

# BMJ Open Cohort profile of the UK Biobank: diagnosis and characteristics of cerebrovascular disease

J Hewitt,<sup>1,2</sup> M Walters,<sup>2</sup> S Padmanabhan,<sup>2</sup> J Dawson<sup>2</sup>

**To cite:** Hewitt J, Walters M, Padmanabhan S, *et al*. Cohort profile of the UK Biobank: diagnosis and characteristics of cerebrovascular disease. *BMJ Open* 2016;**6**:e009161. doi:10.1136/bmjopen-2015-009161

► Prepublication history for this paper is available online. To view these files please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2015-009161>).

Received 29 June 2015  
Revised 2 February 2016  
Accepted 8 February 2016



CrossMark

<sup>1</sup>Department of Geriatric Medicine, Division of Population Medicine, Cardiff University, Cardiff, UK

<sup>2</sup>Institute of Cardiovascular and Medical Sciences, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, UK

## Correspondence to

Dr J Hewitt;  
hewittj2@cardiff.ac.uk

## ABSTRACT

**Purpose:** The UK Biobank is a large-scale biomedical resource, containing sociodemographic and medical information, including data on a previous diagnosis of stroke or transient ischaemic attack (TIA). We described these participants and their medication usage.

**Participants:** We identified participants who either self-reported or were identified from a nurse-led interview, having suffered a stroke or a TIA and compared them against participants without stroke or TIA. We assessed their risk factor burden (sex, age, deprivation, waist to hip ratio (WHR), hypertension, smoking, alcohol intake, diabetes, physical exercise and oral contraception use (oral contraceptive pill, OCP)) and medication usage.

**Findings to date:** We studied 502 650 people (54.41% women), 6669 (1.23%) participants self-reported a stroke. The nurse-led interview identified 7669 (1.53%) people with stroke and 1781 (0.35%) with TIA. Hypertension, smoking, higher WHR, lower alcohol consumption and diabetes were all more common in people with cerebrovascular disease ( $p < 0.0001$  for each). Women with cerebrovascular disease were less likely to have taken the OCP ( $p = 0.0002$ ). People with cerebrovascular disease did more exercise ( $p = 0.03$ ). Antithrombotic medication was taken by 81% of people with stroke (both self-report and nurse-led responders) and 89% with TIA. For self-reported stroke, 63% were taking antithrombotic and cholesterol medications, 54% taking antithrombotic and antihypertensive medications and 46% taking all 3. For the nurse-led interview and TIA, these figures were 65%, 54% and 46%, and 70%, 53% and 45%, respectively.

**Future plans:** The UK Biobank provides a large, generalisable and contemporary data source in a young population. The characterisation of the UK Biobank cohort with cerebrovascular disease will form the basis for ongoing research using this data source.

## INTRODUCTION

The UK Biobank is a national health resource, intended to improve the prevention, diagnosis and treatment of illness.<sup>1</sup> It holds detailed demographic, social and

## Strengths and limitations of this study

- These data represent a very large sample.
- This self-reported data needs additional clinical correlation.

medical information, along with physical measures, such as weight and blood pressure. Participants gave consent to have their future health records linked to their UK Biobank data, a process which is ongoing.

There is comparatively little epidemiological data on younger populations with stroke.<sup>2</sup> The UK Biobank obtained information regarding previous stroke or transient ischaemic attack (TIA) and commonly associated medications (such as blood pressure and lipid-lowering medications) in people aged 40–69 years. It is an order of magnitude larger than previous self-reported estimates of cerebrovascular disease. We aimed to describe the cohort of participants from the UK Biobank who self-reported a diagnosis of stroke and TIA. We aimed to describe the medication data of the cohort.

