Hacking stroke in women: towards aetiology-driven precision prevention

Os, H.J.A. van

17. Summary

In this thesis I aimed to lay the foundation for the precision prevention of stroke in women. In **part I,** the pathophysiology underlying female-specific risk factors for stroke and sex differences in clinical presentation of stroke were discussed. In **part II,** I described how health data – routinely collected in electronic health records (EHR) – can be used to develop prediction models for risk of cardiovascular events and stroke specifically. For this purpose, I used a variety of statistical learning approaches, ranging from traditional regression to complex data driven models.

Summary of part I. Pathophysiology of stroke in women

In **chapter 2,** I tested the hypothesis that increased risk of ischaemic stroke in migraine patients is caused by a higher atherosclerotic burden in the cerebral vasculature, in a cohort (n = 656) derived from the Dutch Acute Stroke Study (DUST). I found no differences in the occurrence of atherosclerosis or stenosis in intracranial in intracranial (51% versus 74%; adjusted risk ratio [aRR]: 0.82; 95% CI: 0.64–1.05) or extracranial vessels (62% versus 79%; aRR: 0.93; 95% CI: 0.77–1.12) between patients with versus without migraine. These findings are important, because they support the hypothesis that the increased risk of ischaemic stroke in migraine patients is, at least for a substantial part, caused by mechanisms other than traditional atherosclerotic processes.

In **chapter 3** the potential association between sex and intra- and extracranial calcifications was shown, using a prospective cohort (n = 1,397) included from the DUST. My results confirm that in the extracranial circulation atherosclerosis is less prevalent in women compared with men (adjusted prevalence ratio [aPR]: 0.86; 95% CI: 0.81–0.92). This suggests that the prevalence of intracranial atherosclerosis is similar in women and in men (aPR: 0.95; 95% CI 0.89–1.01), indicating that the protective effect of oestrogen in women affects the intracranial arteries to a lesser extent compared with the extracranial arteries.

In **chapter 4,** the potential causal role of a history of headache and intrinsic coagulation protein levels in the serum on the risk of ischaemic stroke was investigated. I did this using a case-control study, including 113 women with ischaemic stroke and 598 healthy controls from the RATIO case-control study. My results suggest that a supra-additive effect may exist of the combination of a history of headache and intrinsic coagulation protein antigen levels and -activation on the risk of ischaemic stroke. This effect was most pronounced for kallikrein C1 inhibitor (adjusted odds ratio [aOR] protein alone: 2.2, 95% CI, 1.3–3.8; headache alone: 2.3, 95% CI: 1.3–4.1; in combination: 7.4, 95% CI: 2.9–19) and Factor XI antigen level (aOR protein alone: 1.7, 95% CI: 1.0–2.9; aOR headache alone: 2.0, 95% CI: 1.1–3.7; in combination: 5.2, 95% CI, 2.3–12). Previous research of our group and others gives rise to the hypothesis that sex differences in haemostatic factors can

increase the risk of stroke in women, potentially because of the modulating effect of female sex hormones.[154-156] Interestingly, findings in this chapter may indicate that an interaction exists between intrinsic coagulation factors and headache including migraine. This hypothesis is an alternative to traditional atherosclerotic mechanisms of cerebral infarctions in women.

In **chapters 5.** I assessed the hypothesis that delayed cerebral ischaemia occurs more frequently in aSAH patients with versus without migraine, in a cohort of 582 men and women. The rationale was that the migraine brain is more sensitive to spreading depolarization, which may in a pathological context such as aSAH lead to spreading ischaemia. My results showed that patients with a history of migraine are in general not at increased risk of developing delayed cerebral ischaemia compared with aSAH patients without migraine (adjusted hazards ratio [aHR]: 0.89; 95% CI: 0.56–1.43). However, I could not exclude a possible association in the subgroup of patients under 50 years, because a statistically significant interaction was identified between migraine and age (p-value = 0.075, at an alpha of 0.10 for interaction terms). In addition, younger patients with migraine are presumed to be more sensitive to spreading depolarisation, and, therefore, the association between migraine and delayed cerebral ischaemia may exist exclusively in this subgroup.

Therefore, in **chapter 6**, a follow-up study in patients under 50 years (n = 251) was conducted after additional patients in this age category were included. However, also in this population I did not find an association between a history of migraine and delayed cerebral ischaemia. I, therefore, conclude that a positive history of migraine is not a factor to take into account in treating patients with aSAH at risk of delayed cerebral ischaemia.

In **Chapter 7**, I present the design of the stroke cohort that is part of the Dutch String-of-Pearls Stroke Study cohort, which has resulted in a dataset with uniform and standardised storage of detailed clinical data of all Dutch University Medical Centres. This publication illustrates the potential value of registry-based research, and aids in the comparison between traditional cohort-, registry-, and EHR-based cohorts regarding the research on pathophysiology of stroke in women in this thesis.

In **chapter 8** I offer an overview of the associations between migraine, traditional cardiovascular risk factors and ischaemic stroke aetiologies, stratified for sex. Patients were included from the Dutch String-of-Pearls Stroke Study cohort (n = 2,492). I did not find any associations between a history of migraine and the prevalence of traditional cardiovascular risk factors. I did, however, find that women with a history of migraine had an increased risk of stroke with an onset occurring under 50 years of age, compared with women without migraine. This result confirms previous reports on a younger age at stroke onset in patients with

migraine. Since in these women traditional cardiovascular risk factors did not occur more frequently, results point towards mechanisms other than those mediated by atherosclerosis.

