Session 9

Answering research questions using your own data

Introduction

- Interesting finding in clinic and would like to explore in more detail - but major grant funding unlikely
- Could routinely collected clinic data be used to answer question?
- If so, no need for big funding, could do yourself (with help of your friendly statistician!)

Do women in my clinic have poorer adherence than men? What might be the reasons for this?

A lot of my patients taking drug X seem to be reporting side effects – are people experiencing more side effects with drug X than with drug Y?

How do you get from an idea to published paper?

Example – UK CHIC Study

(1) The idea

- Approached by clinician at a London-based HIV clinic
- Interested in whether treatment toxicities and antiretroviral resistance were more likely to occur in those starting cART with high CD4 counts
- Secondary end-point HPTN 052
- START yet to report
- Would it be possible to investigate this using the routine clinic data collected through UK CHIC?

Ethics approval and informed consent

- UK CHIC Study set up and approved for research purposes
- Pseudonomysed data so doesn’t require informed consent

- Need to check local policies and procedures for obtaining ethical approval before proceeding

(2) Defining the research question

Q: *Are treatment toxicities and antiretroviral resistance more likely to occur in those starting cART with high CD4 counts?*

P Who are our population of interest?

What is our exposure/factor of interest?

I 'Intervention'

C Comparison group

O What is our outcome?

How will we measure exposure and outcome?

(2) Defining the research question

Population

Exposure ('Intervention'/Comparison group)

I =

C=

Outcome

(2) Defining the research question

Population

HIV-positive individuals initiating antiretroviral therapy for the first time

Exposure ('Intervention'/Comparison group)

I =

C=

Outcome

(2) Defining the research question

Population

HIV-positive individuals initiating antiretroviral therapy for the first time

Exposure ('Intervention'/Comparison group)

I = People starting cART with a 'high' CD4 count

C= People starting cART with a 'not high' CD4 count

Outcome

(2) Defining the research question

Population

HIV-positive individuals initiating antiretroviral therapy for the first time

Exposure ('Intervention'/Comparison group)

I = People starting cART with a 'high' CD4 count

C= People starting cART with a 'not high' CD4 count

Outcome

Antiretroviral toxicities

Antiretroviral resistance

(2) Defining the research question

How do we measure exposure and outcome?

- What constitutes ‘high’ CD4 count?

  (continuous / categorical?)

- How do we measure antiretroviral toxicities?

  (treatment discontinuations / reported side-effects?)

- How do we define antiretroviral resistance?

  (new mutations / susceptibility scores?)

(2) Defining the research question

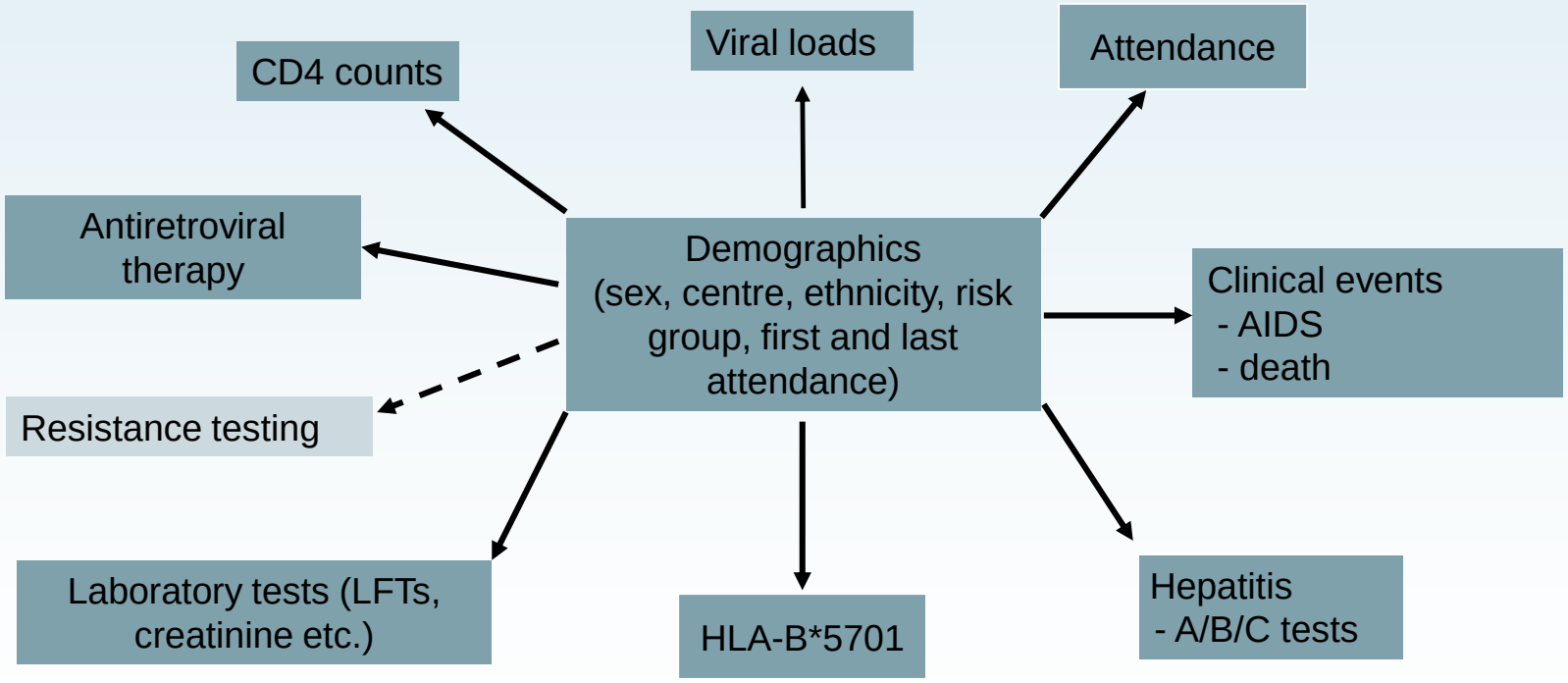
Q: Are treatment toxicities and antiretroviral resistance more likely to occur in those starting cART with high CD4 counts?

Q: Do ART-naïve HIV-positive adults initiating cART with CD4 counts above 350 cells/mm^3 experience a higher rate of antiretroviral toxicities and resistance development on cART than those initiating cART with CD4 counts of 350 cells/mm^3 and below?

What data do you think you would need to answer this research question?

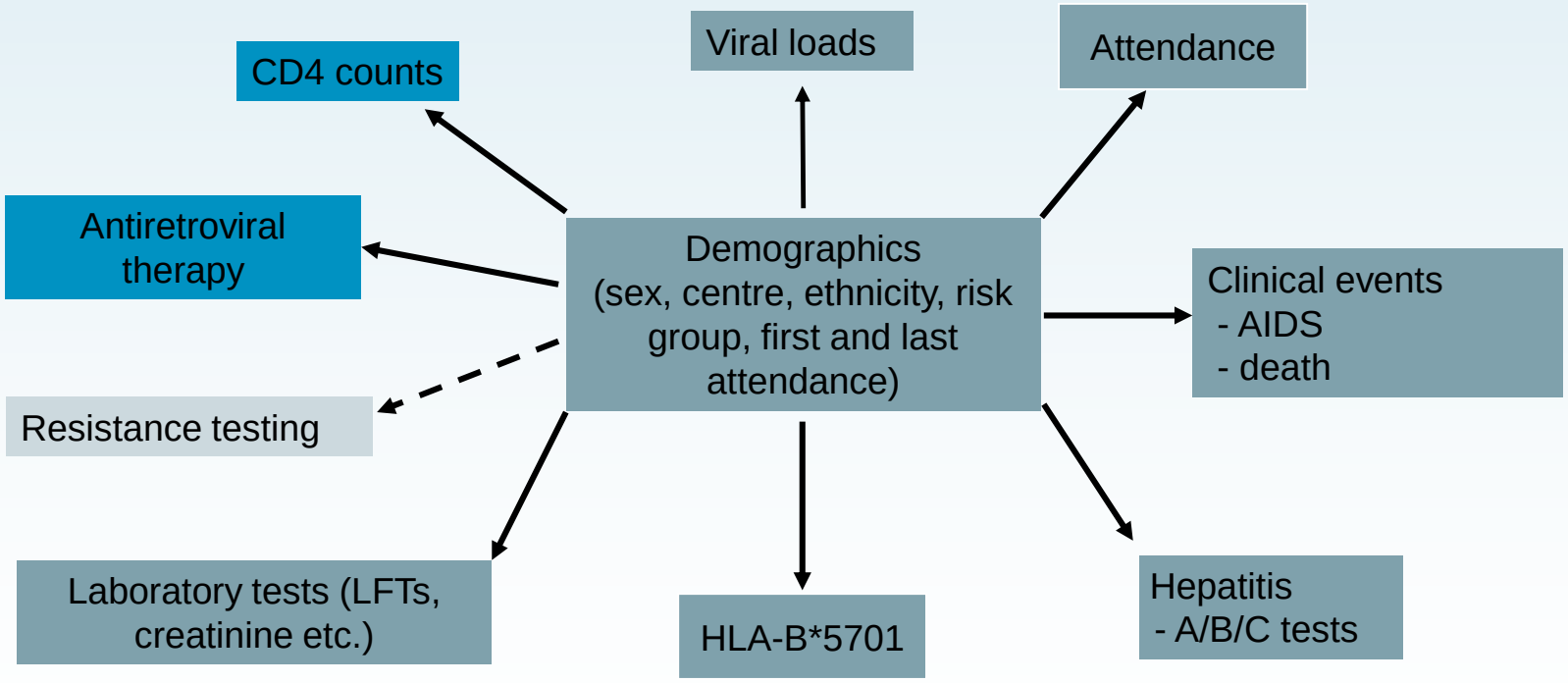
(3) Is the data we need available?

