

2021-3

A Comparison of Risk Factors and Risk Models for Stroke by Age Group Using TILDA Data

Elizabeth Hunter

Technological University Dublin, elizabeth.hunter@tudublin.ie

John D. Kelleher

Technological University Dublin, john.d.kelleher@tudublin.ie

Follow this and additional works at: <https://arrow.tudublin.ie/scschcomcon>



Part of the [Computer Engineering Commons](#)

Recommended Citation

Hunter, E. & Kelleher, J.D. (2021) A Comparison of Risk Factors and Risk Models for Stroke by Age Group Using TILDA Data, *Longitudinal Studies Conference*, Wellcome Genome Campus, UK. 10-12th March 2021.

This Conference Paper is brought to you for free and open access by the School of Computer Sciences at ARROW@TU Dublin. It has been accepted for inclusion in Conference papers by an authorized administrator of ARROW@TU Dublin. For more information, please contact arrow.admin@tudublin.ie, aisling.coyne@tudublin.ie.



This work is licensed under a [Creative Commons Attribution-NonCommercial-Share Alike 4.0 License](#)
Funder: Precise4Q, ADAPT Centre for Digital Content Technology

A Comparison of Risk Factors and Risk Models for Stroke by Age Group Using TILDA Data

Elizabeth Hunter¹ and John D. Kelleher^{1 2}

¹PRECISE4Q Predictive Modelling in Stroke, Technological University Dublin,
Dublin, Ireland

²ADAPT Research Centre, Technological University Dublin, Ireland
`elizabeth.hunter@tudublin.ie`

1 Introduction

Models to predict stroke risk with the aim of stroke prevention often use age as a factor in the model (Choudhury et al., 2015; Conroy et al., 2003; D’Agostino et al., 2008; Wolf et al., 1991). However, stroke risk scores often underestimate risk for specific age groups, particularly younger age groups and the contribution of different risk factors to overall stroke risk changes over time (Boehme et al., 2017; Seshadri et al., 2006). Additionally, because age is a strong predictor of stroke, age can dominate the risk score (Leening et al., 2017). Longitudinal Studies such as the Irish Longitudinal study on Aging (TILDA) allow us to track these change in risk factors (TILDA, 2019). We aim to determine risk factors using an age group specific analysis in order to reduce the underestimation of risk for certain age groups.

2 Methodology

We use data from TILDA to compare the stroke risk factors by age group. In the first year of data collection, out of the 8504 participants, 21 had a stroke before the age of 50, 35 had strokes between 50 and 59, 41 between 60 and 69 and 36 when 70 or older (TILDA, 2019). To investigate if the risk factors vary by age groups, we use Chi-squared tests to compare the distributions of risk factors across and within age groups. The risk factors we consider are sex, smoking and diabetes. To examine how differences by age groups might affect risk score calculations we run a set of five logistic regression models, one for each of the age groups and one with all ages.

3 Risk Factors Across Groups

For each risk factor we compare the distribution of that factor across age groups within the population of those who had a stroke. The p-values for the Chi-squared tests are in Table 1.

Looking at the table we can see that the p-values for the risk factors smoking and diabetes both indicate significance at the 10% level meaning that we can

reject the null hypothesis that there is no difference in the distribution of the risk factors of smoking and diabetes for those who have had a stroke by age group. This leads to the conclusion that the risk factors may contribute differently to the risk of stroke in different age groups.

Table 1. P-values of Chi-squared comparing risk factors across age groups.

Risk Factor	P-value
Sex	0.109
Smoking	0.080
Diabetes	0.010

4 Risk Factors Within Groups

We compare the distribution of the risk factor within the age groups for those with and without a stroke diagnosis. The p-values for the Chi-Squared test can be found in Table 2. We find a significant difference in the distribution of sex between those who had stroke and those who did not for the less than 50 and 70 plus groups. Looking at smokers, we see a significant difference between those who had a stroke and did not in the 50 to 59 and 60 to 69 age. For a diabetes diagnosis, we find the difference is significant in the 60-69 age group and significant at a 10% level for the 50-59 age.

Table 2. P-values of Chi-squared comparing risk factors between those who have had a stroke and those who have not had a stroke by age group.

Risk Factor	Less 50	50-59	60-69	70+
Sex	0.007	0.261	0.456	0.021
Smoking	0.903	0.045	0.023	0.339
Diabetes	1	0.099	0.001	0.446

5 Models by Age Group

After we determined that there is a statistical difference in risk scores both across age groups and within age groups we create a set of logistic regression models: one for each age group and one for all age groups combined. The aim is to compare the model coefficients to determine if using different age groups leads to different coefficients.

