi I, Van Lenthe F, et al. Socioeconomic status and stroke incidence in the US elderly: The role of risk factors in the EPESE study *Stroke* 2006;**37**:1368-1373

8. Gillum RF, Mussolino ME. Education, poverty, and stroke incidence in whites and blacks: The NHANS I epidemiologic follow-up study *J clin Epidemiol* 2003;**56**:188-195

9. Veronesi G, Ferrario MM, Chambless LE, et al. Gender differences in the association between education and the incidence of cardiovascular events in Northern Italy *Eur J Public Health.* 2011;**21**:762-767. doi: 710.1093/eurpub/ckq1155. Epub 2010 Nov 1010.

10. Avendano M, Glymour MM. Stroke disparities in older americans: Is wealth a more powerful indicator of risk than income and education? *Stroke* 2008;**39**:1533-1540

11. Benzeval M, Green MJ, Leyland AH. Do social inequalities in health widen or converge with age? Longitudinal evidence from three cohorts in the West of Scotland *BMC Public Health.* 2011;**11:947.**:10.1186/1471-2458-1111-1947.

12. Dupre ME. Educational differences in age-related patterns of disease: Reconsidering the cumulative disadvantage and age-as-leveler hypotheses *J Health Soc Behav.* 2007;**48**:1-15.

13. Galobardes B, Shaw M, Lawlor DA, et al. Indicators of socioeconomic position (part 1) *J Epidemiol Community Health.* 2006;**60**:7-12.

14. Banks E, Redman S, Jorm L, et al. Cohort profile: The 45 and up study *Int J Epidemiol* 2008;**37**:941-947

15. Centre for Health Record Linkage. http://www.cherel.org.au/data-dictionaries#section6

16. Woodfield R, Grant I, Sudlow CL. Accuracy of electronic health record data for identifying stroke cases in large-scale epidemiological studies: A systematic review from the UK Biobank stroke outcomes group *PLoS One.* 2015;**10**:e0140533. doi: 0140510.0141371/journal.pone.0140533. eCollection 0142015.

17. Kessler RC, Andrews G, Colpe LJ, et al. Short screening scales to monitor population prevalences and trends in non-specific psychological distress *Psychol Med.* 2002;**32**:959-976.

18. Romano PS, Roos LL, Jollis JG. Adapting a clinical comorbidity index for use with ICD-9-CM administrative data: Differing perspectives *J Clin Epidemiol.* 1993;**46**:1075-1079; discussion 1081-1090.

19. Andersson T, Alfredsson L, Kallberg H, et al. Calculating measures of biological interaction *Eur J Epidemiol* 2005;**20**:575-579.

20. Vandenbroucke JP, von Elm E, Altman DG, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): Explanation and elaboration *PLoS Med.* 2007;**4**:e297.

21. Anderson CS, Jamrozik KD, Burvill PW, et al. Ascertaining the true incidence of stroke: Experience from the Perth community stroke study, 1989-1990 *Med J Aust.* 1993;**158**:80-84.

22. Honjo K, Tsutsumi A, Kayaba K. Socioeconomic indicators and cardiovascular disease incidence among Japanese community residents: The Jichi Medical School Cohort Study *Int J Behav Med.* 2010;**17**:58-66. doi: 10.1007/s12529-12009-19051-12527. Epub 12009 Jun 12525.

23. Avendano M, Boshuizen HC, Schellevis FG, et al. Disparities in stroke preventive care in general practice did not explain socioeconomic disparities in stroke *J Clin Epidemiol.* 2006;**59**:1285-1294. Epub 2006 Jul 1211.

24. Murasko JE. Male-female differences in the association between socioeconomic status and atherosclerotic risk in adolescents *Soc Sci Med.* 2008;**67**:1889-1897. doi: 1810.1016/j.socscimed.2008.1809.1018. Epub 2008 Oct 1815.

25. Rosvall M, Ostergren PO, Hedblad B, et al. Occupational status, educational level, and the prevalence of carotid atherosclerosis in a general population sample of middle-aged swedish men and women: Results from the Malmo diet and cancer study *Am J Epidemiol.* 2000;**152**:334-346.

26. Ross CE, Masters RK, Hummer RA. Education and the gender gaps in health and mortality *Demography.* 2012;**49**:1157-1183. doi: 1110.1007/s13524-13012-10130-z.

27. Brown DC, Hummer RA, Hayward MD. The importance of spousal education for the self-rated health of married adults in the United States *Popul Res Policy Rev.* 2014;**33**:127-151.

28. Honjo K, Iso H, Inoue M, et al. Education, social roles, and the risk of cardiovascular disease among middle-aged Japanese women: The JPHC study cohort i *Stroke.* 2008;**39**:2886-2890. doi: 2810.1161/STROKEAHA.2108.514067. Epub 512008 Jul 514024.

29. Kuper H, Adami HO, Theorell T, et al. The socioeconomic gradient in the incidence of stroke: A prospective study in middle-aged women in Sweden *Stroke* 2007;**38**:27-33

30. Kershaw KN, Droomers M, Robinson WR, et al. Quantifying the contributions of behavioral and biological risk factors to socioeconomic disparities in coronary heart disease incidence: The Morgen Study *Eur J Epidemiol.* 2013;**28**:807-814. doi: 810.1007/s10654-10013-19847-10652. Epub 12013 Sep 10614.

31. Ricceri F, Sacerdote C, Giraudo MT, et al. The association between educational level and cardiovascular and cerebrovascular diseases within the EPICOR study: New evidence for an old inequality problem *PLoS One.* 2016;**11**:e0164130. doi: 0164110.0161371/journal.pone.0164130. eCollection 0162016.

32. Jackson C, Mishra G. Depression and risk of stroke in mid-aged women: A prospective longitudinal study. *Stroke* 2013;**44**:1555-1560

33. Thurston RC, Kubzansky LD, Kawachi I, et al. Is the association between socioeconomic position and coronary heart disease stronger in women than in men? *Am J Epidemiol.* 2005;**162**:57-65.

Table 1 Sex-specific age-standardised incidence rates (per 1000 person-years) for stroke, by education level

	Men (N = 116,810			Women (N = 136,847)		
Education level	Person-years	Stroke events, N	Incidence rate, per 1000 person-years* (95% CI)	Person-years	Stroke events, N	Incidence rate, per 1000 person-years* (95% CI)
College/university degree	141,248	352	2.72 (2.43 to 3.02)	143,774	160	2.81 (2.27 to 3.36)
Certificate/diploma/trade or apprenticeship	210,375	784	3.31 (3.06 to 3.56)	174,968	333	3.16 (2.76 to 3.56)
High school certificate	53,3741	200	3.29 (2.82 to 3.76)	64,918	163	3.11 (2.60 to 3.61)
School certificate	82,961	398	3.41 (3.05 to 3.76)	181,673	524	3.17 (2.87 to 3.46)
No qualifications	57,496	297	3.47 (3.02 to 3.92)	79,029	348	3.85 (3.42 to 4.29)

*Age-standardised to the Australian sex-specific standard population

Table 2 Serially adjusted hazard ratios (HRs) for the association between education level and stroke, stratified by sex

	Education level				
	College/ university degree	Certificate/diploma/ trade/ apprenticeship	High school certificate	School certificate	No qualifications
	Reference	HR (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)
MEN (2031 strokes among N = 116,810)					
Model 1[*]	1.00	1.20 (1.06, 1.37)	1.23 (1.04, 1.47)	1.31 (1.13, 1.51)	1.25 (1.07, 1.46)
Model 2[†]	1.00	1.20 (1.06, 1.36)	1.22 (1.03, 1.46)	1.29 (1.12, 1.49)	1.23 (1.05, 1.44)
Model 3[‡]	1.00	1.19 (1.05, 1.35)	1.21 (1.02, 1.44)	1.27 (1.10, 1.47)	1.19 (1.02, 1.39)
Model 4[§]	1.00	1.17 (1.03, 1.33)	1.17 (0.98, 1.39)	1.23 (1.06, 1.42)	1.12 (0.96, 1.32)
Model 5[||]	1.00	1.16 (1.02, 1.32)	1.16 (0.97, 1.38)	1.21 (1.05, 1.40)	1.10 (0.94, 1.30)
WOMEN (1528 strokes among N = 136,847)					
Model 1[*]	1.00	1.21 (1.00, 1.46)	1.21 (0.97, 1.51)	1.23 (1.03, 1.47)	1.41 (1.16, 1.71)
Model 2[†]	1.00	1.22 (1.01, 1.47)	1.23 (0.99, 1.53)	1.25 (1.04, 1.50)	1.44 (1.19, 1.75)
Model 3[‡]	1.00	1.21 (1.00, 1.46)	1.22 (0.98, 1.52)	1.23 (1.03, 1.48)	1.37 (1.13, 1.67)
Model 4[§]	1.00	1.18 (1.98, 1.43)	1.17 (0.94, 1.46)	1.17 (0.98, 1.41)	1.26 (1.04, 1.54)
Model 5[||]	1.00	1.16 (0.96, 1.40)	1.17 (0.93, 1.53)	1.15 (0.96, 1.38)	1.22 (1.00, 1.49)
Model 6[#]	1.00	1.12 (0.90, 1.38)	1.09 (0.85, 1.40)	1.15 (0.94, 1.41)	1.21 (0.97, 1.51)

[*]Adjusted for age
[†]Model 1 + marital status and remoteness
[‡]Model 2 + psychological distress
[§]Model 3 + smoking, alcohol intake, body mass index, physical activity, fruit & vegetable intake, fish consumption
[||]Model 4 + history of: hypertension; heart disease; diabetes; and treatment for raised cholesterol, Charlson comorbidity index and family history of stroke/heart disease
[#]In women, Model 4 + oral contraceptive use, hormone replacement therapy use and menopausal status

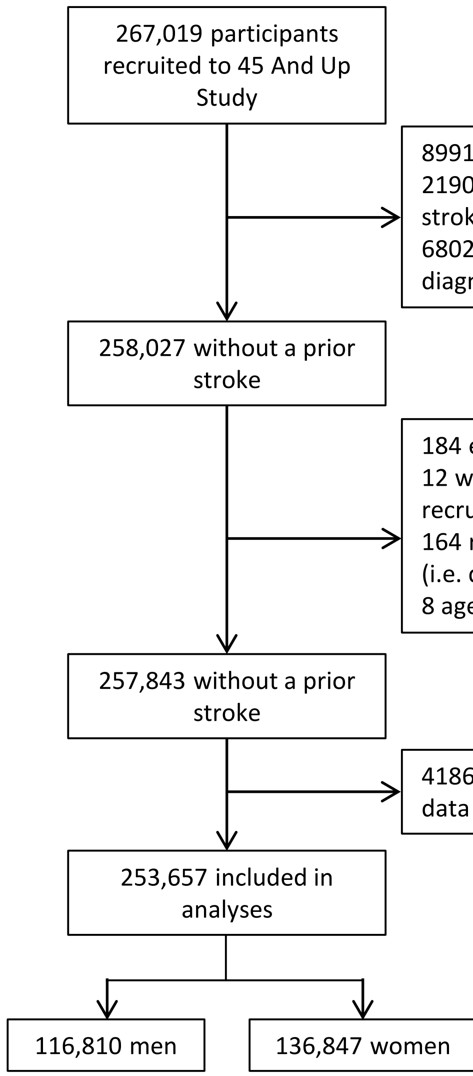
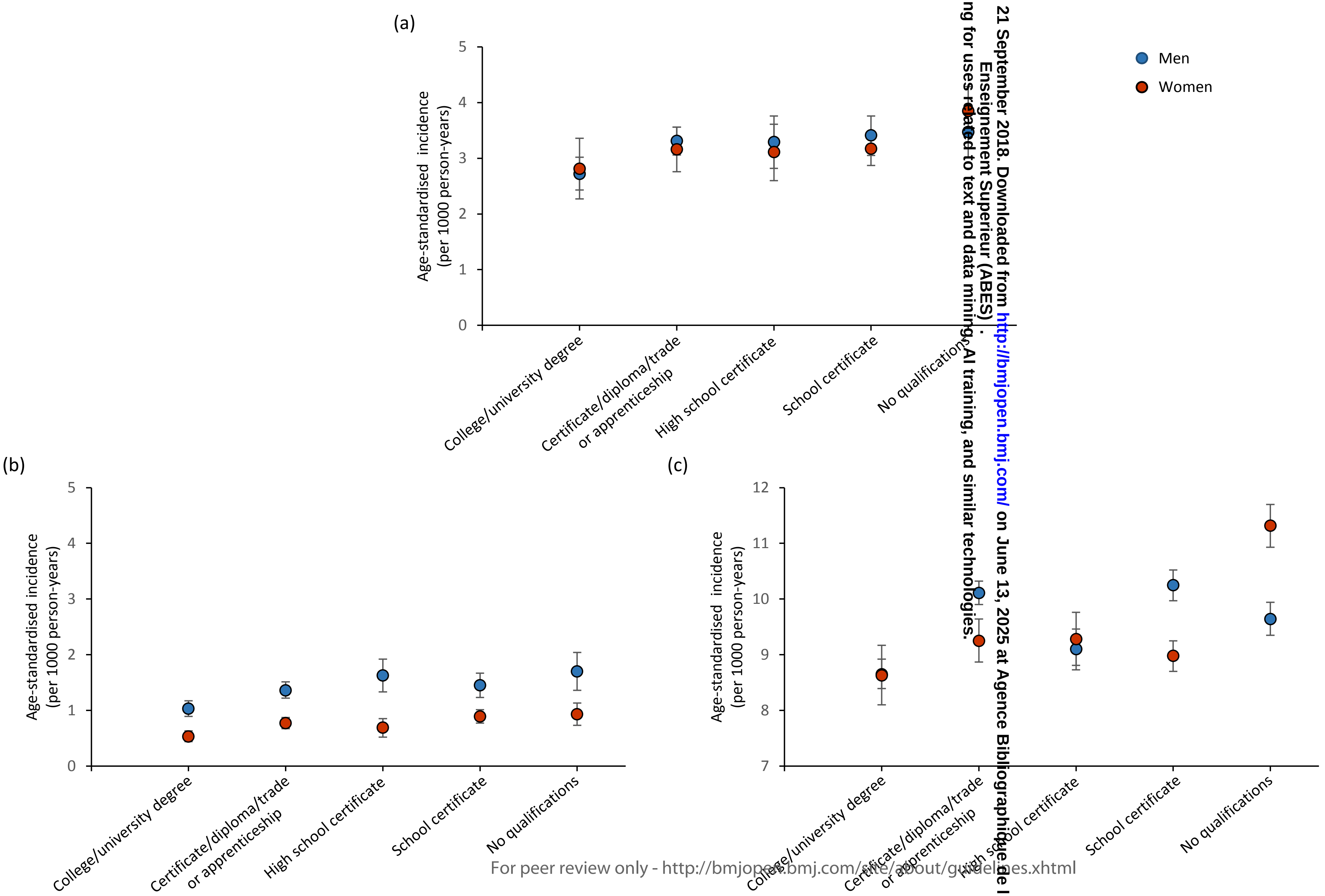

267,019 participants recruited to 45 And Up Study
8991 excluded:
2190 with a prior hospitalised stroke record
6802 who reported a doctor-diagnosed stroke
258,027 without a prior stroke
184 excluded:
12 with date of death prior to recruitment date
164 recruited prior to Jan 2006 (i.e. during pilot phase)
8 aged <45 years at baseline
257,843 without a prior stroke
4186 with missing education data excluded
253,657 included in analyses
116,810 men
136,847 women

Figure 2


(a)
(b)
(c)
Men
Women
Age-standardised incidence (per 1000 person-years)
0
1
2
3
4
5
7
8
9
10
11
12
College/university degree
Certificate/diploma/trade or apprenticeship
High school certificate
School certificate
No qualifications

ONLINE SUPPLEMENT

Education level, sex, and stroke risk: a prospective cohort study

Jackson CA, Sudlow CLM, Mishra GD

Supplementary Table 1 Definitions of covariates included in adjusted cox regression models for the association between psychological distress and incident stroke

Covariates	Definition
Sociodemographic factors	
Marital status	Married/de facto; divorced, separated or widowed; single
Geographical remoteness	Major cities; inner regional; outer regional; and remote or very remote (based on the acessibility/Remoteness Index of Australia)*
Socio-Economic Indexes of Australia (SEIFA) index of relative disadvantage†	Quintiles, from category 1 (least deprived) to 5 (most deprived)
Education	Highest attained qualification: College/university degree; high school certificate; school certificate; certificate/diploma/trade or apprenticeship; or no qualifications
Average annual household income (in Australian Dollars)	Grouped as: ≤19,999; 20,000-29,999; 30,000-39,999; 40,000-49,000; 50,000-69,999; and ≥70,000
Lifestyle factors	
BMI	Body mass index (kg/m^2) created based on self-reported weight and height and included as a continuous variable
Smoking status	From self-reported smoking history, categorical variable created: never; former; current
Alcohol intake	Weekly units of alcohol, defined as: moderate (<21 units/week for men and <14 units/week for women); none; hazardous (21-50 for men and 14-35 for women); and harmful (>50 for men and >35 for women) Participants were asked about weekly number of alcoholic drinks consumed (but were not asked about intake of specific drinks). When calculating weekly alcohol intake we assumed that one alcoholic drink was equivalent to two units of alcohol
Physical activity	Questions on physical activity (from the Active Australia Survey) used to construct a physical activity variable, categorised as: sufficient; insufficient; and sedentary as per the Active Australia Survey recommendations‡
Daily fruit and vegetable intake	Questions on average daily servings of cooked and raw vegetables and fruit or glasses of fruit juice used to create a binary variable of ≥5 or <5 portions of fruit or vegetables
Weekly fish intake	Based on self-reported weekly intake of fish or seafood, and defined as: ≥twice/week; once/week; and never
Physiological and family history variables	
History of hypertension	Self-reported doctor-diagnosed hypertension or self-reported or treatment for high blood pressure in the past month
History of heart disease	Self-reported doctor-diagnosed 'heart disease' or treatment for 'either heart attack or angina' or 'other heart disease' in the past month
Diabetes	Self-reported doctor diagnosed diabetes
Treated cholesterol	self-report of treatment for 'high blood cholesterol' in the past month
Family history of stroke or heart disease	Dichotomous variable based on reported parent and sibling stroke and heart disease history