Q: Are treatment toxicities and antiretroviral resistance more likely to occur in those starting cART with high CD4 counts?



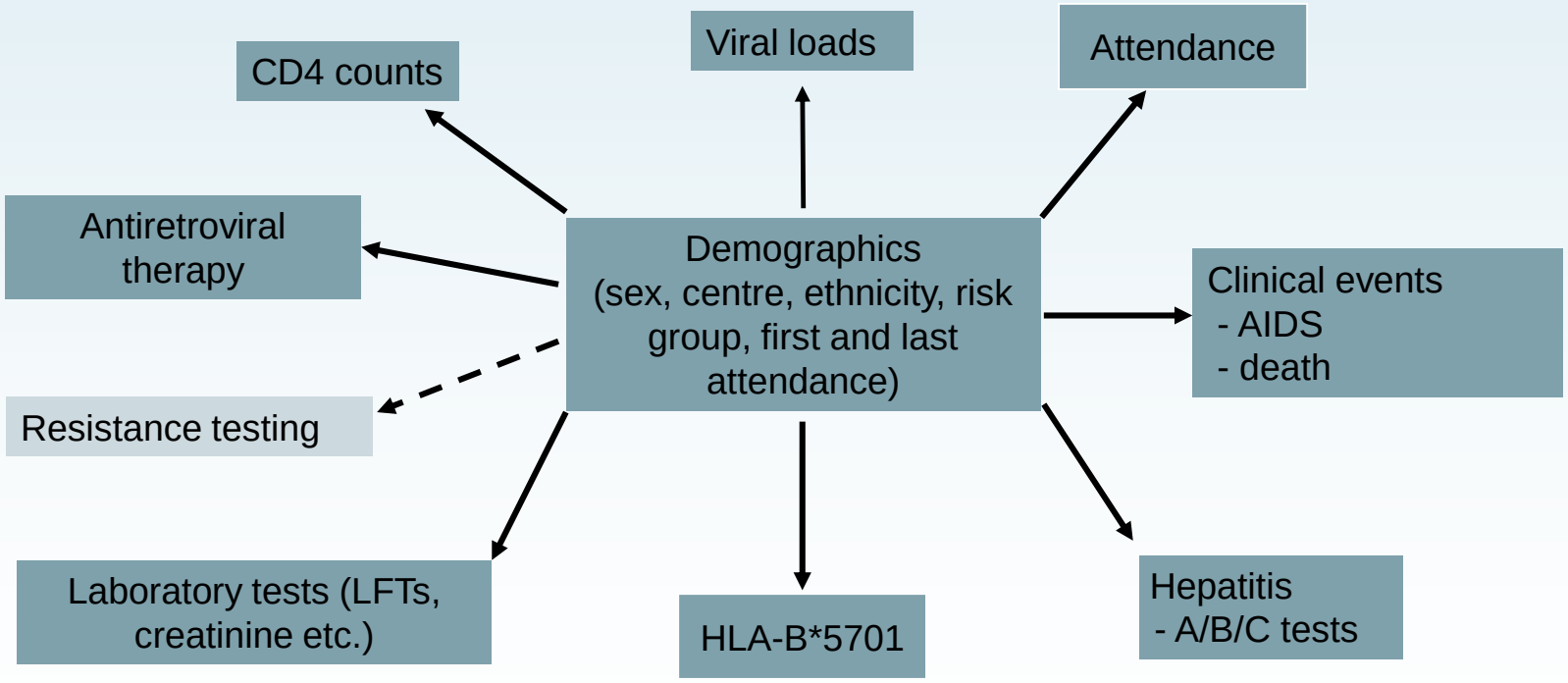
(3) Is the data we need available?

Q: Are treatment toxicities and antiretroviral resistance more likely to occur in those starting cART with high CD4 counts?



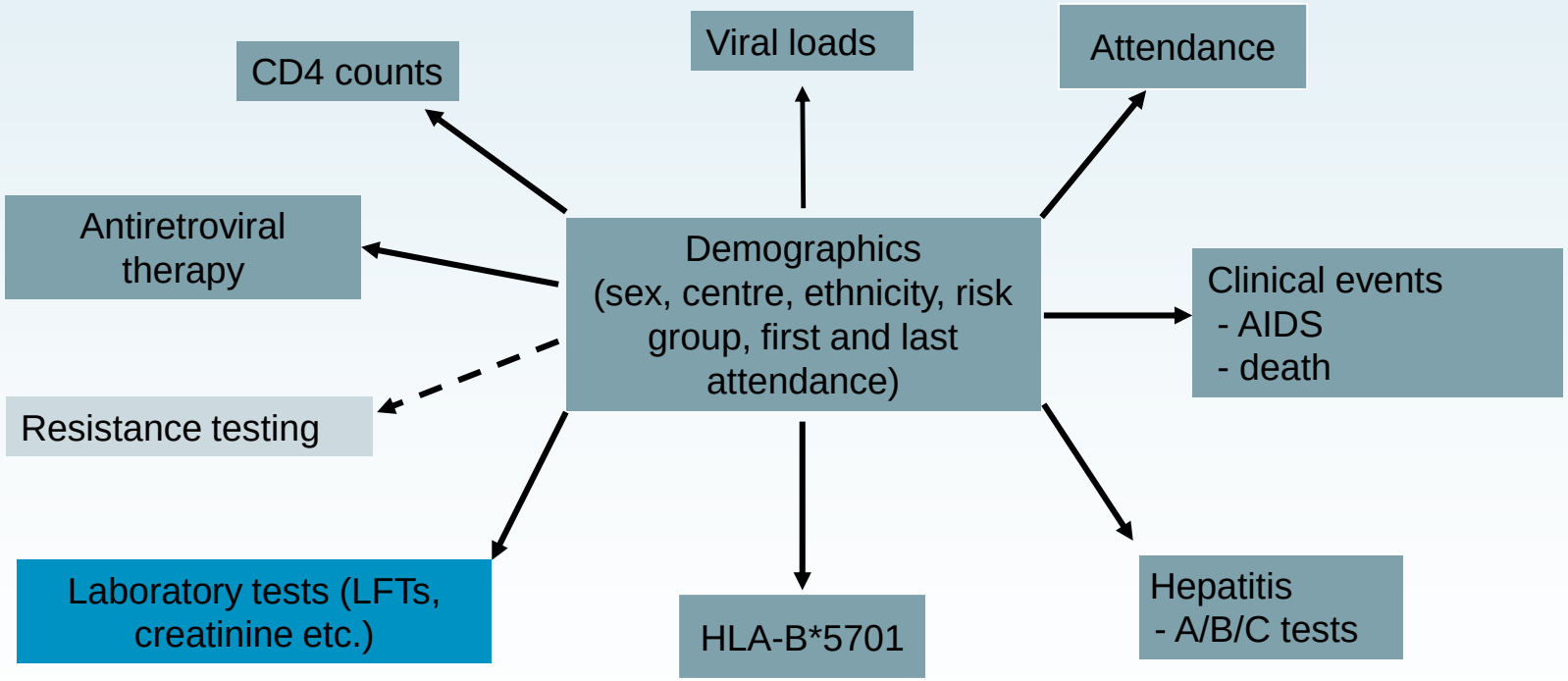
(3) Is the data we need available?

