



WETENSCHAPPELIJK INSTITUUT
VOLKSGEZONDHEID
INSTITUT SCIENTIFIQUE
DE SANTÉ PUBLIQUE



KULeuven



UCLouvain

Ambulatory Care Health
Information Laboratory
ACHIL

ACHIL is funded by the National Institute for Health and Disability Insurance



Quality assessment of automatically extracted data from GPs' EPR

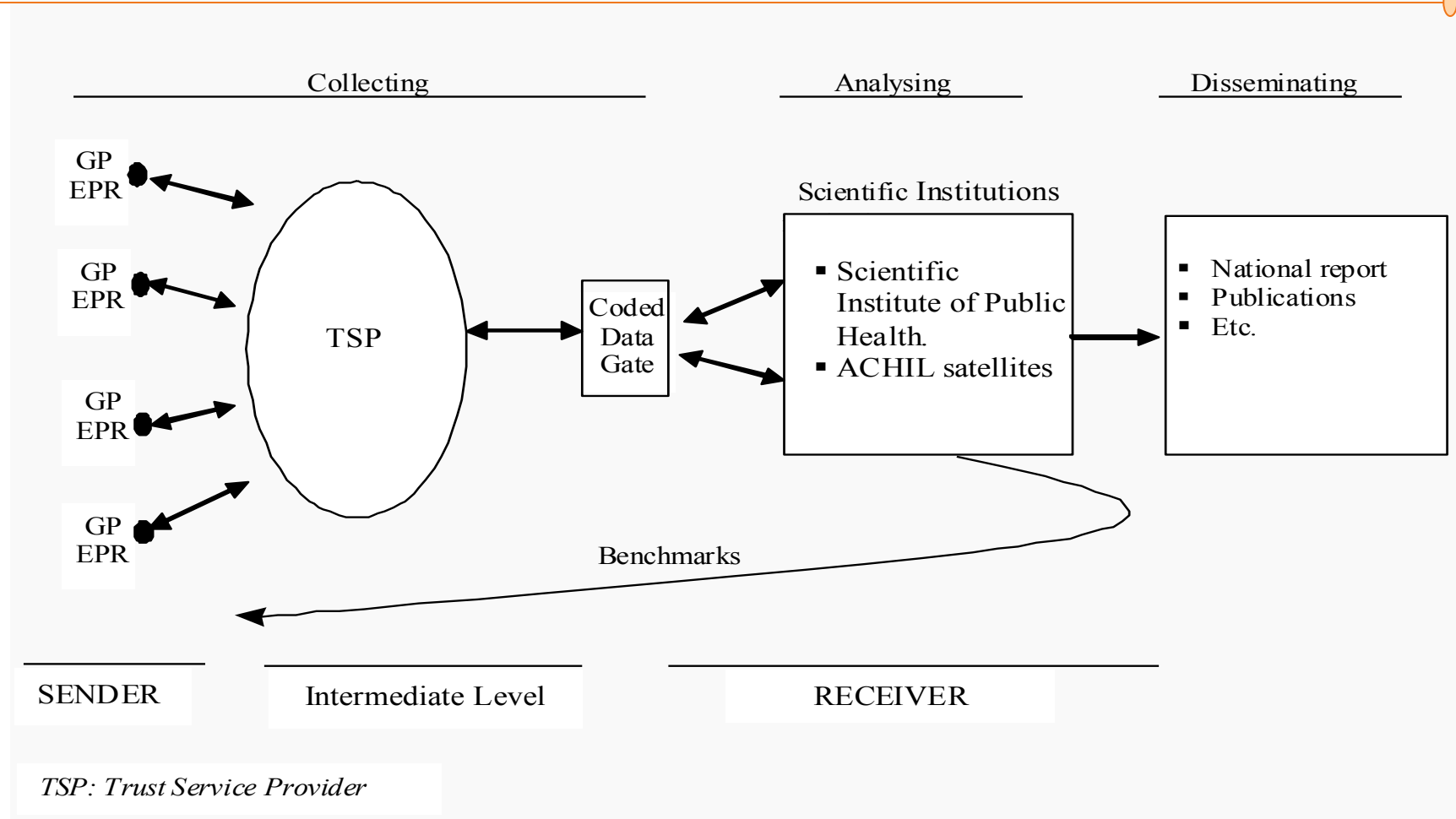
ACHIL research laboratory

*Etienne De Clercq (UCL-IRSS), S. Moreels, V. Van Casteren, N.
Bossuyt (ISP), G. Goderis, S. Bartholomeeusen (KUL)*

(Etienne.DeClercq@UCLouvain.be)

