Appendix e-1 Measurement of repeat exposures

In the main analyses, women were categorised by smoking, alcohol consumption, body mass index and physical activity according to information provided at baseline. To allow for measurement error, regression dilution, and/or changes over time these exposures were re-measured at a later date, and the mean re-measured value in each category was used to score each category for the purposes of trend estimation.

For 59,225 current smokers at baseline who reported excellent to good health, we used the mean number of cigarettes smoked, reported an average of 4 years later (9 cigarettes/day for those reporting <15 cigarettes/day at baseline, and 18 cigarettes/day for those reporting 15+ cigarettes/day at baseline).

For alcohol consumption and physical activity, updated information was obtained for about 30,000 women from a web-based questionnaire completed on average 11 years after baseline.[1] For the three baseline exposure categories for alcohol consumption (≤6, 7-14 and ≥15 drinks per week), the re-measured mean intakes for those in excellent or good health were 7.5, 17.7, 31.7 g of alcohol/ day, and for those reporting fair or poor health the corresponding re-measured values were lower, at 6.1, 16.8, 30.9, respectively. For physical activity, assessed as excess METs tertiles at baseline, the re-measured means were 5.7, 6.8, 8.7 excess METs/ day for those reporting excellent/good and lower, at 5.0, 6.4, 7.5, respectively, for those in fair/poor self-rated health.

For the baseline categories of body mass index (BMI) (<25, 25-29.9, and 30+ kg/m^2) direct measurements were available 6 years later for 3181 women.[2] For women reporting excellent to good health, the mean measured values for each category were 23.6, 28.2 and 33.9 kg/m^2 and for those in fair to poor health, the values were higher at 23.9, 28.8 and 35.6 kg/m^2.

Systematic Review and Meta-analysis

Appendix e-2 ***Search strategy, eligibility criteria and statistical methods***

In August 2017, we searched Pubmed and EMBASE for relevant articles in English published after 1990, using combinations of the following MeSH and free text terms: ischaemic stroke, intracranial haemorrhages, intracerebral haemorrhage, subarachnoid haemorrhage, diabetes mellitus, tobacco use, smoking and cigarettes. Titles and abstracts

of identified articles were initially screened for relevance. Potentially relevant articles were assessed for eligibility using the following inclusion criteria: cohort study, prospective collection of exposure data, relating smoking or diabetes to incident stroke (fatal and non-fatal), simultaneous reporting of separate estimates for all three major stroke types and accounting for potential confounders or effect modifiers (at least age and sex where relevant). In addition we hand searched review articles for other relevant papers.

Our search for prospective studies reporting on diabetes and risk of each of the three types of stroke yielded 733 and 540 potentially relevant studies, from Pubmed and EMBASE, respectively. Screening of titles and abstracts excluded 1225 studies. After assessing full-text articles of 48 studies, a further 46 were excluded (Figure e-1). This left only two eligible studies, one from Europe[3] and one from Asia[4] (Table e-1). In our meta-analysis, relative risk estimates were based on 18,567 stroke events in total: 912 from a study in Japan and 5959 from a study in the UK and a further 11,696 cases from the Million Women Study.

Our search for prospective studies reporting on current smoking and risk of each of the three types of stroke yielded 670 and 328 potentially relevant studies, from Pubmed and EMBASE, respectively. Screening titles and abstracts eliminated 953 and assessing full-text articles eliminated a further 45, which did not fulfil one or more inclusion criteria (Figure e-2). In total, we identified two eligible studies from Europe and North America[5,6] and three from Asia,[7-9] (Table e-2). In our meta-analysis, relative risk estimates were based on 30,903 stroke events in total: 9,052 from studies in Asia and 10,155 from previously published studies in Europe or North America and 11,696 from the Million Women Study.

Where several relative risk estimates were reported in the same publication, we chose the estimate that had adjusted for the largest number of potential confounders, but not for potential mediators, such as adiposity. Where relative risk estimates were published for subgroups of the exposure but not for the whole exposed group (e.g. separate estimates for the amount smoked in current smokers, but not an estimate for all current smokers) the relative risks were combined using generalized least squares.[10]

Study-specific results were combined to calculate a summary estimate by inverse-variance methods.[11] Results for current smoking are shown separately for studies in Asia and Europe or North America, because the health

consequences of the smoking epidemic has not yet had its full effect in Asia and also because the distribution of stroke types differs between populations of Asian and European origin.[12]

References

1. Liu B, Young H, Crowe FL, et al. Development and evaluation of the oxford WebQ, a low-cost, web-based method for assessment of previous 24 h dietary intakes in large-scale prospective studies. *Public Health Nutr*. 2011;14(11):1998-2005.

2. Wright F.L., Green J., Reeves G., Beral V., Cairns B.J. Validity over time of self-reported anthropometric variables during follow-up of a large cohort of UK women. *BMC Medical Research Methodology*. 2015;15 (1).

3. Shah AD, Langenberg C, Rapsomaniki E, et al. Type 2 diabetes and incidence of cardiovascular diseases: A cohort study in 1.9 million people. *Lancet Diabetes Endocrinol*. 2015;3 (2):105-114.

4. Cui R., Iso H., Yamagishi K., et al. Diabetes mellitus and risk of stroke and its subtypes among Japanese: The Japan Public Health Center Study. *Stroke*. 2011;42 (9):2611-2614.

5. Pujades-Rodriguez M, George J, Shah AD, et al. Heterogeneous associations between smoking and a wide range of initial presentations of cardiovascular disease in 1937 360 people in England: Lifetime risks and implications for risk prediction. *Int J Epidemiol*. 2015;44 (1):129-141.

6. Kawachi I, Colditz GA, Stampfer MJ, et al. Smoking cessation and decreased risk of stroke in women. *J Am Med Assoc*. 1993;269(2):232-236.

7. Mannami T, Iso H, Baba S, et al. Cigarette smoking and risk of stroke and its subtypes among middle-aged Japanese men and women: The JPHC study cohort I. *Stroke*. 2004;35 (6):1248-1253.

8. Lawlor DA, Song YM, Sung J, Ebrahim S, Smith GD. The association of smoking and cardiovascular disease in a population with low cholesterol levels: A study of 648,346 men from the Korean National Health System prospective cohort study. *Stroke*. 2008;39 (3):760-767.

9. Hata J, Doi Y, Ninomiya T, et al. Combined effects of smoking and hypercholesterolemia on the risk of stroke and coronary heart disease in Japanese: The Hisayama study. *Cerebrovasc Dis*. 2011;31(5):477-484.

10. Greenland S, Longnecker MP. Methods for trend estimation from summarized dose-response data, with applications to meta-analysis. *Am J Epidemiol*. 1992;135 (11):1301-1309.

11. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med*. 2002;21 (11):1539-1558.

12. Tsai C-, Thomas B, Sudlow CLM. Epidemiology of stroke and its subtypes in Chinese vs white populations. *Neurology*. 2013;81 (3):264-272.