Assessing the representativeness of European HIV cohort participants as compared to HIV surveillance data

G. Vourli, A. Pharris, F. Cazein, D. Costagliola, F. Dabis, J. Del Amo,
V. Delpech, A. Díaz, E. Girardi, A. Gourlay, B. Gunsenheimer-Bartmeyer,
V. Hernando, G. Nikolopoulos, K. Porter, M. Rosińska, C. Sabin,
B. Suligoi, V. Supervie, F. Wit, G. Touloumi

Background

- ECDC collects data through the European Surveillance System (TESSy)
  - Demographic variables such as age group, gender, transmission mode, country of origin are included
  - But data on clinical indicators are often lacking

- EuroCoord cohort data
  - European Network of HIV/AIDS Cohort Studies to Coordinate at European and International Level Clinical Research on HIV/AIDS
  - Data on clinical indicators are available

Rationale

- Explore the possibility of using data from the European HIV cohorts within EuroCoord in settings where surveillance data aren't available or don't cover the outcome of interest

- Allow ECDC, EU Member States, and the European Commission to use cohort data to inform aspects of HIV epidemic generalising cohorts' finding on a national level

Issues

- Cohorts typically include a subset of the HIV diagnosed individuals: those linked to care
- A selected sample that may be different from the whole HIV(+) population
- Specific issues:
  - Geographical coverage
  - Systematic exclusion of specific group(s) of patients (e.g. IDU, migrants)
  - Patient's consent occasionally required, causing further restrictions
  - Patients with advanced disease may have a different probability of being included
- Surveillance systems may have limitations, too.
  - Changes over time regarding geographical coverage
  - Underreporting/ reporting delays
  - Missing data
  - Lack of data on outcome (death) or outmigration of cases notified historically

Cohort vs Surveillance Data

- 11 European countries with an HIV cohort within EuroCoord

- 5 share data with the national surveillance system

- 6 do not share data with the national surveillance system
  - Most of these include data from a subset of HIV diagnosed individuals
  - Started from a specific time point during the epidemic
  - Had substantial changes over time

Project Aim

- Assess the representativeness of data on HIV patients within the European cohorts against persons diagnosed with HIV and reported to TESSy

- Improve understanding of whether and where results from cohort data can be generalised

- Explore and propose methods to improve cohorts' representativeness

Methods

- France, Germany, Greece, Italy, Spain and the UK provided individual cohort data.

- To accommodate countries' specific features, the comparison focused on new cases diagnosed during three time periods [2000-2004], [2005-2009] and [2010-2013].

- Distribution of individuals' age, gender, transmission mode and region of origin were compared.

- Models for the probability of a diagnosed individual to be included in a cohort were applied.

- Weights inversely proportional to the probability of inclusion were generated for each covariate pattern and assigned to each cohort participant.

- Stabilised weights where the denominator represents the cohort's coverage were also produced. Thus, values <1 indicate over-representation and values >1 under-representation in the cohort