**Data collection and follow-up:** Between 2006 and 2010, the UK Biobank collected detailed data on 502 650 people. They used 22 assessment centres based in England, Scotland and Wales. Participants were recruited from National Health Service (NHS) patient registers and contacted if they lived within a reasonable proximity to an assessment centre.<sup>3 4</sup>

Participants provided detailed demographic, socioeconomic and health-related data via a touchscreen questionnaire, including medication history. Participants then underwent a range of physical assessment measures, included repeated blood pressure measurements, height and weight. Finally participants provided blood, urine and saliva samples which have been held in a purpose-built central storage facility, in Stockport, UK, for future analysis. These depletable

samples are due to be characterised by a comprehensive range of biomarkers by the end of 2015, but are currently not readily accessible.

One of the touchscreen questionnaire responses was 'Has a doctor ever told you that you have any of the following conditions? (You can select more than one answer)'. The possible touchscreen responses were 'heart attack, angina, stroke, high blood pressure, none of the above, prefer not to answer'.

Following the questionnaire, a nurse-led interview was performed to address medical history in general. In order to prompt the nurses conducting the questionnaire and improve accuracy, all responses from the touchscreen questionnaire were flagged to the study nurse. Therefore, if a participant had selected 'stroke' from the touchscreen question listed above, then the nurse would have been aware of this response when discussing medical history. The nurses also asked individuals directly about their medical history. Therefore, it is possible that someone may not have responded positively to the touchscreen questionnaire question regarding stroke but may have done so when asked later by the study nurse about medical history. It is also during the nurse-led questionnaire that participants would have been identified as having a medical history of TIA, when questioned about their medical history. Further detailed information regarding conducting the nurse-led questionnaire can be found at the UK Biobank website.<sup>5</sup>

Participants were also asked to self-report via the touchscreen questionnaire their medication history. They were asked specifically if they regularly took 'medication for cholesterol' or 'blood pressure medication' and could answer positively for both responses.

During the nurse-led interview, all medications were recorded as free text and subsequently grouped and coded by UK Biobank. We extracted data on reported antiplatelet or anticoagulant use. Any participant using any single or combination of these drugs was categorised as taking antithrombotic medication.

We also extracted demographic data concerning age, sex and socioeconomic class, measured using the Townsend scale.<sup>6</sup> In order to select the most appropriate potential risk factors for stroke, we used data from the INTERSTROKE study<sup>7</sup> and the American Heart Association's Stroke Councils Scientific Statement Oversight Committee guideline (AHA).<sup>8</sup> Review of these publications identified seven modifiable risk factors that could accurately be reproduced from the UK Biobank. These risk factors represent the majority of the modifiable risk factors contributing to cerebrovascular disease.

From the INTERSTROKE study, we identified hypertension, smoking, waist to hip ratio (WHR), regular physical activity, diabetes mellitus and alcohol intake as risk factors we could study. We could not include psychological factors, diet, cardiac causes and apolipoprotein levels as the data were not available in comparable format. From the AHA guideline, we also identified the oral contraceptive pill (OCP). We used the same

definitions as INTERSTROKE and the AHA guideline as far as possible; we defined a self-reported history of hypertension, smoking (current, ex or never), WHR (divided into tertiles), moderate or strenuous activity for more than 4 h/week and a self-reported history of diabetes mellitus. For women, a history of ever having taken the OCP (yes or no).

Alcohol intake was based on the data recorded in the UK Biobank (alcohol daily or almost daily, alcohol on 3 or 4 days/week, alcohol once or twice a week, alcohol one to three times per month, special occasions only, never drinkers and former drinkers).

We describe the characteristics of the whole population and the population with stroke disease and TIA using descriptive statistics. We compared characteristics of patients who self-reported stroke, those with nurse-reported stroke, and TIA against those who did not using *t* tests for continuous variables and  $\chi^2$  tests for dichotomous variables. The statistical analysis was conducted using STATA V.11 (Stata Corp).

### Cerebrovascular findings to date

UK Biobank released data for use on 502 650 participants. Their average age was 56.52 years, (SD 8.09). There were 273 468 women (54.41%) who were younger than the men (56.35 years (SD 8.00) compared with 56.74 years (SD 8.20),  $p < 0.0001$ ). The mean Townsend Index for all participants was  $-1.29$  (SD 3.10).