Charlson comorbidity index	Based on the modified Charlson Index and using the ICD10-Australian Modification conversion code[§] we identified and weighted 17 chronic disease conditions (excluding stroke from the cerebrovascular disease codes), to give a total comorbidity score, which we then categorised as 0, 1, 2, and ≥3
Menopausal status	Using responses to questions on menopause, hysterectomy and oophorectomy, along with age at which they occurred, a menopausal status variables was created and categorised as: pre-menopause; post-menopause; hysterectomy only; bilateral oophorectomy post-menopause; bilateral oophorectomy (surgical menopause); and 'missing' where information was insufficiently complete or conflicting
Current OCP use	A dichotomous 'current OCP use' variable was created based on questions on ever having used OCP and age at last use
Current HRT use	A dichotomous 'current HRT use' variable was created based on questions on ever having used HRT and age at last use

[*]Australian Department of Health and Aged Care (DHAC). Measuring remoteness: Accessibility/remoteness index of australia (ARIA). Occasional papers: New Series Number 14. Canberra: DHAC; 2001
[†]Australian Bureau of Statistics Socio-economic indexes for areas. http://www.abs.gov.au/websitedbs/censushome.nsf/home/seifa. Accessed 15th February 2017
[‡]Australian Institute of Health and Welfare (AIHW). The Active Australia Survey: a guide and manual for implementation, analysis and reporting. Canberra: AIHW; 2003
[§]Sundararajan V, Henderson T, Perry C, et al. New icd-10 version of the charlson comorbidity index predicted in-hospital mortality. *J Clin Epidemiol.* 2004;57:1288-1294.
BMI = body mass index; HRT = hormone replacement therapy; OCP = oral contraceptive pill

Supplementary Table 2 Comparison of baseline characteristics of included versus excluded participants (i.e. those with versus without education level)

Characteristic	Included (N = 253,657) n (%)	Excluded (N = 4186) n (%)
Age, years (mean ± SD)	62.3 ± 11.0	69.1 ± 12.1
Categorical age, years		
45-59	120,819 (47.6)	1101 (26.3)
60-69	70,865 (27.9)	1035 (24.7)
70-79	38,584 (15.2)	1013 (24.2)
80+	23,389 (9.2)	1037 (24.8)
Female	136,847 (54.0)	2187 (52.3)
SEIFA index of relative disadvantage		
1 (least deprived)	50,472 (19.9)	591 (14.1)
2	49,559 (19.5)	731 (17.5)
3	50,941 (20.1)	904 (21.6)
4	51,635 (20.4)	896 (21.4)
5 (most deprived)	50,846 (20.1)	1062 (25.4)
Missing	*204 (0.1)*	*2 (0.1)*
Geographical remoteness		
Major cities	125,896 (49.6)	2100 (50.2)
Inner regional	73,902 (29.1)	1234 (29.5)
Outer regional	17,882 (7.1)	352 (8.4)
Remote/very remote	602 (0.2)	20 (0.5)
Missing	*35,375 (14.0)*	*480 (11.5)*
Marital status		
Married/de facto	190,834 (75.7)	2718 (64.9)
Divorced/separated/Widowed	46,981 (18.6)	1095 (26.2)
Single	14,394 (5.7)	283 (6.8)
Missing	*1448 (0.6)*	*90 (2.2)*
Psychological distress		
Low	172, 737 (68.1)	2054 (49.1)
Medium	36,719 (14.5)	455 (10.9)
High/very high	16,717 (6.6)	273 (6.5)
Missing	27,484 (10.8)	1404 (33.5)
BMI, kg/m^2 (mean ± SD)	26.9 ± 4.9	26.8 ± 5.1
Missing	*18, 373 (7.2)*	*861 (20.6)*
Smoking status		
Never	145,138 (57.2)	2400 (57.3)
Former	88,910 (35.1)	1410 (33.7)
Current	18,325 (7.2)	323 (7.7)
Missing	*1284 (0.5)*	*53 (1.3)*
Alcohol intake		
Moderate	92,813 (36.6)	1275 (30.5)
None/rarely	81,049 (32.0)	1569 (37.5)
Hazardous	58,437 (23.0)	686 (16.4)
Harmful	16,506 (6.5)	193 (4.6)
Missing	*4852 (1.9)*	*463 (11.1)*
Physical activity		
Sufficiently active	193,391 (76.2)	2432 (58.1)
Insufficiently active	40,750 (16.1)	829 (19.8)
Sedentary	10,429 (4.1)	288 (6.9)
Missing	*9087 (3.6)*	*637 (15.2)*

Fruit and vegetable intake		
≥ 5 portions/week	*169,397 (66.8)*	*2571 (61.4)*
< 5 portions/week	*80,119 (31.6)*	*1435 (34.3)*
Missing	*4141 (1.6)*	*180 (4.3)*
Fish intake		
≥ twice/week	120,327 (47.4)	1931 (46.1)
Once/week	100,951 (39.8)	1463 (35.0)
Never	19,646 (7.8)	333 (8.0)
Missing	*12,733 (5.0)*	*459 (11.0)*
History of hypertension	88,022 (34.7)	1503 (35.9)
History of heart disease	30,514 (12.0)	688 (16.4)
History of diabetes mellitus	21,403 (8.4)	504 (12.0)
Family history of stroke/heart disease	145,306 (57.3)	2123 (50.7)
Missing	*32 (0.01)*	*5 (0.1)*
Treated for high cholesterol	37,238 (14.7)	613 (14.6)
Charlson Comoridity Index		
0	224,857 (88.7)	3338 (79.7)
1	15,918 (6.3)	440 (10.5)
2	6321 (2.5)	194 (4.6)
≥3	6561 (2.6)	214 (5.1)
WOMEN ONLY		
Menopausal status		
Pre-menopausal	24,024 (17.6)	201 (9.2)
Post-menopausal	67,524 (49.3)	1078 (49.3)
Hysterectomy only	26,231 (19.2)	499 (22.8)
Bilateral oophorectomy post-menopause	2274 (1.7)	25. (1.1)
Bilateral oophorectomy (surgical menopause)	9178 (6.7)	151 (6.9)
Missing	*7616 (5.6)*	*233 (10.7)*
Current HRT use	13,911 (10.2)	144 (6.6)
Missing	*548 (0.4)*	*17 (0.8)*
Current OCP use	340 (0.3)	3 (0.1)
Missing	*13,794 (10.1)*	*486 (22.2)*

BMI = body mass index; HRT = hormone replacement therapy; OCP = oral contraceptive; SD = standard deviation

Supplementary Table 3 Baseline characteristics, by education level, among men

	Education level				
Characteristic	College/ university degree N = 29,956 n (%)	Certificate/diploma/ trade/ apprenticeship N = 44,990 n (%)	High school certificate N = 11,421 n (%)	School certificate N = 17,907 n (%)	No qualifications N = 12,536 n (%)
Age, years (mean ± SD)	60.8 ± 10.4	63.3 ± 11.0	62.7± 11.2	65.1 ± 10.9	67.2 ± 10.8
Categorical age, years					
45-59	16,077 (53.7)	19,413 (43.2)	5327 (46.6)	6048 (33.8)	3508 (28.0)
60-69	8104 (27.1)	13,086 (29.1)	3142 (27.5)	5877 (32.8)	3850 (30.7)
70-79	3614 (12.1)	7988 (17.8)	1787 (15.7)	3724 (20.8)	3296 (26.3)
80+	2161 (7.2)	4503 (10.0)	1165 (10.2)	2258 (12.6)	1882 (15.0)
SEIFA index of relative disadvantage					
1 (least deprived)	10,289 (34.4)	7911 (17.6)	2499 (21.9)	2223 (12.4)	940 (7.5)
2	6407 (21.4)	9049 (20.1)	2105 (18.4)	3092 (17.3)	2003 (16.0)
3	5350 (17.9)	9280 (20.6)	2243 (19.6)	3917 (21.9)	2617 (20.9)
4	4462 14.9)	9593 (21.3)	2255 (19.7)	4141 (23.1)	3019 (24.1)
5 (most deprived)	3398 (11.3)	9122 (20.3)	2310 (20.2)	4539 (25.4)	3954 (31.5)
Missing	*50 (0.2)*	*35 (0.1)*	*9 (0.1)*	*5 (0.03)*	*3 (0.02)*
Geographical remoteness					
Major cities	17,651 (58.9)	22,332 (49.6)	5809 (50.9)	7733 (43.2)	5241 (41.8)
Inner regional	6700 (22.4)	13,611 (30.3)	2970 (26.0)	5751 (32.1)	4110 (32.8)
Outer regional	1142 (3.8)	2901 (6.5)	734 (6.4)	1692 (9.5)	1454 (11.6)
Remote/very remote	19 (0.1)	62 (0.1)	30 (0.3)	64 (0.4)	62 (0.5)
Missing	*4444 (14.8)*	*6084 (13.5)*	*1878 (16.4)*	*2667 (14.9)*	*1669 (13.3)*
Marital status					
Married/de facto	25,401 (84.8)	36,754 (81.7)	8973 (78.6)	13,886 (77.6)	9197 (73.4)
Divorced/separated/ Widowed	2733 (9.1)	5761 (12.8)	1488 (13.0)	2532 (14.1)	2122 (16.9)
Single	1576 (5.3)	2104 (4.7)	861 (7.5)	1305 (7.3)	1076 (8.6)
Missing	*246 (0.8)*	371 (0.8)	*99 (0.9)*	*184 (1.0)*	*141 (1.1)*
Psychological distress					
Low	23,360 (78.0)	32,401 (72.0)	8188 (71.7)	11,939 (66.7)	7064 (56.4)
Medium	4043 (13.5)	6096 (13.6)	1579 (13.8)	2382 (13.3)	1728 (13.8)
High/very high	1323 (4.4)	2513 (5.6)	701 (6.1)	1202 (6.7)	1166 (9.3)
Missing	*1230 (4.1)*	*3980 (8.9)*	*953 (8.3)*	*2394 (13.4)*	*2578 (20.6)*

BMI, kg/m^2 (mean ± SD)	26.6 ± 3.9	27.4 ± 4.2	27.2 ± 4.3	27.5 ± 4.4	27.8 ± 4.7
Missing	*1353 (4.5)*	*2858 (6.4)*	*591 (5.2)*	*1161 (6.5)*	*1113 (8.9)*
Smoking status					
Never	18,178 (60.7)	21,162 (47.0)	5130 (44.9)	7812 (43.6)	4753 (37.9)
Former	10,416 (34.8)	20,232 (45.0)	5114 (44.8)	8333 (46.5)	6151 (49.1)
Current	1227 (4.1)	3360 (7.5)	1116 (9.8)	1666 (9.3)	1521 (12.1)
Missing	*135 (0.5)*	*236 (0.5)*	*61. (0.5)*	*106 (0.6)*	*111 (0.9)*
Alcohol intake					
Moderate	14,575 (48.7)	19,679 (43.7)	4794 (42.0)	7382 (41.1)	4550 (36.3)
None/rarely	5644 (18.8)	10,091 (22.4)	2406 (21.1)	4225 (23.6)	4110 (32.8)
Hazardous	7461 (24.9)	10,528 (23.4)	2863 (25.1)	4262 (23.8)	2338 (18.7)
Harmful	2034 (6.8)	4076 (9.1)	1178 (10.3)	1938 (10.8)	1151 (9.2)
Missing	*242 (0.8)*	*616 (1.4)*	*180 (1.6)*	*300 (1.7)*	*387 (3.1)*
Physical activity					
Sufficiently active	24,315 (81.2)	34,789 (77.3)	8599 (75.3)	12,999 (72.6)	8224 (65.6)
Insufficiently active	4534 (15.1)	7180 (16.0)	2024 (17.7)	3256 (18.2)	2450 (19.5)
Sedentary	755 (2.5)	1704 (3.8)	475 (4.2)	899 (5.0)	985 (7.9)
Missing	*352 (1.2)*	*1317 (2.9)*	*323 (2.8)*	*753 (4.2)*	*877 (7.0)*
Fruit and vegetable intake					
≥ 5 portions/week	18,101 (60.4)	*26,077 (58.0)*	6138 (53.7)	9765 (54.5)	*7118 (56.8)*
< 5 portions/week	11.,457 (38.3)	*18,311 (40.7)*	5103 (44.7)	7844 (43.8)	*5159 (41.2)*
Missing	*398 (1.3)*	*602 (1.3)*	*180 (1.6)*	*298 (1.7)*	*259 (2.1)*
Fish intake					
≥ twice/week	14,657 (48.9)	*20,098 (44.7)*	5229 (45.8)	7685 (42.9)	*5283 (42.1)*
Once/week	12,129 (40.5)	19,551 (43.5)	4722 (41.3)	7830 (43.7)	5186 (41.4)
Never	2078 (6.9)	3034 (6.7)	877 (7.7)	1312 (7.3)	1069 (8.5)
Missing	*1092 (3.7)*	*2307 (5.1)*	*593 (5.2)*	*1080 (6.0)*	*998 (8.0)*
History of hypertension	9772 (32.6)	16,728 (37.2)	4232 (37.1)	7306 (40.8)	5285 (42.2)
History of heart disease	2207 (7.4)	4517 (10.0)	1206 (10.6)	2204 (12.3)	1955 (15.6)
History of diabetes mellitus	3954 (13.2)	7455 (16.6)	1711 (15.0)	3335 (18.6)	2678 (21.4)
Family history of stroke/heart disease	16,746 (55.9)	24,298 (54.0)	5941 (52.0)	9693 (54.1)	6556 (52.3)
Missing	*1 (0)*	*7 (0.02)*	*2 (0.02)*	*5 (0.03)*	*5 (0.03)*
Treated for high cholesterol	4566 (15.2)	7086 (15.8)	1766 (15.5)	2999 (16.8)	2139 (17.1)
Charlson Comoridity Index					
0	27,218 (90.9)	39,007 (86.7)	9957 (87.2)	14,941 (83.4)	9890 (78.9)

1	1579 (5.3)	3310 (7.4)	791 (6.9)	1578 (8.8)	1402 (11.2)
2	644 (2.2)	1361 (3.0)	337 (3.0)	681 (3.8)	551 (4.4)
≥3	515 (1.7)	1312 (2.9)	336 (2.9)	707 (4.0)	693 (5.5)

BMI = body mass index; SD = standard deviation

Supplementary Table 4 Baseline characteristics, by education level, among women

	Education level				
Characteristic	College/ university degree N = 30,375 n (%)	Certificate/ diploma/trade/ apprenticeship N = 37,055 n (%)	High school certificate N = 13,826 n (%)	School certificate N = 38,597 n (%)	No qualifications N = 16,994 n (%)
Age, years (mean ± SD)	57.5 ± 9.0	59.8 ± 10.3	61.3 ± 11.5	63.9 ± 10.9	66.5 ± 11.6
Categorical age, years					
45-59	20,662 (68.0)	21,354 (57.6)	7251 (52.4)	15,541 (40.3)	5638 (33.2)
60-69	6699 (22.1)	9433 (25.5)	3434 (24.8)	12,428 (32.2)	4832 (28.4)
70-79	2045 (6.7)	4059 (11.0)	1804 (13.1)	6553 (17.0)	3714 (21.9)
80+	969 (3.2)	2209 (6.0)	1337 (9.7)	4095 (10.6)	2810 (16.5)
SEIFA index of relative disadvantage					
1 (least deprived)	9079 (29.9)	7488 (20.2)	3001 (21.7)	5583 (14.5)	1459 (8.6)
2	6768 (22.3)	7313 (19.7)	2734 (19.8)	7163 (18.6)	2925 (17.2)
3	5525 (18.2)	7482 (20.2)	2590 (18.7)	8293 (21.5)	3644 (21.4)
4	5069 (16.7)	7585 (20.5)	2856 (20.7)	8657 (22.4)	4008 (23.6)
5 (most deprived)	3891 (12.8)	7161 (19.3)	2632 (19.0)	8887 (23.0)	4952 (29.1)
Missing	*43 (0.1)*	26 (0.1)	*13 (0.1)*	*14 (0.04)*	6 (0.04)
Geographical remoteness					
Major cities	16,639 (54.8)	17,454 (47.1)	7146 (51.7)	17,923 (46.4)	7968 (46.9)
Inner regional	7538 (24.8)	11,384 (30.7)	3649 (26.4)	12,615 (32.7)	5574 (32.8)
Outer regional	1516 (5.0)	2601 (7.0)	968 (7.0)	3226 (8.4)	1648 (9.7)
Remote/very remote	32 (0.1)	81 (0.2)	40 (0.3)	126 (0.3)	86 (0.5)
Missing	*4650 (15.3)*	*5535 (14.9)*	*2023 (14.6)*	*4707 (12.2)*	*1718 (10.1)*
Marital status					
Married/de facto	22,017 (72.5)	26,679 (72.0)	9812 (71.0)	27,491 (71.2)	10,624 (62.5)
Divorced/separated/ Widowed	5811 (19.1)	8154 (22.0)	3215 (23.3)	9608 (24.9)	5557 (32.7)
Single	2460 (8.1)	2112 (5.7)	759 (5.5)	1389 (3.6)	752 (4.4)
Missing	*87 (0.3)*	110 (0.3)	*40 (0.3)*	109 (0.3)	61 (0.4)
Psychological distress					
Low	22,064 (78.0)	32,401 (72.0)	8188 (71.7)	11,939 (66.7)	7064 (56.4)
Medium	5037 (13.5)	6096 (13.6)	1579 (13.8)	2372 (13.3)	1728 (13.8)
High/very high	1785 (4.4)	2513 (5.6)	701 (6.1)	1212 (6.7)	1166 (9.3)

