**STATISTICAL ANALYSIS PLAN**

**Association between TGCT polygenic risk score and risk of relapse in low-risk patients under surveillance.**

Responsible: Emilio Ugalde Morales <emilio.ugalde.morales@ki.se>

PI: Fredrik Wiklund <fredrik.wiklund@ki.se>

**Description**

To date 78 established testicular Germ Cell Cancer (TGCT) susceptibility variants have been identified. The aim with this study is to explore if a polygenic risk score (PRS) built on these 78 variants are associated with risk of relapse among patients having surveillance as the primary post-orchidectomy treatment. For that, we seek to pool associations from different populations represented in The Testicular Cancer Consortium.

**Study population**

Target population will consist of men diagnosed with stage I TGCT tumors with surveillance as primary treatment.

**Exposure**

A polygenic risk score (PRS) summarizing the individual-level of genetic predisposition to TGCT. This is calculating as the sum of risk alleles weighted by the effect of association with TGCT. TGCT risk estimates (genetic effect sizes) will be retrieved from the latest meta-analysis of TGCT. Individual-level PRS scores will be computed using Plink software.

**Outcome**

Disease relapses (events) during follow-up. The following variables needed:

* Clinical relapse (yes, no).
* Time to relapse, or end of follow-up, from diagnosis.

**Co-variates**

information on the following phenotypes will be considered in the analysis:

* Age at diagnosis (years).
* Histology (seminoma, non-seminoma).
* Vascular infiltration (yes, no, for non-seminoma) if available.
* First three genetic principal components (PCs).

**Statistical Analysis**

Statistical analysis will consist of:

* Testing the association between the PRS and risk of relapse, overall and stratified by histological groups, using Cox regression models. PRS will be treated both as a continuous (per-1 standard deviation) variable as well as a categorical variable defined by cases-only percentile subgroups (e.g. median and highest tertile vs. lowest tertile of the PRS).
* In sensitivity analysis the same models above will be explored excluding patients with presence of vascular invasion.
* Additional models including age at diagnosis, vascular infiltration, and PCs as co-variates.
* Estimation of absolute risks of relapse (cumulative incidence) by PRS category (lowest tertile, median and highest tertile).
* The results from each cohort will be pooled through a meta-analysis.

**CODEBOOK**

|  |  |  |  |
| --- | --- | --- | --- |
| **Variable** | **Variable name** | **Values** | **Value labels** |
| Clinical relapse | recpro | 0, 1 | No relapse, Relapse |
| Time to relapse | timerec | in years | Time between orchiectomy and relapse, in years |
| Age at diagnosis | age\_dx | in years | na |
| Histology | tumortype | S, NS | Seminoma, non-seminoma |
| Vascular infiltration | vasc | 0, 1 | no, yes |