Ateliers du CISP-Club, Bruxelles (Belgique) 19-21/10/2012

ResoPrim Research Network Architecture



Clinical automatic extracted data

- ❖ New coded and active **diagnosis** (ICPC2, ICD10, Belgian Thesaurus)
(hypertension, diabetes type 2, cardiovascular past event)
- ❖ New coded and active **drug prescription** (ATC code)
(anti-diabetic drugs, anti-hypertension drugs, aspirin, statin)
- ❖ (New referral)
- ❖ **Parameters** (2 most recent values extracted): height, weight, smoking status, syst. & diast. Blood pressure, total & LDL cholesterol

ResoPrim documented care

Documented care = Care * Quality of HRIS

Research Question:

Which are the properties of the HRIS?

Properties of the HRIS

PPV

“proportion of patients with a “gold standard” positive value of those with positive AE data”

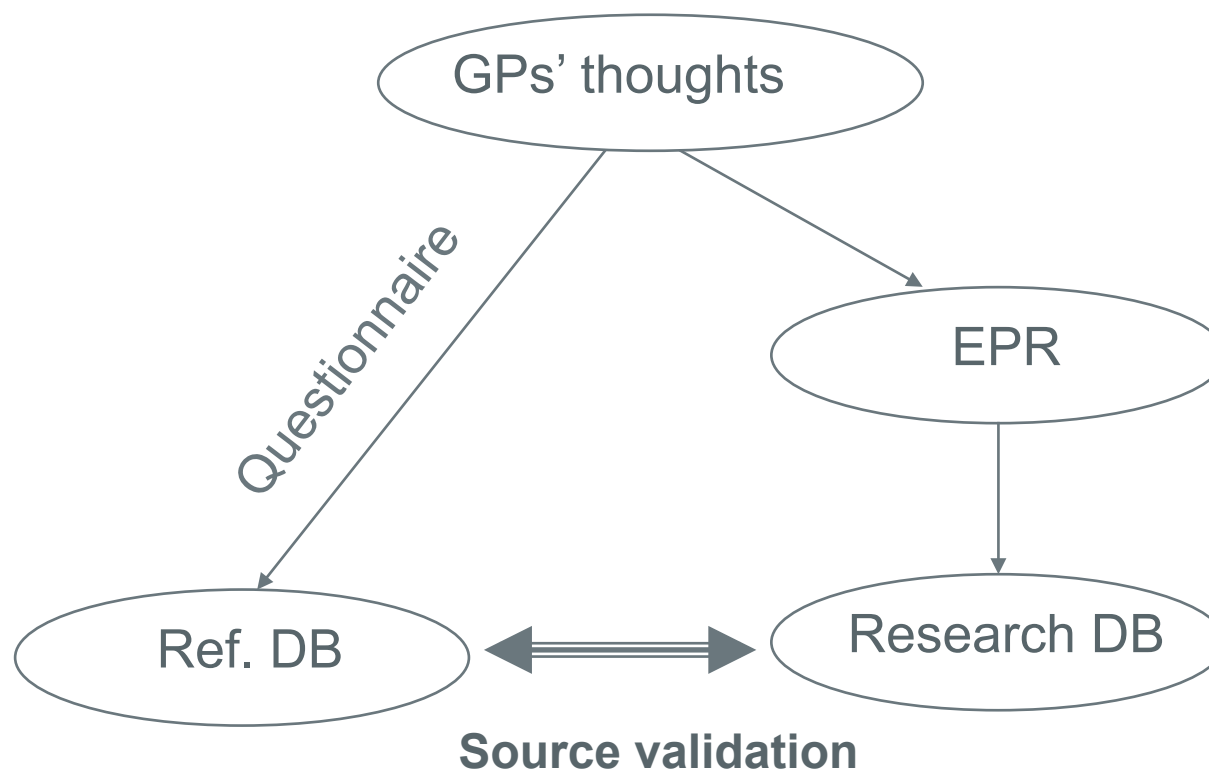
e.g.: proportion of drug codes (extracted from EPR) confirmed by the gold std?

Sensitivity

“proportion of patients with positive AE data of those with a “gold standard” positive value”

e.g. How many patients with drug prescription (gold std) identified by a drug code (extracted from EPR)?

