Basel Institute ceb

for Clinical Epidemiology and Biostatistics

# Making Better Use of Registry Data in Designing Pragmatic Trials

Basel Biometric Society Novartis Campus, January 13, 2016

Heiner C. Bucher MD, MPH,
Basel Institute for Clinical Epidemiology & Biostatistics
University of Basel, University Hospital Basel Switzerland

heiner.bucher@usb.ch
www.ceb-institute.org

University Hospital Basel

# Overview

Innovative approaches for the use of routinely collected data:

- Using registry and cohort data for pragmatic trials
- System-wide resource use studies by privacy preserving anonymous probability-based data linkage of claim data with clinical cohorts

# The future challenges of evidence-based medicine

- Evidence will be based on a more diverse family of data sources and methodologies than the conventional (RCT) study type

- Reshape towards comparative effectiveness analysis of head to head comparisons with real word data
  - Improved methods to (re-) analyse RCT and non-RCT studies (marginal structural models)
  - Data linkage of observational data, resource use data, genetic data, biobanks
  - Indirect comparison, MTA

# Concept of registry-based randomized clinical trial (RRCT)

- High quality registries contain a large and comprehensive set of variables relevant for prognosis and patient outcome
- Comprehensive coverage of patients
  - SWEDEHEART registry of all hospitalized patients with heart problems (PCI, Valve replacement, ICD, etc.) in Sweden
- Patients admitted to hospital are asked to allow for randomisation rather than physician preference for treatment

# Advantages of RRCTs (I)

- A large proportion of less selected patients are available
- Better identification of eligible patients by large scale screening of inclusion and exclusion criteria
- More rapid patient recruitment
- Less costly, relevant data is routinely & prospectively collected

# Advantages of RRCTs (II)

- Higher external validity of RRCTs
- More ballanced research questions (investigator & industry driven)
- More appropriate benefit / harm assessment due to larger number of and less selected patients
- Collection of better health resource use data for cost-effectiveness analysis using the parallel claim data registries

# Disadvantages of RRCTs (I)

- High up front costs for data system development
- Registries may contain large amount of irrelevant data
- Time intensive search strategies for identification of patients and relevant patient parameters
- For drugs or medical devices that require comprehensive safety reporting and strictly defined endpoints the methodology is not different but data collection and monitoring requirements are very high in the context of routinely collected data

# Disadvantages of RRCTs (II)

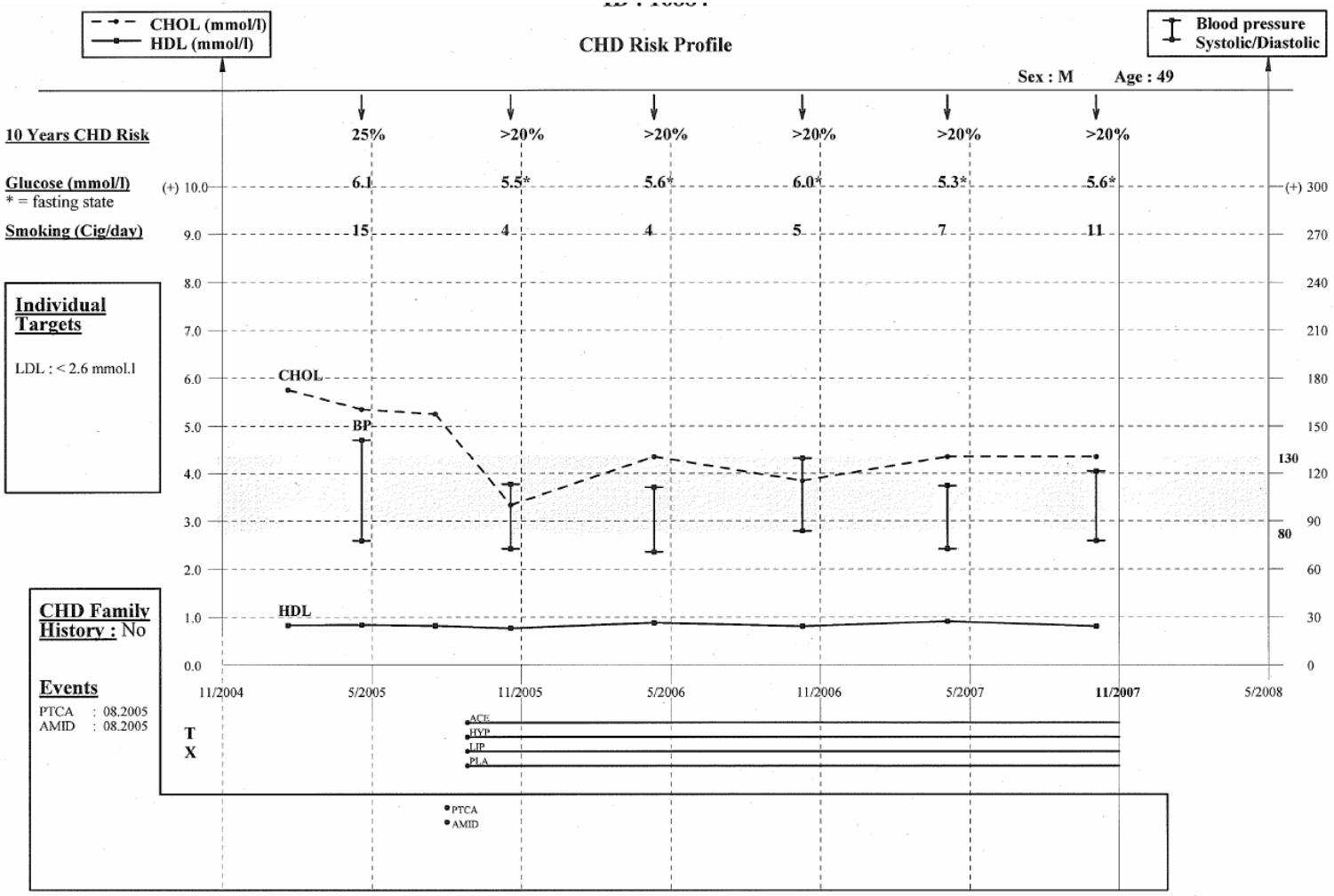
- There are still walls between research and the healthcare setting to enable life-span learning from real world data and shared clinical trial data
- Ethical considerations
- Unresolved issues of:
  - Patient consent
  - Ownership of data
  - Protection of personal data
  - Governance