Missing	1230 (4.1)	3980 (8.9)	953 (8.3)	2394 (13.4)	2578 (20.6)
BMI, kg/m^2 (mean ± SD)	25.9 ± 5.0	26.6 ± 5.3	26.4 ± 5.3	27.1 ± 5.3	27.7 ± 5.7
Missing	*1847 (6.1)*	*2863 (7.7)*	*1084 (7.8)*	*3475 (9.0)*	*2038 (12.0)*
Smoking status					
Never	20,415 (67.2)	22,871 (61.7)	8990 (65.0)	25,446 (65.9)	10,381 (61.1)
Former	8543 (28.1)	11,332 (30.6)	3748 (27.1)	10,275 (26.6)	4766 (28.1)
Current	1295 (4.3)	2684 (7.2)	2688 (7.0)	2688 (7.0)	1742 (10.3)
Missing	*122 (0.4)*	*168 (0.5)*	*188 (0.5)*	*188 (0.5)*	*105 (0.6)*
Alcohol intake					
Moderate	10,954 (36.1)	12,126 (32.7)	4165 (30.1)	11,095 (28.8)	3693 (21.7)
None/rarely	8948 (29.5)	13,915 (37.6)	5384 (38.9)	16,877 (43.7)	9449 (55.6)
Hazardous	8545 (28.1)	8678 (23.4)	3244 (23.5)	8078 (20.9)	2440 (14.4)
Harmful	1563 (5.2)	1704 (4.6)	661 (4.8)	1626 (4.2)	585 (3.4)
Missing	*365 (1.2)*	*632 (1.7)*	*372 (2.7)*	*931 (2.4)*	*827 (4.9)*
Physical activity					
Sufficiently active	25,193 (82.9)	29,522 (79.7)	10,491 (75.9)	28,535 (73.9)	10,724 (63.1)
Insufficiently active	4029 (13.3)	5267 (14.2)	2234 (16.2)	6336 (16.4)	3440 (20.2)
Sedentary	673 (2.2)	1157 (3.1)	546 (4.0)	1830 (4.7)	1405 (8.3)
Missing	*480 (1.6)*	1109 (3.0)	*555 (4.0)*	*1896 (4.9)*	1425 (8.4)
Fruit and vegetable intake					
≥ 5 portions/week	23,460 (77.2)	*28,305 (76.4)*	10,099 (73.0)	28,662 (74.3)	*11,672 (68.7)*
< 5 portions/week	6391 (21.0)	*8147 (22.0)*	3492 (25.3)	9268 (24.0)	*4947 (29.1)*
Missing	*524 (1.7)*	*603 (1.6)*	*235 (1.7)*	*667 (1.7)*	*375 (2.2)*
Fish intake					
≥ twice/week	15,849 (52.2)	*18,698 (50.5)*	7011 (50.7)	18,250 (47.3)	*75676 (44.5)*
Once/week	11,145 (36.7)	18,949 (37.6)	5030 (36.4)	14,989 (38.8)	6420 (37.8)
Never	2356 (7.8)	2849 (7.7)	1142 (8.3)	3202 (8.3)	1727 (10.2)
Missing	*1025 (3.4)*	*1559 (4.2)*	*643 (4.7)*	*2156 (5.6)*	*1280 (7.5)*
History of hypertension	7476 (24.6)	11,316 (30.5)	4357 (31.5)	14,620 (37.9)	6940 (40.8)
History of heart disease	1558 (5.1)	2829 (7.6)	1073 (7.8)	3687 (9.6)	2234 (13.2)
History of diabetes mellitus	1254 (4.1)	2114 (5.7)	904 (6.5)	3026 (7.8)	2016 (11.9)
Family history of stroke/heart disease	18,131 (59.7)	22,349 (60.3)	7749 (56.1)	23,657 (61.3)	10,186 (60.0)
Missing	*1 (0.0)*	*1 (0.0)*	*3 (0.02)*	*2 (0.01)*	*5 (0.03)*
Treated for high cholesterol	3078 (10.1)	4747 (12.8)	1763 (12.8)	5987 (15.5)	3097 (18.2)
Charlson Comoridity Index					

0	28,614 (94.2)	34,038 (91.2)	12,593 (91.1)	34,370 (89.1)	14,229 (83.7)
1	978 (3.2)	1722 (4.7)	660 (4.8)	2348 (6.2)	1520 (8.9)
2	316 (1.0)	587 (1.6)	302 (2.2)	922 (2.4)	620 (3.7)
≥3	467 (1.5)	708 (1.9)	271 (2.0)	927 (2.4)	625 (3.7)
Menopausal status					
Pre-menopause	8083 (26.6)	7454 (20.1)	2777 (20.1)	4369 (11.3)	1341 (7.9)
Post-menopause	15,025 (49.5)	17,861 (48.2)	6769 (49.0)	19,636 (50.9)	8233 (48.5)
Hysterectomy only	3997 (13.2)	6896 (18.6)	2442 (17.7)	8602 (22.3)	4294 (25.3)
Bilateral oophorectomy post-menopause	405 (1.3)	593 (1.6)	213 (1.5)	739 (1.9)	324 (1.9)
Bilateral oophorectomy (surgical menopause)	1319 (4.3)	2374 (6.4)	836 (6.1)	3085 (7.9)	1584 (9.3)
Missing	*1546 (5.1)*	1877 (5.1)	*789 (5.7)*	*2186 (5.7)*	1218 (7.2)
Current HRT use	3097 (10.2)	3878 (10.5)	1284 (9.3)	3934 (10.2)	1698 (10.0)
Missing	*97 (0.3)*	*139 (0.4)*	*60 (0.4)*	*157 (0.4)*	*95 (0.6)*
Current OCP use	106 (0.4)	103 (0.3)	48 (0.4)	57 (0.2)	26 (0.2)
Missing	*1994 (6.6)*	*3193 (8.6)*	*1392 (10.1)*	*4468 (11.6)*	*2747 (16.2)*

BMI = body mass index; HRT = hormone replacement therapy; OCP = oral contraceptive; SD = standard deviation

Supplementary Table 5 Baseline characteristics, by occurrence of stroke, by sex

Characteristic	Men: Stroke (N = 2031)	Men: No stroke (N =114,779)	Women: Stroke (N = 1528)	Women: No stroke (N =135,319)
Age, years (mean ± SD)	74.2 (± 10.4)	63.1 (± 10.9)	76.1 (± 11.5)	61.3 ± 10.8)
Categorical age, years				
40-59	225 (11.1)	50,148 (43.7)	184 (12.0)	70,262 (51.9)
60-69	412 (20.3)	33,647 (29.3)	225 (14.7)	36,581 (27.0)
70-79	651 (32.1)	19,758 (17.2)	392 (25.7)	17,783 (13.1)
80+	743 (36.6)	11,226 (9.8)	727 (47.6)	10,693 (7.9)
Remoteness area*				
Major cities	1079 (53.1)	57,687 (50.3)	824 (53.9)	66,306 (49.0)
Inner regional	570 (28.1)	35,572 (28.4)	403 (26.4)	40,357 (29.8)
Outer regional	132 (6.5)	7791 (6.8)	117 (7.7)	9842 (7.3)
Remote/very remote	**	236 (0.2)	**	363 (0.3)
Missing	*249 (12.3)*	*16,493 (14.4)*	*182 (11.9)*	*18,451 (13.6)*
SEIFA index of relative disadvantage				
1 (least deprived)	440 (21.7)	23,422 (20.4)	300 (19.6)	26,310 (19.4)
2	352 (17.3)	22,304 (19.4)	314 (20.6)	26,589(19.7)
3	397 (19.6)	23,010 (20.1)	295 (19.3)	27,239 (20.1)
4	417 (20.5)	23,043 (20.1)	299 (19.6)	27,876 (20.6)
5 (most deprived)	424 (20.9)	22,899 (20.0)	320 (20.9)	27,203 (20.1)
Missing	1 (0.1)	101 (0.1)	0 (0)	102 (0.1)
Marital status				
Married/de facto	1460 (71.9)	92,751 (80.8)	755 (49.4)	95,868 (70.9)
Divorced/separated/Widowed	424 (20.9)	14,212 (12.4)	669 (43.8)	31,676 (23.4)
Single	129 (6.4)	6793 (5.9)	100 (6.5)	7372 (5.5)
Missing	18 (0.9)	1023 (0.9)	4 (0.3)	403 (0.3)
Psychological distress				
Low	1216 (59.9)	81,736 (71.2)	753 (49.3)	89,032 (65.6)
Moderate	225 (11.1)	15,593 (13.6)	195 (12.8)	20,706 (15.3)
High/very high	123 (6.1)	6782 (5.9)	114 (7.5)	9698 (7.2)
Missing	467 (23.0)	10,668 (9.3)	466 (30.5)	15,883 (11.7)
BMI, kg/m^2 (mean ± SD)	26.5 (± 4.3)	27.2 (± 4.2)	26.0 (± 5.5)	26.7 (± 5.3)
Missing	160 (7.9)	6920 (6.0)	182 (11.9)	11,105 (8.2)

Smoking status				
Never	852 (42.0)	56,183 (49.0)	1074 (70.3)	87,029 (64.3)
Former	1011 (49.8)	49,235 (42.9)	361 (23.6)	38,303 (28.3)
Current	155 (7.6)	8725 (7.6)	84 (5.5)	9361 (6.9)
Missing	*13 (0.6)*	*636 (0.6)*	*9 (0.6)*	*626 (0.5)*
Alcohol intake				
Moderate	843 (41.5)	49,937 (43.5)	339 (22.2)	41,694 (30.8)
None/rarely	534 (26.3)	25,942 (22.6)	770 (50.4)	53,803 (39.8)
Hazardous	450 (22.2)	27,002 (23.5)	291 (19.0)	30,694 (22.7)
Harmful	153 (7.5)	10,224 (809)	41 (2.7)	6088 (4.5)
Missing	*51 (2.5)*	*1674 (1.5)*	*87 (5.7)*	*3010 (2.3)*
Physical activity				
Sufficiently active	1343 (66.1)	87,583 (76.3)	837 (54.8)	103,728 (76.6)
Insufficiently active	449 (22.1)	18,995 (16.6)	378 (24.7)	20,928 (15.5)
Sedentary	135 (6.7)	4683 (4.1)	159 (10.4)	5452 (4.0)
Missing	104 (5.1)	3518 (3.1)	154 (10.1)	5311 (3.9)
Fruit and vegetable intake				
≥ 5 portions/week	*1162 (57.2)*	*66,037 (57.5)*	*392 (25.7)*	*31,853 (23.5)*
< 5 portions/week	*826 (40.7)*	*47,048 (41.0)*	*1093 (71.5)*	*101,105 (74.7)*
Missing	*43 (2.1)*	*1694 (1.5)*	*43 (2.8)*	*2361 (1.7)*
Fish intake				
≥ twice/week	978 (48.2)	51,974 (45.3)	765 (50.1)	66,610 (49.2)
Once/week	785 (38.7)	48,633 (42.4)	531 (34.8)	51,002 (37.7)
Never	129 (6.4)	8241 (7.2)	113 (7.4)	11,163 (8.3)
Missing	*139 (6.8)*	*5931 (5.2)*	*119 (7.8)*	*6544 (4.8)*
History of hypertension	953 (46.9)	42,370 (36.9)	747 (48.9)	43,952 (32.5)
History of heart disease	648 (31.9)	18,485 (16.0)	401 (26.2)	10,980 (8.1)
History of diabetes mellitus	345 (17.0)	11,744 (10.2)	206 (13.5)	9108 (6.7)
Treated for high cholesterol	343 (16.9)	18,213 (15.9)	260 (17.0)	18,422 (13.6)
Family history of CVD	1131 (55.7)	62,103 (54.1)	974 (63.7)	81,098 (59.9)
Missing	*0 (0)*	*20 (0.02)*	*0 (0)*	*12 (0.01)*
Charlson comorbidity index				
0	1432 (70.5)	99,581 (86.8)	1097 (71.8)	122,747 (90.7)
1	262 (12.9)	8398 (7.3)	199 (13.0)	7059 (5.2)
2	149 (7.3)	3425 (3.0)	128 (8.4)	2619 (1.9)
≥ 3	188 (9.3)	3375 (2.9)	104 (6.8)	2894 (2.1)

WOMEN ONLY				
Menopausal status				
Pre-menopausal	-	-	62 (4.1)	23,962 (17.7)
Post-menopausal	-	-	827 (54.1)	66,697 (42.3)
Hysterectomy only	-	-	351 (23.0)	25,880 (19.1)
Bilateral oophorectomy post-menopause	-	-	38 (2.5)	2236 (1.7)
Bilateral oophorectomy (surgical menopause)	-	-	138 (9.0)	9040 (6.7)
Missing	-	-	*112 (7.3)*	*7504 (5.6)*
Current HRT use	-	-	96 (6.2)	13,816 (10.2)
Missing	-	-	*11 (0.7)*	*537 (0.4)*
Current OCP use	-	-	1 (0.07)	*339 (0.3)*
Missing	-	-	*252 (16.5)*	*13,542 (10.0)*

*Number of stroke events not disclosed due to small numbers and potential for de-identification of participants
AUD = Australian dollars; BMI = body mass index; CVD = cardiovascular disease (i.e. stroke or heart disease); HRT = hormone replacement therapy; OCP = oral contraceptive pill; SD = standard deviation; SEIFA = Socio-Economic Indexes of Australia

Supplementary Table 6 Age-standardised incidence rates of stroke (per 1000 person-years), by sex and age-category and age-adjusted hazard ratios for stroke risk using college/university degree as the reference category

Education level	Person-years	Stroke events, N	Incidence rate, per 1000 person-years* (95% CI)	Hazard ratio (95% CI)
Men aged 45-69 (N = 84,432)				
College/university degree	115,340	127	1.03 (0.89 to 1.17)	Reference
Certificate/diploma/trade or apprenticeship	154,6375	244	1.36 (1.22 to 1.50)	1.30 (1.05 to 1.[illegible]2)
High school certificate	40,404	73	1.63 (1.33 to 1.92)	1.57 (1.18 to 2.[illegible])
School certificate	56,542	115	1.45 (1.23 to 1.67)	1.51 (1.17 to 1.[illegible]5)
No qualifications	34,863	78	1.70 (1.36 to 2.04)	1.58 (1.19 to [illegible]0)
Men aged ≥70 (N = 32,378)				
College/university degree	25,908	225	8.66 (8.39 to 8.92)	Reference
Certificate/diploma/trade or apprenticeship	55,737	540	10.10 (9.90 to 10.32)	1.14 (0.98 to 1.[illegible]4)
High school certificate	12,970	127	9.09 (8.73 to 9.46)	1.08 (0.87 to 1.[illegible]4)
School certificate	26,418	283	10.25 (9.97 to 10.52)	1.21 (1.02 to 1.[illegible]4)
No qualifications	22,633	219	9.64 (9.35 to 9.94)	1.12 (0.93 to 1.[illegible]5)
Women aged 45-69 (N = 107,252)				
College/university degree	130,110	64	0.53 (0.44 to 0.63)	Reference
Certificate/diploma/trade or apprenticeship	146,594	116	0.77 (0.67 to 0.87)	1.49 (1.10 to 2.[illegible]3)
High school certificate	50,895	36	0.69 (0.52 to 0.85)	1.31 (0.88 to 1.[illegible]8)
School certificate	133,587	138	0.89 (0.77 to 1.01)	1.67 (1.24 to 2.[illegible]6)
No qualifications	49,999	55	0.93 (0.73 to 1.13)	1.73 (1.20 to 2.[illegible]9)
Women aged ≥70 (N = 29,595)				
College/university degree	13,665	96	8.63 (8.10 to 9.16)	Reference
Certificate/diploma/trade or apprenticeship	28,374	217	9.25 (8.87 to 9.64)	1.03 (0.81 to 1.[illegible]1)
High school certificate	14,023	127	9.28 (8.81 to 9.76)	1.08 (0.83 to 1.[illegible]1)
School certificate	48,086	386	8.98 (8.70 to 9.25)	1.02 (0.81 to 1.[illegible]7)
No qualifications	29,030	293	11.32 (10.93 to 11.70)	1.21 (0.96 to 1.[illegible]2)

*Age-standardised to the Australian sex-specific standard population

Supplementary Table 7 Results of additive interaction calculation showing: hazard ratios for stroke, by education and sex and effect modification by male sex; and hazard ratios for stroke, by education and age and effect modification by age ≥ 70, stratified by sex

Interaction	Male sex	No qualifications	Stroke events/person-years	HR (95%CI)	RERI (95% CI); Synergy index (95% CI)
Male sex by education level*	No	No	160 / 143,774	Reference	-0.17 (-0.52, 0.18); 0.82 (0.57, 1.19)
	Yes	No	348 / 79,029	1.49 (1.23, 1.80)	
	No	Yes	352 / 141,248	1.47 (1.21, 1.79)	
	Yes	Yes	297 / 57,496	1.79 (1.46, 2.18)	

Interaction	Age ≥ 70	No qualifications	Stroke events/person-years	HR (95%CI)	RERI (95% CI); Synergy index (95% CI)
Age by education level* (for men and women separately)	MEN				
	No	No	127 / 115,340	Reference	-0.23 (-0.56, 0.10); 0.30 (0.02, 4.73)
	Yes	No	78 / 34,863	2.03 (1.53, 2.69)	
	No	Yes	225 / 25, 908	7.95 (6.40, 9.89)	
	Yes	Yes	219 / 22,633	8.88 (7.14, 11.05)	
	WOMEN				
	No	No	64 / 130,110	Reference	0.004 (-0.39, 0.40); 1.01 (0.45, 2.27)
	Yes	No	55 / 49,999	2.23 (1.56, 3.20)	
	No	Yes	96 / 70,254	14.34 (10.45, 19.68)	
	Yes	Yes	293 / 29,030	20.61 (15.72, 27.01)	

*Comparing the lowest educational attainment category ('no qualifications') versus the highest educational attainment category (similar results, with more precision, were obtained when the analyses were repeated, but comparing the lowest educational attainment category versus all other educational categories combined)

RERI = relative excess risk due to interaction

Supplementary Table 8 Serially adjusted hazard ratios from the complete-case analysis for the association between education distress and stroke, stratified by gender

	Education				
	College/university degree	Certificate/diploma/ trade/ apprenticeship	High school certificate	School/intermediate certificate	No qualifications
MEN (1151 strokes among N = 77,785)					
Model 1*	1.00	1.19 (1.02, 1.39)	1.09 (0.86, 1.38)	1.24 (1.[illegible], 1.50)	1.29 (1.05, 1.59)
Model 2†	1.00	1.19 (1.01, 1.39)	1.08 (0.86, 1.37)	1.22 (1.01, 1.48)	1.27 (1.03, 1.57)
Model 3‡	1.00	1.17 (1.00, 1.37)	1.07 (0.85, 1.35)	1.20 (0.99, 1.45)	1.23 (1.00, 1.52)
Model 4§	1.00	1.15 (0.98, 1.35)	1.02 (0.81, 1.30)	1.15 (0.95, 1.40)	1.16 (0.94, 1.44)
Model 5‖	1.00	1.14 (0.97, 1.34)	1.01 (0.80, 1.28)	1.13 (0.94, 1.37)	1.14 (0.92, 1.41)
WOMEN (606 strokes among N =77,525)					
Model 1*	1.00	1.36 (1.02, 1.79)	1.32 (0.94, 1.84)	1.30 (0.99, 1.70)	1.53 (1.13, 2.06)
Model 2†	1.00	1.37 (1.04, 1.81)	1.32 (0.95, 1.85)	1.32 (1.01, 1.74)	1.57 (1.16, 2.12)
Model 3‡	1.00	1.36 (1.03, 1.80)	1.31 (0.94, 1.84)	1.30 (0.99, 1.71)	1.42 (1.05, 1.93)
Model 4§	1.00	1.32 (1.00, 1.75)	1.24 (0.88, 1.73)	1.22 (0.92, 1.60)	1.26 (0.92, 1.71)
Model 5‖	1.00	1.29 (0.97, 1.71)	1.22 (0.87, 1.71)	1.18 (0.89, 1.55)	1.20 (0.88, 1.64
Model 6#	1.00	1.28 (0.97, 1.70)	1.21 (0.86, 1.70)	1.16 (0.88, 1.54)	1.19 (0.87, 1.62)

