RESIDUAL RISK ESTIMATION FOR CARRIER SCREENING IN RARE DISEASES: AN ACCURATE APPROACH BASED ON DISEASE ALLELE FREQUENCIES OF PATHOGENIC VARIANTS

Torriani Camilla[1], Ciccone Roberto[2,3], Ferraro Ottavia Eleonora[1], Novara Francesca[3], Villani Simona[1], Monti Maria Cristina[1]

[1] *Department of Public Health, Experimental and Forensic Medicine, Unit of Biostatistics and Clinical Epidemiology, University of Pavia*
[2] *Department of Molecular Medicine, University of Pavia*
[3] *Medical Genetics Unit - NextLab Italy*

Introduction

Carrier Screening (CS) is a genetic test performed on healthy individuals to detect their carrier status of recessive conditions [1, 2, 4] and, hence, their risk of having affected children (conditioned to a mating type including two carriers). Being aware of own carrier status means reducing the risk of transmitting the disease/s to children [3]. However, an individual who tests negative for a specific condition still bears a nonzero possibility of being a carrier for a pathogenic variant not detected by an assay. We refer to this probability as Residual Risk (RR).

Objective/s

The aim of the present study was to provide an accurate approach to estimate RR for carrier screening in rare diseases by evaluating pathogenic variants in certain detectable genomic region by an assay. As a proof of concept, our method was applied on some genes (PEX 6, IDS and DHODH) using Disease Allele Frequencies (DAFs) found in available literature.

Methods

An observational pilot study was performed based on genetic screening results. A summary of the different steps [4] involved is reported in Figure 1.

Gene-specific disease selection and genetic model selection

Genetic conditions of interest and the specific pathogenic variants to be tested are selected in collaboration with a genetic counselor.

The genetic model of inheritance for the selected gene is determined through a literature review. The focus is on autosomal recessive and X-linked recessive inheritance models.

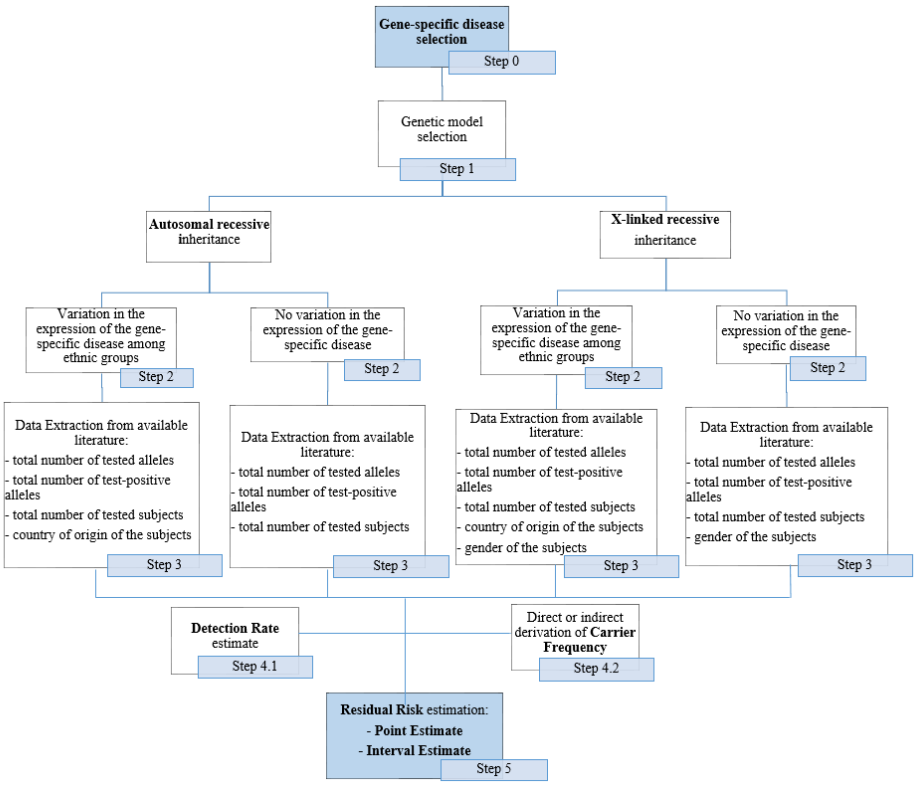
Data collection

Relevant data for each gene is extracted from published literature. Databases are searched for studies reporting clinical diagnosis and genetic information on affected individuals from different ethnicities. The number of positive test alleles and total tested alleles for each pathogenic variant, as well as the geographic origin and gender of the subjects, are recorded.

