**Template for RESPOND Statistical Analysis Plan**

1. Research objectives

*List research objectives and outcomes concisely and clearly*

1. Study design

Cross sectional, longitudinal, nested case control etc.

1. Study population

Where possible, please use the standard inclusion and exclusion criteria, used in previous RESPOND analyses, as listed below. If these criteria are not applied, a justification for should be included here.

* 1. Inclusion criteria
* aged 18 years or older at baseline
* have a CD4 count and viral load measurement either 1 year prior to or within 12 weeks after baseline
	1. Exclusion criteria
* Missing information on gender
* For analyses focusing on clinical events, individuals from cohorts with low event reporting

Please ask your internal RESPOND statistician to provide a copy of the most recent inclusion of key variables and events document

* 1. Baseline and censoring date
	2. Key stratification factors

Planned sub-group analyses should be defined a-priori or a test for interaction performed. To check the robustness of results analysis are often conducted stratified by key sub-groups including gender (male/female) or antiretroviral naïve/experienced

1. Definitions
	1. Exposure
	2. Outcomes
	3. Comparison groups
	4. Confounders, predictors and effect modifier

*This table is an example and should be modified per analysis to reflect variables considered in the proposed analysis. Where possible, please use these definitions or categorisations to be consistent with previous RESPOND work. If alternative definitions are to be used, please justify.*

 *Please check that all relevant variables are included here, with relevant definitions included. Please reference the time point at which variables are defined (for example, at baseline), and whether the variable will be fitted as continuous, categorical, at baseline or as time-updated. For categorical variables, include details of categorisations; for continuous variables, include details of how the variable will be fitted and whether any transformations will be used.*

Confounders, predictors and effect modifiers

|  |  |  |
| --- | --- | --- |
| Variable | Categories | Comments / definitions |
| Calendar year |  |  |
| Age |  | If adjusting as a covariate use ln(age) |
| Gender\* | Male; female | Reflecting gender as reported by cohorts without reference to self identified gender. |
| HIV risk group | MSM; IDU; heterosexual sex; other; unknown or missing |  |
| Ethnicity | White; Black; other; unknown or missing | Some cohorts were prohibited from reporting ethnicity |
| Time since HIV diagnosis | Categorical or continuous |  |
| Body mass index | <18.5 kg/m2; 18.5-<25 kg/m2; 25-30 kg/m2; >30 kg/m2; missing |  |
| Smoking status | Past; current; never; unknown or missing |  |
| CD4 cell nadir  | Continuous  | Taken as the lowest CD4 cell count prior to baseline. If no CD4 cell count was measured, the first measurement within 12 weeks after baseline was used |
| CD4 cell count  | Continuous  | Taken as the most recent CD4 cell count within 12 months prior to baseline. If no CD4 cell count was measured, the first measurement within 12 weeks after baseline was used |
| Viral suppression status  | <200; >200 | Taken as the most recent VL within 12 months before baseline. If no VL was measured prior to baseline, the first measurement within 12 weeks after baseline should be used |
| Prior ART-experience and Viral suppression status  | ART-naïve; ART-experienced with VL <200 copies/mL; ART-experienced with VL ≥200 copies/mL | ART experience should be defined based on whether first ART start date is before or after baseline. Individuals who are ART naïve at baseline with a VL < 200 copies/mL at first ART start date should be excluded, as it likely indicates missing ART data.VL taken as the most recent VL within 12 months before baseline. If no VL was measured prior to baseline, the first measurement within 12 weeks after baseline should be used |
| Viral hepatitis C | No; yes; unknown | Defined by use of anti-HCV medication, a positive HCV antibody test, a positive HCV RNA qualitative test, HCV RNA >615 IU/mL, and/or a positive genotype test |
| Viral hepatitis B | No; yes; unknown | Defined by a positive HBV surface antigen test and/or HBV DNA >357 IU/mL |
| Hypertension | No; yes; unknown | Confirmed by use of anti-hypertensives at any time before baseline or if the most recent systolic or diastolic blood pressure measurement before baseline is higher than 140 or 90 mmHg, respectively |
| Diabetes | No; yes; unknown | Defined as reported diagnosis, use of anti-diabetic medication, glucose ≥11.1 mmol/L, and/or HbA1c ≥6.5% or ≥48 mmol/mol |
| AIDS (non-cancer) | No; yes; unknown | Composite diagnosis as defined by the CDC list of AIDS-defining conditions  |
| AIDS cancer | No; yes; unknown | Composite diagnosis of Kaposi’s sarcoma, non-Hodgkin lymphoma, cervical cancer |
| Non-AIDS defining cancer | No; yes; unknown | Any non-AIDS cancer, excluding skin cancers (expect malignant melanoma) and pre-cancers  |
| End stage liver disease | No; yes; unknown | Composite diagnosis of ascites (where extrahepatic reasons are excluded), hepatic encephalopathy grade III-IV, hepatorenal syndrome, endoscopically verified variceal bleeding, spontaneous bacterial peritonitis liver transplantation, hepatocellular carcinoma  |
| End stage renal disease | No; yes; unknown | Composite diagnosis of dialysis more than three months or kidney transplant  |
| Cardiovascular disease | No; yes; unknown | Composite diagnosis of myocardial infarction, stroke or invasive cardiovascular procedure  |
| Chronic kidney disease | No; yes; unknown | Two consecutive measurements of eGFR measured at least 3 months apart <60 mL/min if the first eGFR was >60 mL/min or a 25% decline if first eGFR was <60 mL/min. eGFR was calculated using the CKD-EPI creatinine equation in RESPOND  |
| Fractures | No; yes; unknown |  |
| Dyslipidaemia | No; yes; unknown | Defined as total cholesterol>239.4mg/dL or HDL cholesterol<34.7mg/dL or triglyceride >203.55mg/dL or use of lipid lowering treatments  |
| Geographical region | Western Europe; Northern Europe/Australia; Southern Europe; Eastern/East Central Europe; USA | Due to low numbers, Australia is usually combined with Northern Europe in analysis models, Eastern Central Europe is combined with Eastern Europe, and USA is combined with Western Europe.  |
| Exposure to protease inhibitors | No; yes; or cumulative exposure, or recent exposure, or past exposure | Recent exposure usually defined as on drug within the previous 6 months |
| Exposure to integrase strand transfer inhibitors | No; yes; or cumulative exposure, or recent exposure, or past exposure | Recent exposure usually defined as on drug within the previous 6 months |
| Exposure to non-nucleoside reverse transcriptase inhibitors | No; yes; or cumulative exposure, or recent exposure, or past exposure | Recent exposure usually defined as on drug within the previous 6 months |
| Exposure to nucleoside reverse transcriptase inhibitors | No; yes; or cumulative exposure, or recent exposure, or past exposure | Recent exposure usually defined as on drug within the previous 6 months |
| CD4 cell count  | Continuous | Ln2(CD4) if a covariate |
| Systolic blood pressure | Continuous | Continuous if a covariate |
| Total cholesterol | Continuous | Continuous if a covariate |
| High-density lipoprotein | Continuous | Continuous if a covariate |