Table 3 provides the coefficients for each of the 5 logistic regression models. Note there are not enough people in the data set under the age of 50 with diabetes to include this as a factor in the less than 50 model. Looking at the coefficients of the logistic regression models we see variation between the age groups. For example, the coefficient for sex varies from -2.41 to 0.08 and the intercept from -0.53 to 2.31.

Table 3. Coefficients for logistic regression models by age group

	Less than 50	50 to 59	60 to 69	70 Plus	All Ages
Intercept	2.31	0.14	-0.53	0.10	0.23
Sex-Female	-2.41	-0.77	0.08	-0.85	-0.84
Diabetes	-	-0.71	1.27	0.07	0.19
Smoke	-0.81	0.60	0.78	0.63	0.09

6 Conclusions

We have shown that the distribution of risk scores varies by age group and when risk models are created for each age group the coefficients differ across models. This shows that models for distinct age groups maybe a useful tool in helping to predict individual stroke risk. Additionally, it highlights the importance of using longitudinal studies to investigate stroke risk. We conclude that treating age groups separately may lead to more accurate predictive models for stroke risk, however, more work needs to be done comparing the discriminatory ability of the age models versus the model for all ages as well as the calibration of the models. Going forward we also aim to look at different time horizons for different age groups and potentially include different risk factors in each of the age group models depending on the factors that are most predictive for that age.

Funding This project received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No. 777107, and by the ADAPT Centre for Digital Content Technology funded under the SFI Research Centres Programme (Grant 13/RC/2106) and co-funded under the European Regional Development Funds.

References

- Boehme, A. K., Esenwa, C., & Elkind, M. S. V. (2017, February). Stroke Risk Factors, Genetics, and Prevention. *Circulation research*, *120*(3), 472–495. Retrieved 2021-08-09, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5321635/> doi: 10.1161/CIRCRESAHA.116.308398
- Choudhury, M. J. H., Chowdhury, M. T. I., Nayeem, A., & Jahan, W. A. (2015, April). Modifiable and Non-Modifiable Risk Factors of Stroke: A Review Update. *Journal of National Institute of Neurosciences Bangladesh*, *1*(1), 22–26. Retrieved 2021-08-09, from <https://www.banglajol.info/index.php/JNINB/article/view/22944> (Number: 1) doi: 10.3329/jninb.v1i1.22944
- Conroy, R. M., Group, o. b. o. t. S. p., Pyörälä, K., Group, o. b. o. t. S. p., Fitzgerald, A. P., Group, o. b. o. t. S. p., ... Group, o. b. o. t. S. p. (2003, June). Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. *European Heart Journal*, *24*(11), 987–1003. Retrieved 2020-11-05, from <https://academic.oup.com/eurheartj/article/24/11/987/427645> (Publisher: Oxford Academic) doi: 10.1016/S0195-668X(03)00114-3
- D’Agostino, R. B., Vasan, R. S., Pencina, M. J., Wolf, P. A., Cobain, M., Massaro, J. M., & Kannel, W. B. (2008, February). General Cardiovascular Risk Profile for Use in Primary Care: The Framingham Heart Study. *Circulation*, *117*(6), 743–753. Retrieved 2020-11-05, from <https://www.ahajournals.org/doi/10.1161/CIRCULATIONAHA.107.699579> doi: 10.1161/CIRCULATIONAHA.107.699579
- Leening, M. J., Cook, N. R., & Ridker, P. M. (2017, May). Should we reconsider the role of age in treatment allocation for primary prevention of cardiovascular disease? *European Heart Journal*, *38*(20), 1542–1547. Retrieved 2021-08-09, from <https://doi.org/10.1093/eurheartj/ehw287> doi: 10.1093/eurheartj/ehw287
- Seshadri, S., Beiser, A., Kelly-Hayes, M., Kase, C. S., Au, R., Kannel, W. B., & Wolf, P. A. (2006, February). The Lifetime Risk of Stroke: Estimates From the Framingham Study. *Stroke*, *37*(2), 345–350. Retrieved 2020-10-28, from <https://www.ahajournals.org/doi/10.1161/01.STR.0000199613.38911.b2> doi: 10.1161/01.STR.0000199613.38911.b2
- TILDA. (2019). The irish longitudinal study on ageing (tilda) wave 1, 2009–2011. [dataset]. *Irish Social Science Data Archive*.(Version 1.9). Retrieved from www.ucd.ie/issda/data/tilda/wave1
- Wolf, P. A., D’Agostino, R. B., Belanger, A. J., & Kannel, W. B. (1991, March). Probability of stroke: a risk profile from the Framingham Study. *Stroke*, *22*(3), 312–318. Retrieved 2020-11-05, from <https://www.ahajournals.org/doi/10.1161/01.STR.22.3.312> doi: 10.1161/01.STR.22.3.312