Q: Are treatment toxicities and antiretroviral resistance more likely to occur in those starting cART with high CD4 counts?



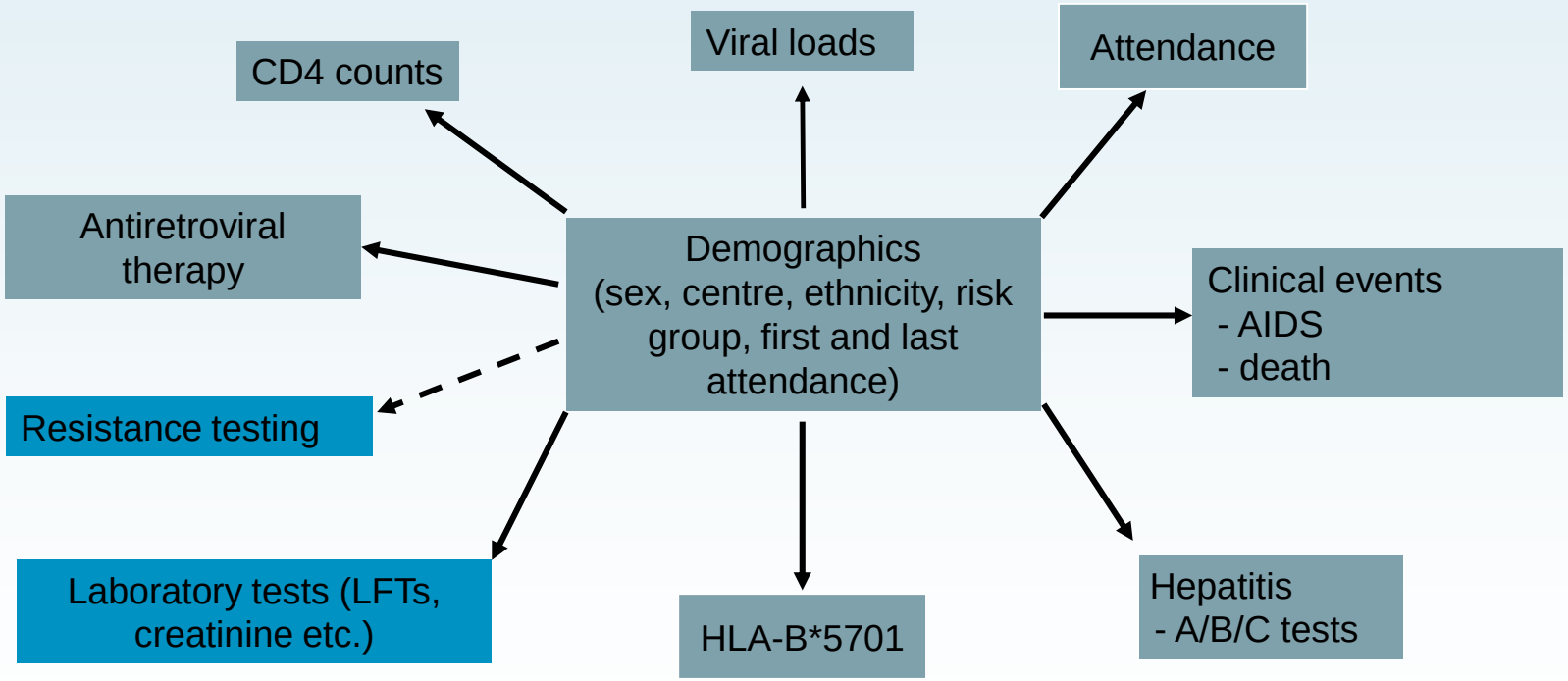
(3) Is the data we need available?

Q: Are treatment toxicities and antiretroviral resistance more likely to occur in those starting cART with high CD4 counts?



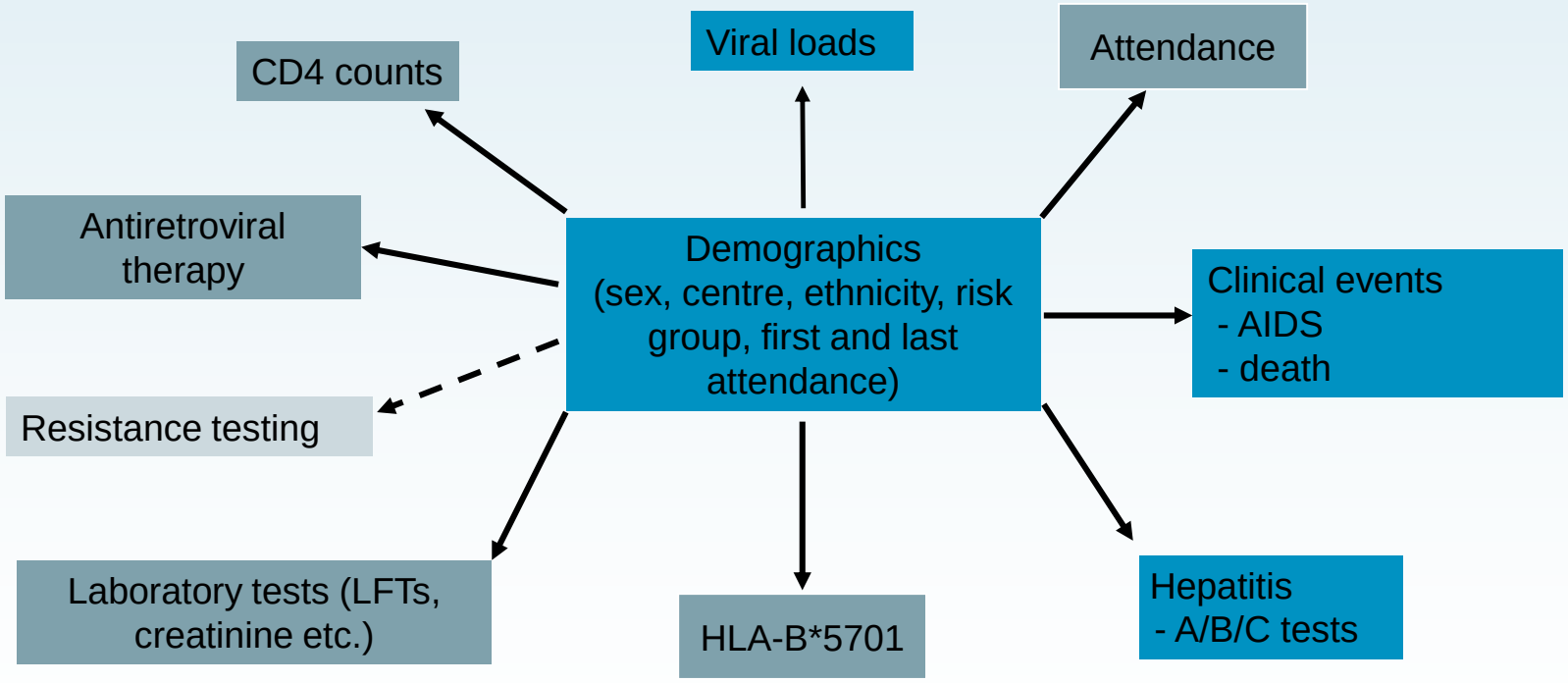
(3) Is the data we need available?

Q: Are treatment toxicities and antiretroviral resistance more likely to occur in those starting cART with high CD4 counts?



(3) Is the data we need available?

Q: Are treatment toxicities and antiretroviral resistance more likely to occur in those starting cART with high CD4 counts?