\* RESPOND also includes a small number of transgender individuals, typically too small to include as a separate category.

Abbreviations: MSM – men who have sex with men; IDU – injecting drug use; VL – viral load; HCV – hepatitis C; RNA - ribonucleic acid; HBV – hepatitis B; CKD – chronic kidney disease; eGFR - estimated glomerular filtration rate; CKD-EPI - Chronic Kidney Disease Epidemiology Collaboration; CG – Cockcroft-Gault; HDL - high-density lipoproteins; USA – United States of America

1. Data analysis

*Provide a detailed description of a proposed analysis for each objective, linked to the objectives above. The data analysis plan should note*

1. *descriptive statistics and methods for comparing groups in any descriptive analyses*
2. *methodology/models to be used and model outputs; (eg incidence rate ratios, 95% confidence intervals etc; see methodology document* [*here*](file:///%5C%5Cregionh.top.local%5Cdfs%5CLogget%5CLovbeskyttetMapper%5CRESPOND-STAT%5CRESPOND_SAP%5CMethodology%20document)*)*
3. *model building strategies including criteria for inclusion or exclusion of specific variables*
4. *assumptions for methodology; eg LOCF, left censoring, time-lagging*
5. *criteria for inclusion and exclusion of specific variables and/or cohort participation depending on completeness of data over the duration of the study*
6. *Protocol for handling missing or incomplete data*
7. *A priori subgroup analyses and/or interaction testing including consideration of multiple testing*
8. *Sensitivity analyses*

*If the outcome of the analysis is cardiovascular disease, cancer, end stage liver disease, end stage renal disease, fractures, or mortality, please include a sensitivity analysis analysing validated events only.*

1. *Sample size or precision estimates for effect measures*
2. *Potential biases; how will they be addressed and potential impact on findings*
3. Timelines

This section is intended to provide key steps in the analysis when results should be circulated for review by the statistical team. A full set of results should be circulated to the statistical team and the writing group for review prior to the circulation of any abstract or manuscript

|  |  |
| --- | --- |
| Document | Timeline |
| Project proposal and SAP |  |
| Analysis results |  |
| Conference abstract |  |
| Draft manuscript |  |