# Randomized trial of a computerized coronary heart disease risk assessment tool in HIV-infected patients receiving combination antiretroviral therapy

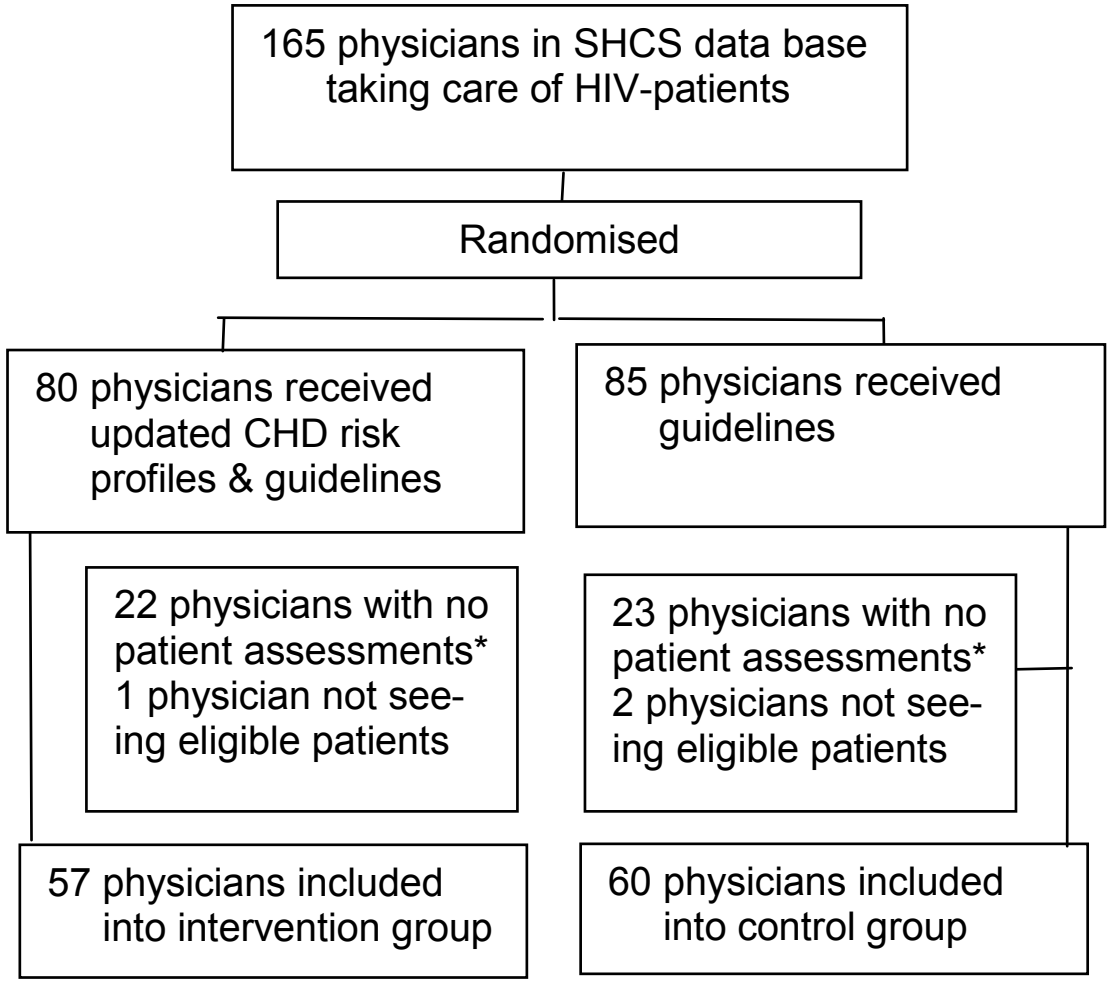
Bucher HC Antivir Ther 2010;15:3

- Objective
  - To investigate whether systematic provisions of CHD risk profiles and evidence-based guidelines to physicians improves in HIV-infected patients
    - Total cholesterol (primary endpoint)
    - Framingham risk scores, systolic & diastolic blood pressure (secondary endpoints)
- Design
  - Cluster RCT nested into the Swiss HIV Cohort Study