(4) Refining the research question

Population

HIV-positive adults initiating cART for first time

Also need to have laboratory tests, resistance tests, CD4 counts and viral load available

Exclusions – pregnancy
cART started before 2000

(4) Refining the research question

Exposure

I = People starting cART with CD4 count ≥500 or 351-499 cells/mm^3

C= People starting cART with CD4 count ≤350 cells/mm^3

(4) Refining the research question

Outcome (1)

Laboratory abnormalities

Grade 3/4 adverse events[1] in any of the following laboratory markers:

Liver	Renal	Blood	Other
ALT	Creatinine	Haemoglobin	Amylase
AST		Platelet count	Cholesterol
ALP			Glucose
Albumin			

[1] Division of AIDS Table for Grading the Severity of Adult and Paediatric Adverse Events

(4) Refining the research question

Outcome (2)

Antiretroviral resistance mutations

New PI, NNRTI or NRTI resistance mutations (compared to baseline) upon virological failure

(5) Refining the research question

Q: Are treatment toxicities and antiretroviral resistance more likely to occur in those starting cART with high CD4 counts?

Q1: Do ART-naïve HIV-positive adults initiating cART with CD4 counts above 350 cells/mm^3 experience a higher rate of laboratory-defined adverse events on cART than those initiating cART with CD4 counts of 350 cells/mm^3 and below?

(5) Refining the research question

Q: Are treatment toxicities and antiretroviral resistance more likely to occur in those starting cART with high CD4 counts?

Q2: Are ART-naïve HIV-positive adults initiating cART with CD4 counts above 350 cells/mm^3 more likely to develop antiretroviral resistance than those initiating cART with CD4 counts of 350 cells/mm^3 and below?

(6) What biases/limitations do we need to consider when using routinely collected data?

- *Confounding*
  - Who starts cART with high CD4 counts in routine practice?
  - Likely to differ from those starting cART below 350 cells/mm^3 threshold
  - Do we collect data on important confounders?
  - Can we adjust for these in our analyses?
  - Unmeasured confounding

(6) What biases/limitations do we need to consider when using routinely collected data?

		CD4 count at start of ART (cells/mm^3)		
		≤350	351-499	≥500
Sex, n (%)	Male	**6147 (78.2)**	**958 (87.2)**	**406 (90.8)**
Ethnicity, n (%)	White	4586 (58.4)	787 (71.6)	334 (74.7)
	Black African	1861 (23.7)	131 (11.9)	42 (9.4)
	Black other	404 (5.1)	52 (4.7)	17 (3.8)
	Other/unknown	1009 (12.8)	129 (11.7)	54 (12.1)
Mode of HIV acquisition, n (%)	Sex between men	4518 (57.5)	801 (72.9)	347 (77.6)
	Heterosexual	2739 (34.9)	212 (19.3)	68 (15.2)
	Other/unknown	603 (7.7)	86 (7.8)	32 (7.2)
Regimen type, n (%)	2 NRTI + PI (/r)	1893 (24.1)	311 (28.3)	186 (41.6)
	2 NRTI + NNRTI	5559 (70.7)	718 (65.3)	236 (52.8)
	≥ 3 NRTI	173 (2.2)	22 (2.0)	8 (1.8)
	Other Combination	235 (3.0)	48 (4.4)	17 (3.8)

(6) What biases/limitations do we need to consider when using routinely collected data?

		CD4 count at start of ART (cells/mm^3)		
		≤350	351-499	≥500
Sex, n (%)	Male	6147 (78.2)	958 (87.2)	406 (90.8)
Ethnicity, n (%)	White	**4586 (58.4)**	**787 (71.6)**	**334 (74.7)**
	Black African	1861 (23.7)	131 (11.9)	42 (9.4)
	Black other	404 (5.1)	52 (4.7)	17 (3.8)
	Other/unknown	1009 (12.8)	129 (11.7)	54 (12.1)
Mode of HIV acquisition, n (%)	Sex between men	**4518 (57.5)**	**801 (72.9)**	**347 (77.6)**
	Heterosexual	2739 (34.9)	212 (19.3)	68 (15.2)
	Other/unknown	603 (7.7)	86 (7.8)	32 (7.2)
Regimen type, n (%)	2 NRTI + PI (/r)	1893 (24.1)	311 (28.3)	186 (41.6)
	2 NRTI + NNRTI	5559 (70.7)	718 (65.3)	236 (52.8)
	≥ 3 NRTI	173 (2.2)	22 (2.0)	8 (1.8)
	Other Combination	235 (3.0)	48 (4.4)	17 (3.8)

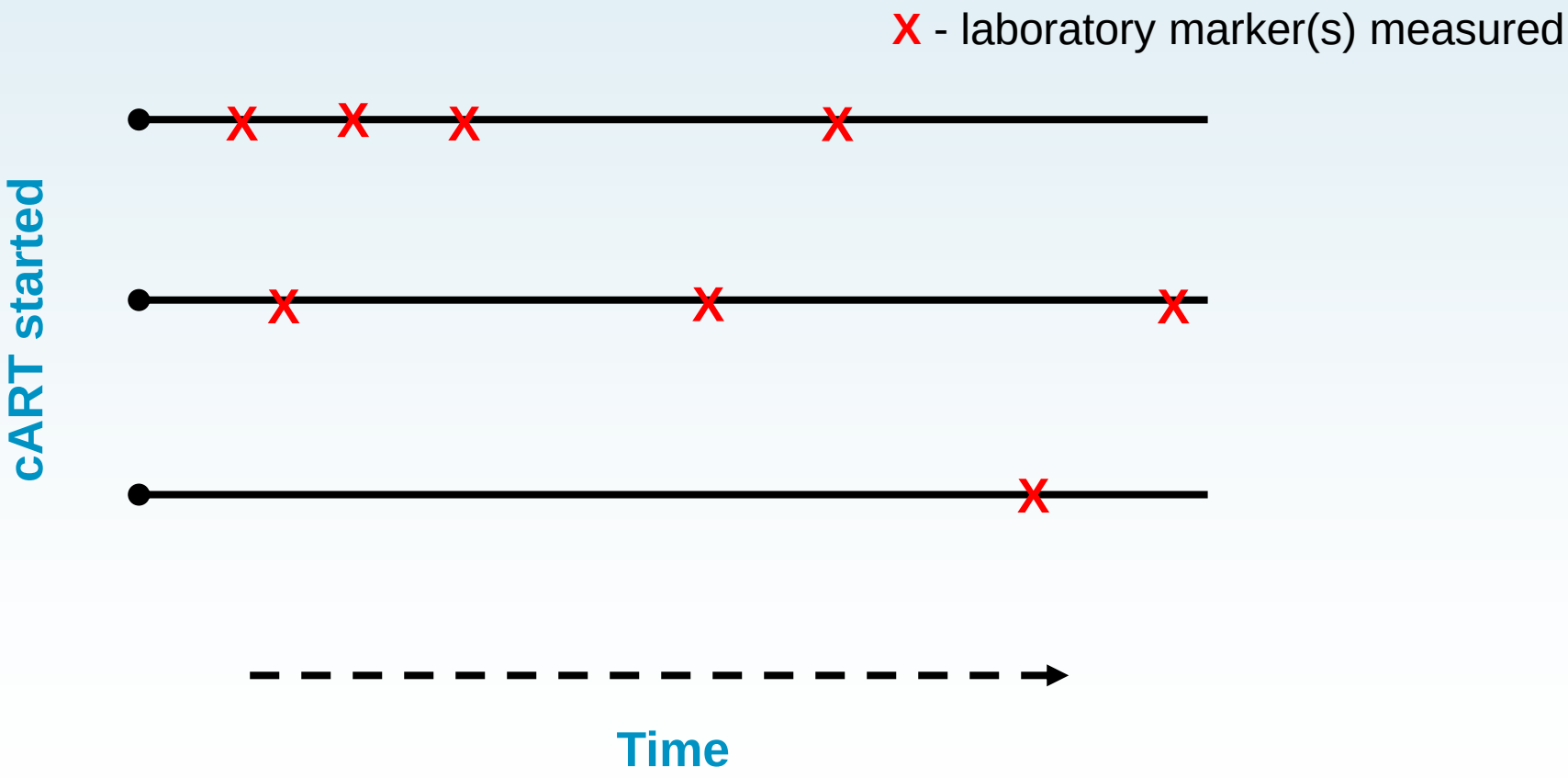
(6) What biases/limitations do we need to consider when using routinely collected data?

