Reevaluation of risk factors for time to subsequent events after first stroke occurrence using a new weighted all-cause effect measure

Ann-Kathrin Ozga (✉ a.ozga@uke.de)

University Medical Center Hamburg-Eppendorf <https://orcid.org/0000-0002-9501-7301>

Bernhard Rauch

Stiftung Institut für Herzinfarktforschung

Frederick Palm

HELIOS Klinikum Schwerin

Christian Urbanek

Department of Neurology, University of Heidelberg

Armin Grau

Department of Neurology, University of Heidelberg

Heiko Becher

Institute of Medical Biometry and Epidemiology, University Medical Center Hamburg-Eppendorf

Geraldine Rauch

Charité Universitätsmedizin Berlin

Research article

Keywords: stroke, death, composite endpoint, time-to-event, risk factors, weighted all-cause hazard ratio

Posted Date: December 17th, 2019

DOI: <https://doi.org/10.21203/rs.2.19014/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at BMC Public Health on June 1st, 2020. See the published version at <https://doi.org/10.1186/s12889-020-08971-4>.

Abstract

Background : Risk factors for stroke include atrial fibrillation, hypertension, diabetes mellitus, smoking, and high cholesterol. However, the role of these factors on subsequent cardiovascular events or death is less clear due to therapeutic measures. We therefore aim to get insights into the persistence of known risk factors on subsequent stroke or death one year after the first stroke and to illustrate how the new weighted all-cause hazard ratio can ease the interpretation of competing time-to-event endpoints with different clinical relevance.

Methods : This study evaluates the one year follow-up of 470 first ever stroke cases identified in the area of Ludwigshafen, Germany, with 23 deaths and 34 subsequent stroke events. The recently introduced weighted all-cause hazard ratio was used which allows a weighting of the competing endpoints in a composite endpoint. We extended this approach to allow adjustment for covariates. The investigated risk factors were atrial fibrillation, hypertension, diabetes mellitus, smoking, and hypercholesterolemia adjusting for age and sex.

Results : None of these meanwhile treated risk factors of which some have been modified after first stroke remained to be associated with subsequent death or stroke. Cause-specific effects point sometimes into opposite directions.

Conclusions : Using the new weighted hazard ratio, we can support that well established risk factors for the occurrence of an index stroke are no good predictors of further disease progress defined by death or recurrent stroke. It has been demonstrated that the new weighted hazard ratio provides interpretation advantages over the common all-cause hazard ratio.

Background

Stroke is the second leading cause of death worldwide [1, 2]. Studies showed a risk of dying one year after stroke around 40 percent [3, 4]. Common known risk factors for first ever stroke like atrial fibrillation, hypertension, diabetes mellitus, smoking, or high cholesterol are well accepted and explain about 50–70 percent of all strokes [5, 6]. For subsequent strokes after a first non-fatal stroke event, risk factors were also analyzed but more heterogeneous results were gained with indicators including diabetes mellitus, coronary artery disease, hypertension, a high baseline score on the National Institute of Health Stroke Scale, intracranial arterial stenosis, or hyperlipidemia [7–9]. Although a recent meta-analysis suggested that hypertension, diabetes mellitus, atrial fibrillation, and coronary heart disease might be risk factors for subsequent stroke occurrence, they did not consider time-to-event analysis [10]. Even less analyzed are risk factors for "time to death" after first ever stroke which might differ from the common risk factors for "time to first ever stroke". Although the rate of dying as a direct consequence of stroke was reported to be high it seems to be difficult to quantify additional risk factors that influence the disease progress after a first stroke [3, 4]. Heterogeneous groups of patients may be one reason for this observation, as different stroke etiologies markedly influence the risk of subsequent events. Moreover, the consequent

implementation of secondary prevention strategies - especially protective medication - markedly reduces the rate of subsequent clinical events [11]. Therefore patients at special risk in secondary prevention after a first non-fatal stroke represent a special selection and need to be newly defined in general, but also in all day care. Furthermore, concrete efforts are made in medical research to prevent subsequent strokes after a first stroke [12].