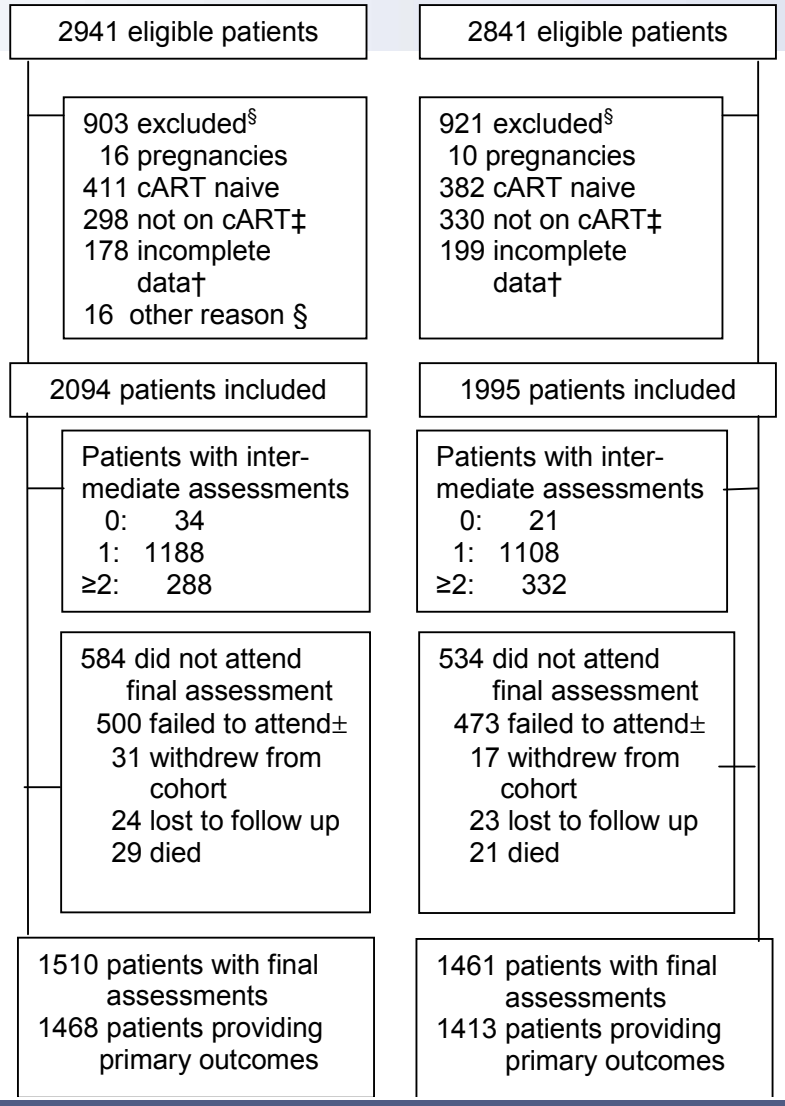
# SHCS CHD risk profiles for charts



# Results


165 physicians in SHCS data base taking care of HIV-patients
Randomised
80 physicians received updated CHD risk profiles & guidelines
85 physicians received guidelines
22 physicians with no patient assessments*
1 physician not see-ing eligible patients
23 physicians with no patient assessments*
2 physicians not see-ing eligible patients
57 physicians included into intervention group
60 physicians included into control group

# Results


2941 eligible patients
903 excluded§
16 pregnancies
411 cART naive
298 not on cART‡
178 incomplete data†
16 other reason §
2094 patients included
Patients with inter-mediate assessments
0: 34
1: 1188
≥2: 288
584 did not attend final assessment
500 failed to attend±
31 withdrew from cohort
24 lost to follow up
29 died
1510 patients with final assessments
1468 patients providing primary outcomes
2841 eligible patients
921 excluded§
10 pregnancies
382 cART naive
330 not on cART‡
199 incomplete data†
1995 patients included
Patients with inter-mediate assessments
0: 21
1: 1108
≥2: 332
534 did not attend final assessment
473 failed to attend±
17 withdrew from cohort
23 lost to follow up
21 died
1461 patients with final assessments
1413 patients providing primary outcomes

# Baseline characteristics of patients

| | Intervention n=2094 | Control n=1995 |
|---|---|---|
| Median years of age [IQR] | 44 [39-51] | 44 [39-50] |
| Female (%) | 30 | 30 |
| Current smokers (%) | 45 | 46 |
| Median systolic blood pressure [IQR] | 125 [115-135] | 121 [112-133] |
| On antihypertensive medication (%) | 14 | 13 |
| Median total cholesterol [IQR] | 4.9 [4.2-5.7] | 5.0 [4.3-5.7] |
| Diagnosed as diabetic (%) | 5 | 5 |
| Family history of CVD (%) | 12 | 12 |
| Framingham risk $\geq$ 10% (%) | 26 | 25 |

