Translational Stroke Research (2019) 10:458–459
https://doi.org/10.1007/s12975-018-0681-4

COMMENTARY

# Commentary to: Masoli et al. Clinical Outcomes of CADASIL-Associated NOTCH3 mutations in 451,424 European Ancestry Community Volunteers. (Translational Stroke Research Oct 2018)

J. W. Rutten[1] · E. B. van den Akker[2] · S. A. J. Lesnik Oberstein[1]

Received: 16 November 2018 / Accepted: 6 December 2018 / Published online: 18 December 2018
© The Author(s) 2018

Dear Editor,

We read with great interest the study by Masoli et al. in which the incidence of cerebrovascular outcomes in individuals with a CADASIL-associated *NOTCH3* mutation in the UK Biobank (UKB) volunteer study is described [1]. Through imputation of array data and subsequent filtering on imputation quality and putative impact on protein function, two *NOTCH3* variants were nominated for further study, namely p.Arg1231Cys and p.Ala1020Pro, with a prevalence of respectively 0.04% and 0.96% in UKB. The p.Arg1231Cys variant was found to be associated with a higher percentage of incident stroke or TIA, whereas p.Ala1020Pro was not. We would like to further discuss several aspects of this important study in the context of recent developments in NOTCH3/CADASIL research.

More than 200 CADASIL-associated *NOTCH3* missense variants have been described in literature, almost exclusively leading to a cysteine alteration in one of the 34 epidermal growth factor-like repeat (EGFr) domains of the NOTCH3 protein [2]. The p.Ala1020Pro variant does not fall into this category and, moreover, has repeatedly been described as a non-pathogenic variant [2–4]. Hence, it is reassuring, though not surprising, that Masoli et al. found that this variant is not associated with increased risk of incident stroke or TIA.

This leaves the p.Arg1231Cys variant as the only detected CADASIL-associated variant in UKB, with a prevalence of 0.4 in 1000. The p.Arg1231Cys variant is located in EGFr domain 31 of the NOTCH3 protein and has been reported as causative in multiple CADASIL pedigrees [5, 6]. We have recently described that CADASIL patients with a cysteine-altering variant in one of EGFr domains 7-34 have a milder phenotype than patients with a variant in one of EGFr domains 1-6 [7]. The fact that EGFr 7-34 variants are milder likely explains why these variants predominate in the general population, whereas the more severe EGFr 1-6 mutations predominate in CADASIL pedigrees. The data of Masoli et al. are in line with this, as they describe that individuals with the p.Arg1231Cys variant had a substantially increased risk of incident stroke or TIA, but nonetheless the number of incident strokes or TIAs during follow-up is much lower than would be expected in CADASIL patients [8].

Furthermore, we would also like to delineate our concerns with respect to the suitability of imputed array data for the detection and subsequent analyses of rare variants. Array data predominantly target variants that are common in the general population. Imputation approaches use these common variants to predict the genotypic status of neighbouring variants not directly measured by the array. Hence, this implies that the genotypic status of rare variants in imputed data is generally the result of a prediction, rather than of a direct measurement. The prediction of the correct genotypic status becomes increasingly difficult with increasing rarity of the variant, as was recently illustrated by Mitt et al. using a study of 500 individuals with HRC imputed array data and directly measured genotypes coming from Whole Exome Sequencing (WES) data. While 17.0% of the predicted genotypes of rare variants (population frequency of ≤ 0.5%.) was false positive, a staggering 58.1% of the directly observed non-reference genotypes in WES were altogether missed by the imputed array data [9]. Comparisons between the results of Masoli et al. and gnomAD, a genome variant database containing sequencing data of 141,456 individuals [10], similarly suggests that UKB systematically underreports the prevalence

✉ J. W. Rutten
j.w.rutten@lumc.nl

[1] Department of Clinical Genetics, Leiden University Medical Center, 2333 ZA Leiden, The Netherlands

[2] Department of Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands

of CADASIL-associated rare variants in *NOTCH3*. The reported prevalence of 0.04% for the only detected CADASIL-associated variant p. Arg1231Cys in the UKB is identical to the prevalence of this specific variant in the European subpopulation in gnomAD [11]. However, in gnomAD, we also reported an additional 41 distinct more rare CADASIL-associated variants with a total prevalence of CADASIL-associated variants of 3.2/1000. Notably, UKB contains the genetic data of more than three times the individuals of gnomAD. Collectively, this suggests that many of the rare variants in UKB have not been detected, due to the fact that imputation of array data was used, rather than exome- or genome sequencing data.