In this work, we aim to deliver a reevaluation of potential risk factors for premature death or stroke after a first stroke event. Taking into account that death acts as a competing event to a subsequent non-fatal stroke, the composite endpoint with the two components "time to subsequent death" or "time to subsequent stroke" was taken as primary endpoint. These two endpoints are clearly of different clinical relevance, as a fatal event always represents the worst outcome. This has to be addressed in the analysis of composite endpoints. However, common statistical methods for composite endpoints are usually based on the all-cause hazard ratio, which ignores the different impact of the components. To overcome this problem, we recently introduced the weighted all-cause hazard ratio to simplify interpretation of a composite endpoint with competing components of different clinical relevance [13, 14]. In the present study, this new methodologic approach was extended for a situation to incorporate adjustment for confounders. This allows to adequately model a complex time-to-event framework, in which two competing events with different clinical relevance are of primary interest.

The primary aim of this exploratory study is to (re-)evaluate the effect of common risk factors for a first stroke on "time to subsequent death or stroke" using a new weighted effect measure. As only 60 percent of patients survive more than one year after a first stroke, it is important to characterize the disease progress after a first stroke event [3, 4]. The second aim is to illustrate how the weighted all-cause hazard ratio helps to better understand the effect of an endpoint combining competing events of different clinical relevance.

Methods

Study population

For the present analysis we used the 470 cases from the population-based case-control study on ischemic stroke called GENESIS ("Genetic, Socio-economic and Inflammatory Determinants of Ischemic Stroke and their Interdependence"). Details of the study design are given in the publication by Becher et al. and Palm et al. [15, 16]. Briefly, we examined incident cases of ischemic stroke occurring in the population of Ludwigshafen am Rhein (South-West Germany) from 2007 to 2012. The diagnosis of "stroke" was based on the definition of the World Health Organization [17]. Age- and sex- matched controls were recruited from the same population. Cases and controls underwent a personal interview with trained interviewers and provided a blood sample. The study population was further contacted by phone around one, three, and twelve month after an index stroke. At the follow-up time points with exact dates it was only recorded whether they had a subsequent stroke in the time since the last visit but not

when. Thus, no exact event time points are given for the event subsequent stroke whereas the exact death times are known. Severity of second stroke was not assessed.

Statistical methods

Baseline characteristics of patients are summarized with descriptive statistical methods. Continuous variables are described by means and standard deviations. Categorical data are summarized as absolute and relative frequencies.

To analyze the risk factors that might affect the time to death or subsequent stroke we applied the weighted all-cause hazard ratio thereby taking into account the occurrence of competing events [13, 14]. Thereby, different "relevance weights" for the two competing components "death" and "subsequent stroke" were applied. For choosing the weights, we followed the recommendations for finding a weighting scheme as described by Ozga and Rauch which are also explained in the additional file [AdditionalFile.pdf, 14]. For the first step we identified that death is the most relevant component and thus we choose a weight of 1 for death. Afterwards, the event "subsequent stroke" has to be set into relation to the most clinical relevant event by answering the theoretical question "how many strokes are as equally harmful as one death?". The clinical consideration that stroke is still a severe event on the one hand and the assumption that there are more strokes than deaths in the observed time period on the other hand led us to choose a weight of 0.7 for the event stroke. Thereby, grading was based on the disability weights described in the Global Burden of Disease study 2016 [18]. Although the definition of those weights is not directly comparable to our approach (i.e. they are chosen dependent on the medical records, e.g. speech impairment, being confined to bed or wheelchair), they also set the event stroke into relation to death. The disability weights for ischemic stroke ranged from 0.019 to 0.588 with an upper confidence limit of 0.744 for the latter. The weighted all-cause hazard ratio was originally introduced for a two group comparison. For our analysis we extended this effect measure to adjust for confounders [see AdditionalFile.pdf for methodological details]. In addition to the weighted all-cause hazard ratio, we also provide results for the non-weighted all-cause hazard ratio resulting from the Cox proportional hazards model and for the cause-specific hazard ratios for each component separately. Along with the analysis using other weighting schemes, these later analyses are interpreted as sensitivity analyses. All hazard ratios are reported along with 95%-confidence intervals. For the weighted all-cause hazard ratio, these are estimated via bootstrap sampling with 10000 runs. Risk factors that were evaluated at baseline were atrial fibrillation, hypertension, diabetes mellitus, smoking, and hypercholesterolemia. Furthermore, we adjusted for age and sex.