# Effects of the intervention on primary & secondary endpoints

| | | Unweighted analysis | | Weighted analysis | |
|---|---|---|---|---|---|
| | n (%) | Estimate | 95% CI | Estimate | 95% CI |
| **Primary outcome** | | | | | |
| Total cholesterol mmol/l | 2881 (70) | -0.02 | -0.10 to 0.06 | -0.03 | -0.12 to 0.05 |
| **Secondary outcomes** | | | | | |
| Systolic blood pressure mm Hg | 2935 (72) | -0.5 | -1.8 to 0.8 | -0.4 | -1.7 to 0.9 |
| Diastolic blood pressure mm Hg | 2935 (72) | -0.7 | -1.8 to 0.5 | -0.4 | -1.6 to 0.7 |
| Framingham risk score % | 2829 (69) | -0.1 | -0.3 to 0.1 | -0.1 | -0.3 to 0.1 |

# Change in drug management and CV events in patients with VL< 50 copies/ml at baseline

| | Framingham risk < 10% at baseline | | Framingham risk > 10% at baseline | |
|---|---|---|---|---|
| | Intervention n = 772 | Control n = 675 | Intervention n = 233 | Control n = 231 |
| Started new cART component (%) | 33 | 39 | 31 | 31 |
| Started abacavir (%) | 4 | 6 | 4 | 6 |
| Started atazanavir (%) | 5 | 3 | 5 | 6 |
| Started any drug that reduces CV risk (%) | 10 | 7 | 17 | 16 |
| Stopped any PI (%) | 10 | 12 | 9 | 11 |
| Experienced CV event (%) | 1 | 0 | 3 | 1 |

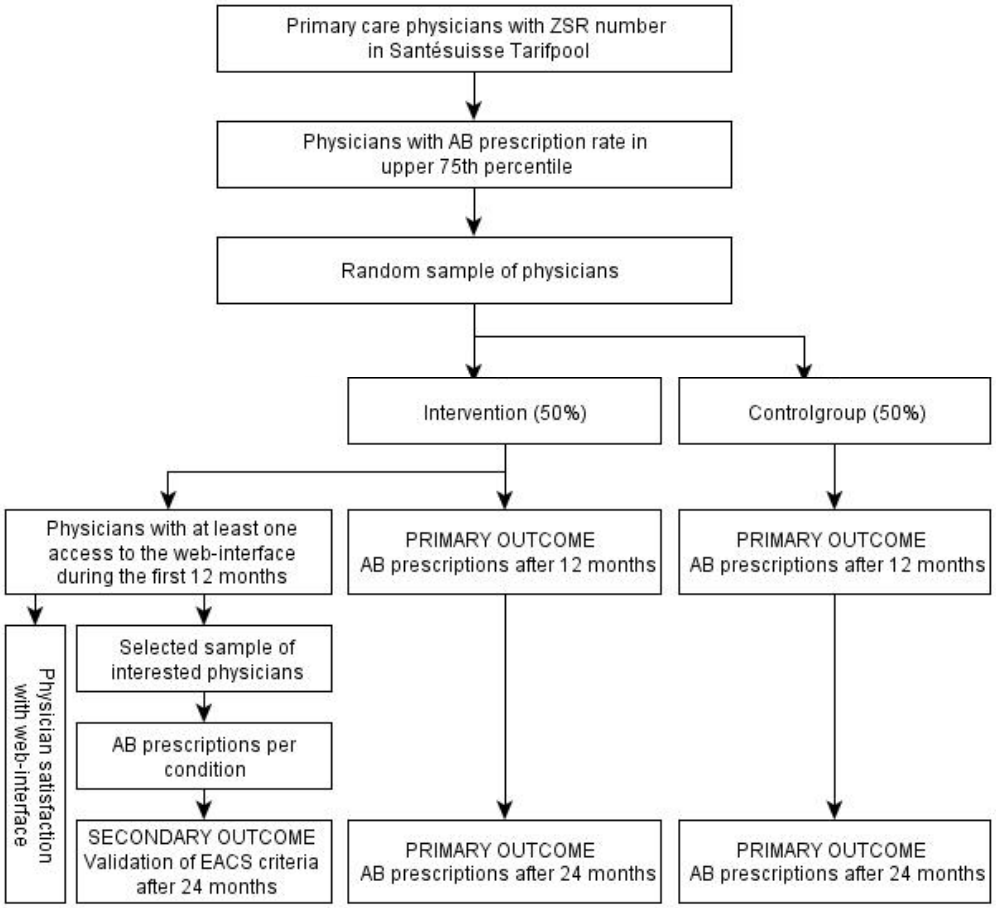
# Benchmark trial to lower antibiotic prescription nationwide in primary care in Switzerland