More participants self-reported stroke during the nurse-led interview ( $n=7669$  (1.53%)) than on the touchscreen questionnaire ( $n=6699$  (1.33%)). TIA was reported via the verbal questionnaire in 1781 (0.35%) of participants. Patients self-reporting stroke and TIA were older and more likely to be female ( $p < 0.0001$ ) in comparison to patients who did not. Higher levels of deprivation were seen in people self-reporting cerebrovascular disease via both the touchscreen and nurse-led questionnaire ( $p < 0.0001$ ), but this was not the case for reports of TIA ( $p=0.13$ ). The full results are shown in [table 1](#).

The cardiovascular risk factors in patients self-reporting stroke and TIA are shown in [table 2](#) along with patients who reported no stroke or TIA.

Self-reported hypertension was recorded in 27.01% of the whole UK Biobank population. It was over 50% for each of the patients with a history of self-reported stroke, stroke from the nurse-led questionnaire and TIA. People with any cerebrovascular history were more likely to be former or current smokers. Diabetes was reported in 5.25% of the UK Biobank population but over 15% of patients with stroke and 10% of patients with TIA. Each of these associations was significant ( $p < 0.001$ , for each).

Self-reported alcohol intake was lower in people with cerebrovascular disease in comparison to the population without cerebrovascular disease. In people who drank less than weekly, including never and former drinkers, there was a higher proportion of cerebrovascular disease

**Table 1** Self-report for stroke and transient ischaemic attack (TIA) with demographics

	Sex		Age (SD)	Deprivation (SD)	
	Men	Women			
Self-reported stroke via touchscreen questionnaire					
Yes	6699 (1.33)	3983 (1.74)	2716 (0.99)	60.8 (6.78)	-0.25 (3.54)
No	495 951 (98.67)	225 199 (98.26)	270 458 (99.01)	57.4 (6.78)	-1.25 (3.11)
Self-reported stroke via nurse interview					
Yes	7669 (1.53)	4516 (1.97)	3153 (1.15)	60.89 (6.73)	-0.42 (3.48)
No	494 051 (98.29)	224 205 (97.83)	269 852 (98.68)	56.4 (8.09)	-1.31 (3.08)
Missing (or not completed)	930 (0.19)	461 (0.2)	469 (0.17)		
Self-reported TIA via nurse interview					
Yes	1781 (0.35)	958 (0.42)	823 (0.3)	62.19 (5.87)	-1.31 (3.05)
No	499 939 (99.46)	227 763 (99.38)	272 176 (99.53)	57.19 (7.9)	-1.22 (3.12)
Missing (or not completed)	930 (0.19)	461 (0.2)	469 (0.17)		
Number (%)	502 650 (100)	229 182 (100)	273 468 (100)		

recorded ( $p < 0.001$ ) compared with those who drank more frequently.

For women, there were 81.05% who had previously taken the OCP. For self-reported stroke and nurse-led stroke responses, these figures were 73.83% and 74.22%, respectively ( $p < 0.001$ ), and for TIA, this figure was 75.55% ( $p = 0.0002$ ).

There were 23 823 (4.74%) people who reported doing at least 4 or more hours of moderate or strenuous physical exercise per week. Of the 303 897 people with complete data for this response 302 (6.4%) of those who had a previous stroke identified via the touchscreen

questionnaire responded positively to this question (vs 17 458 (5.84%),  $p = 0.1$ , for those without a previous stroke). For the nurse-led self-report question, these figures were 357 (6.47%) and 23 463 (5.78%),  $p = 0.03$ , and for self-reported TIA, they were 91 (6.45%) vs 17 669 (5.84%),  $p = 0.32$ .