Building our “gold standard”





UCLouvain

Pilot Site

ResoPrim phase-2 pilot, (Summer 2007, +/- 7 weeks)
data gathered from:

- 43 GP practices
- 4 software systems
- 10 307 patients
- 13 372 contacts



Questionnaire: missing & incoherent values

Missing & incoherent values in the questionnaire			
	mean	Min – Max (question)	Min – Max (practice)
missing	4.2%	0.5% - 5.7% (Diab.: 30%)	0% - 18.4 %
incoherence	5.3%	1.5% - 14%	0% - 14.3%

→ 90% seems an acceptable value for Sensitivity and PPV



Automatic extracted diagnoses

Sensitivity and Positive Predictive Value (PPV) of automatic extraction vs questionnaire

<u>Diagnoses</u>	Sensitivity	PPV
Hypertension	45.5%	82.6%
Diabetes	47.9%	71.7%
Past Cardio-Vasc. Event	29.9%	57.5%
Family PCVE	1.2%	(50.0%)



DM 2 patients identification

Sensitivity and Positive Predictive Value (PPV) of automatic extraction vs questionnaire

<u>DM2 patients</u>	ICPC	ICPC-ATC	ICPC-ATC-HbA1C
PPV	71.7%	75.8%	72.1%
Sensitivity	47.9%	77.8%	79.6%

Automatic extracted drugs

Sensitivity and Positive Predictive Value (PPV) of automatic extraction vs questionnaire

<u>Drugs</u>	Sensitivity	PPV
Anti-HT drugs	73.6%	97.8%
Anti-diab. Drugs	68.9%	96.3%
Aspirin	44.3%	91.9%
Statin	52.9%	90.4%

Automatic extracted parameters

Sensitivity and Positive Predictive Value (PPV) of automatic extraction (most recent value) vs questionnaire

<u>Parameters</u>	Missing (AE)	Sensitivity (AE)	PPV (AE)
Hypercholesterolemia	19.5%	59.1%	42.9%
(Smoking)	43.8%	(34.1%)	(82.2%)
Blood pressure <140/90	6.1%	65.9%	92.5%
BMI>25	36.9%	91.7%	73.8%

Care & Documented care

$$P_Q = P_{AE} * PPV / \text{Sens.}$$

	PPV	Sens.	Prevalence Statin		
			AE	Q	Calculated
Year 1	90.4%	52.9%	22.39%	37.84%	38.26%
Year 2	90.4%	84.3%	35.00%	???	37.53%
Year 2bis	90.4%	52.9%	25.00%	???	42.72%



Lessons learned

Drugs: PPV ↗, Sensitivity ↘

Diagnoses : PPV ↘, Sensitivity ↘

Parameters : ??



BUT

Lessons learned (continued)

- ❖ Not representative GP sample!
- ❖ Great variations (missing, PPV, Sens.) by
 - Practice
 - Software system
- ❖ Completeness of the extraction!
- ❖ Study restricted to coded and structured data
- ❖ Robustness of the properties?
- ❖ Easier “data quality” measures (e.g.: global indicator)?



UCLouvain

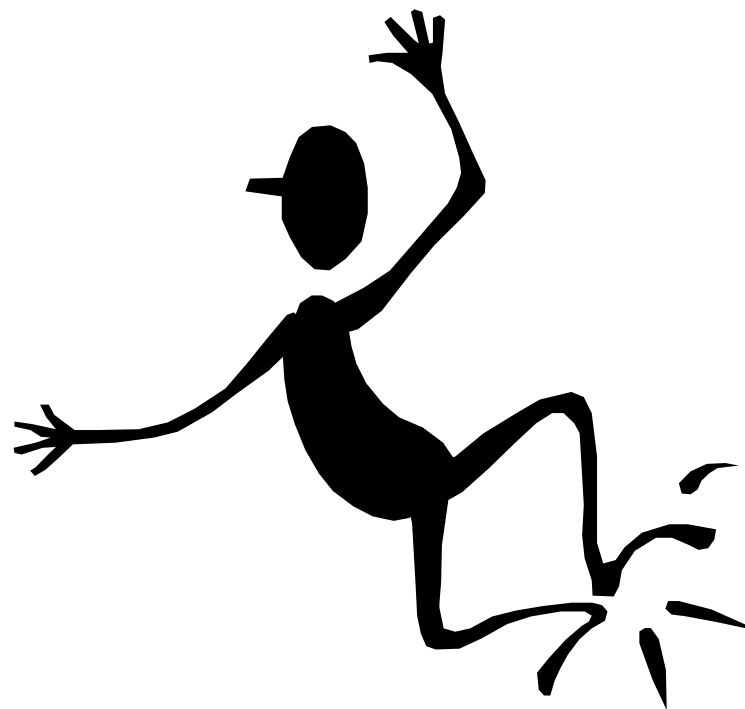
Our message ...

Before any secondary usage of data extracted from EPR, we strongly advice assessing properties of the Health Research Information System





UCLouvain



That's all ...!