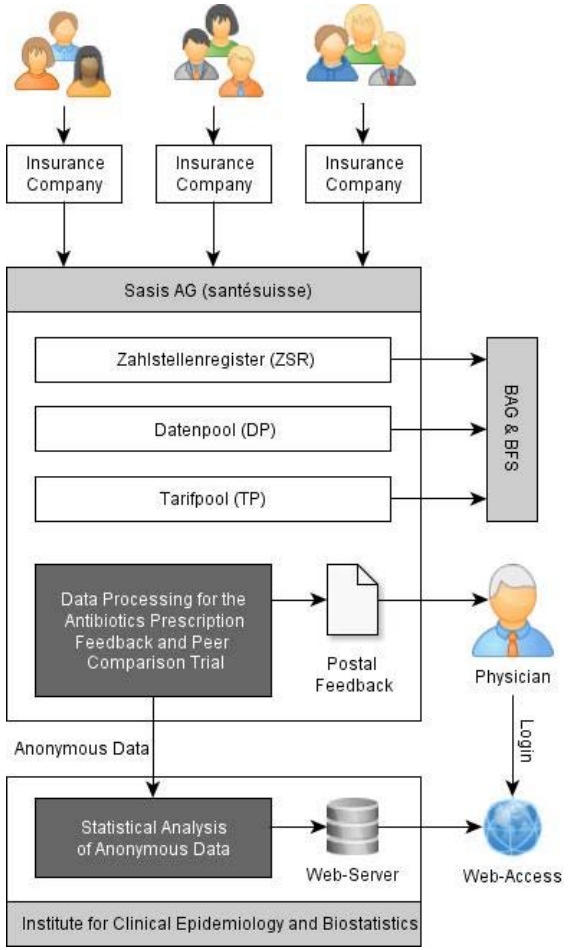
- Population
  - 2400 Board certified primary care physicians in Switzerland with high antibiotic prescription rates (above 75$^{th}$ percentile of antibiotic prescriptions)
- Intervention
  - Quarterly feedback on prescription rates (web-based, email reminder and letters) for 24 months
  - Evidence-based guidelines on the use of antibiotics in primary care
- Control
  - No intervention (not informed about trial)

# Outline Benchmark Trial

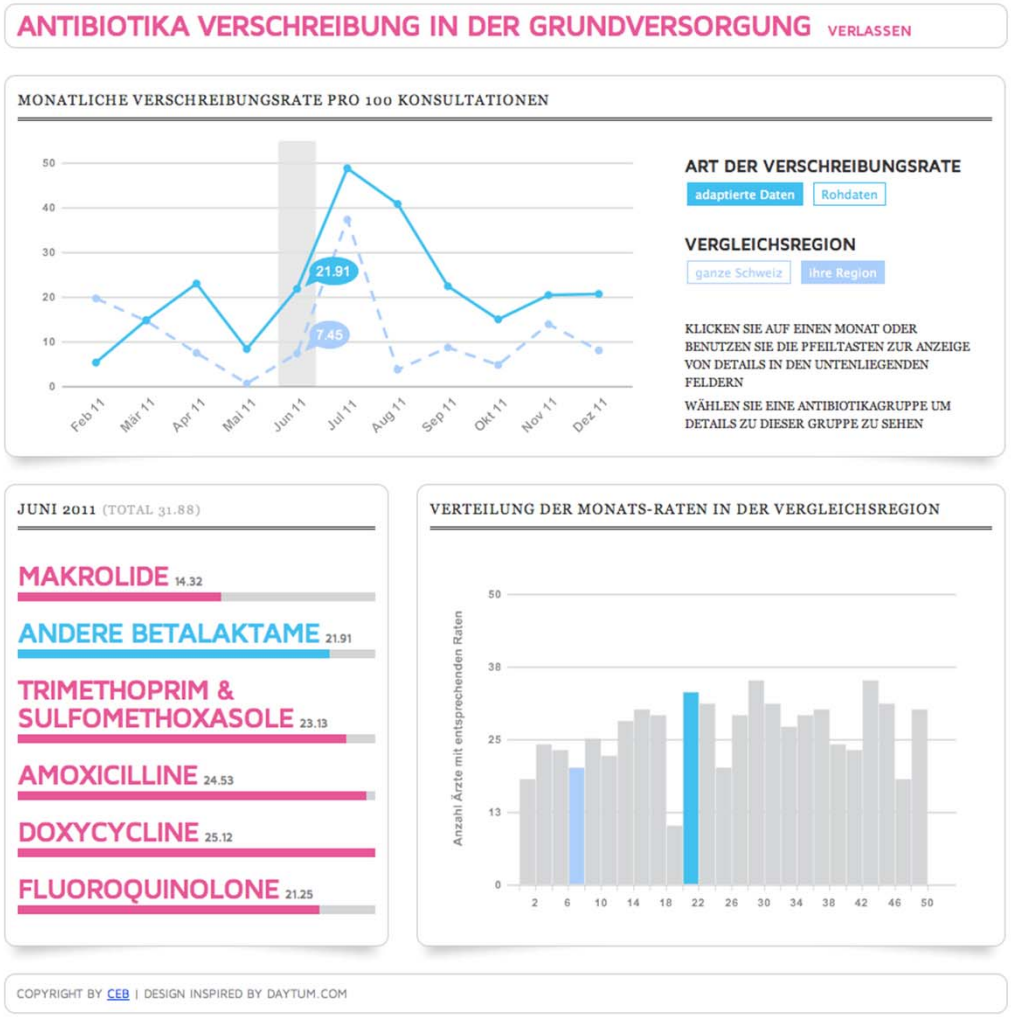
- Outcome
  - Primary: antibiotic prescription rate
- Data
  - Nationwide reimbursement data of health insurers (Tarifpool, SASIS Santésuisse)