In total, 76 397 (15.20%) people who were identified as taking an antithrombotic medication, 104 028 (20.70%) of people self-reported as receiving a blood pressure medication and 86 907 (17.29%) of people self-reported receiving a cholesterol medication. For each of these responses participants who reported

**Table 2** Descriptive characteristics ( $p < 0.001$ , unless stated)

	Total UK Biobank	Stroke diagnosed by touchscreen questionnaire	Stroke diagnosed by nurse-led questionnaire	Transient ischaemic attack (nurse-led questionnaire)
	N=502 650 (100%)	N=6699	N=7669	N=1781
Hypertension (self-reported)	135 787 (27.06)	3770 (56.34)	4257 (55.51)	912 (51.21)
Smoking				
Current	52 989 (10.92)	1103 (16.48)	1223 (15.95)	191 (10.72)
Former	173 203 (34.52)	2851 (42.60)	3250 (42.38)	731 (41.04)
Never	275 565 (54.92)	2738 (40.91)	3196 (41.67)	859 (48.23)
Waist to hip ratio (tertiles)				
Lowest	NA	1181 (17.63)	1375 (17.93)	399 (22.4)
Middle	NA	1987 (29.66)	2308 (30.10)	574 (32.23)
Highest	NA	3470 (51.80)	3913 (51.02)	801 (44.97)
Self-report diabetes				
Yes	26 408 (5.26)	1042 (15.57)	1151 (15.01)	178 (9.99)
Alcohol frequency				
Daily or almost daily	101 794 (20.29)	1217 (18.19)	1437 (18.74)	389 (21.84)
3–4 times/week	115 459 (23.01)	1100 (16.44)	1280 (16.69)	340 (19.09)
1–2/week	129 325 (25.77)	1560 (23.31)	1800 (23.47)	406 (22.8)
Upto 3/month	55 871 (11.14)	713 (10.65)	821 (10.71)	200 (11.23)
Special occasions	58 032 (11.57)	1010 (15.09)	1132 (14.76)	233 (13.08)
Never	22 551 (4.49)	476 (7.11)	529 (6.9)	108 (6.06)
Former	18 115 (3.61)	597 (8.92)	651 (8.49)	104 (5.84)
Oral contraceptive pill (N=272 047)				
Yes	220 501 (81.05)	1989 (73.83)	2326 (74.22)	621 (75.55) ( $p = 0.002$ )

NA, not applicable.

**Table 3** Medication history

	Total UK	Stroke diagnosed by touchscreen	Stroke diagnosed by nurse-led	Transient ischemic attack (nurse-led questionnaire)
	Biobank	questionnaire	questionnaire	
	N=502 650 (100%)	N=6699	N=7669	N=1781
Antithrombotic medications	76 397 (15.20)	5414 (80.82)	6243 (81.41)	1577 (88.55)
Cholesterol medications	86 907 (17.29)	4728 (70.58)	5466 (71.27)	1303 (73.16)
Blood pressure (BP) medications	104 027 (20.70)	4107 (61.31)	4681 (61.04)	996 (55.92)
Antithrombotic and cholesterol	45 525 (9.06)	4267 (63.70)	4958 (64.65)	1238 (69.51)
Antithrombotic and BP	42 650 (8.48)	3582 (53.47)	4107 (53.55)	930 (52.22)
Antithrombotic, cholesterol and BP	32 374 (6.44)	3063 (45.72)	3532 (46.06)	794 (44.58)

stroke or TIA disease all received more of these medications than the population as a whole, every response was significant (<0.0001). The full results are given in [table 3](#).

## DISCUSSION

The objective of this study was to characterise the UK Biobank participants with a history of stroke and TIA and to gain insights into the treatment of patients at population level in the UK.

There are surprisingly few studies documenting the prevalence of stroke and TIA disease. For example, a 2012 review<sup>2</sup> identified only five studies of stroke prevalence in the Western world. They estimated that for an adult population, including older populations, the prevalence of stroke ranged from 0.15% in Italy to between 1.7% and 2.6% in the UK and the USA. When these figures were assessed with respect to sex, there was a suggestion that prevalence was higher in men. Age is by far the largest non-modifiable risk factor for stroke making the findings in our younger cohort (1.74% for men and 0.99% for women) comparable.