		CD4 count at start of ART (cells/mm^3)		
		≤350	351-499	≥500
Sex, n (%)	Male	6147 (78.2)	958 (87.2)	406 (90.8)
Ethnicity, n (%)	White	4586 (58.4)	787 (71.6)	334 (74.7)
	Black African	1861 (23.7)	131 (11.9)	42 (9.4)
	Black other	404 (5.1)	52 (4.7)	17 (3.8)
	Other/unknown	1009 (12.8)	129 (11.7)	54 (12.1)
Mode of HIV acquisition, n (%)	Sex between men	4518 (57.5)	801 (72.9)	347 (77.6)
	Heterosexual	2739 (34.9)	212 (19.3)	68 (15.2)
	Other/unknown	603 (7.7)	86 (7.8)	32 (7.2)
Regimen type, n (%)	2 NRTI + PI (/r)	**1893 (24.1)**	**311 (28.3)**	**186 (41.6)**
	2 NRTI + NNRTI	5559 (70.7)	718 (65.3)	236 (52.8)
	≥ 3 NRTI	173 (2.2)	22 (2.0)	8 (1.8)
	Other Combination	235 (3.0)	48 (4.4)	17 (3.8)

(6) What biases/limitations do we need to consider when using routinely collected data?

- *Infrequent monitoring*
  - Laboratory tests not performed at regular intervals in all patients
  - More likely to have a test if sick or displaying symptoms
  - Resistance testing only performed if virological failure

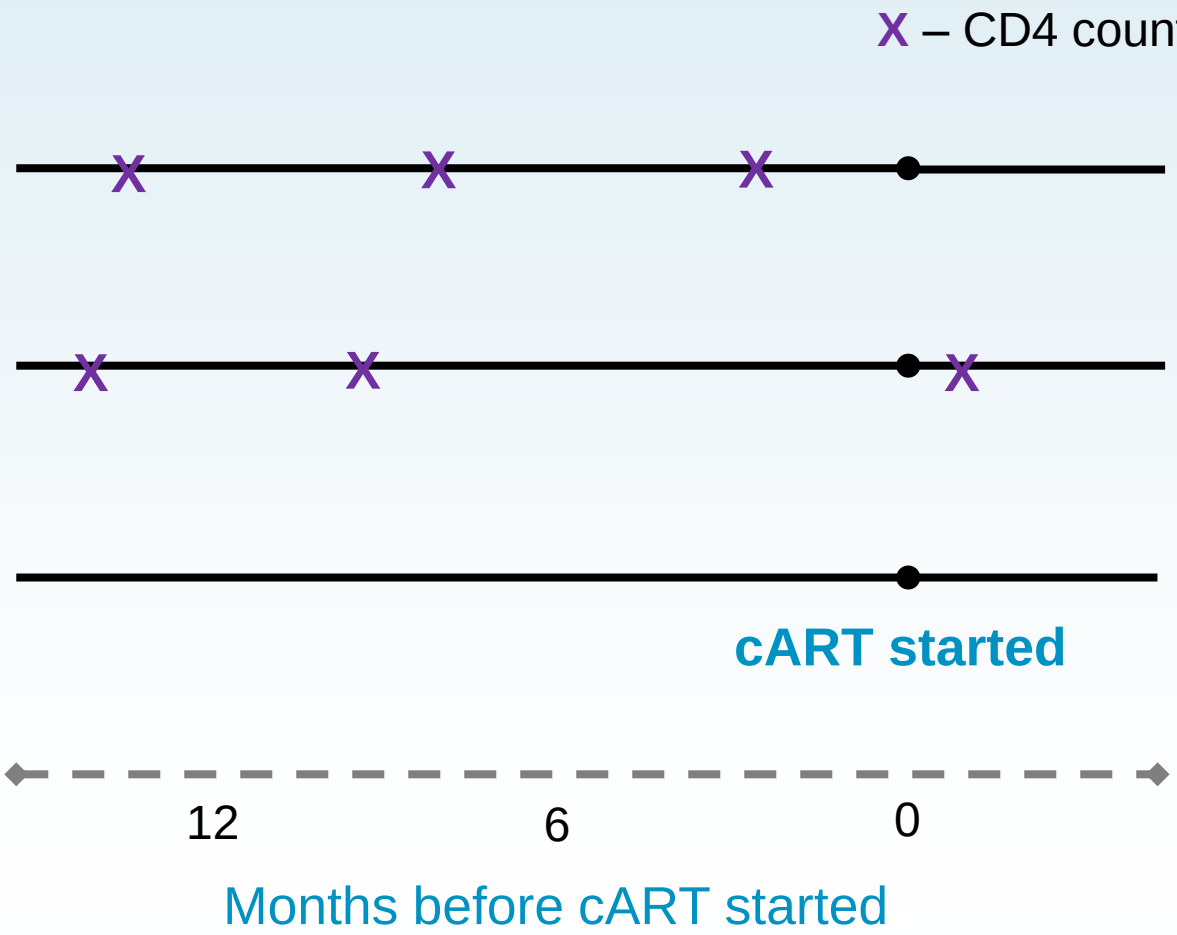
(6) What biases/limitations do we need to consider when using routinely collected data?



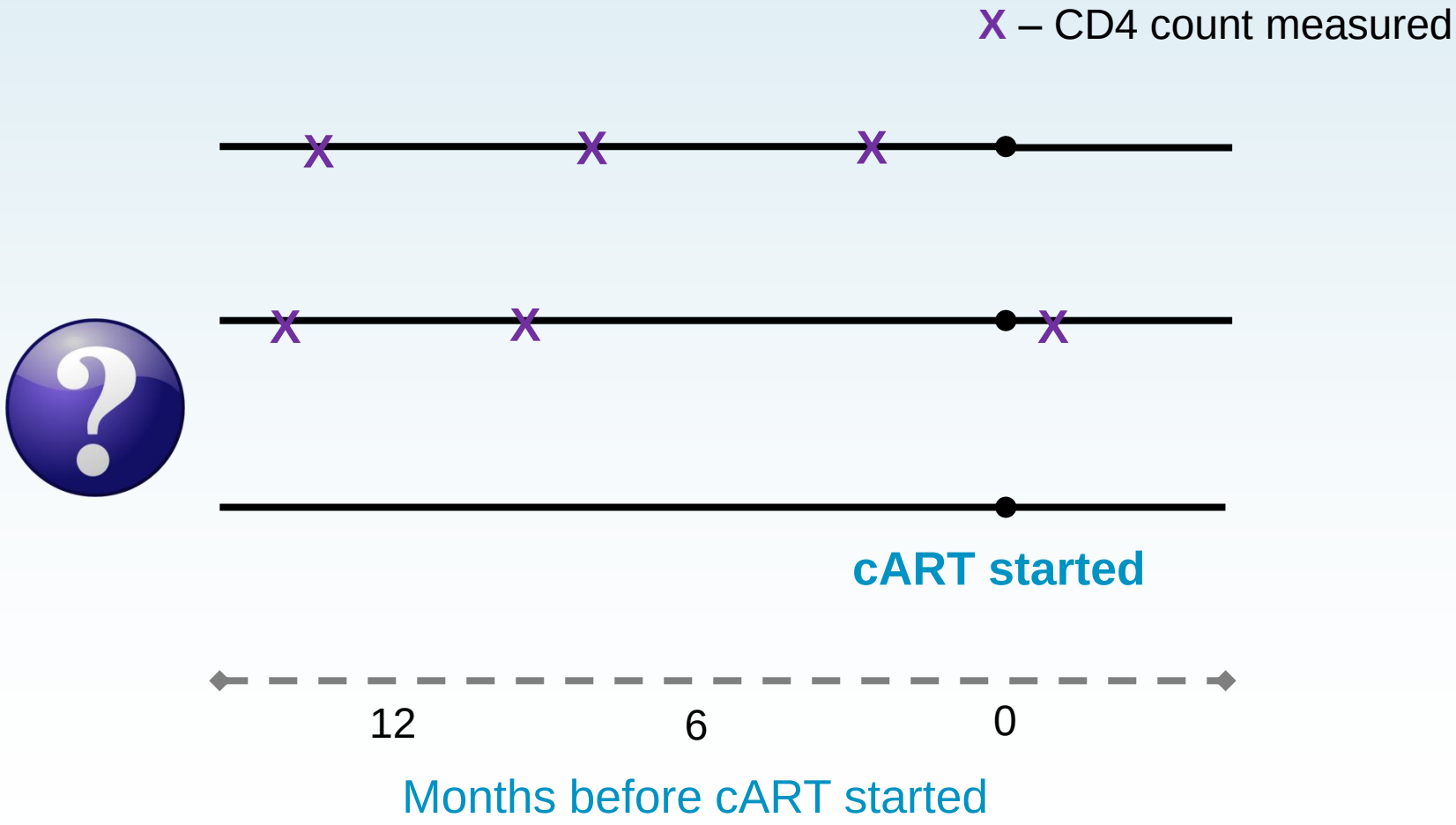
(6) What biases/limitations do we need to consider when using routinely collected data?