*Adjusted for age

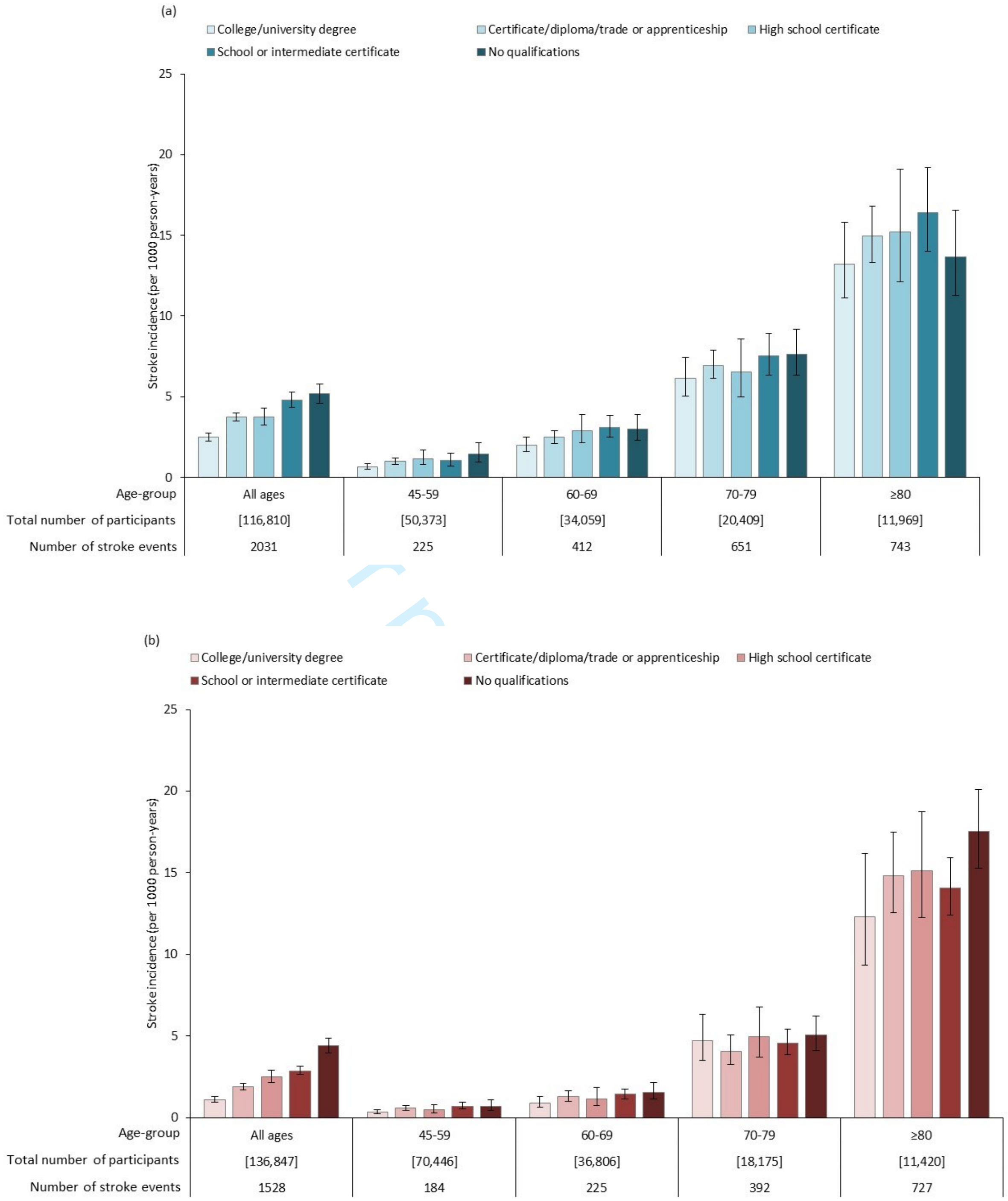
†Adjusted for age, marital status, remoteness

‡Adjusted for age, marital status, remoteness, psychological distress

§Adjusted for age, marital status, remoteness, psychological distress, smoking, alcohol intake, body mass index, physical activity, fruit & vegetable intake, fish consumption

‖Adjusted for age, marital status, remoteness, psychological distress, smoking, alcohol intake, body mass index, physical activity, fruit & vegetable intake, fish consumption, history of hypertension, history of heart disease, history of diabetes, history of treatment for raised cholesterol, Charlson comorbidity index and family history of stroke or heart disease

#In women, adjusted for age, marital status, remoteness, psychological distress, smoking, alcohol intake, body mass index, physical activity, fruit & vegetable intake, fish consumption, history of hypertension, history of heart disease, history of diabetes, history of treatment for raised cholesterol, Charlson comorbidity index, family history of stroke or heart disease, oral contraceptive use, hormone replacement therapy use and menopausal status



Supplementary Figure 1 Stroke incidence, per 1000 person-years, by age-group in (a) men and (b) women

STROBE Statement—Checklist of items that should be included in reports of ***cohort studies***

	Item No	Recommendation	Page number
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1,2
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4, paragraphs 1 & 2
Objectives	3	State specific objectives, including any prespecified hypotheses	4, final paragraph
Methods			
Study design	4	Present key elements of study design early in the paper	5,6
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	5
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	5
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed	NA
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	5, 6
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	5-7 & supplementary Table 1
Bias	9	Describe any efforts to address potential sources of bias	6
Study size	10	Explain how the study size was arrived at	5
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	6 & supplementary Table 1
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	6,7
		(*b*) Describe any methods used to examine subgroups and interactions	7
		(*c*) Explain how missing data were addressed	6
		(*d*) If applicable, explain how loss to follow-up was addressed	NA
		(*e*) Describe any sensitivity analyses	7
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	7 & Figure 1
		(b) Give reasons for non-participation at each stage	Figure 1
		(c) Consider use of a flow diagram	Figure 1
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	8 & supplementary tables 3 & 4
		(b) Indicate number of participants with missing data for each variable of interest	Supplementary tables 3 & 4

		(c) Summarise follow-up time (eg, average and total amount)	8, paragraph 3
Outcome data	15*	Report numbers of outcome events or summary measures over time	8, paragraph 3
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	9, Table 2
		(*b*) Report category boundaries when continuous variables were categorized	Tables
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	8, final paragraph & Table 1
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	9, Paragraphs 2 & 3; 10 first paragraph
Discussion			
Key results	18	Summarise key results with reference to study objectives	Page 10, paragraph 2
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	Page 10, paragraph 3
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	11-12
Generalisability	21	Discuss the generalisability (external validity) of the study results	10, paragraph 3
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	3

*Give information separately for exposed and unexposed groups.

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 http://www.strobe-statement.org.

BMJ Open

Education, sex, and risk of stroke: a prospective cohort study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-024070.R1
Article Type:	Research
Date Submitted by the Author:	16-Jul-2018
Complete List of Authors:	Jackson, Caroline A.; University of Edinburgh, Usher Institute of Population Health Sciences & Informatics Sudlow, Cathie; University of Edinburgh, Centre for Clinical Brain Sciences Mishra, Gita; University of Queensland, Faculty of Medicine, School of Public Health
Primary Subject Heading:	Epidemiology
Secondary Subject Heading:	Public health
Keywords:	SOCIAL MEDICINE, Stroke < NEUROLOGY, PUBLIC HEALTH, EPIDEMIOLOGY

Education, sex, and risk of stroke: a prospective cohort study

Caroline A Jackson (PhD)[1,2*], Cathie LM Sudlow (DPhil)[1,3], Gita D Mishra (PhD)[2]

[1]Institute of Population Health Sciences & Informatics

University of Edinburgh

Nine Bioquarter

9 Little France Road

Edinburgh

EH16 4UX

[2]School of Public Health

University of Queensland

Herston

Brisbane

QLD 4006

[3]Centre for Clinical Brain Sciences

University of Edinburgh

Chancellor's Building

49 Little France Crescent

Edinburgh

EH16 4SB

*Corresponding author details:

Email: caroline.jackson@ed.ac.uk

Tel: +44 (0) 131 651 7872

Word count: 3653

ABSTRACT

Objective: to determine whether: the association between highest educational attainment and stroke differed by sex and age; and whether potential mediators of observed associations differ by sex.

Design: prospective cohort study

Setting: population based, New South Wales, Australia

Participants: 253,657 stroke-free participants from the New South Wales 45 and Up Study.

Outcome measures: first-ever stroke events, identified through linkage to hospital and mortality records

Results: During mean follow-up of 4.7 years, 2031 and 1528 strokes occurred among men and women, respectively. Age-standardised stroke rate was inversely associated with education level, with the absolute risk difference between the lowest and highest education group greater among women than men. In relative terms, stroke risk was slightly more pronounced in women than men when comparing low versus high education (age-adjusted HRs: 1.41, 95% CI 1.16, 1.71 and 1.25, 95% CI 1.07, 1.46, respectively), but there was no clear evidence of statistical interaction. This association persisted into older age, but attenuated. Much of the increased stroke risk was explained by modifiable lifestyle factors, in both men and women.

Conclusion: Low education is associated with increased stroke risk in men and women, and may be marginally steeper in women than men. This disadvantage attenuates but persists into older age, particularly for women. Modifiable risk factors account for much of the excess risk from low education level. Public health policy and governmental decision-making should reflect the importance of education, for both men and women, for positive health throughout the life-course.

Keywords: education; socioeconomic disparities; stroke; cohort study; data linkage

ARTICLE SUMMARY

Strengths and limitations of this study

- A strength is the large study population and large number of outcome events, allowing stratification by sex
- This is one of just a few studies on this topic to have included both men and women and both fatal and non-fatal stroke outcome events
- Breadth of data enabled adjustment of many potential confounders/mediators, allowing identification of candidate mediators of the observed association
- Limitations include the 'healthy cohort' effect which may mean findings are less generalizable to the general population
- We could not identify non-fatal strokes for which participants were not admitted to hospital

FUNDING

This work was supported by a University of Queensland Early Career Researcher Grant awarded to CAJ (grant number 2013002357).

CONFLICTS OF INTEREST

None declared.

INTRODUCTION

Socioeconomic disparities in health are well-recognised, with lower socioeconomic position (SEP) associated with greater risk of mortality and disease, including cerebrovascular disease.[1] Based on cause-specific mortality studies, absolute socioeconomic inequalities in health appear to be greater for men than women.[2-4] There is, however, accumulating evidence that the SEP-cardiovascular disease relationship might actually be steeper in women than men.[5, 6] This has been investigated far less in relation to stroke, with the few, generally small, existing studies having reported conflicting findings.[7-9] Furthermore, few studies have explored whether the potential underlying mechanisms of the SEP-stroke association might differ by sex.

Age differences in the SEP-stroke association have also been rarely studied.[7, 10] There are conflicting views as to whether SEP disparities in health persist into older age. Some believe that the inequality gap narrows with age, whilst others propose that this is an artefact of mortality selection.[11, 12] It is important to not only determine whether lower SEP poses a particular disadvantage to women's health in terms of stroke risk, but also to ascertain whether such an effect, if it exists, persists over the life-course in order to inform approaches to reducing health inequalities.

Valid examination of sex and age differences in SEP inequalities in disease relies on careful measurement of SEP in both sexes. Educational attainment is a particularly useful SEP measure since it is easy to measure, elicits a high response rate, is relevant regardless of age and employment status and performs well regardless of sex.[13] In addition, since education is generally completed in young adulthood and is strongly related to parental characteristics, it also partly captures early life SEP.[13] Importantly, this also minimises potential for reverse causation between SEP and health outcomes (and their risk factors).

Using data from the 45 and Up study, a large Australian prospective cohort study, we aimed to determine whether: the association between highest educational attainment and stroke differed by sex and age; and whether potential mediators of observed associations differ by sex.

METHODS

Study population

We included participants from the Sax Institute's 45 and Up Study, a prospective cohort recruited between 2006 and 2009 from New South Wales (NSW), Australia, general population aged 45 years or over. Recruitment methods are described in detail elsewhere.[14] Briefly, potential participants were randomly sampled from the Department of Human Services (formerly Medicare Australia) enrolment database and mailed a self-administered questionnaire and information leaflet. Participants gave informed consent, including for follow-up via linkage to routinely collected health datasets. For this study, the cohort was linked to the NSW Admitted Patient Data Collection, the Australian Capital Territory (ACT) Admitted Patient Collection and the Australian Bureau of Statistics Death Data, with linkage performed by the Centre for Health Record Linkage.[15] We excluded participants with a previous hospitalised stroke record or self-reported stroke at baseline.

The conduct of the 45 and Up Study was approved by the University of NSW Human Research Ethics Committee. Ethical approval for the present study was obtained from the NSW Population and Health Services Research Ethics Committee, the ACT Health Human Research Ethics Committee and the University of Queensland Institutional Human Research Ethics Committee.

Education

Education was self-reported, with participants provided with the following list and asked to provide their highest attained qualification: college/university degree; certificate/diploma/trade/ apprenticeship; higher school or leaving certificate (or equivalent); school or intermediate certificate (or equivalent), typically awarded at the end of year 10 when children are aged 14, with successful completion required for the completion of the higher school certificate); or no qualifications.

Stroke

We identified incident stroke from hospital admission discharge records and mortality records and defined all strokes based on ICD-10 codes (I60, I61, I63 and I64). Ischaemic stroke was defined using I63 and I64 (since the majority of 'undetermined' strokes coded as I64 will be ischaemic).[16]

Covariates

Definitions of all included covariates are given in supplementary Table 1. We adjusted for demographic and other socioeconomic measures (including marital status and geographical remoteness). Whilst area-based deprivation was available, we did not adjust for it since in univariate analyses it was not associated with stroke risk (perhaps because area is not measured at a small enough level). We also did not analyse or adjust for average household income, for two reasons. Firstly, household income is a less reliable marker of SES, particularly among retired people (a substantial proportion of the cohort under study) because although actual household income might be low in this group, they often possess other financial assets, information on which was not collected. Secondly, average household income was missing in a substantial proportion of people (21%), with higher rates of missing data among women and older people. Thus we did not include income in our models, focusing instead on education level. We also adjusted for: psychological distress (as measured by the Kessler Psychological Distress scale[17]), a range of lifestyle factors (including fruit, vegetable and fish intake); history of disease/stroke risk factors (including physical comorbidity based on a modified Charlson comorbidity index[18] using hospital admission data in the five years prior to recruitment); and, among women, reproductive factors.

Statistical analyses

We performed analyses using Stata version 12. We calculated stroke rates by sex, education level and age-group, age-standardised to the Australian sex-specific standard population. We initially categorised age as follows: 45-59; 60-69; 70-79 and ≥80 years. For the purpose of investigating interactions, we later collapsed the two lowest and two highest groups, to dichotomise age into <70 and ≥70 years.

Overall, the frequency of missing values was less than 5% for most covariates, with the exception of geographical remoteness, which was less complete. Missing values were widely spread across participants, with 33% of men and 43% of women having missing values for at least one variable.

Missing data patterns indicated that data were likely to be missing at random but not missing completely at random. We therefore used multiple imputation by chained equations to impute missing values of included covariates. We imputed data for men and women separately, since we had additional sex-specific covariates for women, performing 35 imputations for men and 45 imputations for women.

We used Cox regression to obtain hazard ratios (HRs) with 95% confidence intervals (CIs) for the association between education level and stroke. We used age as the time axis, following participants from recruitment date and censoring for stroke event date, non-stroke death and end of follow-up (31st Dec 2012). We obtained sex-specific unadjusted HRs, before serially adjusting for groups of covariates including: other sociodemographic factors; psychological distress; lifestyle factors; and disease history. Thus we treated all covariates as common sources (confounders) in our analyses. Since some of these covariates might actually lie on a possible causal pathway between education level and stroke, we acknowledge that this assumption may not be valid and discuss the implications of this in our discussion. Our primary analysis included the multiple imputed dataset. In sensitivity analyses, we restricted our outcome to ischaemic strokes and also performed a complete case analysis. There was no clear violation of the proportional hazards assumption for education or any covariates in any analyses.

We investigated effect modification by sex and age by testing for multiplicative interaction (through testing for statistical interaction in the models). Since additive interaction is more important for understanding public health, we also investigated supra-additive interaction[19] between education (comparing the lowest versus highest education categories) and stroke, by calculating the relative excess risk of interaction (RERI) and synergy index with accompanying 95% CIs. A relative excess risk of interaction (RERI) of greater/lesser than 0 and a synergy index of greater/lesser than 1 indicate that the combined effects of each exposure is more or less than expected from adding the individual effects.

Results are reported in accordance with the Strengthening the Reporting of Observational Studies (STROBE) statement.[20]

Patient and public involvement

We did not include patient or public involvement in the conduct of the present study.

RESULTS

Among 257,843 participants eligible for inclusion, we excluded 4186 (1.6%) with missing information on education, ultimately including 253,657 participants (Figure 1). Compared to included participants, excluded participants were older and more likely to be: female; from areas of higher deprivation; and less healthy (Supplementary Table 2).

We included 116,810 men and 136,847 women (mean ± SD age in years 63.3 ± 11.0 and 61.4 ± 10.9 respectively). Sex-specific cross-tabulations of characteristics by education level are given in Supplementary Tables 3 and 4. Distribution of education level differed by sex and age. The proportion with a college or university qualification was slightly higher in men than women (25.6% versus 22.2%) and the proportion with no qualifications slightly higher in women than men (12.4% versus 10.7%). The proportion with no qualifications or a school/intermediate certificate was higher among older age groups (and was greater in women than men), whereas the proportion with a college/university degree was higher among younger age groups (with far less disparity between women and men). Lower educational status was associated with poorer lifestyle behaviour and clinical stroke risk factors, with similar patterns observed in men and women. Similarly, most characteristics were associated with stroke risk (Supplementary Table 5).

Absolute and relative stroke risk by education, sex and age

Mean follow-up was 4.70 (± SD 0.98) years. Among men, 2031 strokes (1696 ischaemic) occurred during 545,543 person-years of follow-up and among women 1528 strokes (1225 ischaemic) occurred during 644,362 person-years.