Our findings noted a difference in reporting of prevalence between the self-reported and nurse-led figures for cerebrovascular disease. There were 970 more people identified with stroke disease following the nurse-led questionnaire. Every person who responded positively on the touchscreen questionnaire to having stroke disease will have been challenged by the nurse during the nurse-led interview. All of these people were included in the nurse-led responses as having a stroke; therefore, these data have been at least partially verified by the nurse. The nurse also further identified people with cerebrovascular disease via questions on medical history. We would advocate that the figure (7669) is the most accurate estimation of cerebrovascular disease within the UK Biobank.

Reduced socioeconomic status was associated with both self-reported and nurse-reported stroke and TIA. Deprivation and stroke are associated with an increase in mortality across a range of countries<sup>9</sup> and also a Scottish

meta-analysis linked deprivation with stroke incidence even after allowing for cardiovascular risk factors.<sup>10</sup> However, our data are an order of magnitude larger than the studies included in that review and add a contemporary UK-wide assessment of stroke and deprivation.

When selecting the risk factors to study from the UK Biobank, we attempted to use the risk factors characterised in other studies. The advantages of this approach are that the reproducibility of our study is increased and our results can be compared directly to those generated from the high-income countries included in their studies. We were able to reproduce six of the risk factors used in INTERSTROKE study and the AHA guideline.

Hypertension was reported by 27% of the whole cohort but in over 50% of the participants with cerebrovascular disease. These figures are consistent with other epidemiological studies.<sup>11</sup> Our findings reflect both the causative effect of blood pressure and a heightened awareness of hypertension treatment in stroke survivors. Unsurprisingly, our results show an increased number of current smokers and former smokers and a reduced rate of never-smokers in those reporting cerebrovascular disease.

UK Biobank WHR data can be closely compared with the INTERSTROKE study data. This is because we were able to conduct our analysis using tertiles and also because while weight may have altered following a diagnosis of cerebrovascular disease, it is likely to be one of the most challenging of modifiable risk factors to change following diagnosis, although data on poststroke weight change are sparse.<sup>12</sup> In the INTERSTROKE study, the incident rates were 23%, 33% and 43% for the increasing tertiles of WHR. These results are comparable to our results for both self-reported TIA (22%, 33% and 45%, respectively) and stroke (18%, 30% and 51%). The high figure of 51% implies either an increased weight gain following stroke or a lower prevalence of obesity in the INTERSTROKE study, which included data from both high and low socioeconomic countries.



The self-reported history of diabetes mellitus was 5.25%. For a cohort of middle aged, predominantly Caucasian Europeans, this is what we would have anticipated.<sup>13</sup> Estimates put diabetes prevalence between 10% and 25% in people with existing cerebrovascular disease depending on population and method used to diagnose diabetes,<sup>14</sup> and our results (15% for stroke and 10% for TIA) compare closely to a self-reported diagnosis in a young population.

Our results suggest that people who undertook at least 4 h of physical activity per week were more likely to have had a cerebrovascular event. Four hours is a relatively large amount and studies have consistently shown benefit in increased physical activity in reducing cerebrovascular disease.<sup>15</sup> Our results are likely to be due to reverse causation; either increased activity in cerebrovascular sufferers to improve their general health or the residual physical disability being reflected in normal activities of daily living being deemed at least moderately strenuous.

In people with cerebrovascular disease who drank three times or more per month, their frequency of drinking was below people without cerebrovascular disease. Further, in people who have never drunk or no longer drink, there was more cerebrovascular disease, which may reflect either a reduced frequency of drinking following diagnosis or an increased prevalence of disease; the well-described 'J-shaped curve' of alcohol-related illness.<sup>16</sup>

The OCP is associated with ischaemic stroke, particularly at higher doses.<sup>17 18</sup> We found that use of the oral contraceptive was lower for our cerebrovascular outcomes. Hence, it is possible that some of our participants with a medical history of cerebrovascular disease were not prescribed the OCP, explaining the findings.

