Niko Kallio, Mervi Kinnunen, Kati Kristiansson, Minttu Marttila, Heidi Marjonen, Helena Kääriäinen, Markus Perola

Polygenic risk scores for venous thromboembolism (VTE)

National Institute for Health and Welfare (THL), Helsinki, Finland

P4 healthcare + Omics

Research on omics, including genomics and metabolomics, has identified thousands of genomic regions and several blood metabolites linked to common human diseases and features. This new knowledge combined with samples collected to biobanks could provide new tools for predictive, preventive, personalized and participatory (P4) healthcare. So far the clinical utilization of omics is fairly small and unvalidated.

P4 + Population health = P5 study

The aim of the P5 study is to test the value of returning genetic and metabolomics risk information to participants. The aim is to study:

- Does risk information have effect on health behaviors and selected health outcomes
- How people react to genetic risk vs metabolic risk information
- Can risk information be shared through an internet portal

Participants

P5 is a continuation of FinHealth 2017 Study involving thousands of Finns from around the country. All participants are re-consented and no new sample are needed as we will utilize information, samples, and measurements obtained in the FinHealth Study.

Venous Thromboembolism (VTE)

The risk for venous thromboembolism is a combination of genetic and environmental factors and can be influenced the lifestyle choices made by the participant.

Genetic risk

Prior to estimating the genetic risk for P5 participants, we validated the value of the genetic testing in another Finnish population sample, FINRISK (n=20 865).

Single clinical variants (SCV)

There are two known SCVs that confer higher risk for VTE (table 1).
We examined whether FINRISK participan ts carry these SCVs and
estimated the combined risk of traditional risk factors, SCVs and polygenic risk.

Table1. Selected single clinical variants.

SCV	Risk allele/non-risk allele	Minor allele Frequency
FV Leiden	T/C	0.02
F2 (prothrombin) G20210A	A/G	0.008

Polygenic risk score (PRS)

In a PRS, genetic information across the genome is summarized to a genetic risk score. (Figure 1) The polygenic risk for VTE has so far been evaluated only in few studies. We compared the most recent published scores in the FINRISK cohort and calculated individual risk estimates based on the most valid score. From FINRISK cohort we used imputed genotypes and calculated PRSs from allele dosages. We evaluated
a score from de Haan's study [2] (28 SNPs), Kim's study [1] (16 SNPs) and a third PRS which consisted of the five most commonly used SNPs of multiple PRS studies. PRSs were normalized to facilitate comparison.

Results

We examined all three polygenic risk scores with Cox proportional hazards model in a sample with 255 VTE cases. First, we evaluated gender, BMI, cancer, smoking, diabetes, rheumatoid arthritis and liver disease with Cox proportional hazard model for their significant association with VTE events. Only smoking, BMI and rheumatoid arthritis were significant with HR 1.40, 1.04 and 2.60 and p-value 0.03, 0.0007 and $4.2x10^{-5}$, respectively.

Leiden factor V mutation showed predictive value for VTE risk assessment and although prothrombin mutation did seem to associate with higher risk, the result was not significant. (Table 2)

Table2. Single clinical variants.

SCV	Hazards Ratio	CI	p-value
Factor V Leiden	2.52	1.68-3.80	$8.80x10^{-6}$
Prothrombin	1.68	0.62-4.56	0.31

We then examined all three PRSs with Cox models adjusted for the significant traditional risk factors in a sample of 255 VTE cases (127 males, 128 females). All three scores showed predictive value (table 3) for VTE risk assessment.

Table 3. Cox proportional hazards models with polygenic risk scores.

PRS	Hazards Ratio*	CI	p-value
Kim	1.26	1.14-1.38	$3.23x10^{-6}$
Five most common	1.20	1.07-1.33	0.001
DeHaan	1.20	1.11-1.31	$1.08x10^{-5}$

*Adjusted with smoking, BMI, rheumatoid arthritis, geographical area, genotyping batch and first 10 pricipal components

When adjusting the Cox's model with PRSs for the Leiden mutation, the effect of Leiden mutation was not significant but the Kim score remained borderline significantly associated with increased risk for VTE. This indicates that most of the effect of the score is accounted by the Leiden mutation also included in the Kim's score but residual effect remains for the additional variants.

We then evaluated reclassification of risk induced by adding PRS to the model with traditional risk factors using Kaplan-Mayer net reclassification method. The result indicated that although adding PRS reclassified VTE cases significantly better, it did not do so for non-cases.

Conclusion

Kim's PRS has the best predictive value in our sample but did not bring additional value for genetic risk assessment for VTE when Factor V Leiden was included in the model. Nowadays it is possible to build a whole genome wide risk scores with millions of variants and these might be useful for clinical applications. Hormone replacement therapy was not included in our models and it may improve our models and reclassification.

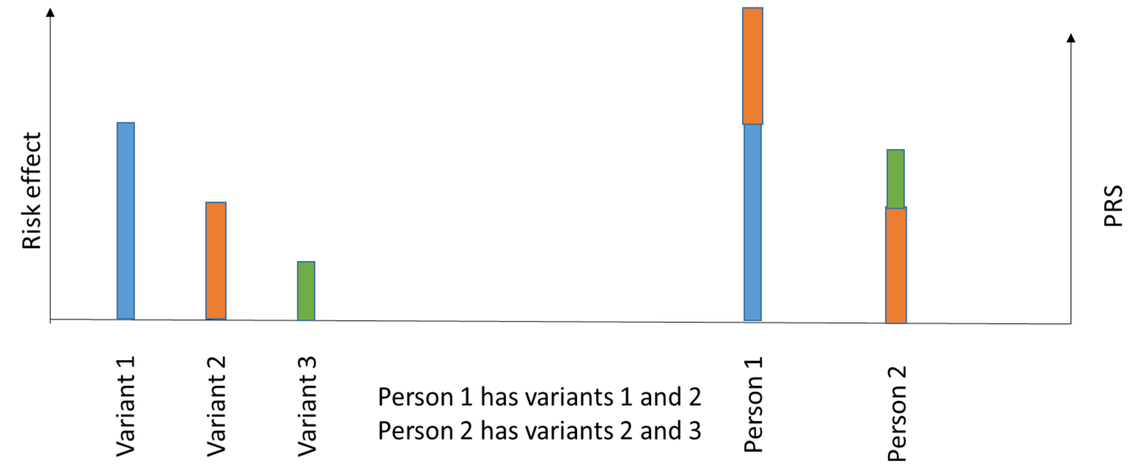

Figure 1.Building a polygenic risk score

[1] Kim J, Kraft P, Hagan KA, et al. Interaction of a genetic risk score with physical activity, physical inactivity, and body mass index in relation to venous thromboembolism risk. *Genet Epidemiol*. 2018;42:354–365.
[2] de Haan, H. G., Bezemer, I. D., Doggen, C. J., Le Cessie, S., Reitsma, P. H., Arellano, A. R., Tong, C. H., Devlin, J. J., Bare, L. A., Rosendaal, F. R., & Vossen, C. Y. (2012). Multiple SNP testing improves risk prediction of first venous thrombosis. *Blood*, 120(3), 656-663.