- *Missing data*
  - CD4 count at cART start
  - Censor patient follow-up when lost to follow-up

(6) What biases/limitations do we need to consider when using routinely collected data?



(6) What biases/limitations do we need to consider when using routinely collected data?



(7) Feasibility analyses

- Need to have enough people starting cART at CD4 counts above 350 cells/mm^3
- Patients need to have:
  - laboratory measures available
  - resistance test results available at cART start and virological failure
- Will our sample size be large enough?
- Will we observe enough outcome events?

(8) Writing the concept sheet

- 1-1.5 pages

- Title
- Hypothesis/background
- Aim(s)
- Inclusion/exclusion
- Proposed analysis
- Variables needed
- Possible limitations

(8) Writing the concept sheet

- 1-1.5 pages

- Title ✓
- Hypothesis/background ✓
- Aim(s) ✓
- Inclusion/exclusion ✓
- *Proposed analysis*
- Variables needed ✓
- Possible limitations ✓

(9) Analysis plan

- Research question and choice of exposure and outcome variables will determine most appropriate statistical tests/models to use
- Should be outlined in concept sheet
- Sensitivity and sub-group analyses may be needed – should be planned a priori

But . . . .

- May have to be more flexible than in RCTs
- As project develops, analysis plan may change

(9) Analysis plan

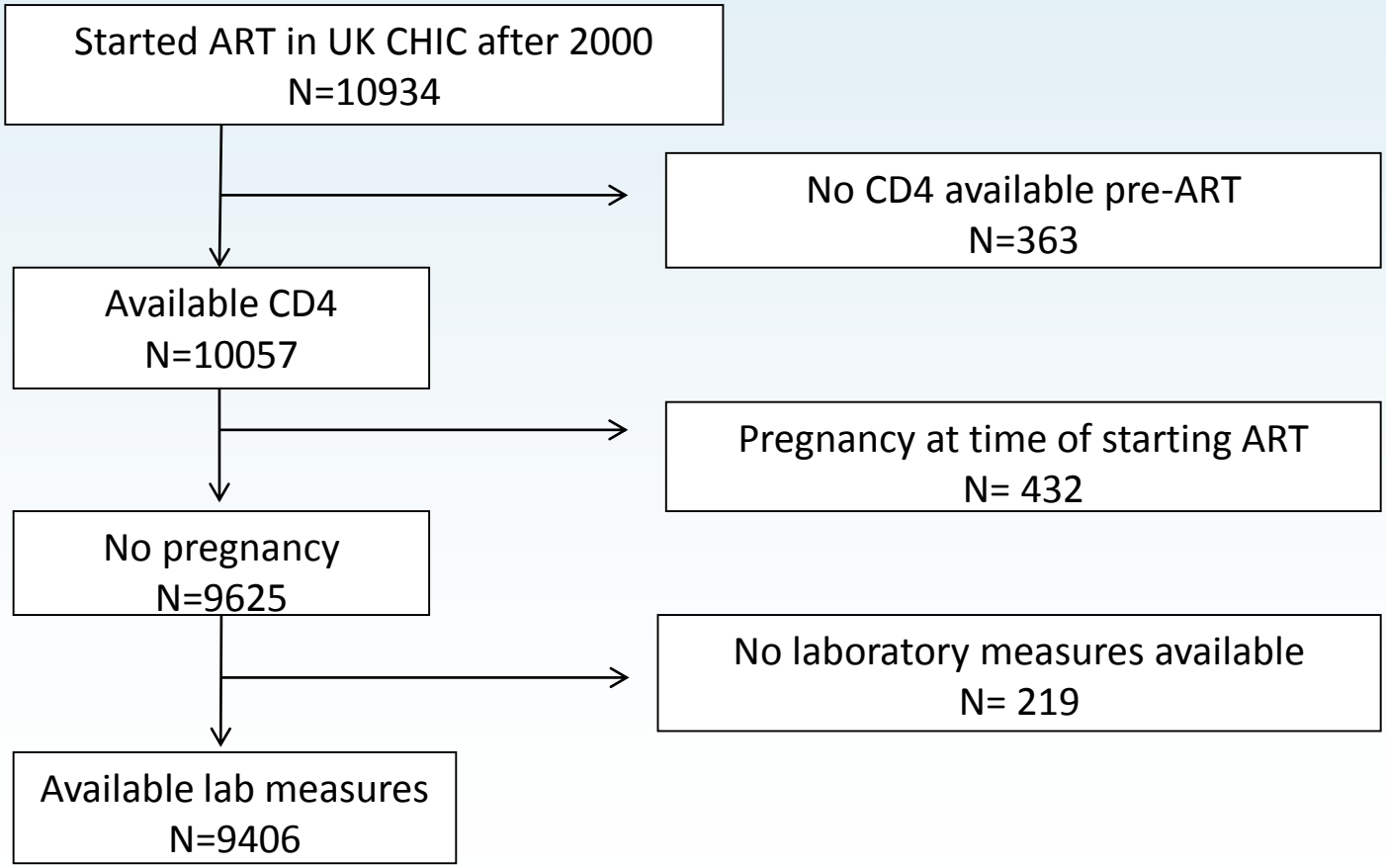
- Descriptive analyses
  - Get to know your data!
  - Identify differences in exposure groups and potential confounders
- Exploratory analyses
  - Not your main end-point
  - Provide some insight/aids interpretation of main results
- Main analysis
  - Analysis of primary end-point
  - Adjusting for confounders (regression models)
- Sensitivity and sub-group analyses
  - Are methods valid?

Descriptive analyses

- Patient flow-chart through study
  - Numbers eligible, excluded, analysed
  - Compare those included/excluded