Results
Cohorts' Coverage
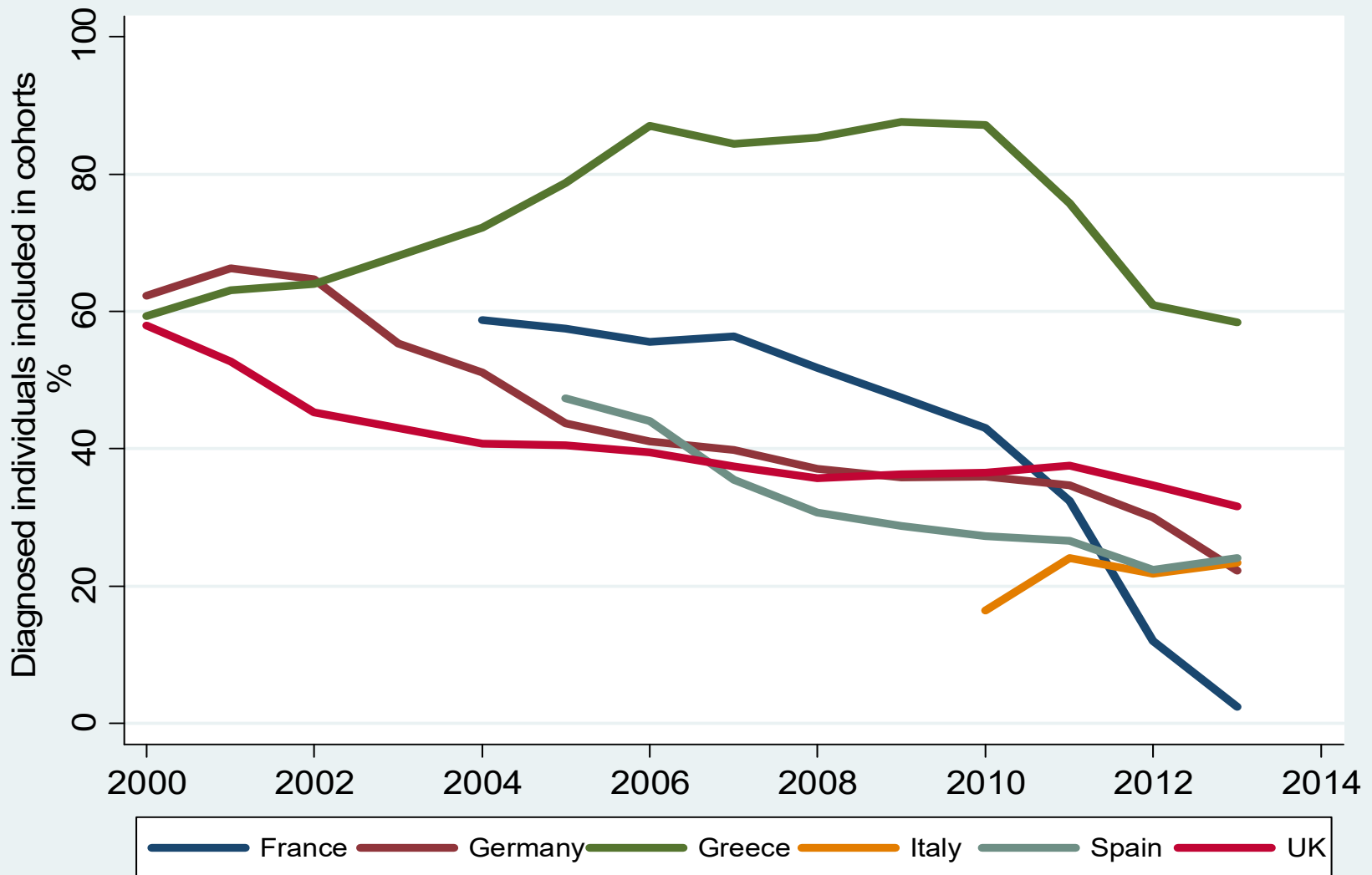
Inclusion Weights
Probability of inclusion

All patients	France	Germany	Greece	Italy	Spain	UK
2000-2004	1.7	1.7	1.5	-	-	2.2
2005-2009	1.9	2.5	1.2	-	2.9	2.6
2010-2013	4.4	3.3	1.4	4.7	4.0	2.8

- Irrespectively of her/his characteristics, each cohort participant should contribute 1.2-4.7 copies of herself/himself to reproduce the population of diagnosed individuals in her/his country in each of the study periods

- Subgroups of HIV diagnosed individuals may have different chances of being included in the cohorts