It is also possible that the result detected with cerebrovascular disease, physical activity, alcohol consumption and OCP use reflects bias associated with the cross-sectional nature of the data and this need to be a stated limitation.

A high proportion of participants with self-reported and nurse-reported stroke (81%) and TIA (89%) were receiving an antithrombotic medication. The figure of 81% would seem appropriate for stroke, because up to 20% of people may have suffered intracerebral haemorrhage, thus restricting use and antithrombotic medication is not suitable for every individual. Our findings compare favourably to the Post Stroke Rehabilitation Outcome Project.<sup>19</sup> Their study in 2005 of 1161 people with both ischaemic and haemorrhagic stroke identified 62% of people as receiving antithrombotic medication.

While stroke disease does appear to have appropriate rates of antithrombotic prescription, TIA does appear to be undertreated for antithrombotic medication. TIA is almost exclusively a disease that results from infarction, embolic or otherwise, rather than haemorrhage. While antithrombotic will not be suitable for all of these people,

it is likely that 11% of untreated TIA disease represents suboptimal management.

We also demonstrated that large numbers of positive responders were taking cholesterol and blood pressure medications, including combinations of these medications. Not all of these medications will be suitable for each person with cerebrovascular disease; however, only about 45% of participants were receiving all three medications, suggesting these medications are underprescribed overall. Neither did we have information on whether treatment targets had been adequately achieved with these medications, something which is also likely to be suboptimal.<sup>20</sup> It is also likely that this young population of healthy volunteer participants are more likely to take and adhere to medication than less conscious populations.

It is important to highlight what we have not done in this study. First, we did not consider all of the well-recognised risk factors in as much detail as possible, for example, psychological disease. Smith *et al*<sup>21</sup> recently proposed criteria for probable major depression and probable bipolar disorder within the UK Biobank. Currently these criteria are not directly available from UK Biobank. As these and other criteria become established, researchers will be able to test hypotheses directly between them and cerebrovascular disease in the UK Biobank. Second, we only studied our exposure and outcome variables in a comparatively simple form, when other possible combinations could have been considered. For example, we recorded self-reported blood pressure but blood pressure was also recorded at the assessment visit. Neither did we perform detailed cross-sectional analysis of each variable, such as smoking or alcohol and in particular the pharmacotherapeutic data. Both were conscious decisions. To have done so would have been outside the remit of our objectives, which were to provide an overview of the data. Each risk factor would in itself form a separate detailed analysis.

Third, we have not looked at many of the known but less well-established risk factors. There were two reasons for not doing so. Often this was where the exact question did not exist within the UK Biobank, for example, recreational drug use. The second reason was where clinical information was available it would be best studied in relation to future linked data. For example, will the people who self-reported an asymptomatic carotid stenosis incur a different rate of incident cerebrovascular disease? Finally and perhaps most importantly, these are self-reported data. As such it is likely that it will underestimate the true prevalence of self-reported and nurse-reported stroke disease, as well as TIA.

Despite the limitations of the data, this sample of half a million people still represents the largest self-reported estimate of cerebrovascular disease in a population of this or any age. It adds to the available data on young populations with self-reported stroke and TIA disease which are less well described than older populations. The clinical relevance of the data is to highlight that

many people were not receiving all of the secondary prevention medication that would have been expected.

**Contributors** JH conceived the project and performed all data analysis. JH and JD wrote the first draft. JH, JD, SP and MW developed the project, commented on results and contributed to all subsequent drafts.

**Funding** University of Glasgow. 10.13039/501100000853.

**Competing interests** None declared.

**Ethics approval** North East MREC, UK.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data sharing statement** No additional data are available.