The pattern of absolute stroke risk was not linear with education level. For both men and women, absolute stroke rates were lowest in the group with highest educational attainment. The stroke rates for each of the other education categories were higher, but generally similar to each other, particularly for men. Among women the stroke rates in the three middle categories of educational attainment are very similar, with the rate in the group with no qualifications even higher (Figure 2a and Table 1). In contrast, among men, in comparison to the highest education category, stroke rates for each of the other education categories were higher, but to a similar extent, with confidence intervals for estimates overlapping. If anything, the stroke rate in those with no qualifications was slightly lower than in those with a trade/apprenticeship or school certificate (Figure 2c and Supplementary Table 6). Among those with no qualifications, the stroke rate was slightly greater in women than men and the absolute risk difference between lowest and highest education categories was greater in women than men (Table 1). After stratifying by age group, a similar pattern was observed for those aged under 70 years (Figure 2b and supplementary Figure 1) and, among women, for those aged 70 years or over.

In relative terms, having no qualifications was significantly associated with increased stroke risk in men and women, but was slightly more pronounced in women than men (age-adjusted HRs 1.41, 95% CI 1.16 to 1.71 and 1.25, 95% CI 1.07 to 1.46, respectively; Table 2). Among women, a similar magnitude of effect was observed for the middle three categories of education level versus the highest level, each of which was associated with about a 20% increased risk of stroke (Table 2, model 1). This pattern persisted when stratifying by age-group, although effect estimates were larger in those aged 45-69 than 70 years or over for both sexes. The association between education and stroke among those aged 45-69 was slightly stronger in women than men (HR lowest versus highest education level: 1.73, 95% CI 1.20 to 2.49 and 1.58, 95% CI 1.19 to 2.10, respectively). The association between education and stroke was somewhat weaker in men aged 70 years or over, but persisted among women in this age group, with having no qualifications associated with a 21% increased risk of stroke (HR 1.21, 95% CI 0.96 to 1.52), although this was not statistically significant

(Supplementary Table 6). However, having only a school certificate was statistically significantly associated with increased stroke risk in men aged 70 or over (HR 1.21, 95% CI 1.02, 1.44). Similar associations were found when we restricted our analyses to ischaemic strokes only.

Sex and age interactions with education

Although the effect of education on stroke risk appeared to be slightly greater among women than men there was no evidence of statistical interaction on either the multiplicative or additive scale. Although the RERI and synergy index were less than 0 and 1, respectively, confidence intervals did not exclude the possibility of no interaction (Supplementary Table 7). Similarly, there was no clear evidence of statistical interaction between age and education for men or women. (Supplementary Table 7).

Fully-adjusted education-stroke associations

Adjusting for confounding by marital status and geographical remoteness had little effect on the associations (Table 2). Following additional adjustment for behavioural factors and disease history, which could confound and/or possibly mediate a potentially causal association between education level and stroke risk, the magnitude of effect was attenuated and, for almost all education categories, was no longer statistically significant (fully-adjusted HRs for lowest versus highest education level in men and women: 1.10, 95% CI 0.94 to 1.30 and 1.21, 95% CI 0.97 to 1.51, respectively; Table 2). Whilst the same pattern of association was observed in the complete-case analysis, when compared to the analysis of the imputed dataset, the magnitude of effect was slightly greater for all education categories (Supplementary Table 8).

DISCUSSION

We found an inverse, albeit non-linear, association between education level and stroke risk in both men and women, which was slightly more pronounced in women than men. The association weakened somewhat in older ages for men and, to a lesser extent, for women. Modifiable lifestyle factors explained much of the increased stroke risk in both sexes.

Our study benefits from key strengths. We included a very large stroke-free study population, among which a large number of strokes occurred, providing sufficient power to stratify by sex. The breadth of data collected allowed us to adjust for a wide range of potential confounders/mediators, including those less commonly adjusted for in previous studies, to identify candidate mediators through which education level might affect stroke risk. There are however, some limitations. First, the participation rate in the 45 and Up Study is about 18% and given the 'healthy cohort effect', it is unlikely to be representative of the general NSW population aged 45 and over.[14] However, the cohort is heterogeneous across collected variables. Thus, whilst people with a low education level may well be under-represented in this cohort, given the cohort size and heterogeneity, this is unlikely to have had an effect on internal comparisons of exposure and outcome.[14] Second, it is important to recognise that the meaning of education differs by birth cohort. The consequences of having no high school education may differ among those born in the 1950s versus the 1920s for example. However, we did take this into account by examining the association between education and stroke risk within different age-groups. Third, we were unable to identify non-fatal strokes for which people were not admitted to hospital (estimated to be about 15% of all strokes, but likely higher among older people[21]) or strokes that occurred outside NSW and the ACT (which will have been few in number). Fourth, misclassification of stroke diagnosis may have occurred within hospital and mortality records. However, a recent systematic review of the accuracy of hospital and mortality records suggests that the use of appropriately selected, stroke-specific codes (rather than broad cerebrovascular disease codes) yields positive predictive values of greater than 70% in most studies and greater than 90% in some studies.[16] Finally, since we do not have time-varying information on covariates, we cannot be certain that all covariates are indeed confounders and not mediators in the relationship between education and stroke. We may therefore have over-adjusted our analyses by including possible mediators, thereby underestimating the association between education and stroke.

Our finding that education level is associated with stroke risk is consistent with the findings from many existing studies.[7, 9, 22-26] We did not however observe an incremental increase in stroke risk with decreasing education level. This is in keeping with some other studies where education was classified into more than three groups and where a non-linear pattern of association with stroke was observed.[7, 24] This may reflect a true non-linear relationship between education level and stroke in some settings. In the present study, the distribution of other stroke risk factors, such as history of diabetes, hypertension and key lifestyle factors was very similar among those with a diploma/trade/apprenticeship, higher school certificate or school/intermediate certificate. This might help explain why similar stroke rates were seen in these groups. Also, among men, the similarity in stroke risk between each of these education groups and the group with no qualifications is intriguing. As discussed below, this might reflect that men are able to compensate for having no qualifications to a greater extent than women.

Our study makes an important contribution to the small body of existing literature on SEP-sex interactions on stroke risk, and indeed circulatory disease in general. A recent systematic review and meta-analysis concluded that there are sex differences in the relationship between education and coronary heart disease, but not stroke.[5] However, the authors included studies of stroke mortality and not just stroke incidence. Results of the latter studies tended to suggest greater education-stroke inequalities in men than women. They also included studies from low and middle-income settings, where the SEP-stroke gradient itself might differ from high-income settings.

Findings from the very few studies that have explored interactions by age and sex on stroke risk are mixed. One reported no difference by sex[27], whereas a second found an association between low education and increased stroke risk among men but not women.[23] Consistent with our findings, two studies found the education-stroke association was weaker[8] or absent[9] in men.

There is biological plausibility for a different SEP-stroke association among women compared to men. Sex differences in the association between household income and atherosclerotic processes in adolescence were observed in one study, which were partially explained by a stronger income-adiposity association in females.[28] Similarly, the associations between SEP and carotid atherosclerosis in mid-age have been reported to be greater among men than women.[29] Furthermore, based on the theory of resource substitution (which states that resources can substitute for one another), education may be more important to women's health compared to men's because they have fewer socioeconomic resources of other kinds (such as income, power, authority and wealth) to draw upon.[30] In our study, income was lower among women than men across educational categories, but there was substantial missing data on income, especially among women, which precluded more detailed analyses. Whilst we were able to adjust for marital status, we were unable to adjust for spousal educational status, which is also thought to play a role in individuals' health.[31]

We also demonstrated that the adverse effect of low education on stroke risk attenuated but did persist in older age groups. The slightly weaker association between education and stroke among elderly men may reflect mortality selection[12], since educational disparities in mortality are greater among men than women. Our findings therefore support the view that socioeconomic inequalities in health persist into older age.[11] Two previous studies on education and stroke risk stratified by age group, but they did not stratify by sex. In one, low education was associated with increased stroke risk in those aged 65-74 years, but a decreased risk in those aged 75 years or over.[7] In contrast, the other study found no association among those aged 65-74 years, but a trend towards reduced risk in those with high versus low education aged 50-64 years.[10] Two studies of women only, with mean age less than 50 years, reported a magnitude of effect comparable to that among women aged less than 70 years in our study.[25, 32]

Our results suggest that lifestyle factors predominantly account for the observed increased stroke risk but that psychological distress and disease history are also important, perhaps more so for women than men. It is difficult to compare our sex-specific findings to similar studies of education and stroke, given that so few studies stratified by sex. A formal mediation study, using a path analysis approach, suggested that behavioural factors account for much of the education-cardiovascular disease gradient[33], which concurs with our findings. There is evidence that, when stratifying by sex, lifestyle behaviours may account for all of the education-cardiovascular association in men, but not women.[34] Our findings support this to some extent in that the slight excess stroke risk seen in women does persist after full adjustment for confounders/potential mediators. In addition, previous studies have posited a possible role for psychosocial factors as partial mediators of the association between education and stroke[35] and cardiovascular disease[36] in women. In line with this, adjustment for psychological distress in our study notably attenuated the effect estimates for women, but not men.

Since this is one of the few studies to date to examine the association between education and stroke risk by sex, within the same study population, further research is needed to confirm or refute the possibility of steeper socioeconomic-stroke disparities in women compared to men. Further investigation is also needed to shed more detailed understanding on the underlying mechanisms of the relationship between socioeconomic status and stroke risk and determine where these differ between men and women. Mediation analyses to unpick these mechanisms should stratify by sex and include non-conventional risk factors (such as psychosocial factors) as well as traditional cardiovascular risk factors. Such studies will inform tailored prevention strategies aimed at reducing health inequalities. Meanwhile, ongoing public health investment is needed to facilitate healthy lifestyle behaviour and reduce uptake of poor health behaviours, particularly in vulnerable socioeconomic groups.

Conclusion

Our study suggests that the education-stroke relationship is present in both men and women, but may be marginally stronger for women than men, with low education a disadvantage to women throughout the life-course. The critical importance of education, of both men and women, for positive health throughout the life-course should be reflected in public health and educational policy and governmental decision-making.

FIGURE LEGENDS

Figure 1 Flow diagram of included participants from the 45 and Up study

Figure 2 Absolute stroke risk by education level among men and women, showing age-standardised stroke incidence (per 1000 person-years) for (a) all ages; (b) age 45-69 years and (c) age 70 years or over (at baseline)

AUTHOR CONTRIBUTIONS

CAJ designed the study, CLMS and GDM contributed to the analytical strategy and interpretation of the results. CAJ drafted the manuscript and CLMS and GDM commented on and approved the final draft.

ACKNOWLEDGEMENTS

This research was completed using data collected through the 45 and Up Study (www.saxinstitute.org.au). The 45 and Up Study is managed by the Sax Institute in collaboration with major partner Cancer Council NSW; and partners: the National Heart Foundation of Australia (NSW Division); NSW Ministry of Health; NSW Government Family & Community Services – Ageing, Carers and the Disability Council NSW; and the Australian Red Cross Blood Service. We thank the many thousands of people participating in the 45 and Up Study.

DATA SHARING STATEMENT

The 45 And Up study data is accessible to any bona fide researcher who has a scientifically sound and feasible research proposal, has ethics approval for the study, and can cover the licence fees and costs incurred by accessing the data through the Secure Unified Research Environment provided by the New South Wales Sax Institute (https://www.saxinstitute.org.au/our-work/45-up-study/for-researchers/). Researchers accessing the data do not retain copies of the dataset and are not able to share these data with anyone else.

REFERENCES

1. Marshall IJ, Wang Y, Crichton S, et al. The effects of socioeconomic status on stroke risk and outcomes. *Lancet Neurol* 2015;**14**:1206-1218. doi: 1210.1016/S1474-4422(1215)00200-00208.

2. Huisman M, Kunst AE, Bopp M, et al. Educational inequalities in cause-specific mortality in middle-aged and older men and women in eight Western European populations. *Lancet* 2005;**365**:493-500.

3. Mackenbach JP, Hu Y, Artnik B, et al. Trends in inequalities in mortality amenable to health care in 17 European countries. *Health Aff* 2017;**36**:1110-1118. doi: 1110.1377/hlthaff.2016.1674.

4. Mustard CA, Etches J. Gender differences in socioeconomic inequality in mortality. *J Epidemiol Community Health* 2003;**57**:974-980.

5. Backholer K, Peters SAE, Bots SH, et al. Sex differences in the relationship between socioeconomic status and cardiovascular disease: A systematic review and meta-analysis. *J Epidemiol Community Health* 2017;**71**:550-557. doi: 510.1136/jech-2016-207890. Epub 202016 Dec 207814.

6. Ernstsen L, Bjerkeset O, Krokstad S. Educational inequalities in ischaemic heart disease mortality in 44,000 Norwegian women and men: The influence of psychosocial and behavioural factors. The Hunt Study. *Scand J Public Health* 2010;**38**:678-685. doi: 610.1177/1403494810380300. Epub 1403494810382010 Aug 1403494810380310.

7. Avendano M, Kawachi I, Van Lenthe F, et al. Socioeconomic status and stroke incidence in the US elderly: The role of risk factors in the EPESE study. *Stroke* 2006;**37**:1368-1373

8. Gillum RF, Mussolino ME. Education, poverty, and stroke incidence in whites and blacks: The NHANES I epidemiologic follow-up study. *J clin Epidemiol* 2003;**56**:188-195

9. Veronesi G, Ferrario MM, Chambless LE, et al. Gender differences in the association between education and the incidence of cardiovascular events in Northern Italy. *Eur J Public Health* 2011;**21**:762-767. doi: 710.1093/eurpub/ckq1155. Epub 2010 Nov 1010.

10. Avendano M, Glymour MM. Stroke disparities in older Americans: Is wealth a more powerful indicator of risk than income and education? *Stroke* 2008;**39**:1533-1540

11. Benzeval M, Green MJ, Leyland AH. Do social inequalities in health widen or converge with age? Longitudinal evidence from three cohorts in the West of Scotland. *BMC Public Health* 2011;**11:947.**:10.1186/1471-2458-1111-1947.

12. Dupre ME. Educational differences in age-related patterns of disease: Reconsidering the cumulative disadvantage and age-as-leveler hypotheses. *J Health Soc Behav* 2007;**48**:1-15.

13. Galobardes B, Shaw M, Lawlor DA, et al. Indicators of socioeconomic position (part 1). *J Epidemiol Community Health* 2006;**60**:7-12.

14. Banks E, Redman S, Jorm L, et al. Cohort profile: The 45 and Up Study. *Int J Epidemiol* 2008;**37**:941-947

15. Centre for Health Record Linkage. http://www.cherel.org.au/data-dictionaries#section6. Last accessed 1st March 2017.

16. Woodfield R, Grant I, Sudlow CL. Accuracy of electronic health record data for identifying stroke cases in large-scale epidemiological studies: A systematic review from the UK Biobank stroke outcomes group. *PLoS One* 2015;**10**:e0140533. doi: 0140510.0141371/journal.pone.0140533. eCollection 0142015.

17. Kessler RC, Andrews G, Colpe LJ, et al. Short screening scales to monitor population prevalences and trends in non-specific psychological distress. *Psychol Med* 2002;**32**:959-976.

18. Romano PS, Roos LL, Jollis JG. Adapting a clinical comorbidity index for use with ICD-9-CM administrative data: Differing perspectives. *J Clin Epidemiol* 1993;**46**:1075-1079.

19. Andersson T, Alfredsson L, Kallberg H, et al. Calculating measures of biological interaction. *Eur J Epidemiol* 2005;**20**:575-579.

20. Vandenbroucke JP, von Elm E, Altman DG, et al. Strengthening the reporting of observational studies in epidemiology (STROBE): Explanation and elaboration. *PLoS Med* 2007;**4**:e297.

21. Anderson CS, Jamrozik KD, Burvill PW, et al. Ascertaining the true incidence of stroke: Experience from the Perth Community Stroke Study, 1989-1990. *Med J Aust* 1993;**158**:80-84.

22. Andersen KK, Steding-Jessen M, Dalton SO, et al. Socioeconomic position and incidence of ischemic stroke in Denmark 2003-2012. A nationwide hospital-based study. *J Am Heart Assoc* 2014;**3(4).**e000762. doi: 000710.001161/JAHA.000113.000762.

23. Avendano M, Boshuizen HC, Schellevis FG, et al. Disparities in stroke preventive care in general practice did not explain socioeconomic disparities in stroke. *J Clin Epidemiol* 2006;**59**:1285-1294. Epub 2006 Jul 1211.

24. Jackson C, Jones M, Mishra G. Educational and homeownership inequalities in stroke incidence: A population-based longitudinal study of mid-aged women. *European Journal of Public Health* 2013;**24**:231-236. doi: 10.1093/eurpub/ckt073

25. Kuper H, Adami HO, Theorell T, et al. The socioeconomic gradient in the incidence of stroke: A prospective study in middle-aged women in Sweden. *Stroke* 2007;**38**:27-33

26. Liu L, Xue F, Ma J, et al. Social position and chronic conditions across the life span and risk of stroke: A life course epidemiological analysis of 22,847 American adults in ages over 50. *Int J Stroke* 2013;**8**:50-55. doi: 10.1111/j.1747-4949.2012.00927.x. Epub 02012 Dec 00911.

27. Honjo K, Tsutsumi A, Kayaba K. Socioeconomic indicators and cardiovascular disease incidence among japanese community residents: The Jichi Medical School Cohort Study. *Int J Behav Med* 2010;**17**:58-66. doi: 10.1007/s12529-12009-19051-12527. Epub 12009 Jun 12525.

28. Murasko JE. Male-female differences in the association between socioeconomic status and atherosclerotic risk in adolescents. *Soc Sci Med* 2008;**67**:1889-1897. doi: 1810.1016/j.socscimed.2008.1809.1018. Epub 2008 Oct 1815.

29. Rosvall M, Ostergren PO, Hedblad B, et al. Occupational status, educational level, and the prevalence of carotid atherosclerosis in a general population sample of middle-aged Swedish men and women: Results from the Malmo diet and cancer study. *Am J Epidemiol* 2000;**152**:334-346.

30. Ross CE, Masters RK, Hummer RA. Education and the gender gaps in health and mortality. *Demograph* 2012;**49**:1157-1183. doi: 1110.1007/s13524-13012-10130-z.

31. Brown DC, Hummer RA, Hayward MD. The importance of spousal education for the self-rated health of married adults in the United States. *Popul Res Policy Rev* 2014;**33**:127-151.