The following definitions were considered:

The following definitions were considered:

- *atrial fibrillation*: persistent and paroxysmal
- *hypertension*: values 140/90 mmHg at rest at three time points

- *diabetes mellitus*: fasting-plasma-glucose >125 mg/dl (>7 mmol/l) or peak blood sugar >200 mg/dl (>11 mmol/l) or two hours value with oral test for glucose tolerance > 200 mg/dl (>11 mmol/l)
- *smoking*: having smoked on a regular daily basis a minimum of one cigarette per day or five cigarettes per week or one pack per month or two cigars or two pipes per week over six months or more months during life
- *hypercholesterolemia*: fasting cholesterol value in the serum >200 mg/dl

As mentioned above the exact time of an event only has been delivered for all deaths but not for non-fatal strokes. In the latter cases only time intervals were presented ("interval censoring"). For unadjusted univariable analyses, some methods were proposed to account for this, such as different score tests or a parametric model [19–22]. The parametric model can be extended easily to a model adjusting for confounders. However, since we have no information about the underlying event-time distributions and a misspecification of the distribution can introduce serious bias, we decided against a parametric approach and in favour of a naive approach. If the event time is known to lay in an interval, the naive approach assumes that the event occurred at the upper boundary of the interval, thus at the next observed time point. To assess the robustness of this naive approach, we repeated the analyses with the event times replaced by the lower boundary of the interval and by the middle of the interval. We used the statistical software R Version 3.5.1 for all analyses [23]. The algorithm Mersenne Twister is used for random number generating needed to calculate the bootstrap confidence intervals [24].

Results

Baseline and data characteristics

Baseline characteristics of the study population are outlined in Table 1. In total, 23 (4.90%) patients died and 34 (7.20%) experienced a subsequent stroke within one year. Thus, 34 event times were interval-censored. The median time-to-event (either death or stroke) was 3.19 months (1st -3rd quantile: 1.48–11.53). The median time to death was 4.11 months (1st -3rd quantile: 2.45–8.71) and median time to subsequent stroke was 3.02 months (1st -3rd quantile: 1.12–11.56), respectively. The remaining 413 patients were censored at 12 months plus 35 days (395 days). Table 2 summarizes the distribution of events for categorical data (categorical data collected at index stroke).

Table 1
Baseline characteristics.

Characteristic	
Age (years) number of patients (% valid) mean (standard deviation)	470 (100) 66.29 (10.86)
Male sex – number of patients (%)	282/470 (60.0)
Body-mass-index (kg/m ²) number of patients (% valid) mean (standard deviation)	470 (100) 28.77 (5.17)
Diabetes mellitus – number of patients (%)	145/470 (30.9)
Atrial fibrillation – number of patients (%)	92/470 (19.6)
Hypercholesterolemia – number of patients (%)	342/470 (72.8)
Hypertension – number of patients (%)	407/470 (86.6)
Smoking – number of patients (%)	145/470 (30.9)
Physical activity – number of patients (%) no or little activity 1–2 days per week 3–7 days per week	315/465 (67.7) 67/465 (14.4) 83/465 (17.8)
Education level – number of patients (%) no/primary school middle school high school	360/470 (76.6) 55/470 (11.7) 55/470 (11.7)

Table 2
Distribution of events for categorical data.