Descriptive analyses

- Patient flow-chart through study

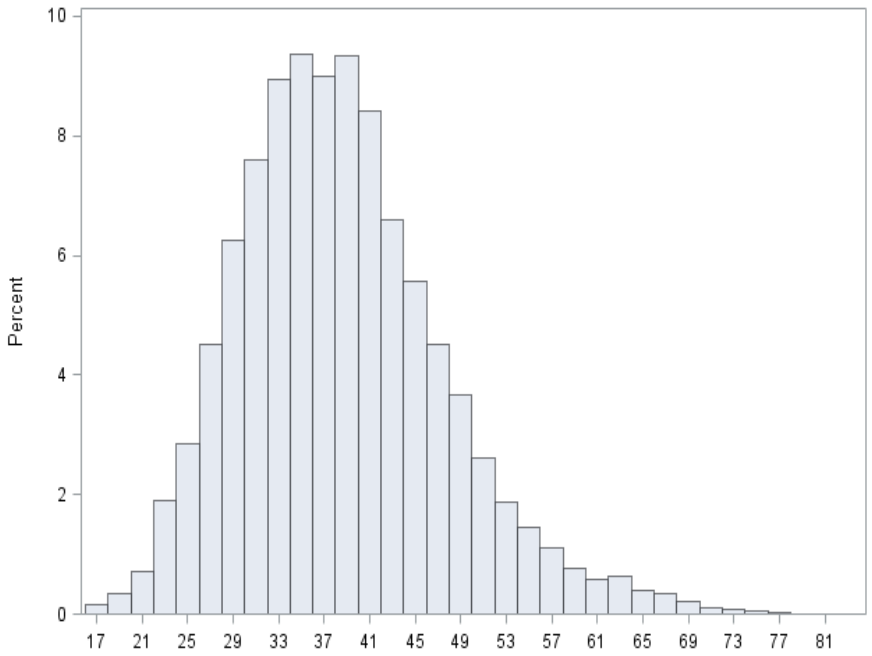


Descriptive analyses

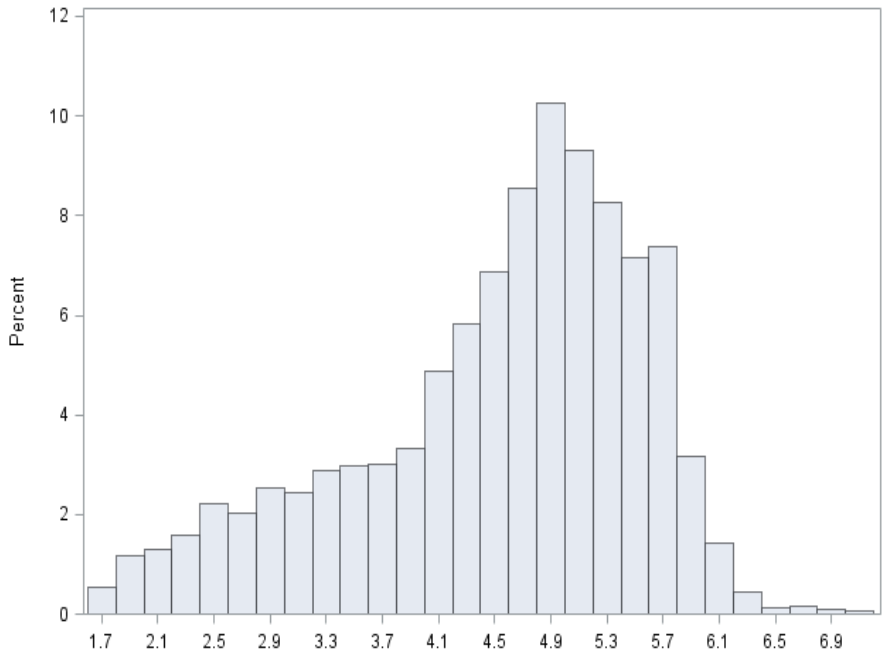
- Baseline characteristics
  - Exposure groups
  - Demographics, HIV markers, AIDS events, cART regimen, calendar year, HBV/HCV co-infection etc.
  - Data checks (errors, outliers, normal distribution, missing data etc.)
  - Comparison by exposure of interest (CD4 count groups)
  - Univariate tests (chi-square, t-test, ANOVA etc.)

Descriptive analyses

- Data checks



Age at cART start



Log_{10} VL at cART start

Descriptive analyses

- Baseline characteristics

		CD4 count at start of ART (cells/mm^3)			P-value
		≤350	351-499	≥500	
Sex, n (%)	Male	6147 (78.2)	958 (87.2)	406 (90.8)	<0.001
Ethnicity, n (%)	White	4586 (58.4)	787 (71.6)	334 (74.7)	<0.001
	Black African	1861 (23.7)	131 (11.9)	42 (9.4)	
	Black other	404 (5.1)	52 (4.7)	17 (3.8)	
	Other/unknown	1009 (12.8)	129 (11.7)	54 (12.1)	
Mode of HIV acquisition, n (%)	Sex between men	4518 (57.5)	801 (72.9)	347 (77.6)	<0.001
	Heterosexual	2739 (34.9)	212 (19.3)	68 (15.2)	
	Other/unknown	603 (7.7)	86 (7.8)	32 (7.2)	
Regimen type, n (%)	2 NRTI + PI (/r)	1893 (24.1)	311 (28.3)	186 (41.6)	<0.001
	2 NRTI + NNRTI	5559 (70.7)	718 (65.3)	236 (52.8)	
	≥ 3 NRTI	173 (2.2)	22 (2.0)	8 (1.8)	
	Other Combination	235 (3.0)	48 (4.4)	17 (3.8)	

Descriptive analyses

- Patient follow-up
  - Time under follow-up
  - Frequency of monitoring (laboratory tests, resistance tests)
  - Comparison by exposure of interest (CD4 count groups)
  - Univariate tests

Descriptive analyses

- Patient follow-up

		CD4 count at start of ART (cells/mm^3)		
		≤350	351-499	≥500
Patient follow up, years	Sum	24628.0	2330.4	902.3
	Median (IQR)	2.5 (0.0, 11)	1.4 (0.0, 10.6)	1.0 (0.0, 10.6)
Average number of laboratory tests/ year	Median (IQR)	2 (0,3)	2 (0,3)	3 (0,5)

Descriptive analyses

- Outcome
  - Crude rate of grade 3/4 laboratory adverse events
  - Number of resistance mutations at viral failure**
  - Comparison by exposure of interest (CD4 count groups)
  - Univariate tests

Descriptive analyses

- Outcome

Baseline CD4 count, cells/mm^3	N	Virological rebound, n (%)	Resistance test, n (%)	New resistance mutation, n (%)
≤350	6514	488 (7.5)	260 (53.3)	107 (41.2)
351-499	996	46 (4.6)	20 (43.5)	**3 (15.0)**
≥500	408	30 (7.4)	10 (33.3)	**1 (10.0)**
Total	7918	564 (7.1)	290 (51.4)	111 (38.3)
p-value		0.005	0.056	0.012

Exploratory analyses

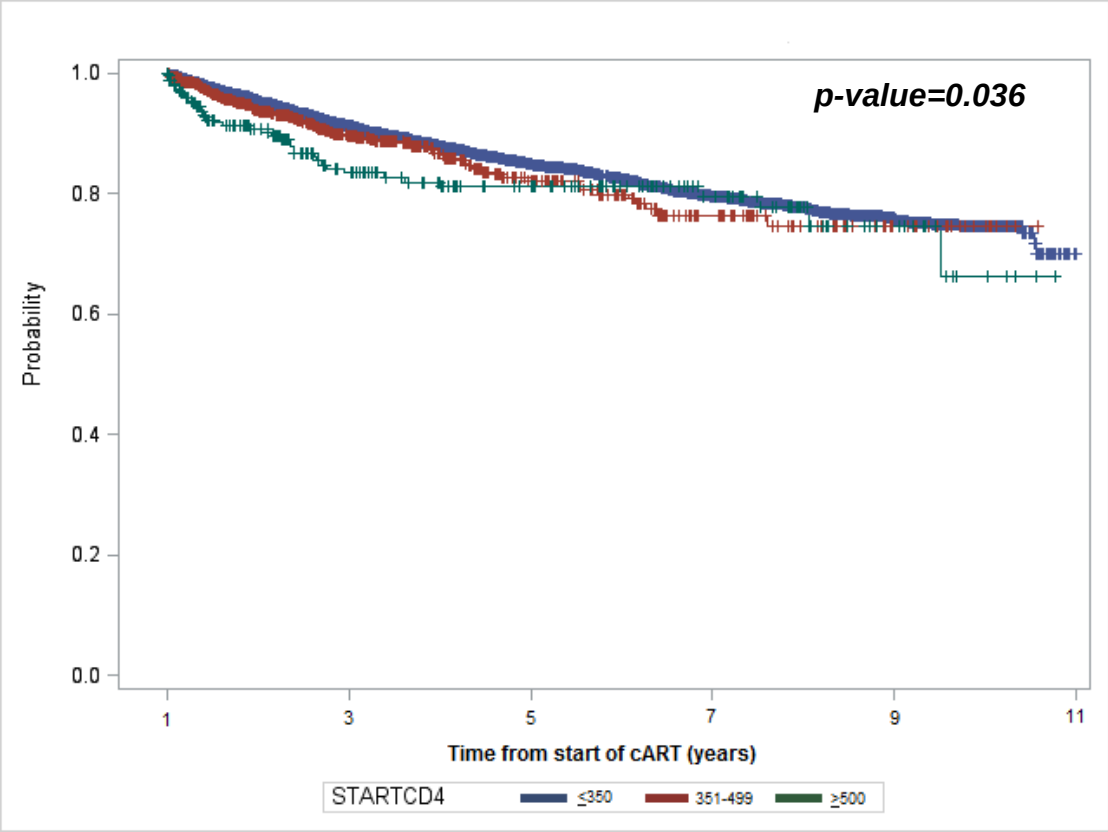
Aware of large limitation - reasons for starting cART with high CD4 counts not known

Undertook range of preliminary analyses to understand differences between 3 CD4 count groups

- Predictors of starting cART at high CD4 count
- How was CD4 count associated with following outcomes?
  - Virological suppression
  - Virological rebound
  - Treatment switching
  - Discontinuation of cART