# Selection of physicians and intervention Benchmark Trial



# Prototype Web-Application
# Benchmark Trial



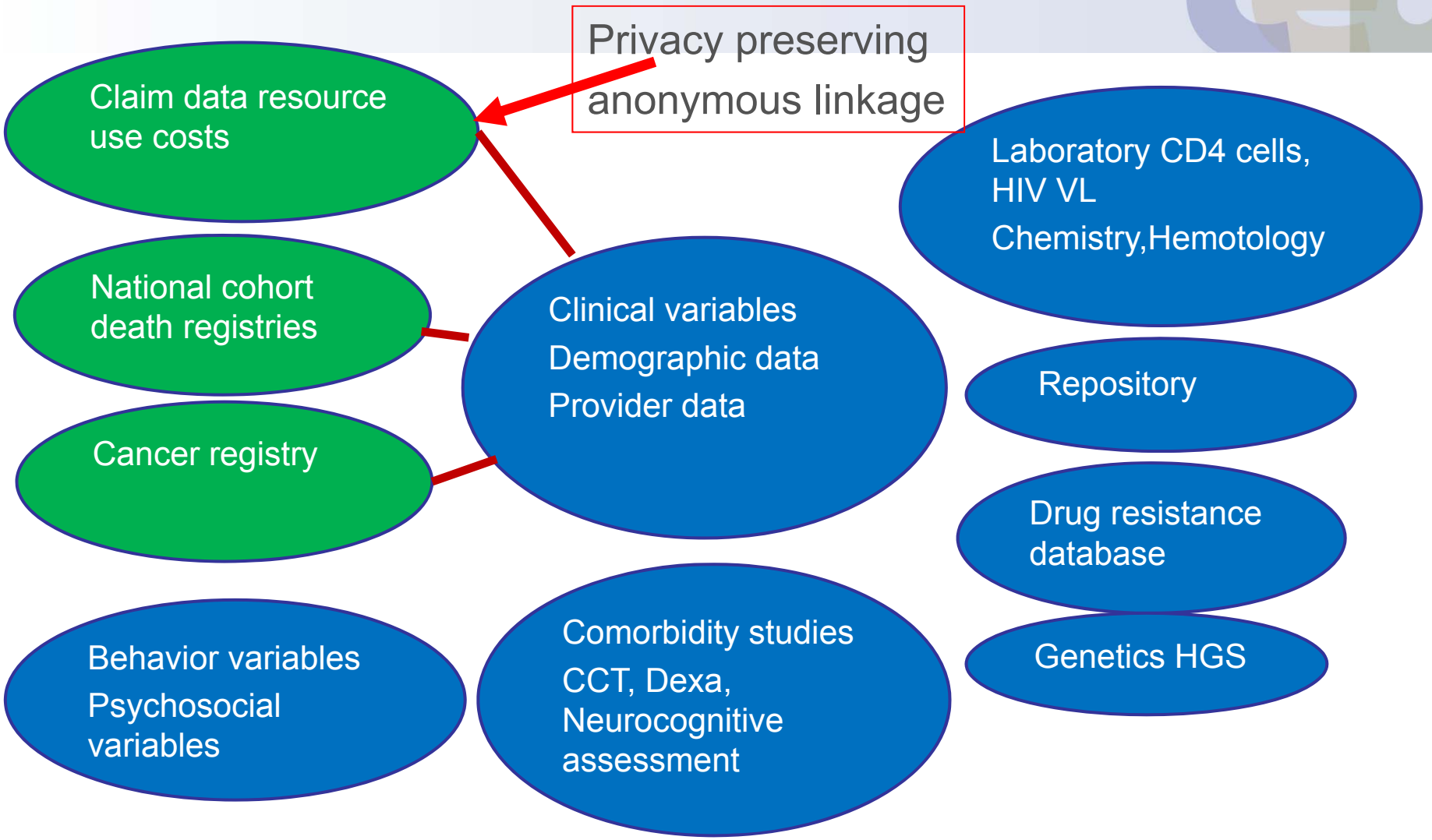
# Opportunities
# Benchmark Trial

- Generalisability
  - High external validity through inclusion of primary care physicians
- Potential impact
  - Directed at high prescribers nationwide
- Novelty
  - Little evidence whether benchmarking, monitoring and guideline provision reduces physicians‘ prescription behaviour
  - Use of routinely collected health care data for an intervention trial at population scale in Switzerland

# Privacy preserving probability-based data linkage of claim data with prospectively collected cohort data

- An example from
- The Swiss HIV Cohort Study
- Helsana, largest health insurer in Switzerland covering 20% of the Swiss population

# Data structure of the Swiss HIV Cohort Study



# Costs of HIV and non-HIV related comorbidity

- Late presentation (CD4 cells < 350 cells/µL) is the most important reason for HIV related morbidity

- HIV infected individuals are at higher risk of CVD, end stage renal disease, liver related comorbidity (HCV, HBV) and cancer than non-HIV infected individuals
- Resource use and costs of late presentation and HIV and non-HIV related comorbidity are not well known