Cognisance of the full-phenotypic spectrum associated with genomic variants is essential to enable a correct interpretation and prognosis for individuals in whom these variants will be detected as 'incidental' findings in the era of whole genome- and exome sequencing. CADASIL, as we know it, reflects only 1% of the total number of individuals with a cysteine altering *NOTCH3* variant in the population [11]. In view of this, and in view of the emerging broad phenotypic spectrum of *NOTCH3* cysteine altering variants, we suggest that the term '*NOTCH3* disease spectrum' may be more appropriate, reserving CADASIL for the severe end of this spectrum.

To conclude, the work by Masoli et al. nicely illustrates how large population-based datasets with longitudinal follow-up can shed light on the full phenotypic spectrum of variants previously only known to be associated with highly penetrant severe disease. With the advent of whole genome sequencing to be performed in 50,000 UKB participants, we look forward to the results of directly measured CADASIL-associated variants and their cerebrovascular phenotypes.

## Compliance with Ethical Standards

**Conflict of Interest** The authors declare that they have no conflict of interest.

**Ethical Approval** This article does not contain any studies with human participants performed by any of the authors.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## References

1. Masoli JAH, Pilling LC, Kuchel GA, Melzer D. Clinical outcomes of CADASIL-associated NOTCH3 mutations in 451,424 European Ancestry Community Volunteers. Transl Stroke Res. 2018. https://doi.org/10.1007/s12975-018-0671-6.
2. Rutten JW, Haan J, Terwindt GM, van Duinen SG, Boon EM, Lesnik Oberstein SA. Interpretation of NOTCH3 mutations in the diagnosis of CADASIL. Expert Rev Mol Diagn. 2014;14(5):593–603. https://doi.org/10.1586/14737159.2014.922880.
3. Ungaro C, Mazzei R, Conforti FL, Sprovieri T, Servillo P, Liguori M, et al. CADASIL: extended polymorphisms and mutational analysis of the NOTCH3 gene. J Neurosci Res. 2009;87(5):1162–7. https://doi.org/10.1002/jnr.21935.
4. Schmidt H, Zeginigg M, Wiltgen M, Freudenberger P, Petrovic K, Cavalieri M, et al. Genetic variants of the NOTCH3 gene in the elderly and magnetic resonance imaging correlates of age-related cerebral small vessel disease. Brain. 2011;134(Pt 11):3384–97. https://doi.org/10.1093/brain/awr252.
5. Abou Al-Shaar H, Qadi N, Al-Hamed MH, Meyer BF, Bohlega S. Phenotypic comparison of individuals with homozygous or heterozygous mutation of NOTCH3 in a large CADASIL family. J Neurol Sci. 2016;367:239–43. https://doi.org/10.1016/j.jns.2016.05.061.
6. Joutel A, Vahedi K, Corpechot C, Troesch A, Chabriat H, Vayssiere C, et al. Strong clustering and stereotyped nature of Notch3 mutations in CADASIL patients. Lancet. 1997;350(9090):1511–5. https://doi.org/10.1016/S0140-6736(97)08083-5.
7. Rutten JW, Van Eijsden BJ, Duering M, Jouvent E, Opherk C, Pantoni L, et al. The effect of NOTCH3 pathogenic variant position on CADASIL disease severity: NOTCH3 EGFr 1–6 pathogenic variant are associated with a more severe phenotype and lower survival compared with EGFr 7–34 pathogenic variant. Gened Med. 2018. https://doi.org/10.1038/s41436-018-0088-3.
8. Puy L, De Guio F, Godin O, Duering M, Dichgans M, Chabriat H, et al. Cerebral microbleeds and the risk of incident ischemic stroke in CADASIL (cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy). Stroke. 2017;48(10):2699–703. https://doi.org/10.1161/STROKEAHA.117.017839.
9. Mitt M, Kals M, Parn K, Gabriel SB, Lander ES, Palotie A, et al. Improved imputation accuracy of rare and low-frequency variants using population-specific high-coverage WGS-based imputation reference panel. Eur J Hum Genet. 2017;25(7):869–76. https://doi.org/10.1038/ejhg.2017.51.
10. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature. 2016;536(7616):285–91. https://doi.org/10.1038/nature19057.
11. Rutten JW, Dauwerse HG, Gravesteijn G, van Belzen MJ, van der Grond J, Polke JM, et al. Archetypal NOTCH3 mutations frequent in public exome: implications for CADASIL. Ann Clin Transl Neurol. 2016;3(11):844–53. https://doi.org/10.1002/acn3.344.