32. Honjo K, Iso H, Inoue M, et al. Education, social roles, and the risk of cardiovascular disease among middle-aged Japanese women: The JPHC Study Cohort I. *Stroke* 2008;**39**:2886-2890. doi: 2810.1161/STROKEAHA.2108.514067. Epub 512008 Jul 514024.

33. Kershaw KN, Droomers M, Robinson WR, et al. Quantifying the contributions of behavioral and biological risk factors to socioeconomic disparities in coronary heart disease incidence: The Morgen Study. *Eur J Epidemiol* 2013;**28**:807-814. doi: 810.1007/s10654-10013-19847-10652. Epub 12013 Sep 10614.

34. Ricceri F, Sacerdote C, Giraudo MT, et al. The association between educational level and cardiovascular and cerebrovascular diseases within the EPICOR study: New evidence for an old inequality problem. *PLoS One* 2016;**11**:e0164130. doi: 0164110.0161371/journal.pone.0164130. eCollection 0162016.

35. Jackson C, Mishra G. Depression and risk of stroke in mid-aged women: A prospective longitudinal study. *Stroke* 2013;**44**:1555-1560

36. Thurston RC, Kubzansky LD, Kawachi I, et al. Is the association between socioeconomic position and coronary heart disease stronger in women than in men? *Am J Epidemiol* 2005;**162**:57-65.

Table 1 Sex-specific age-standardised incidence rates (per 1000 person-years) for stroke, by education level

	Men (N = 116,810)			Women (N = 136,847)		
Education level	Person-years	Stroke events, N	Incidence rate, per 1000 person-years* (95% CI)	Person-years	Stroke events, N	Incidence rate, per 1000 person-years* (95% CI)
College/university degree	141,248	352	2.72 (2.43 to 3.02)	143,774	160	2.81 (2.27 to 3.36)
Certificate/diploma/trade or apprenticeship	210,375	784	3.31 (3.06 to 3.56)	174,968	333	3.16 (2.76 to 3.56)
Higher school certificate	53,3741	200	3.29 (2.82 to 3.76)	64,918	163	3.11 (2.60 to 3.61)
School/intermediate certificate	82,961	398	3.41 (3.05 to 3.76)	181,673	524	3.17 (2.87 to 3.46)
No qualifications	57,496	297	3.47 (3.02 to 3.92)	79,029	348	3.85 (3.42 to 4.29)

*Age-standardised to the Australian sex-specific standard population

Table 2 Serially adjusted hazard ratios (HRs) for the association between education level and stroke, stratified by sex

	Education level				
	College/ university degree	Certificate/diploma/ trade/ apprenticeship	Higher school certificate	School/intermediate certificate	No qualifications
	Reference	HR (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)
MEN (2031 strokes among N = 116,810)					
Model 1[*]	1.00	1.20 (1.06, 1.37)	1.23 (1.04, 1.47)	1.31 (1.13, 1.51)	1.25 (1.07, 1.46)
Model 2[†]	1.00	1.20 (1.06, 1.36)	1.22 (1.03, 1.46)	1.29 (1.12, 1.49)	1.23 (1.05, 1.44)
Model 3[‡]	1.00	1.19 (1.05, 1.35)	1.21 (1.02, 1.44)	1.27 (1.10, 1.47)	1.19 (1.02, 1.39)
Model 4[§]	1.00	1.17 (1.03, 1.33)	1.17 (0.98, 1.39)	1.23 (1.06, 1.42)	1.12 (0.96, 1.32)
Model 5[||]	1.00	1.16 (1.02, 1.32)	1.16 (0.97, 1.38)	1.21 (1.05, 1.40)	1.10 (0.94, 1.30)
WOMEN (1528 strokes among N = 136,847)					
Model 1[*]	1.00	1.21 (1.00, 1.46)	1.21 (0.97, 1.51)	1.23 (1.03, 1.47)	1.41 (1.16, 1.71)
Model 2[†]	1.00	1.22 (1.01, 1.47)	1.23 (0.99, 1.53)	1.25 (1.04, 1.50)	1.44 (1.19, 1.75)
Model 3[‡]	1.00	1.21 (1.00, 1.46)	1.22 (0.98, 1.52)	1.23 (1.03, 1.48)	1.37 (1.13, 1.67)
Model 4[§]	1.00	1.18 (1.98, 1.43)	1.17 (0.94, 1.46)	1.17 (0.98, 1.41)	1.26 (1.04, 1.54)
Model 5[||]	1.00	1.16 (0.96, 1.40)	1.17 (0.93, 1.53)	1.15 (0.96, 1.38)	1.22 (1.00, 1.49)
Model 6[#]	1.00	1.12 (0.90, 1.38)	1.09 (0.85, 1.40)	1.15 (0.94, 1.41)	1.21 (0.97, 1.51)

[*] Adjusted for age
[†] Model 1 + marital status and remoteness
[‡] Model 2 + psychological distress
[§] Model 3 + smoking, alcohol intake, body mass index, physical activity, fruit & vegetable intake, fish consumption
[||] Model 4 + history of: hypertension; heart disease; diabetes; and treatment for raised cholesterol, Charlson comorbidity index and family history of stroke/heart disease
[#] In women, Model 4 + oral contraceptive use, hormone replacement therapy use and menopausal status

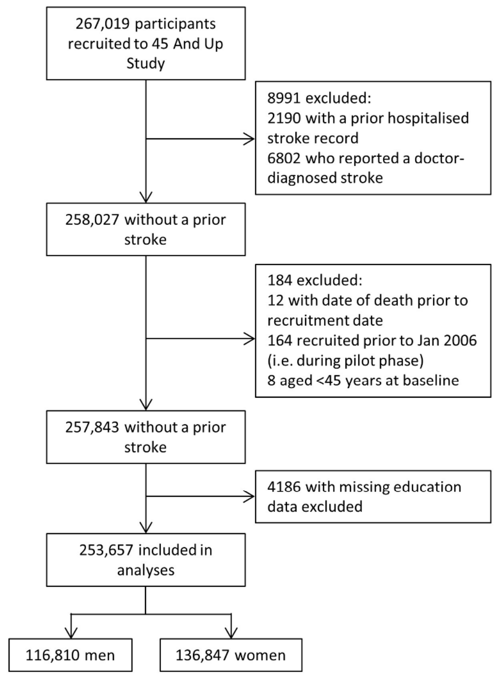


Figure 1 Flow diagram of included participants from the 45 and Up study

190x254mm (300 x 300 DPI)

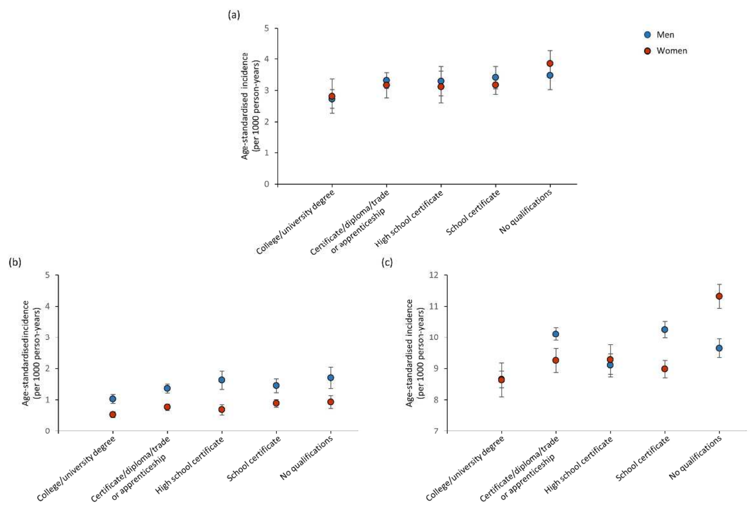


Figure 2 Absolute stroke risk by education level among men and women, showing age-standardised stroke incidence (per 1000 person-years) for (a) all ages; (b) age 45-69 years and (c) age 70 years or over (at baseline)

338x190mm (300 x 300 DPI)

ONLINE SUPPLEMENT

Education level, sex, and stroke risk: a prospective cohort study

Jackson CA, Sudlow CLM, Mishra GD

Supplementary Table 1 Definitions of covariates included in adjusted cox regression models for the association between psychological distress and incident stroke

Covariates	Definition
Sociodemographic factors	
Marital status	Married/de facto; divorced, separated or widowed; single
Geographical remoteness	Major cities; inner regional; outer regional; and remote or very remote (based on the acessibility/Remoteness Index of Australia)*
Socio-Economic Indexes of Australia (SEIFA) index of relative disadvantage†	Quintiles, from category 1 (least deprived) to 5 (most deprived)
Education	Highest attained qualification: College/university degree; high school certificate; school certificate; certificate/diploma/trade or apprenticeship; or no qualifications
Average annual household income (in Australian Dollars)	Grouped as: ≤19,999; 20,000-29,999; 30,000-39,999; 40,000-49,000; 50,000-69,999; and ≥70,000
Lifestyle factors	
BMI	Body mass index (kg/m^2) created based on self-reported weight and height and included as a continuous variable
Smoking status	From self-reported smoking history, categorical variable created: never; former; current
Alcohol intake	Weekly units of alcohol, defined as: moderate (<21 units/week for men and <14 units/week for women); none; hazardous (21-50 for men and 14-35 for women); and harmful (>50 for men and >35 for women) Participants were asked about weekly number of alcoholic drinks consumed (but were not asked about intake of specific drinks). When calculating weekly alcohol intake we assumed that one alcoholic drink was equivalent to two units of alcohol
Physical activity	Questions on physical activity (from the Active Australia Survey) used to construct a physical activity variable, categorised as: sufficient; insufficient; and sedentary as per the Active Australia Survey recommendations‡
Daily fruit and vegetable intake	Questions on average daily servings of cooked and raw vegetables and fruit or glasses of fruit juice used to create a binary variable of ≥5 or <5 portions of fruit or vegetables
Weekly fish intake	Based on self-reported weekly intake of fish or seafood, and defined as: ≥twice/week; once/week; and never
Physiological and family history variables	
History of hypertension	Self-reported doctor-diagnosed hypertension or self-reported or treatment for high blood pressure in the past month
History of heart disease	Self-reported doctor-diagnosed 'heart disease' or treatment for 'either heart attack or angina' or 'other heart disease' in the past month
Diabetes	Self-reported doctor diagnosed diabetes
Treated cholesterol	self-report of treatment for 'high blood cholesterol' in the past month
Family history of stroke or heart disease	Dichotomous variable based on reported parent and sibling stroke and heart disease history

Charlson comorbidity index	Based on the modified Charlson Index and using the ICD10-Australian Modification conversion code[§] we identified and weighted 17 chronic disease conditions (excluding stroke from the cerebrovascular disease codes), to give a total comorbidity score, which we then categorised as 0, 1, 2, and ≥3
Menopausal status	Using responses to questions on menopause, hysterectomy and oophorectomy, along with age at which they occurred, a menopausal status variables was created and categorised as: pre-menopause; post-menopause; hysterectomy only; bilateral oophorectomy post-menopause; bilateral oophorectomy (surgical menopause); and 'missing' where information was insufficiently complete or conflicting
Current OCP use	A dichotomous 'current OCP use' variable was created based on questions on ever having used OCP and age at last use
Current HRT use	A dichotomous 'current HRT use' variable was created based on questions on ever having used HRT and age at last use

[*]Australian Department of Health and Aged Care (DHAC). Measuring remoteness: Accessibility/remoteness index of australia (ARIA). Occasional papers: New Series Number 14. Canberra: DHAC; 2001
[†]Australian Bureau of Statistics Socio-economic indexes for areas. http://www.abs.gov.au/websitedbs/censushome.nsf/home/seifa. Accessed 15th February 2017
[‡]Australian Institute of Health and Welfare (AIHW). The Active Australia Survey: a guide and manual for implementation, analysis and reporting. Canberra: AIHW; 2003
[§]Sundararajan V, Henderson T, Perry C, et al. New icd-10 version of the charlson comorbidity index predicted in-hospital mortality. *J Clin Epidemiol.* 2004;57:1288-1294.
BMI = body mass index; HRT = hormone replacement therapy; OCP = oral contraceptive pill

Supplementary Table 2 Comparison of baseline characteristics of included versus excluded participants (i.e. those with versus without education level)

Characteristic	Included (N = 253,657) n (%)	Excluded (N = 4186) n (%)
Age, years (mean ± SD)	62.3 ± 11.0	69.1 ± 12.1
Categorical age, years		
45-59	120,819 (47.6)	1101 (26.3)
60-69	70,865 (27.9)	1035 (24.7)
70-79	38,584 (15.2)	1013 (24.2)
80+	23,389 (9.2)	1037 (24.8)
Female	136,847 (54.0)	2187 (52.3)
SEIFA index of relative disadvantage		
1 (least deprived)	50,472 (19.9)	591 (14.1)
2	49,559 (19.5)	731 (17.5)
3	50,941 (20.1)	904 (21.6)
4	51,635 (20.4)	896 (21.4)
5 (most deprived)	50,846 (20.1)	1062 (25.4)
Missing	*204 (0.1)*	*2 (0.1)*
Geographical remoteness		
Major cities	125,896 (49.6)	2100 (50.2)
Inner regional	73,902 (29.1)	1234 (29.5)
Outer regional	17,882 (7.1)	352 (8.4)
Remote/very remote	602 (0.2)	20 (0.5)
Missing	*35,375 (14.0)*	*480 (11.5)*
Marital status		
Married/de facto	190,834 (75.7)	2718 (64.9)
Divorced/separated/Widowed	46,981 (18.6)	1095 (26.2)
Single	14,394 (5.7)	283 (6.8)
Missing	*1448 (0.6)*	*90 (2.2)*
Psychological distress		
Low	172, 737 (68.1)	2054 (49.1)
Medium	36,719 (14.5)	455 (10.9)
High/very high	16,717 (6.6)	273 (6.5)
Missing	27,484 (10.8)	1404 (33.5)
BMI, kg/m^2 (mean ± SD)	26.9 ± 4.9	26.8 ± 5.1
Missing	*18, 373 (7.2)*	*861 (20.6)*
Smoking status		
Never	145,138 (57.2)	2400 (57.3)
Former	88,910 (35.1)	1410 (33.7)
Current	18,325 (7.2)	323 (7.7)
Missing	*1284 (0.5)*	*53 (1.3)*
Alcohol intake		
Moderate	92,813 (36.6)	1275 (30.5)
None/rarely	81,049 (32.0)	1569 (37.5)
Hazardous	58,437 (23.0)	686 (16.4)
Harmful	16,506 (6.5)	193 (4.6)
Missing	*4852 (1.9)*	*463 (11.1)*
Physical activity		
Sufficiently active	193,391 (76.2)	2432 (58.1)
Insufficiently active	40,750 (16.1)	829 (19.8)
Sedentary	10,429 (4.1)	288 (6.9)
Missing	*9087 (3.6)*	*637 (15.2)*

Fruit and vegetable intake		
≥ 5 portions/week	*169,397 (66.8)*	*2571 (61.4)*
< 5 portions/week	*80,119 (31.6)*	*1435 (34.3)*
Missing	*4141 (1.6)*	*180 (4.3)*
Fish intake		
≥ twice/week	120,327 (47.4)	1931 (46.1)
Once/week	100,951 (39.8)	1463 (35.0)
Never	19,646 (7.8)	333 (8.0)
Missing	*12,733 (5.0)*	*459 (11.0)*
History of hypertension	88,022 (34.7)	1503 (35.9)
History of heart disease	30,514 (12.0)	688 (16.4)
History of diabetes mellitus	21,403 (8.4)	504 (12.0)
Family history of stroke/heart disease	145,306 (57.3)	2123 (50.7)
Missing	*32 (0.01)*	*5 (0.1)*
Treated for high cholesterol	37,238 (14.7)	613 (14.6)
Charlson Comoridity Index		
0	224,857 (88.7)	3338 (79.7)
1	15,918 (6.3)	440 (10.5)
2	6321 (2.5)	194 (4.6)
≥3	6561 (2.6)	214 (5.1)
WOMEN ONLY		
Menopausal status		
Pre-menopausal	24,024 (17.6)	201 (9.2)
Post-menopausal	67,524 (49.3)	1078 (49.3)
Hysterectomy only	26,231 (19.2)	499 (22.8)
Bilateral oophorectomy post-menopause	2274 (1.7)	25. (1.1)
Bilateral oophorectomy (surgical menopause)	9178 (6.7)	151 (6.9)
Missing	*7616 (5.6)*	*233 (10.7)*
Current HRT use	13,911 (10.2)	144 (6.6)
Missing	*548 (0.4)*	*17 (0.8)*
Current OCP use	340 (0.3)	3 (0.1)
Missing	*13,794 (10.1)*	*486 (22.2)*

BMI = body mass index; HRT = hormone replacement therapy; OCP = oral contraceptive; SD = standard deviation