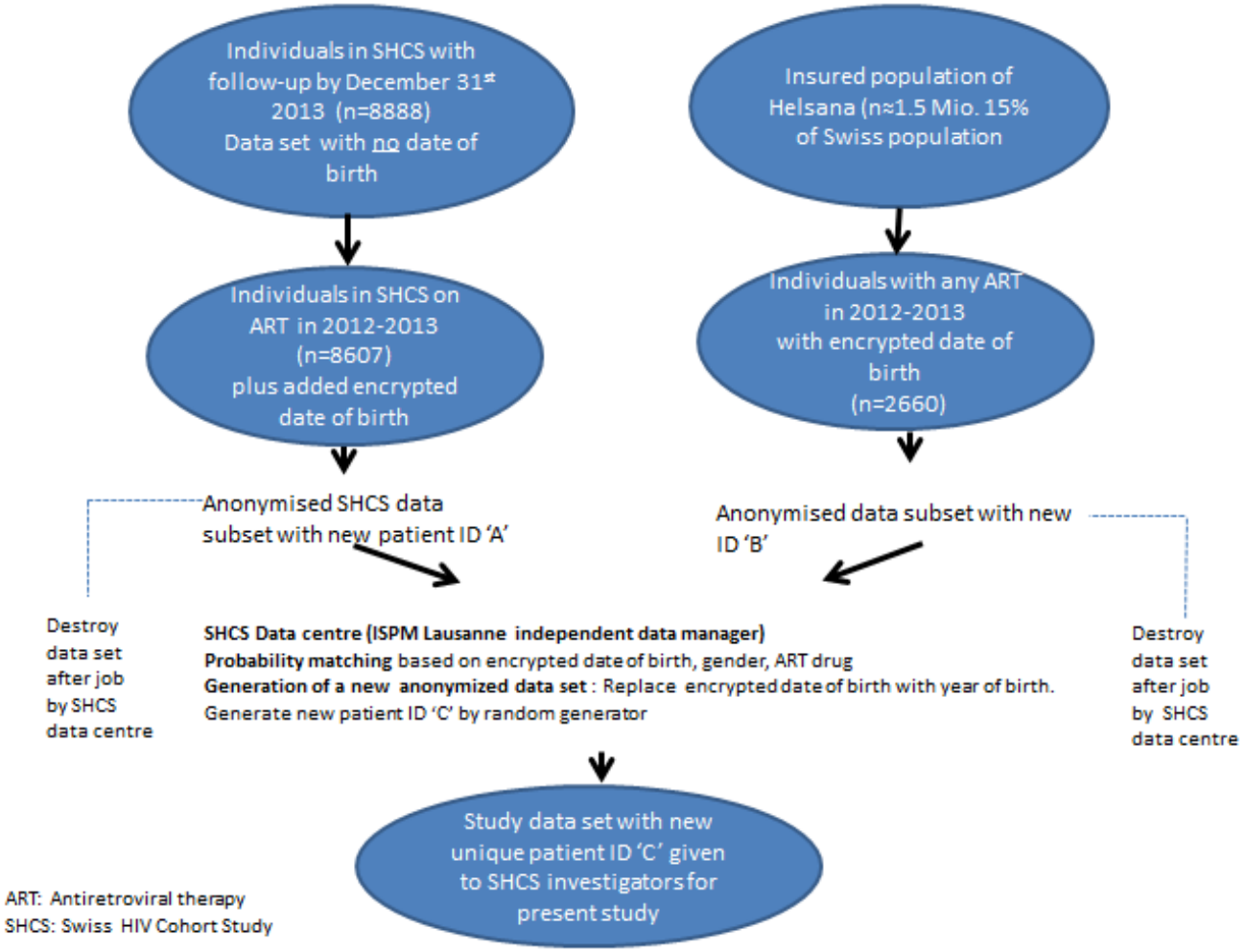
- Information on non ART drug use in the SHCS is limited

# Goals of the pilot study

- Evaluate the feasibiltiy and the validity of privacy preserving anonymous matching in the SHCS for claim data
- Collection of resource use data for future cost-effectiveness analyses
- Evaluate possibility for pharmacoepidemiological studies
- Evaluate whether pilot can be extended to
  - To include more health insurers
  - Other cohorts (Swiss Transplant Cohort Study)