Variable	Number of events (%)	Number of deaths (%)	Number of strokes (%)
Diabetes mellitus	13/57 (22.8)	10/23 (43.5)	3/34 (8.8)
Atrial fibrillation	10/57 (17.5)	6/23 (26.1)	4/34 (11.8)
Hypercholesterolemia	45/57 (78.9)	20/23 (87.0)	25/34 (73.5)
Hypertension	49/57 (86.0)	20/23 (87.0)	29/34 (85.3)
Smoking	24/57 (42.1)	9/23 (39.1)	15/34 (44.1)

Analyzing risk factors

Figure 1 shows the adjusted weighted all-cause hazard ratio (HR) at time point 395 days with the 95% bootstrap confidence intervals (CI). Figure 2 depicts the adjusted unweighted all-cause hazard ratio and

the adjusted cause-specific hazard ratios for each event type separately. In all models, adjustment is made for all other independent variables. In Figure 1 it can be seen that for diabetic patients the weighted average risk for a subsequent event is significantly reduced with a weighted hazard ratio of 0.39 (95%-CI=(0.20, 0.88)) whereas none of the other known risk factors for first ever stroke show a statistically significant association.

Furthermore, when investigating the cause-specific hazard ratios it can be seen in Fig. 2 that the effect of the different event types can point into opposite directions: e.g. for diabetes mellitus, the cause-specific hazard ratio for "death" suggests a higher mortality risk (1.75; 95%-CI=(0.74, 4.12)), but for "subsequent stroke" the event risk is lower (cause-specific hazard ratio 0.23; 95%-CI=(0.07, 0.75)).

For the other covariates no statistically significant effects were observed.

The results for other weighting schemes can be found in Fig. 3. The differences in the estimated effects between the analyses for the unweighted all-cause effect (Fig. 2) and the analysis with weight 1 for both event types are due to the differences in the estimated baseline hazard: For the weighted approach different cause-specific baseline hazards are assumed whereas the model for the unweighted estimation implies the same baseline hazard for the components. The analysis using the lower or mid interval time points for the event type stroke showed similar results [see AdditionalFile.pdf]. This sensitivity analyses suggest that our chosen weights and interval time point provide reasonably stable results.

Discussion

We presented the association of known major cardiovascular risk factors for a first ever stroke with the combined endpoint "time to death" and "time to second stroke" during follow-up by using a recently published method for weighted composite endpoints [13, 14]. The presented results show that under the conditions of current treatment and secondary prevention after a first stroke, the rates of death or subsequent non-fatal strokes during the first follow-up year are not associated anymore with the well accepted and individually ascertained cardiovascular risk factors. These data support the hypothesis that treatment and secondary prevention measures addressing these risk factors and risk diseases are effective in the prevention of subsequent serious clinical events at least during the first year of follow-up. Notwithstanding, the presented data do not allow to extrapolate these observations to longer time periods, as the actual treatment options may only be effective to postpone secondary clinical events but not to completely prevent them. These events observed during the first year after a first non-fatal stroke therefore need to be evaluated in more detail to get more information on causes and deleterious mechanisms. For example, additional confounders might influence the association seen for diabetic patients, like diet or physical activity. Furthermore, the degree of organ damage also influences the disease progress but could not be explicitly incorporated in the model due to missing information. The time point at which therapy for any of the risk disease has being started might also play a role in event occurrence, e.g. younger patients who are treated will have a lower event risk with higher age than those whose therapy starts late in life. However, in our study population the course of therapeutic measures could not be determined.

The study has some limitations: The sample size and thus the number of events is small. The mortality in the study group is low due to the case selection for the underlying case-control study, in which strokes with severe impairment were excluded. However, similar stroke rates were found in other studies [3, 8, 25]. The low event number also results in a limited accuracy of the estimates as expressed by wide confidence intervals. Furthermore, secondary stroke records were based on patient information which was not verified by medical supervision. Severity of second stroke was not assessed which would have allowed individual weights. Additionally, we did not incorporate a possible change of patients' conditions over time, e.g. if patients stopped smoking. Further, there was no documentation of the causes of death.