Supplementary Table 3 Baseline characteristics, by education level, among men

	Education level				
Characteristic	College/ university degree N = 29,956 n (%)	Certificate/diploma/ trade/ apprenticeship N = 44,990 n (%)	Higher school certificate N = 11,421 n (%)	School/intermediate certificate N = 17,907 n (%)	No qualifications N = 12,536 n (%)
Age, years (mean ± SD)	60.8 ± 10.4	63.3 ± 11.0	62.7± 11.2	65.4 ± 10.9	67.2 ± 10.8
Categorical age, years					
45-59	16,077 (53.7)	19,413 (43.2)	5327 (46.6)	6048 (33.8)	3508 (28.0)
60-69	8104 (27.1)	13,086 (29.1)	3142 (27.5)	5877 (32.8)	3850 (30.7)
70-79	3614 (12.1)	7988 (17.8)	1787 (15.7)	3724 (20.8)	3296 (26.3)
80+	2161 (7.2)	4503 (10.0)	1165 (10.2)	2258 (12.6)	1882 (15.0)
SEIFA index of relative disadvantage					
1 (least deprived)	10,289 (34.4)	7911 (17.6)	2499 (21.9)	2223 (12.4)	940 (7.5)
2	6407 (21.4)	9049 (20.1)	2105 (18.4)	3092 (17.3)	2003 (16.0)
3	5350 (17.9)	9280 (20.6)	2243 (19.6)	3917 (21.9)	2617 (20.9)
4	4462 14.9)	9593 (21.3)	2255 (19.7)	4131 (23.1)	3019 (24.1)
5 (most deprived)	3398 (11.3)	9122 (20.3)	2310 (20.2)	4539 (25.4)	3954 (31.5)
Missing	*50 (0.2)*	*35 (0.1)*	*9 (0.1)*	*5 (0.03)*	*3 (0.02)*
Geographical remoteness					
Major cities	17,651 (58.9)	22,332 (49.6)	5809 (50.9)	7733 (43.2)	5241 (41.8)
Inner regional	6700 (22.4)	13,611 (30.3)	2970 (26.0)	5751 (32.1)	4110 (32.8)
Outer regional	1142 (3.8)	2901 (6.5)	734 (6.4)	1692 (9.5)	1454 (11.6)
Remote/very remote	19 (0.1)	62 (0.1)	30 (0.3)	64 (0.4)	62 (0.5)
Missing	*4444 (14.8)*	*6084 (13.5)*	*1878 (16.4)*	*2667 (14.9)*	*1669 (13.3)*
Marital status					
Married/de facto	25,401 (84.8)	36,754 (81.7)	8973 (78.6)	13,886 (77.6)	9197 (73.4)
Divorced/separated/ Widowed	2733 (9.1)	5761 (12.8)	1488 (13.0)	2532 (14.1)	2122 (16.9)
Single	1576 (5.3)	2104 (4.7)	861 (7.5)	1305 (7.3)	1076 (8.6)
Missing	*246 (0.8)*	371 (0.8)	*99 (0.9)*	*184 (1.0)*	*141 (1.1)*
Psychological distress					
Low	23,360 (78.0)	32,401 (72.0)	8188 (71.7)	11,939 (66.7)	7064 (56.4)
Medium	4043 (13.5)	6096 (13.6)	1579 (13.8)	2372 (13.3)	1728 (13.8)
High/very high	1323 (4.4)	2513 (5.6)	701 (6.1)	1202 (6.7)	1166 (9.3)
Missing	*1230 (4.1)*	*3980 (8.9)*	*953 (8.3)*	*2394 (13.4)*	*2578 (20.6)*

BMI, kg/m² (mean ± SD)	26.6 ± 3.9	27.4 ± 4.2	27.2 ± 4.3	27.5 ± 4.4	27.8 ± 4.7
Missing	*1353 (4.5)*	*2858 (6.4)*	*591 (5.2)*	*1171 (6.5)*	*1113 (8.9)*
Smoking status					
Never	18,178 (60.7)	21,162 (47.0)	5130 (44.9)	7812 (43.6)	4753 (37.9)
Former	10,416 (34.8)	20,232 (45.0)	5114 (44.8)	8333 (46.5)	6151 (49.1)
Current	1227 (4.1)	3360 (7.5)	1116 (9.8)	1656 (9.3)	1521 (12.1)
Missing	*135 (0.5)*	*236 (0.5)*	*61. (0.5)*	*106 (0.6)*	*111 (0.9)*
Alcohol intake					
Moderate	14,575 (48.7)	19,679 (43.7)	4794 (42.0)	7362 (41.1)	4550 (36.3)
None/rarely	5644 (18.8)	10,091 (22.4)	2406 (21.1)	4225 (23.6)	4110 (32.8)
Hazardous	7461 (24.9)	10,528 (23.4)	2863 (25.1)	4262 (23.8)	2338 (18.7)
Harmful	2034 (6.8)	4076 (9.1)	1178 (10.3)	1938 (10.8)	1151 (9.2)
Missing	*242 (0.8)*	*616 (1.4)*	*180 (1.6)*	*300 (1.7)*	*387 (3.1)*
Physical activity					
Sufficiently active	24,315 (81.2)	34,789 (77.3)	8599 (75.3)	12,999 (72.6)	8224 (65.6)
Insufficiently active	4534 (15.1)	7180 (16.0)	2024 (17.7)	3256 (18.2)	2450 (19.5)
Sedentary	755 (2.5)	1704 (3.8)	475 (4.2)	899 (5.0)	985 (7.9)
Missing	*352 (1.2)*	*1317 (2.9)*	*323 (2.8)*	*753 (4.2)*	*877 (7.0)*
Fruit and vegetable intake					
≥ 5 portions/week	18,101 (60.4)	*26,077 (58.0)*	6138 (53.7)	9765 (54.5)	*7118 (56.8)*
< 5 portions/week	11.,457 (38.3)	*18,311 (40.7)*	5103 (44.7)	7844 (43.8)	*5159 (41.2)*
Missing	*398 (1.3)*	*602 (1.3)*	*180 (1.6)*	*298 (1.7)*	*259 (2.1)*
Fish intake					
≥ twice/week	14,657 (48.9)	*20,098 (44.7)*	5229 (45.8)	7685 (42.9)	*5283 (42.1)*
Once/week	12,129 (40.5)	19,551 (43.5)	4722 (41.3)	7830 (43.7)	5186 (41.4)
Never	2078 (6.9)	3034 (6.7)	877 (7.7)	1312 (7.3)	1069 (8.5)
Missing	*1092 (3.7)*	*2307 (5.1)*	*593 (5.2)*	*1080 (6.0)*	*998 (8.0)*
History of hypertension	9772 (32.6)	16,728 (37.2)	4232 (37.1)	7306 (40.8)	5285 (42.2)
History of heart disease	2207 (7.4)	4517 (10.0)	1206 (10.6)	2204 (12.3)	1955 (15.6)
History of diabetes mellitus	3954 (13.2)	7455 (16.6)	1711 (15.0)	3335 (18.6)	2678 (21.4)
Family history of stroke/heart disease	16,746 (55.9)	24,298 (54.0)	5941 (52.0)	9693 (54.1)	6556 (52.3)
Missing	*1 (0)*	*7 (0.02)*	*2 (0.02)*	*5 (0.03)*	*5 (0.03)*
Treated for high cholesterol	4566 (15.2)	7086 (15.8)	1766 (15.5)	2999 (16.8)	2139 (17.1)
Charlson Comoridity Index					
0	27,218 (90.9)	39,007 (86.7)	9957 (87.2)	14,941 (83.4)	9890 (78.9)

1	1579 (5.3)	3310 (7.4)	791 (6.9)	1[illegible]78 (8.8)	1402 (11.2)
2	644 (2.2)	1361 (3.0)	337 (3.0)	6[illegible]1 (3.8)	551 (4.4)
≥3	515 (1.7)	1312 (2.9)	336 (2.9)	7[illegible]7 (4.0)	693 (5.5)

BMI = body mass index; SD = standard deviation

Supplementary Table 4 Baseline characteristics, by education level, among women

	Education level				
Characteristic	College/ university degree N = 30,375 n (%)	Certificate/ diploma/trade/ apprenticeship N = 37,055 n (%)	Higher school certificate N = 13,826 n (%)	School intermediate certificate N = 38,597 n (%)	No qualifications N = 16,994 n (%)
Age, years (mean ± SD)	57.5 ± 9.0	59.8 ± 10.3	61.3 ± 11.5	63.9 ± 10.9	66.5 ± 11.6
Categorical age, years					
45-59	20,662 (68.0)	21,354 (57.6)	7251 (52.4)	15,541 (40.3)	5638 (33.2)
60-69	6699 (22.1)	9433 (25.5)	3434 (24.8)	12,408 (32.2)	4832 (28.4)
70-79	2045 (6.7)	4059 (11.0)	1804 (13.1)	6553 (17.0)	3714 (21.9)
80+	969 (3.2)	2209 (6.0)	1337 (9.7)	4095 (10.6)	2810 (16.5)
SEIFA index of relative disadvantage					
1 (least deprived)	9079 (29.9)	7488 (20.2)	3001 (21.7)	5583 (14.5)	1459 (8.6)
2	6768 (22.3)	7313 (19.7)	2734 (19.8)	7163 (18.6)	2925 (17.2)
3	5525 (18.2)	7482 (20.2)	2590 (18.7)	8293 (21.5)	3644 (21.4)
4	5069 (16.7)	7585 (20.5)	2856 (20.7)	8657 (22.4)	4008 (23.6)
5 (most deprived)	3891 (12.8)	7161 (19.3)	2632 (19.0)	8887 (23.0)	4952 (29.1)
Missing	*43 (0.1)*	26 (0.1)	*13 (0.1)*	*14 (0.04)*	6 (0.04)
Geographical remoteness					
Major cities	16,639 (54.8)	17,454 (47.1)	7146 (51.7)	17,923 (46.4)	7968 (46.9)
Inner regional	7538 (24.8)	11,384 (30.7)	3649 (26.4)	12,615 (32.7)	5574 (32.8)
Outer regional	1516 (5.0)	2601 (7.0)	968 (7.0)	3226 (8.4)	1648 (9.7)
Remote/very remote	32 (0.1)	81 (0.2)	40 (0.3)	126 (0.3)	86 (0.5)
Missing	*4650 (15.3)*	*5535 (14.9)*	*2023 (14.6)*	*4707 (12.2)*	*1718 (10.1)*
Marital status					
Married/de facto	22,017 (72.5)	26,679 (72.0)	9812 (71.0)	27,491 (71.2)	10,624 (62.5)
Divorced/separated/ Widowed	5811 (19.1)	8154 (22.0)	3215 (23.3)	9608 (24.9)	5557 (32.7)
Single	2460 (8.1)	2112 (5.7)	759 (5.5)	1389 (3.6)	752 (4.4)
Missing	*87 (0.3)*	110 (0.3)	*40 (0.3)*	109 (0.3)	61 (0.4)
Psychological distress					
Low	22,064 (78.0)	32,401 (72.0)	8188 (71.7)	11,439 (66.7)	7064 (56.4)
Medium	5037 (13.5)	6096 (13.6)	1579 (13.8)	2372 (13.3)	1728 (13.8)
High/very high	1785 (4.4)	2513 (5.6)	701 (6.1)	1202 (6.7)	1166 (9.3)

Missing	1230 (4.1)	3980 (8.9)	953 (8.3)	23[illegible]4 (13.4)	2578 (20.6)
BMI, kg/m^2 (mean ± SD)	25.9 ± 5.0	26.6 ± 5.3	26.4 ± 5.3	2[illegible].1 ± 5.3	27.7 ± 5.7
Missing	*1847 (6.1)*	*2863 (7.7)*	*1084 (7.8)*	*3[illegible]55 (9.0)*	*2038 (12.0)*
Smoking status					
Never	20,415 (67.2)	22,871 (61.7)	8990 (65.0)	25,[illegible]46 (65.9)	10,381 (61.1)
Former	8543 (28.1)	11,332 (30.6)	3748 (27.1)	10,[illegible]75 (26.6)	4766 (28.1)
Current	1295 (4.3)	2684 (7.2)	2688 (7.0)	2[illegible]88 (7.0)	1742 (10.3)
Missing	*122 (0.4)*	*168 (0.5)*	*188 (0.5)*	*1[illegible]8 (0.5)*	*105 (0.6)*
Alcohol intake					
Moderate	10,954 (36.1)	12,126 (32.7)	4165 (30.1)	1[illegible],[illegible]95 (28.8)	3693 (21.7)
None/rarely	8948 (29.5)	13,915 (37.6)	5384 (38.9)	16,[illegible]77 (43.7)	9449 (55.6)
Hazardous	8545 (28.1)	8678 (23.4)	3244 (23.5)	[illegible]078 (20.9)	2440 (14.4)
Harmful	1563 (5.2)	1704 (4.6)	661 (4.8)	1[illegible]16 (4.2)	585 (3.4)
Missing	*365 (1.2)*	*632 (1.7)*	*372 (2.7)*	*9[illegible]1 (2.4)*	*827 (4.9)*
Physical activity					
Sufficiently active	25,193 (82.9)	29,522 (79.7)	10,491 (75.9)	2[illegible],[illegible]35 (73.9)	10,724 (63.1)
Insufficiently active	4029 (13.3)	5267 (14.2)	2234 (16.2)	63[illegible]6 (16.4)	3440 (20.2)
Sedentary	673 (2.2)	1157 (3.1)	546 (4.0)	1[illegible]30 (4.7)	1405 (8.3)
Missing	*480 (1.6)*	1109 (3.0)	*555 (4.0)*	*1[illegible]96 (4.9)*	1425 (8.4)
Fruit and vegetable intake					
≥ 5 portions/week	23,460 (77.2)	*28,305 (76.4)*	10,099 (73.0)	28,[illegible]62 (74.3)	*11,672 (68.7)*
< 5 portions/week	6391 (21.0)	*8147 (22.0)*	3492 (25.3)	92[illegible]8 (24.0)	*4947 (29.1)*
Missing	*524 (1.7)*	*603 (1.6)*	*235 (1.7)*	*6[illegible]7 (1.7)*	*375 (2.2)*
Fish intake					
≥ twice/week	15,849 (52.2)	*18,698 (50.5)*	7011 (50.7)	18,[illegible]50 (47.3)	*75676 (44.5)*
Once/week	11,145 (36.7)	18,949 (37.6)	5030 (36.4)	14,[illegible]89 (38.8)	6420 (37.8)
Never	2356 (7.8)	2849 (7.7)	1142 (8.3)	3[illegible]02 (8.3)	1727 (10.2)
Missing	*1025 (3.4)*	*1559 (4.2)*	*643 (4.7)*	*2[illegible]56 (5.6)*	*1280 (7.5)*
History of hypertension	7476 (24.6)	11,316 (30.5)	4357 (31.5)	14,[illegible]10 (37.9)	6940 (40.8)
History of heart disease	1558 (5.1)	2829 (7.6)	1073 (7.8)	3[illegible]87 (9.6)	2234 (13.2)
History of diabetes mellitus	1254 (4.1)	2114 (5.7)	904 (6.5)	3[illegible]26 (7.8)	2016 (11.9)
Family history of stroke/heart disease	18,131 (59.7)	22,349 (60.3)	7749 (56.1)	23,[illegible]57 (61.3)	10,186 (60.0)
Missing	*1 (0.0)*	*1 (0.0)*	*3 (0.02)*	*[illegible] (0.01)*	*5 (0.03)*
Treated for high cholesterol	3078 (10.1)	4747 (12.8)	1763 (12.8)	59[illegible]7 (15.5)	3097 (18.2)
Charlson Comoridity Index					

0	28,614 (94.2)	34,038 (91.2)	12,593 (91.1)	34,[illegible]70 (89.1)	14,229 (83.7)
1	978 (3.2)	1722 (4.7)	660 (4.8)	2[illegible]78 (6.2)	1520 (8.9)
2	316 (1.0)	587 (1.6)	302 (2.2)	9[illegible]2 (2.4)	620 (3.7)
≥3	467 (1.5)	708 (1.9)	271 (2.0)	9[illegible]7 (2.4)	625 (3.7)
Menopausal status					
Pre-menopause	8083 (26.6)	7454 (20.1)	2777 (20.1)	43[illegible]9 (11.3)	1341 (7.9)
Post-menopause	15,025 (49.5)	17,861 (48.2)	6769 (49.0)	1[illegible]36 (50.9)	8233 (48.5)
Hysterectomy only	3997 (13.2)	6896 (18.6)	2442 (17.7)	[illegible]2 (22.3)	4294 (25.3)
Bilateral oophorectomy post-menopause	405 (1.3)	593 (1.6)	213 (1.5)	7[illegible]9 (1.9)	324 (1.9)
Bilateral oophorectomy (surgical menopause)	1319 (4.3)	2374 (6.4)	836 (6.1)	3[illegible]65 (7.9)	1584 (9.3)
Missing	*1546 (5.1)*	1877 (5.1)	*789 (5.7)*	*2[illegible]86 (5.7)*	1218 (7.2)
Current HRT use	3097 (10.2)	3878 (10.5)	1284 (9.3)	[illegible]4 (10.2)	1698 (10.0)
Missing	*97 (0.3)*	*139 (0.4)*	*60 (0.4)*	*1[illegible]7 (0.4)*	*95 (0.6)*
Current OCP use	106 (0.4)	103 (0.3)	48 (0.4)	[illegible]7 (0.2)	26 (0.2)
Missing	*1994 (6.6)*	*3193 (8.6)*	*1392 (10.1)*	*4[illegible]68 (11.6)*	*2747 (16.2)*

BMI = body mass index; HRT = hormone replacement therapy; OCP = oral contraceptive; SD = standard deviation

Supplementary Table 5 Baseline characteristics, by occurrence of stroke, by sex

Characteristic	Men		Women	
	Stroke (N = 2031)	No stroke (N =114,779)	Stroke (N = 1528)	No stroke (N =135,319)
Age, years (mean ± SD)	74.2 (± 10.4)	63.1 (± 10.9)	76.1 (± 11.5)	61.3 ± 10.8)
Categorical age, years				
40-59	225 (11.1)	50,148 (43.7)	184 (12.0)	70,262 (51.9)
60-69	412 (20.3)	33,647 (29.3)	225 (14.7)	36,581 (27.0)
70-79	651 (32.1)	19,758 (17.2)	392 (25.7)	17,783 (13.1)
80+	743 (36.6)	11,226 (9.8)	727 (47.6)	10,693 (7.9)
Remoteness area*				
Major cities	1079 (53.1)	57,687 (50.3)	824 (53.9)	66,306 (49.0)
Inner regional	570 (28.1)	35,572 (28.4)	403 (26.4)	40,357 (29.8)
Outer regional	132 (6.5)	7791 (6.8)	117 (7.7)	9842 (7.3)
Remote/very remote	**	236 (0.2)	**	363 (0.3)
Missing	*249 (12.3)*	*16,493 (14.4)*	*182 (11.9)*	*18,451 (13.6)*
SEIFA index of relative disadvantage				
1 (least deprived)	440 (21.7)	23,422 (20.4)	300 (19.6)	26,310 (19.4)
2	352 (17.3)	22,304 (19.4)	314 (20.6)	26,589(19.7)
3	397 (19.6)	23,010 (20.1)	295 (19.3)	27,239 (20.1)
4	417 (20.5)	23,043 (20.1)	299 (19.6)	27,876 (20.6)
5 (most deprived)	424 (20.9)	22,899 (20.0)	320 (20.9)	27,203 (20.1)
Missing	1 (0.1)	101 (0.1)	0 (0)	102 (0.1)
Marital status				
Married/de facto	1460 (71.9)	92,751 (80.8)	755 (49.4)	95,868 (70.9)
Divorced/separated/Widowed	424 (20.9)	14,212 (12.4)	669 (43.8)	31,676 (23.4)
Single	129 (6.4)	6793 (5.9)	100 (6.5)	7372 (5.5)
Missing	18 (0.9)	1023 (0.9)	4 (0.3)	403 (0.3)
Psychological distress				
Low	1216 (59.9)	81,736 (71.2)	753 (49.3)	89,032 (65.6)
Moderate	225 (11.1)	15,593 (13.6)	195 (12.8)	20,706 (15.3)
High/very high	123 (6.1)	6782 (5.9)	114 (7.5)	9698 (7.2)
Missing	467 (23.0)	10,668 (9.3)	466 (30.5)	15,883 (11.7)
BMI, kg/m^2 (mean ± SD)	26.5 (± 4.3)	27.2 (± 4.2)	26.0 (± 5.5)	26.7 (± 5.3)
Missing	160 (7.9)	6920 (6.0)	182 (11.9)	11,105 (8.2)