In **chapter 9,** I assessed the joint effect of migraine and combined oral contraceptive (COC) use on the risk of ischaemic stroke. I defined a nested case-control study using data from a population-based cohort, including 617 cases and 6,170 age-matched controls. Next, I integrated these data with previously published evidence using a systematic review and meta-analysis. In my case-control study, I found a significant increase in risk of ischaemic stroke (aOR: 6.83; 95% CI: 3.95–11.7) for both migraine and COC use combined versus neither factor. In migraine patients who both smoked and used COC versus women without migraine who did not smoke or use COC, the risk of stroke was increased substantially (aOR: 30.2; 95% CI: 4.22–610). After the synthesis of these new data with previously published data, I found that in young women with migraine who use COC compared with women without migraine and COC use, the increase in the risk of ischaemic stroke may be supra-additive, also in those using COCs with a low estrogen dose. I also showed that interactions between female-specific (COC use, migraine) and traditional cardiovascular (smoking) may lead to substantially increased risks for ischaemic stroke.

In **chapter 10,** I performed a meta-analysis to investigate whether there are sex differences in the clinical presentation of acute stroke or transient ischaemic attack (TIA). My meta-analysis shows, for the first time, substantial differences in women versus men with respect to both focal symptoms (facial weakness; OR 1.12; 95% CI: 1.02–1.24; based on 6 studies) and non-focal symptoms (for example minor change in level of consciousness or mental status change: OR 1.29, 95% CI 1.08–1.54; based on 13 studies; coma or stupor: OR 1.31; 95% CI: 1.16–1.49; 15 studies). I advise that clinicians should be aware of these differences, because the more frequently occurring non-focal stroke symptoms in women could result in an increased risk of misdiagnosis and possible undertreatment of stroke in women.

In **chapter 11,** the pathophysiology of headache as a presenting symptom of acute ischaemic stroke is investigated in the DUST population (n = 284). I found that headache occurred less frequently in patients with versus patients without atherosclerosis in the extracranial anterior circulation (35% versus 48%; RR 0.72; 95% CI: 0.54–0.97). This finding supports the hypothesis that vascular wall elasticity is a necessary contributing factor for the occurrence of headache during acute ischaemic stroke. I found no sex differences in the incidence of headache in this population. However, because this study helps in better characterising headache as a presenting symptom of ischaemic stroke, the results contribute the recognition

of ischaemic stroke based on non-focal symptoms. As such, this chapter ties in with the conclusions from Chapter 10.

Summary of part II. Prediction of stroke in women

My aim in **chapter 12** was to quantify the impact of different choices regarding the preparation of EHR-derived data on the predictive performance of models. As a case study, we focussed on the estimation of cardiovascular risk in the Dutch ELAN primary care cohort. In total, 89,491 patients were included, of whom 6,736 suffered from a first-ever cardiovascular event during a median follow-up of eight years. The definition of the outcome that was solely based on diagnosis codes resulted in a systematic underestimation of the risk of cardiovascular events (calibration curve intercept: 0.84; 95% CI: 0.83–0.84). Contrarily, complete case analysis led to overestimation of the risk (calibration curve intercept: -0.52; 95% CI: -0.53–-0.51). With these results, I showed that data preparation choices regarding the definition of the outcome or methods to handle missing values can substantially impact model calibration. This in turn may hamper reliable clinical decision support. It is, therefore, essential that methodological choices are transparently presented in prediction research and are motivated to safeguard model transportability from one EHR context to another clinical setting.

In **chapter 13,** I compared the predictive performance between multiple complex data-driven and simple regression models, with respect to functional and reperfusion outcome after ischaemic stroke. For this purpose, I included 1,383 patients of the MR CLEAN Registry cohort, which consists of ischaemic stroke patients who underwent endovascular treatment. I hypothesised that complex data-driven models would outperform logistic regression models with respect to discrimination between good and poor radiological or functional outcome, potentially because of more efficient processing of non-linear relationships and complex interactions between variables. No clinically relevant differences were found between all model, irrespective of the method used for predictor selection. The added value of this chapter is the publication of a fully automated analysis data preparation and model analysis pipeline, with models covering much of the statistical learning spectrum, and a nested validation procedure to account for overfitting. This pipeline was reused in Chapters 14 and 15.

In **chapter 14,** the potential added value of female-specific and psychosocial factors compared with only traditional cardiovascular factors for the prediction of risk of stroke was assessed in the Dutch population-based STIZON cohort of women under 50 years (n = 409,026). Analyses were stratified by three age groups of 20–29, 30–39 and 40–49 years at baseline. Stroke occurred in 2,751 women during a median of 11 years of follow-up. The incidence rate of stroke was 6.9 (6.6–7.2) per 10,000

person years. Adding female-specific and psychosocial risk factors to traditional cardiovascular predictors improved discriminatory performance of prediction models for women under 50 years, notably in 30–39 and 40–49 year age groups (Δ C-statistic: 0.023 and 0.029) compared with the reference models. The 'stroke risk age' tool I developed and presented in this chapter can be used to support risk communication in primary care, which – after validation and implementation research – could lead to motivation for a healthier lifestyle.

In **chapter 15,** the aim was to develop a prediction model for first-ever cardiovascular event risk in primary care patients aged 30–49 years, for which I also used the STIZON cohort. In total, 542,147 patients without cardiovascular disease or prescription of statins prior to baseline were included. Sex-specific EHR-derived prediction models for first-ever cardiovascular events were found to have moderate discriminatory performance and are well calibrated. Data-driven predictor selection leads to identification of non-traditional cardiovascular predictors which increase discriminatory performance of models and correct reclassification of events, mostly in women. In women, the Cox PH model including 50 most important predictors resulted in an increase in C-statistic compared with the reference model of 0.03, and a net correct reclassification of events of 3.7%. These models can be used to identify women within the primary care practice population, whose absolute risk reaches the 2.5% risk cut-off, which could lead to an early initiation of preventive treatment.