As a methodological result, our analysis emphasizes the importance to take all event types simultaneously into consideration to derive meaningful interpretations. Thereby, the different clinical relevance of competing event types should be taken into account which can be done with the weighted all-cause hazard ratio which yields a weighted average effect. This methodology considerably strengthens the basis of associations measured in clinical studies and registries. Serving as a model for this biometrical approach the present study clearly shows that the associations of well-known risk factors for the occurrence of a first stroke do not apply anymore during the period of secondary prevention. By this new methodologic approach our study differs from other studies which examined exclusively either the risk for recurrent stroke or death with regard to other predictors [7–9].

Conclusion

In summary, we, for the first time, extended the weighted all-cause hazard ratio to adjust for confounders. Thus, it is now generally possible to assess several risk factors for a time-to-event analysis for a composite endpoint which incorporates competing components of different clinical relevance.

List Of Abbreviations

CI: confidence interval; GENESIS: Genetic, Socio-economic and Inflammatory Determinants of Ischemic Stroke and their Interdependence; HR: hazard ratio

Declarations

Ethics approval and consent to participate. The study was approved by the ethics committee of the Landesärztekammer Rheinland-Pfalz (837.333.05(4991)).

Consent for publication. Not applicable.

Availability of data and materials. The datasets used and analysed during the current study are available from the corresponding author on reasonable request.

Competing interests. The authors declare that they have no competing interests.

Funding. This work was supported by the German Research Foundation [Grant RA 2347/1-2]. HB was partly funded by Stiftung Charité. The German Research Foundation and Stiftung Charité had no influence on any of the research, i.e. study design, analysis, interpretation, or writing, done in this article.

Authors' contributions. AO implemented the statistical analysis, produced the results and wrote the first draft of the manuscript. GR and HB contributed to all parts of the manuscript. All authors read and approved the final manuscript. BR, FP, CU, and, AG were contributors in writing the manuscript

Acknowledgements. This work was supported by the German Research Foundation [Grant RA 2347/1-2]. We thank Annette Aigner.

References

1. Lozano R, Naghavi M, Lim S, Shibuya K, Aboyans V, Abraham J, et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet*. 2012; 380: 2095-2128.
2. Murray CJ, Vos T, Lozano R, Naghavi M, Flaxman AD, Michaud C, et al. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet*. 2012; 380: 2197-2223.
3. Hankey GJ, Jamrozik K, Broadhurst RJ, Forbes S, Burvill PW, Anderson CS, et al. Five-year survival after first-ever stroke and related prognostic factors in the Perth Community Stroke Study. *Stroke*. 2000; 31:2080-2086.
4. Brønnum-Hansen, H; Davidsen, M; Thorvaldsen, P; Danish MONICA Study Group. Long-term survival and causes of death after stroke. *Stroke*. 2001; 32:2131-2136.
5. Becher H, Grau A, Steindorf K, Bugge F, Hacke W. Previous infection and other risk factors for acute cerebrovascular ischaemia: attributable risks and the characterisation of high risk groups. *J Epidemiol Biostat*. 2000; 5:277-283.
6. Gorelick PB. Stroke prevention therapy beyond antithrombotics: unifying mechanisms in ischemic stroke pathogenesis and implications for therapy: an invited review. *Stroke*. 2002; 33:862-875.
7. Fu GR, Yuan W, Du W, Yang Z, Fu N, Zheng H, et al. Risk Factors Associated with Recurrent Strokes in Young and Elderly Patients: A Hospital-based Study. *Int J Gerontol*. 2015; 9:63-66.
8. Wangqin R, Wang X, Wang Y, Xian Y, Zhao X, Liu L, et al. Risk factors associated with 90-day recurrent stroke in patients on dual antiplatelet therapy for minor stroke or high-risk TIA: a subgroup analysis of the CHANCE trial. *Stroke Vasc Neurol*. 2017; 2:176-183.
9. Leoo T, Lindgren A, Petersson J, von Arbin M. Risk factors and treatment at recurrent stroke onset: results from the Recurrent Stroke Quality and Epidemiology (RESQUE) Study. *Cerebrovasc Dis*. 2008; 25:254-260.
10. Zheng S, Yao B. Impact of risk factors for recurrence after the first ischemic stroke in adults: A systematic review and meta-analysis. *J Clin Neurosci*. 2019; 60:24-30.