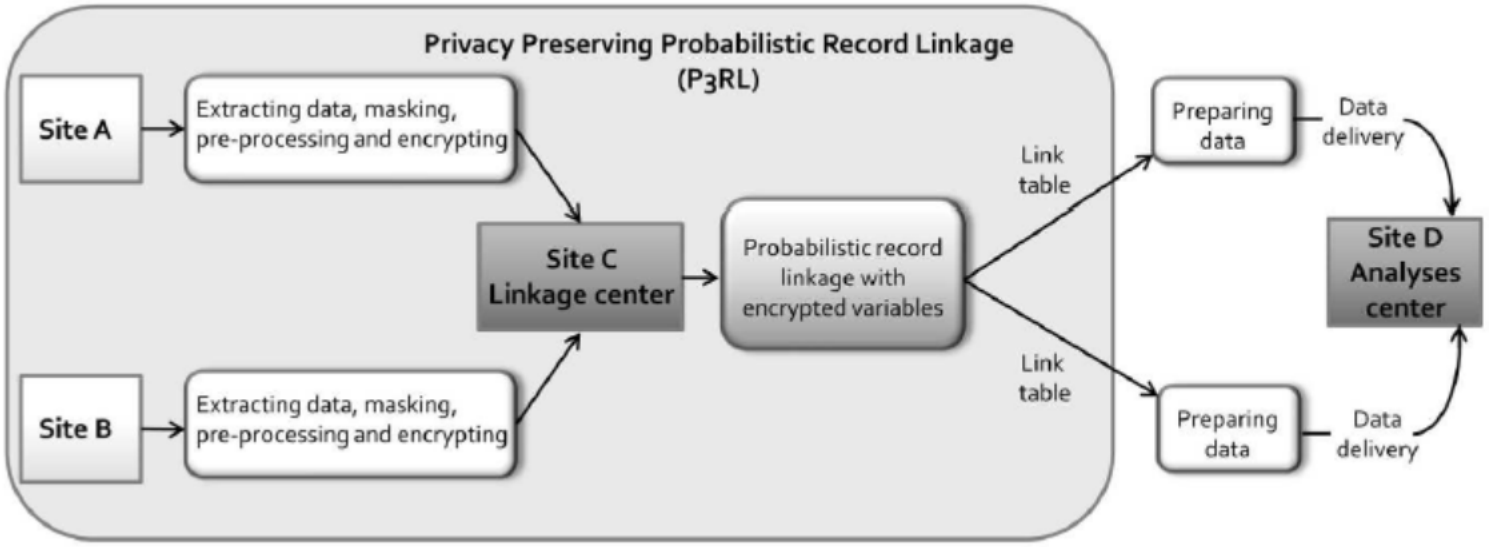
- Evaluate whether routine annually mergers can be established

# Set-up of a pilot study for a merger of prospective clinical and cost claim data: Swiss HIV Cohort Study



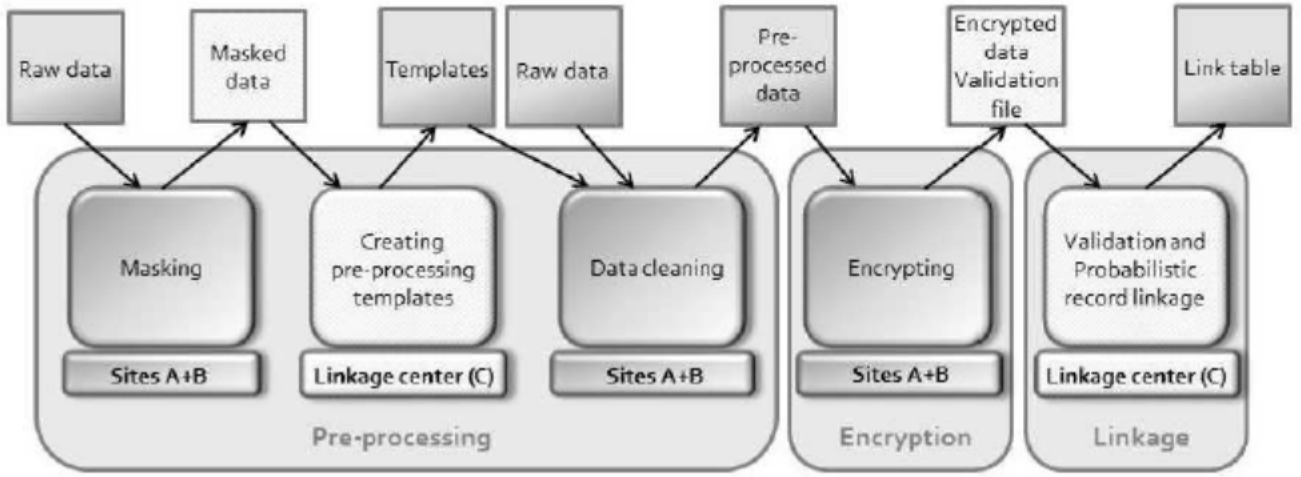

ART: Antiretroviral therapy
SHCS: Swiss HIV Cohort Study

# Privacy preserving probabilistic record linkage




*Schmidlin K .2015* BMC Med Res Methodol. 2015;15:46

# Data masking, encrypting and probabilistic linkage




*Schmidlin K .* BMC Med Res Methodol. 2015;15:46

## Conclusions:

- RRCTs are an interesting option for head to head comparisons in settings with registry data of high quality exist
  - May reduce cost
  - Facilitate rapid recruitment
- Privacy preserving anonymised matching may allow to considerable enrich observational data research
  - Matching is resource intense
  - Possibility of routine linkage
  - Interesting possibilities for monitoring, health economic studies

# Thanks to:

Collaborators:
Jim Young, Juliane Schäfer, Ramon Saccilotto, Lars Hemkens, Selene Leon Reyes, Dominik Glinz
Organisations:
Swiss HIV Cohort Study
Helsana
Funders:
Stiftung Institut für klinische Epidemiologie
Swiss National Science Foundation, Gottfried und Julia Bangerter-Rhyner Stiftung