Statistical analysis

Variation of gene expression among ethnic groups: the prevalence of the gene is examined within different ethnic groups to assess if there are variations in risk based on ethnicity.

Detection Rate (DR) estimate is calculated as the sum of Disease Allele Frequencies (DAFs) divided by the total tested alleles. DR represents the proportion of positive test results for each pathogenic variant.
Direct or indirect Carrier Frequency (CF) derivation is obtained either directly from published articles or public databases, or indirectly derived based on gene/disease prevalence using the Hardy-Weinberg Law. If no specific carrier frequency is available, a default value may be applied.
RR estimate is calculated by multiplying the CF by one minus the DR. It represents the likelihood of being a carrier despite a negative screening result. RR can be expressed as a point estimate or as a ratio. Confidence intervals for RR can also be derived based on the interval estimation of DR. Overall, the test aims to provide an estimation of the RR of genetic diseases by considering genetic models of inheritance, ethnic variations, DRs, and CF.
All analyses were performed using R software, 2021. (Main libraries and functions: quantile function, PropCIs library, exactci function, pie function).
Figure 1- Analytical Process Flowchart.



Results

PEX6

311 patients with Zellweger spectrum disorder and 85 pathogenic variants of interest were identified.
Each patient may have one or more pathogenic variant and has inherited two disease-causing alleles independently. The total number of tested alleles was twice the number of subjects included. Only 31.5% of the pathogenic alleles were detected by an assay. The estimated risk of being a carrier of a pathogenic variant in the general population is 1 in 234 individuals, given a Carrier Screening negative result. There is a 95% confidence that the true RR estimate is between 1 in 213-269.

IDS

1 533 patients with Hunter syndrome and 602 pathogenic variants of interest were identified.
Men contributed one disease-causing allele, while women contributed two. For some studies, data on country of origin and gender were recorded. Results are reported at the general population level, but were also detailed by ethnicity: 66.6% of the pathogenic alleles were detected by an assay. The estimated RR indicates that 166 336 healthy individuals need to be screened to find a carrier. There is a 95% confidence that the true RR estimate is between 1 in 134 859 and 226 193 people.

DHODH

No sufficient information is available from the literature.
The laboratory report may use assumptions: CF equal to 1:500 and DR never less than 10%.
The result could be useful as a theoretical convention to indicate a RR even without detailed information.

Conclusions

The results obtained align with the expected RR, indicating consistency in the estimates. However, there still are several critical issues related to the calculation of the RR metric: the lack of a standardized process to report RR; different methodological approaches for calculating sensitivity of the test; challenges in identifying individuals affected by the genetic condition accurately, and the complexity of classifying patients' ancestry into ethnic groups. To further investigate the variability of parameters and as future perspectives, it would be beneficial to systematically collect RR results for each gene from different sources. This would allow for a clearer understanding of the results network and enable differentiation based on the methods used to calculate the metrics. The impact lies in its relevance to public health, as the calculation of RR is crucial for evaluating the reliability of genetic screening and estimating the incidence of new cases based on our ability to identify carriers in a given population. Additionally, by establishing a reference source and promoting clear and common conventions, it has the potential to reduce the variability of RR results among laboratories.

References

1. Azimi, M., Schmaus, K., Greger, V., Neitzel, D., Rochelle, R., & Dinh, T. (2016). Carrier screening by next-generation sequencing: health benefits and cost effectiveness. Molecular Genetics & Genomic Medicine, 4(3), 292–302. https://doi.org/10.1002/mgg3.204

2. Chung, C. C. Y., Chu, A. T. W., & Chung, B. H. Y. (2022). Rare disease emerging as a global public health priority. Frontiers in Public Health, 10. https://doi.org/10.3389/fpubh.2022.1028545
3. Dungan, J. (2018). Expanded carrier screening: what the reproductive endocrinologist needs to know. Fertility and Sterility, 109(2), 183–189. https://doi.org/10.1016/j.fertnstert.2017.11.030
4. Leung, M. L., McAdoo, S., Watson, D., Stumm, K., Harr, M., Wang, X., Chung, C. H., Mafra, F., Nesbitt, A. I., Hakonarson, H., & Santani, A. (2021). A Transparent Approach to Calculate Detection Rate and Residual Risk for Carrier Screening. The Journal of Molecular Diagnostics, 23(1), 91–102. https://doi.org/10.1016/j.jmoldx.2020.10.009