11. Feigin VL, Norrving B, Mensah GA. Global Burden of Stroke. *Semin Neurol*. 2018; 38: 208-211.
12. Oza R, Rundell K, Garcellano M. Recurrent Ischemic Stroke: Strategies for Prevention. *Am Fam Physician*. 2017; 96: 436-440.
13. Rauch G, Kunzmann K, Kieser M, Wegscheider K, König J, Eulenburg C. A weighted combined effect measure for the analysis of a composite time-to-first-event endpoint with components of different clinical relevance. *Stat Med*. 2018; 37:749-767.
14. Ozga AK, Rauch G. Introducing a new estimator and test for the weighted all-cause hazard ratio. *BMC Med Res Methodol*. 2019; 19:1-16.
15. Becher H, Palm F, Aigner A, Safer A, Urbanek C, Buggle F, et al. Socioeconomic Conditions in Childhood, Adolescence, and Adulthood and the Risk of Ischemic Stroke. *Stroke*. 2016; 47:173-179.
16. Palm F, Urbanek C, Rose S, Buggle F, Bode B, Hennerici MG, et al. Stroke Incidence and Survival in Ludwigshafen am Rhein, Germany: the Ludwigshafen Stroke Study (LuSSt). *Stroke*. 2010; 41:1865-1870.
17. Truelsen T, Begg S, Methers C. The global burden of cerebrovascular disease.
https://www.who.int/healthinfo/statistics/bod_cerebrovasculardiseasestroke.pdf. Accessed 8 March 2019
18. GBD 2016 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet*. 2017; 390:1211-1259.
19. Fay MP. Comparing Several Score Tests for Interval Censored Data. *Stat Med*. 1999; 18:273-285.
20. Finkelstein DM. A Proportional Hazards Model for Interval-Censored Failure Time Data. *Biometrics*. 1986; 42:845-854.
21. Finkelstein MD, Wolfe RA. A Semiparametric Model for Regression Analysis of Interval-Censored Failure Time Data. *Biometrics*. 1985; 41:933-945.
22. Lindsey JC, Ryan LM. Tutorial in biostatistics methods for interval-censored data. *Stat Med*. 1998; 17:219-238.
23. R Core Team. R: A Language and Environment for Statistical Computing (Version 3.5.1).
<http://www.R-project.org/>. Accessed March 2019
24. Matsumoto M, Nishimura T. Mersenne twister. A 623-dimensionally equidistributed uniform pseudorandom number generator. *ACM Trans Model Comput Simul*. 1998; 8:3-30.
25. Amerenco P, Lavallée PC, Labreuche J, Albers GW, Bornstein NM, Canhão P, et al. One-Year Risk of Stroke after Transient Ischemic Attack or Minor Stroke. *N Engl J Med*. 2016; 374:1533-1542.

Additional File

The additional file [AdditionalFile.pdf] contains the description of the “Adjusted weighted all-cause hazard ratio” as well as general considerations for choosing the weights, an illustration of correlation between

independent variables, and further results for the estimated weighted hazard ratio using the lower or mid interval time points for the event “stroke” and the weights 0.7 for “stroke” and 1 for “death”.

Figures

Figure 1

Adjusted weighted hazard ratios with weights 1 for death and 0.7 for stroke. All effects are adjusted for all other independent variables. Confidence intervals were produced via bootstrap sampling. For age the hazard ratio is for a ten years difference.

Figure 2

Adjusted unweighted all-cause hazard ratio and cause-specific hazard ratios for sensitivity analysis. All effects are adjusted for all other independent variables. Confidence intervals were produced via bootstrap sampling. For age the hazard ratio is for a ten years difference.

Figure 3

Adjusted weighted hazard ratios with weight 1 for death and weights for stroke ranging between 0 and 1 for sensitivity analysis. All effects are adjusted for all other independent variables. Confidence intervals were produced via bootstrap sampling. For age the hazard ratio is for a ten years difference.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [AdditionalFile.pdf](#)