**Open Access** This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

## REFERENCES

- Collins R. What makes UK Biobank special? *Lancet* 2012;379:1173–4.
- Zhang Y, Chapman AM, Plested M, *et al.* The incidence, prevalence, and mortality of stroke in France, Germany, Italy, Spain, the UK, and the US: a literature review. *Stroke Res Treat* 2012;2012:436125.
- Allen N, Sudlow C, Downey P, *et al.* UK Biobank: current status and what it means for epidemiology. *Health Policy Technol* 2012;1:123–6.
- <http://www.ukbiobank.ac.uk>
- <http://biobank.ctsu.ox.ac.uk/crystal/refer.cgi?id=100235>
- Phillimore P, Beattie A, Townsend P. Widening inequality of health in northern England, 1981–91. *BMJ* 1994;308:1125–8.
- O'Donnell MJ, Xavier D, Liu L, *et al.* Risk factors for ischaemic and intracerebral haemorrhagic stroke in 22 countries (the interstroke study): a case-control study. *Lancet* 2010;376:112–23.
- Goldstein LB, Adams R, Alberts MJ, *et al.* Primary prevention of ischemic stroke: a guideline from the American Heart Association/American Stroke Association stroke council: cosponsored by the atherosclerotic peripheral vascular disease interdisciplinary working group; cardiovascular nursing council; clinical cardiology council; nutrition, physical activity, and metabolism council; and the quality of care and outcomes research interdisciplinary working group: the American Academy of Neurology affirms the value of this guideline. *Stroke* 2006;37:1583–633.
- Wu SH, Woo J, Zhang XH. Worldwide socioeconomic status and stroke mortality: an ecological study. *Int J Equity Health* 2013;12:42.
- Kerr GD, Slavin H, Clark D, *et al.* Do vascular risk factors explain the association between socioeconomic status and stroke incidence: a meta-analysis. *Cerebrovasc Dis* 2011;31:57–63.
- Kannel WB. Blood pressure as a cardiovascular risk factor: prevention and treatment. *JAMA* 1996;275:1571–6.
- Jönsson AC, Lindgren I, Norrvig B, *et al.* Weight loss after stroke: a population-based study from the Lund Stroke Register. *Stroke* 2008;39:918–23.
- Group DS. Age- and sex-specific prevalences of diabetes and impaired glucose regulation in 13 European cohorts. *Diabetes Care* 2003;26:61–9.
- Hewitt J, Castilla Guerra L, Fernández-Moreno M, *et al.* Diabetes and stroke prevention: a review. *Stroke Res Treat* 2012;2012:673187.
- Lee CD, Folsom AR, Blair SN. Physical activity and stroke risk: a meta-analysis. *Stroke* 2003;34:2475–81.
- Ferreira MP, Weems MK. Alcohol consumption by aging adults in the United States: health benefits and detriments. *J Am Diet Assoc* 2008;108:1668–76.
- Peragallo Urrutia R, Coeytaux RR, McBroom AJ, *et al.* Risk of acute thromboembolic events with oral contraceptive use: a systematic review and meta-analysis. *Obstet Gynecol* 2013;122:380–9.
- Zeitoun K, Carr BR. Is there an increased risk of stroke associated with oral contraceptives? *Drug* 1999;20:467–73.
- Zorowitz RD, Smout RJ, Gassaway JA, *et al.* Antiplatelet and anticoagulant medication usage during stroke rehabilitation: the post-stroke rehabilitation outcomes project (PSROP). *Top Stroke Rehabil* 2005;12:11–19.
- McAlister FA, Majumdar SR, Padwal RS, *et al.* Case management for blood pressure and lipid level control after minor stroke: prevention randomized controlled trial. *CMAJ* 2014;186:577–84.
- Smith DJ, Nicholl BI, Cullen B, *et al.* Prevalence and characteristics of probable major depression and bipolar disorder within UK Biobank: cross-sectional study of 172,751 participants. *PLoS ONE* 2013;8:e75362.