To complete an analysis looking at the association between poor immune function and cancer risk in

**adult** (≥18 years old) people with **HIV (PWH)** following enrollment.

**immune function**

* **Poor immune function** is defined as having an **AIDS-defining illness** or having **a CD4 count less than 200** prior to enrollment.

This association likely differs between those who are able to maintain HIV viral suppression and those who cannot.

* **Viral suppression** is defined **as having <200 copies/mL** at enrollment.

Therefore, the exposure of **immune function** must be stratified into four groups:

1. Poor immune function and virally suppressed
2. Poor immune function and not virally suppressed
3. Good immune function and virally suppressed
4. Good immune function and not virally suppressed

Individuals should be followed from enrollment date to the first of: 1) cancer diagnosis; 2) loss-to-follow-up or administrative censoring (12/31/2020). Please create this exposure variable using the datasets provided and create a final analytic file named “**analytic\_file.csv**” that includes one row per patient with the following variables:

|  |  |  |
| --- | --- | --- |
| **Variable** | **Description** | **Format** |
| ID | patient ID | Numeric (4 digits) |
| Immune\_function | 4 level immune function | 0=poor immune function and virally suppressed; 1=poor immune function and not virally suppressed;  2=good immune function and virally suppressed; 3=good immune function and not virally suppressed |
| t | Time from **enrollment date** to first of **cancer diagnosis** or **loss to follow up** | Numeric, measured in years (with fractions of a year rounded to the thousandth decimal – e.g., 48 days should be coded 0.131) |
| Cancer | Indicator for cancer diagnosis | 1=patient diagnosed with cancer during the study period (**prior to loss to follow up**);  0=patient was censored before cancer diagnosis |
| Age | age at **enrollment** (years) | Numeric two digit age |