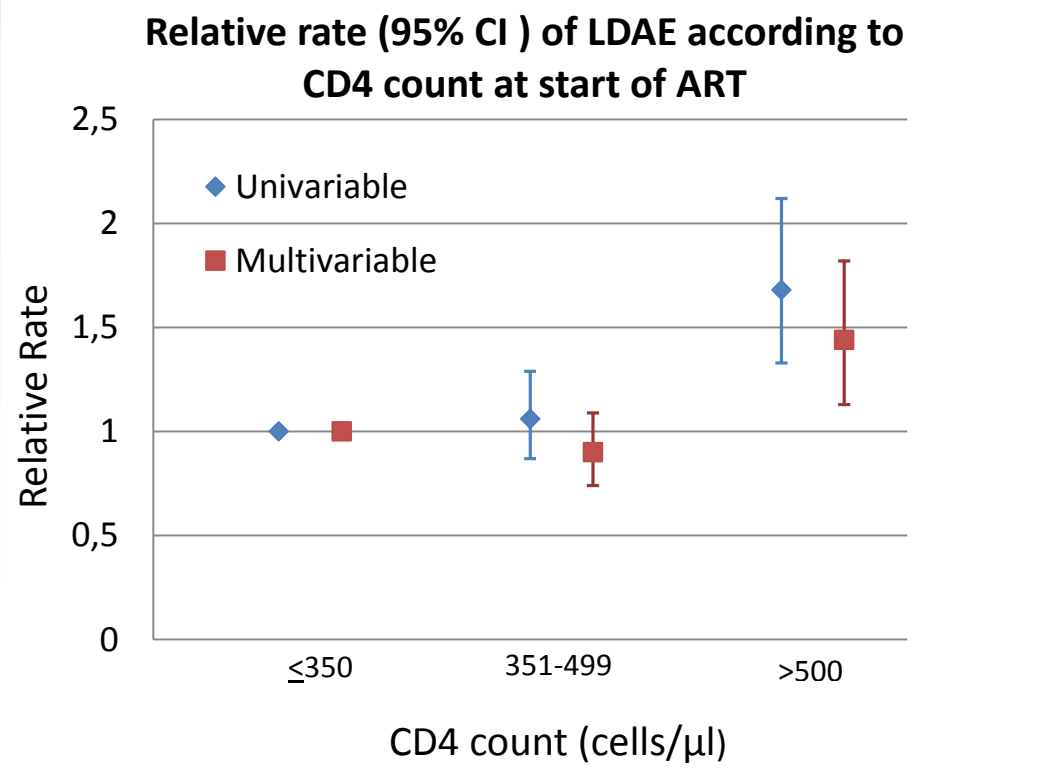
Exploratory analyses

Kaplan-Meier graph of time to discontinuation of cART, according to CD4 count at start of cART



Main analysis

- Should answer research question
- Provide estimates that are adjusted for measured confounders (regression models)



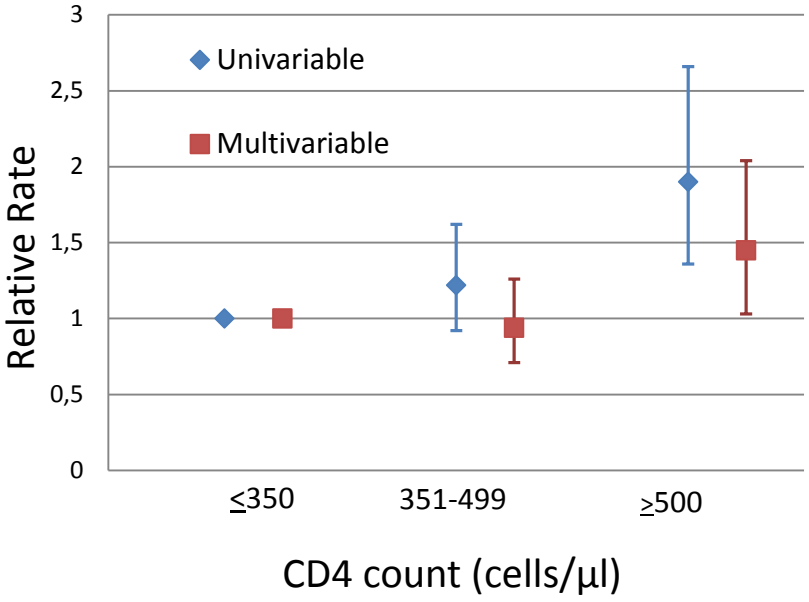
Sensitivity/sub-group analyses

- If we change variable definitions, do our conclusions remain unchanged?
  - LDAE grouped by type (LFTS, renal function, blood, other) and analysed separately
  - Considered absolute change in laboratory measures

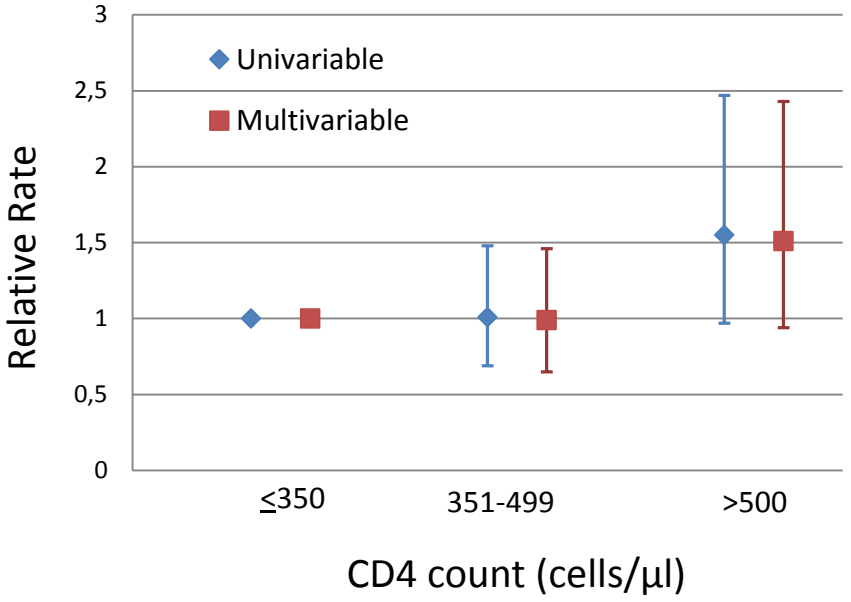
- If we change the population studied, do our conclusions remain unchanged?
  - Excluded those with HBV co-infection at cART start

Sensitivity/sub-group analyses

Relative rate (95% CI) of liver-related LDAE according to CD4 count at start of ART



Relative rate (95% CI) of blood-related LDAE according to CD4 count at start of ART



(10) Write the paper!

CONCISE COMMUNICATION

Laboratory adverse events and discontinuation of therapy according to CD4$^+$ cell count at the start of antiretroviral therapy

Sophie Jose[a], Killian Quinn[b], Teresa Hill[a], Clifford Leen[c], John Walsh[b], Phillip Hay[d], Martin Fisher[e], Frank Post[f], Mark Nelson[g], Mark Gompels[h], Margaret Johnson[i], David Chadwick[j], Richard Gilson[k], Caroline Sabin[a], Sarah Fidler[b], on behalf of the UK CHIC Steering Committee

Objective: Few data describe antiretroviral treatment (ART)-related adverse events when treatment is initiated at CD4 cell counts more than 350 cells/μl. We compared rates of laboratory-defined adverse events (LDAEs) according to CD4 cell count at ART initiation.

Design: Analysis of on-going cohort study.

Methods: ART-naive persons initiating ART from 2000 to 2010 were included. Chi-square, analysis of variance (ANOVA) and Kruskal–Wallis tests compared character-

Summary

- It is possible to answer a research question using data from your own clinic and without the need for big programme grants
- Need to be aware of potential biases and limitations that are present
- Lots of background work/analyses that don’t make it to the finished article
- Get statistical advice from the start!

Example

- You want to know whether starting patients on a single tablet regimen leads to better outcomes than multi-tablet regimens . . . . . .

Example

- You want to know whether starting patients on a single tablet regimen leads to better outcomes than multi-tablet regimens . . . . . .

- **Could you investigate this using your own data?**

- **What limitations/bias would you need to consider?**

- **What descriptive analyses would you undertake?**