- According to logistic regression models people injecting drugs, those born in another country and those with low CD4 counts at diagnosis were less likely to be included in almost all cohorts. Women and older individuals were also under-represented occasionally
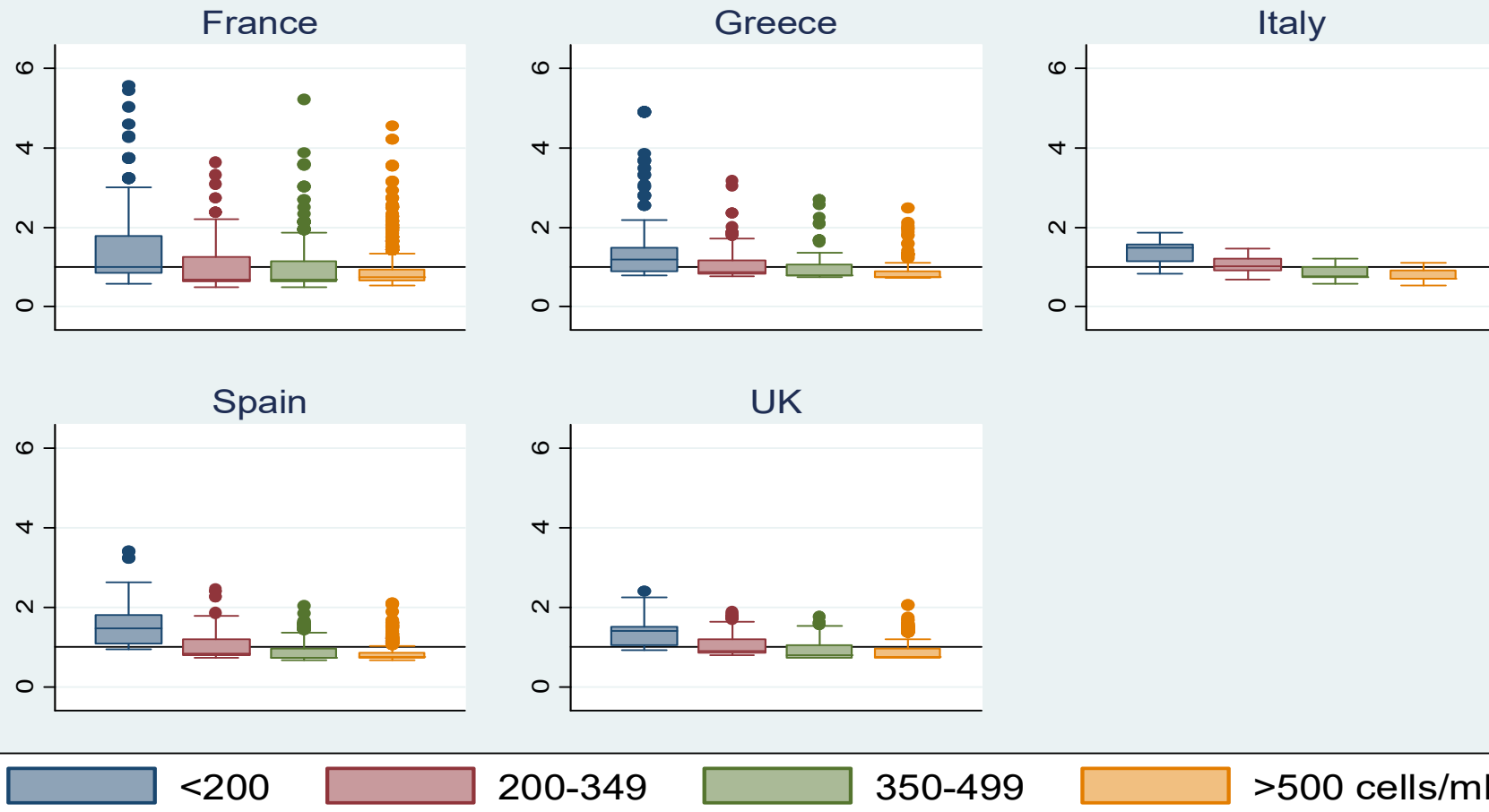