Smoking status				
Never	852 (42.0)	56,183 (49.0)	1074 (70.3)	87,029 (64.3)
Former	1011 (49.8)	49,235 (42.9)	361 (23.6)	38,303 (28.3)
Current	155 (7.6)	8725 (7.6)	84 (5.5)	9361 (6.9)
Missing	*13 (0.6)*	*636 (0.6)*	*9 (0.6)*	*626 (0.5)*
Alcohol intake				
Moderate	843 (41.5)	49,937 (43.5)	339 (22.2)	41,694 (30.8)
None/rarely	534 (26.3)	25,942 (22.6)	770 (50.4)	53,803 (39.8)
Hazardous	450 (22.2)	27,002 (23.5)	291 (19.0)	30,694 (22.7)
Harmful	153 (7.5)	10,224 (809)	41 (2.7)	6088 (4.5)
Missing	*51 (2.5)*	*1674 (1.5)*	*87 (5.7)*	*3010 (2.3)*
Physical activity				
Sufficiently active	1343 (66.1)	87,583 (76.3)	837 (54.8)	103,728 (76.6)
Insufficiently active	449 (22.1)	18,995 (16.6)	378 (24.7)	20,928 (15.5)
Sedentary	135 (6.7)	4683 (4.1)	159 (10.4)	5452 (4.0)
Missing	104 (5.1)	3518 (3.1)	154 (10.1)	5311 (3.9)
Fruit and vegetable intake				
≥ 5 portions/week	*1162 (57.2)*	*66,037 (57.5)*	*392 (25.7)*	*31,853 (23.5)*
< 5 portions/week	*826 (40.7)*	*47,048 (41.0)*	*1093 (71.5)*	*101,105 (74.7)*
Missing	*43 (2.1)*	*1694 (1.5)*	*43 (2.8)*	*2361 (1.7)*
Fish intake				
≥ twice/week	978 (48.2)	51,974 (45.3)	765 (50.1)	66,610 (49.2)
Once/week	785 (38.7)	48,633 (42.4)	531 (34.8)	51,002 (37.7)
Never	129 (6.4)	8241 (7.2)	113 (7.4)	11,163 (8.3)
Missing	*139 (6.8)*	*5931 (5.2)*	*119 (7.8)*	*6544 (4.8)*
History of hypertension	953 (46.9)	42,370 (36.9)	747 (48.9)	43,952 (32.5)
History of heart disease	648 (31.9)	18,485 (16.0)	401 (26.2)	10,980 (8.1)
History of diabetes mellitus	345 (17.0)	11,744 (10.2)	206 (13.5)	9108 (6.7)
Treated for high cholesterol	343 (16.9)	18,213 (15.9)	260 (17.0)	18,422 (13.6)
Family history of CVD	1131 (55.7)	62,103 (54.1)	974 (63.7)	81,098 (59.9)
Missing	*0 (0)*	*20 (0.02)*	*0 (0)*	*12 (0.01)*
Charlson comorbidity index				
0	1432 (70.5)	99,581 (86.8)	1097 (71.8)	122,747 (90.7)
1	262 (12.9)	8398 (7.3)	199 (13.0)	7059 (5.2)
2	149 (7.3)	3425 (3.0)	128 (8.4)	2619 (1.9)
≥ 3	188 (9.3)	3375 (2.9)	104 (6.8)	2894 (2.1)

WOMEN ONLY				
Menopausal status				
Pre-menopausal	-	-	62 (4.1)	23,962 (17.7)
Post-menopausal	-	-	827 (54.1)	66,697 (42.3)
Hysterectomy only	-	-	351 (23.0)	25,880 (19.1)
Bilateral oophorectomy post-menopause	-	-	38 (2.5)	2236 (1.7)
Bilateral oophorectomy (surgical menopause)	-	-	138 (9.0)	9040 (6.7)
Missing	-	-	*112 (7.3)*	*7504 (5.6)*
Current HRT use	-	-	96 (6.2)	13,816 (10.2)
Missing	-	-	*11 (0.7)*	*537 (0.4)*
Current OCP use	-	-	1 (0.07)	*339 (0.3)*
Missing	-	-	*252 (16.5)*	*13,542 (10.0)*

[*]Number of stroke events not disclosed due to small numbers and potential for de-identification of participants
AUD = Australian dollars; BMI = body mass index; CVD = cardiovascular disease (i.e. stroke or heart disease); HRT = hormone replacement therapy; OCP = oral contraceptive pill; SD = standard deviation; SEIFA = Socio-Economic Indexes of Australia

Supplementary Table 6 Age-standardised incidence rates of stroke (per 1000 person-years), by sex and age-category and age-adjusted hazard ratios for stroke risk using college/university degree as the reference category

Education level	Person-years	Stroke events, N	Incidence rate, per 1000 person-years* (95% CI)	Hazard ratio (95% CI)
Men aged 45-69 (N = 84,432)				
College/university degree	115,340	127	1.03 (0.89 to 1.17)	Reference
Certificate/diploma/trade or apprenticeship	154,6375	244	1.36 (1.22 to 1.50)	1.30 (1.05 to 1.[illegible]2)
Higher school certificate	40,404	73	1.63 (1.33 to 1.92)	1.57 (1.18 to 2.[illegible]1)
School/intermediate certificate	56,542	115	1.45 (1.23 to 1.67)	1.51 (1.17 to 1.[illegible]5)
No qualifications	34,863	78	1.70 (1.36 to 2.04)	1.58 (1.19 to [illegible].[illegible]0)
Men aged ≥70 (N = 32,378)				
College/university degree	25,908	225	8.66 (8.39 to 8.92)	Reference
Certificate/diploma/trade or apprenticeship	55,737	540	10.10 (9.90 to 10.32)	1.14 (0.98 to 1.[illegible]4)
Higher school certificate	12,970	127	9.09 (8.73 to 9.46)	1.08 (0.87 to 1.[illegible]4)
School/intermediate certificate	26,418	283	10.25 (9.97 to 10.52)	1.21 (1.02 to 1.[illegible]4)
No qualifications	22,633	219	9.64 (9.35 to 9.94)	1.12 (0.93 to 1.[illegible]5)
Women aged 45-69 (N = 107,252)				
College/university degree	130,110	64	0.53 (0.44 to 0.63)	Reference
Certificate/diploma/trade or apprenticeship	146,594	116	0.77 (0.67 to 0.87)	1.49 (1.10 to 2.[illegible]3)
Higher school certificate	50,895	36	0.69 (0.52 to 0.85)	1.31 (0.88 to 1.[illegible]8)
School/intermediate certificate	133,587	138	0.89 (0.77 to 1.01)	1.67 (1.24 to 2.[illegible]6)
No qualifications	49,999	55	0.93 (0.73 to 1.13)	1.73 (1.20 to 2.[illegible]9)
Women aged ≥70 (N = 29,595)				
College/university degree	13,665	96	8.63 (8.10 to 9.16)	Reference
Certificate/diploma/trade or apprenticeship	28,374	217	9.25 (8.87 to 9.64)	1.03 (0.81 to 1.[illegible]1)
Higher school certificate	14,023	127	9.28 (8.81 to 9.76)	1.08 (0.83 to 1.[illegible]1)
School/intermediate certificate	48,086	386	8.98 (8.70 to 9.25)	1.02 (0.81 to 1.[illegible]7)
No qualifications	29,030	293	11.32 (10.93 to 11.70)	1.21 (0.96 to 1.[illegible]2)

*Age-standardised to the Australian sex-specific standard population

Supplementary Table 7 Results of additive interaction calculation showing: hazard ratios for stroke, by education and sex and effect modification by male sex; and hazard ratios for stroke, by education and age and effect modification by age ≥ 70, stratified by sex

Interaction	Male sex	No qualifications	Stroke events/person-years	HR (95%CI)	RERI (95% CI); Synergy index (95% CI)
Male sex by education level*	No	No	160 / 143,774	Reference	
	Yes	No	348 / 79,029	1.49 (1.23, 1.80)	-0.17 (-0.[illegible]2, 0.18); 0.82 (0.57, 1.19)
	No	Yes	352 / 141,248	1.47 (1.21, 1.79)	
	Yes	Yes	297 / 57,496	1.79 (1.46, 2.18)	

Interaction	Age ≥ 70	No qualifications	Stroke events/person-years	HR (95%CI)	RERI (95% CI); Synergy index (95% CI)
Age by education level* (for men and women separately)	MEN				
	No	No	127 / 115,340	Reference	
	Yes	No	78 / 34,863	2.03 (1.53, 2.69)	-0.23 (-0.56, 0.10); 0.30 (0.02, 4.73)
	No	Yes	225 / 25, 908	7.95 (6.40, 9.89)	
	Yes	Yes	219 / 22,633	8.88 (7.14, 11.05)	
	WOMEN				
	No	No	64 / 130,110	Reference	
	Yes	No	55 / 49,999	2.23 (1.56, 3.20)	0.004 (-0.39, 0.40); 1.01 (0.45, 2.27)
	No	Yes	96 / 70,254	14.34 (10.45, 19.68)	
	Yes	Yes	293 / 29,030	20.61 (15.72, 27.01)	

*Comparing the lowest educational attainment category ('no qualifications') versus the highest educational attainment category (similar results, with more precision, were obtained when the analyses were repeated, but comparing the lowest educational attainment category versus all other educational categories combined)

RERI = relative excess risk due to interaction

Supplementary Table 8 Serially adjusted hazard ratios from the complete-case analysis for the association between education distress and stroke, stratified by gender

	Education				
	College/university degree	Certificate/diploma/ trade/ apprenticeship	Higher school certificate	School/intermediate certificate	No qualifications
MEN (1151 strokes among N = 77,785)					
Model 1*	1.00	1.19 (1.02, 1.39)	1.09 (0.86, 1.38)	1.24 (1.03, 1.50)	1.29 (1.05, 1.59)
Model 2†	1.00	1.19 (1.01, 1.39)	1.08 (0.86, 1.37)	1.22 (1.01, 1.48)	1.27 (1.03, 1.57)
Model 3‡	1.00	1.17 (1.00, 1.37)	1.07 (0.85, 1.35)	1.20 (0.99, 1.45)	1.23 (1.00, 1.52)
Model 4§	1.00	1.15 (0.98, 1.35)	1.02 (0.81, 1.30)	1.15 (0.95, 1.40)	1.16 (0.94, 1.44)
Model 5‖	1.00	1.14 (0.97, 1.34)	1.01 (0.80, 1.28)	1.13 (0.94, 1.37)	1.14 (0.92, 1.41)
WOMEN (606 strokes among N =77,525)					
Model 1*	1.00	1.36 (1.02, 1.79)	1.32 (0.94, 1.84)	1.30 (0.99, 1.70)	1.53 (1.13, 2.06)
Model 2†	1.00	1.37 (1.04, 1.81)	1.32 (0.95, 1.85)	1.32 (1.01, 1.74)	1.57 (1.16, 2.12)
Model 3‡	1.00	1.36 (1.03, 1.80)	1.31 (0.94, 1.84)	1.30 (0.99, 1.71)	1.42 (1.05, 1.93)
Model 4§	1.00	1.32 (1.00, 1.75)	1.24 (0.88, 1.73)	1.22 (0.92, 1.60)	1.26 (0.92, 1.71)
Model 5‖	1.00	1.29 (0.97, 1.71)	1.22 (0.87, 1.71)	1.18 (0.89, 1.55)	1.20 (0.88, 1.64
Model 6#	1.00	1.28 (0.97, 1.70)	1.21 (0.86, 1.70)	1.16 (0.88, 1.54)	1.19 (0.87, 1.62)

*Adjusted for age

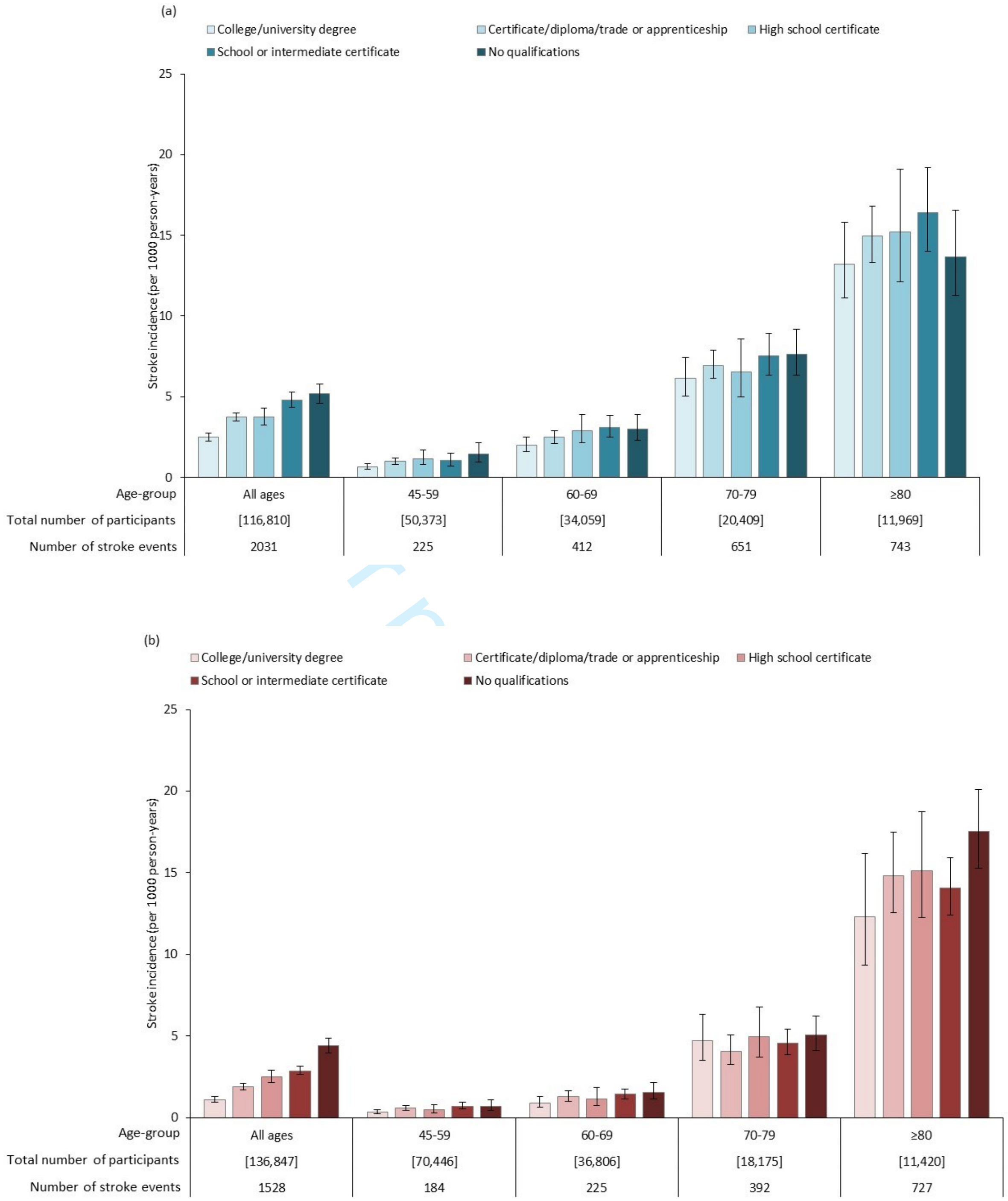
†Adjusted for age, marital status, remoteness

‡Adjusted for age, marital status, remoteness, psychological distress

§Adjusted for age, marital status, remoteness, psychological distress, smoking, alcohol intake, body mass index, physical activity, fruit & vegetable intake, fish consumption

‖Adjusted for age, marital status, remoteness, psychological distress, smoking, alcohol intake, body mass index, physical activity, fruit & vegetable intake, fish consumption, history of hypertension, history of heart disease, history of diabetes, history of treatment for raised cholesterol, Charlson comorbidity index and family history of stroke or heart disease

#In women, adjusted for age, marital status, remoteness, psychological distress, smoking, alcohol intake, body mass index, physical activity, fruit & vegetable intake, fish consumption, history of hypertension, history of heart disease, history of diabetes, history of treatment for raised cholesterol, Charlson comorbidity index, family history of stroke or heart disease, oral contraceptive use, hormone replacement therapy use and menopausal status



Supplementary Figure 1 Stroke incidence, per 1000 person-years, by age-group in (a) men and (b) women

STROBE Statement—Checklist of items that should be included in reports of ***cohort studies***

	Item No	Recommendation	Page number
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1,2
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4, paragraphs 1 & 2
Objectives	3	State specific objectives, including any prespecified hypotheses	4, final paragraph
Methods			
Study design	4	Present key elements of study design early in the paper	5,6
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	5
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	5
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed	NA
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	5, 6
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	5-7 & supplementary Table 1
Bias	9	Describe any efforts to address potential sources of bias	6
Study size	10	Explain how the study size was arrived at	5
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	6 & supplementary Table 1
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	6,7
		(*b*) Describe any methods used to examine subgroups and interactions	7
		(*c*) Explain how missing data were addressed	6
		(*d*) If applicable, explain how loss to follow-up was addressed	NA
		(*e*) Describe any sensitivity analyses	7
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	7 & Figure 1
		(b) Give reasons for non-participation at each stage	Figure 1
		(c) Consider use of a flow diagram	Figure 1
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	8 & supplementary tables 3 & 4
		(b) Indicate number of participants with missing data for each variable of interest	Supplementary tables 3 & 4

		(c) Summarise follow-up time (eg, average and total amount)	8, paragraph 3
Outcome data	15*	Report numbers of outcome events or summary measures over time	8, paragraph 3
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	9, Table 2
		(*b*) Report category boundaries when continuous variables were categorized	Tables
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	8, final paragraph & Table 1
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	9, Paragraphs 2 & 3; 10 first paragraph
Discussion			
Key results	18	Summarise key results with reference to study objectives	Page 10, paragraph 2
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	Page 10, paragraph 3
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	11-12
Generalisability	21	Discuss the generalisability (external validity) of the study results	10, paragraph 3
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	3

*Give information separately for exposed and unexposed groups.

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 http://www.strobe-statement.org.