Stabilised weights according to transmission mode
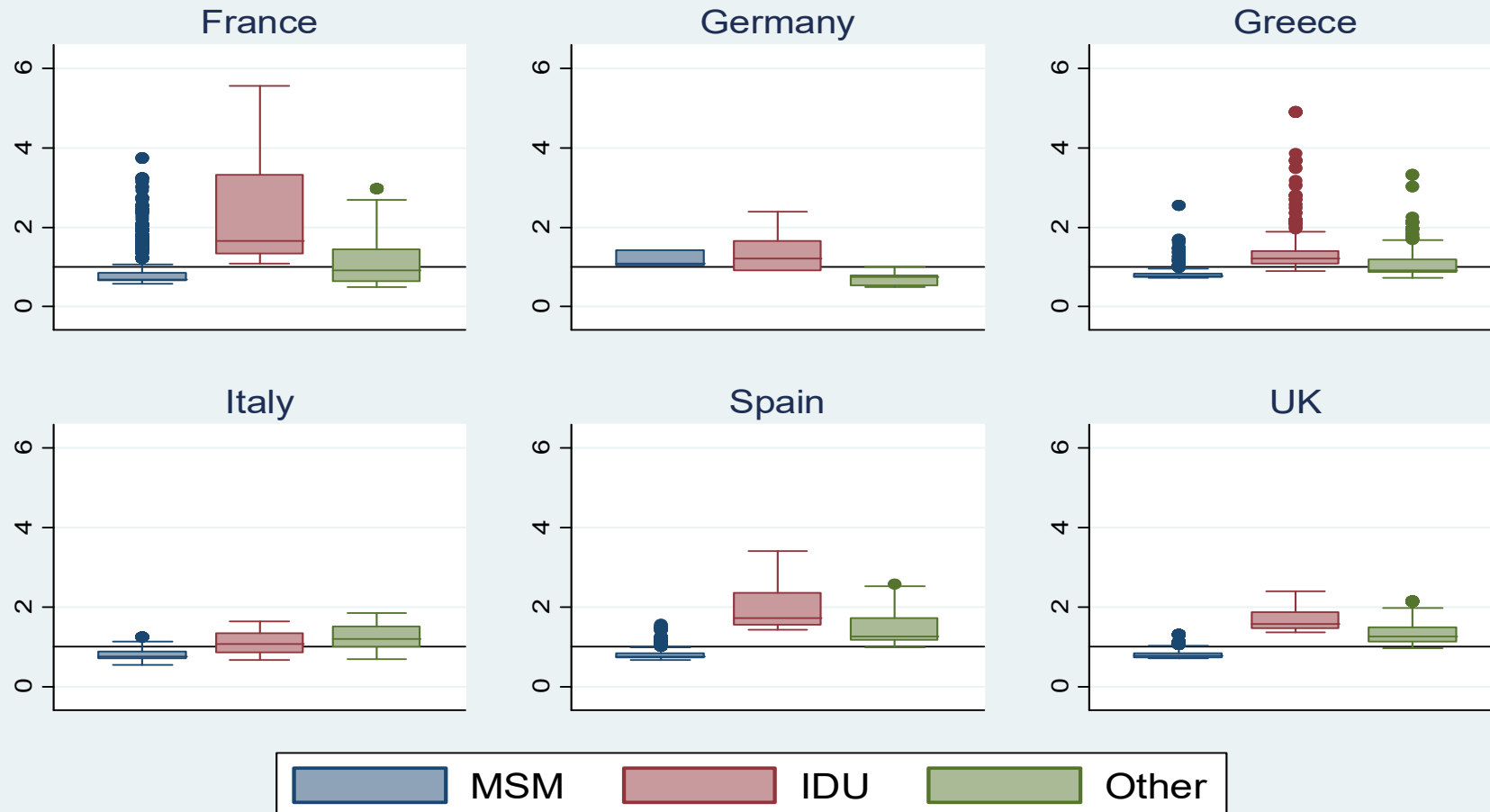
Stabilised weights according to CD4 category
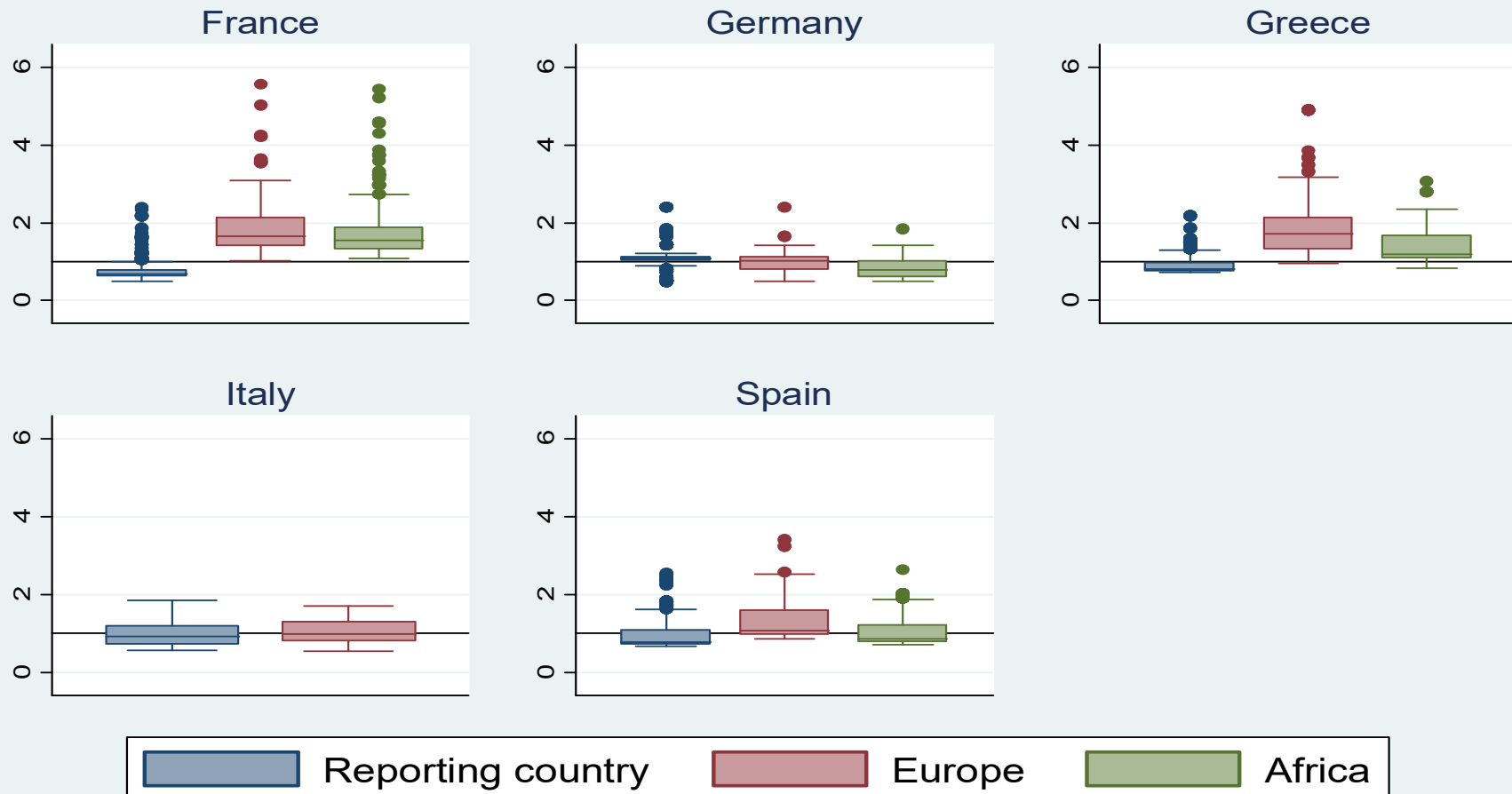

CD4 category
2010-2013

France · Greece · Italy · Spain · UK

Legend: <200 · 200-349 · 350-499 · >500 cells/ml
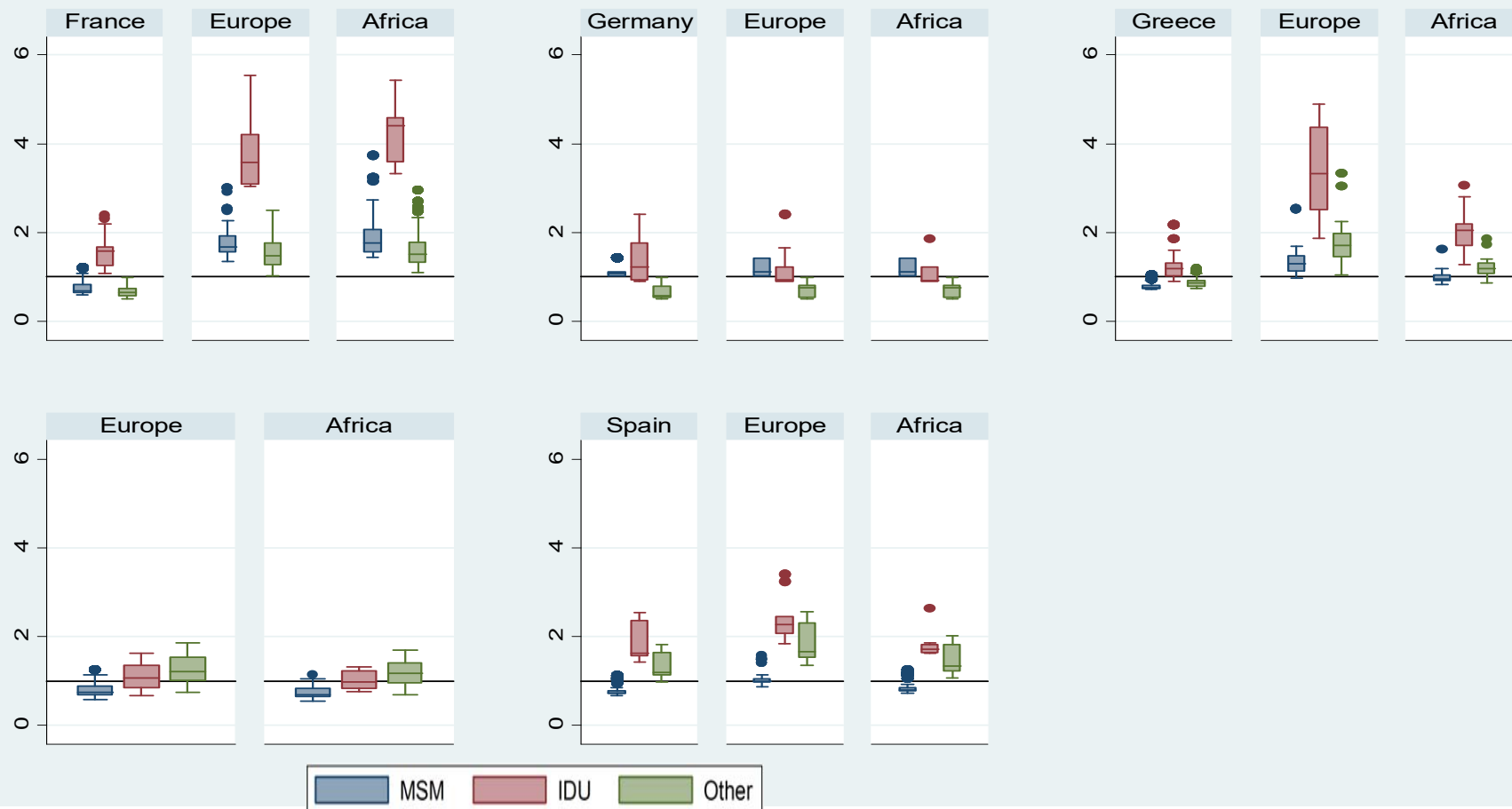
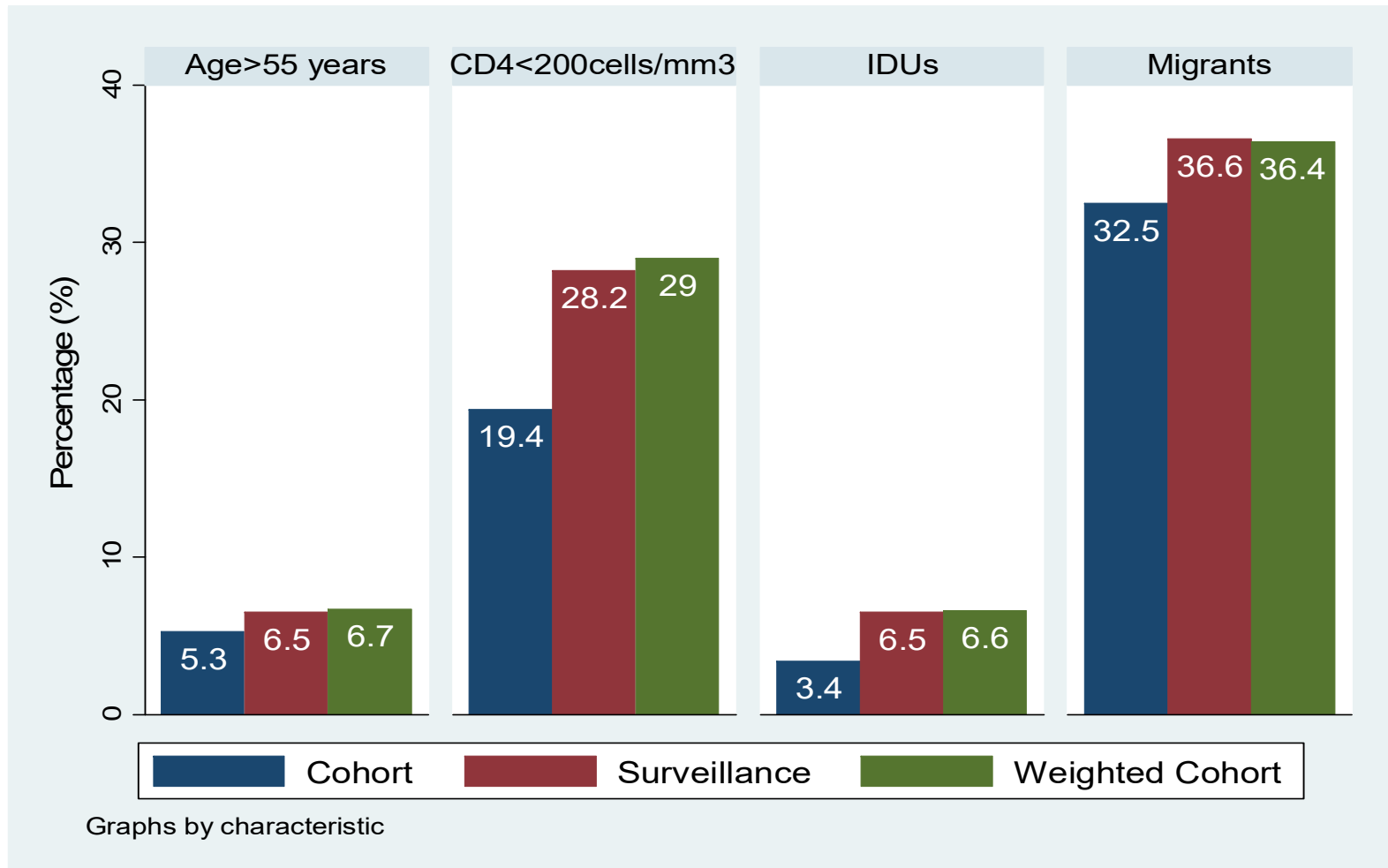# Stabilised weights according to region of origin

Region of Origin
2010-2013

Stabilised weights according to transmission mode by region of origin

Transmission Mode by Origin
2010-2013

Weights' application

Graphs by characteristic

Although there are differences in the distribution of individual characteristics in cohort and surveillance data, variables' distributions approaches the corresponding distributions in TESSy data after applying the inclusion weights.

Conclusions

- European cohorts capture a rather representative sample of the population of HIV diagnosed individuals

- Vulnerable HIV diagnosed individuals are most likely to be under-represented in the cohorts

- Weighting can be applied to correct for mis-representation of subgroups of patients in the cohorts

- Results of this project could be used to more effectively triangulate HIV surveillance and EuroCoord data for public health action

- Weighting can be applied to other analyses and cohorts

- Limitation: Issues concerning surveillance system (e.g. under-reporting, reporting delays) were not taken into account