

Validation of National Quality Registries

Aldana Rosso, PhD

Epidemiology and Registry Center South, ERC South

Aldana.Rosso@skane.se, AR@rcsyd.se

2015-03-13



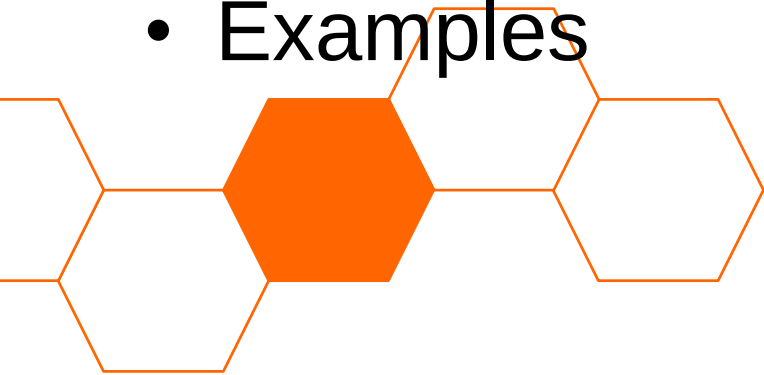
Epidemiology and Register Centre South (ERC Syd)

- It started 2014-01-01 as a fusion between Epi-Centrum and RC Syd Lund.
- Two teams: Epidemiology and Quality Registries.
- The Quality Registries team has also another team in Karlskrona (RC Syd Karlskrona).
- <http://www.skane.se/ercsyd>



Agenda

- Introduction
- Data quality
- Consequences of errors in the quality registries
- An introduction to sampling theory
- Examples



Data Quality

- ISO 8402-1986: "The totality of features and characteristics of an entity that bears on in its ability to satisfy stated and implied needs".
- Valideringshandboken: "Data av hög kvalitet är relevanta, fullständiga, korrekta och konsistenta."

A decorative graphic consisting of several orange-outlined hexagons of varying sizes, some of which are solid orange, arranged in a cluster at the bottom left of the slide.

<http://www.kvalitetsregister.se/sekundarnavigering/valideringshandbok.4.42e2e0a7143003c9eed8a30.html>

Registerpopulation

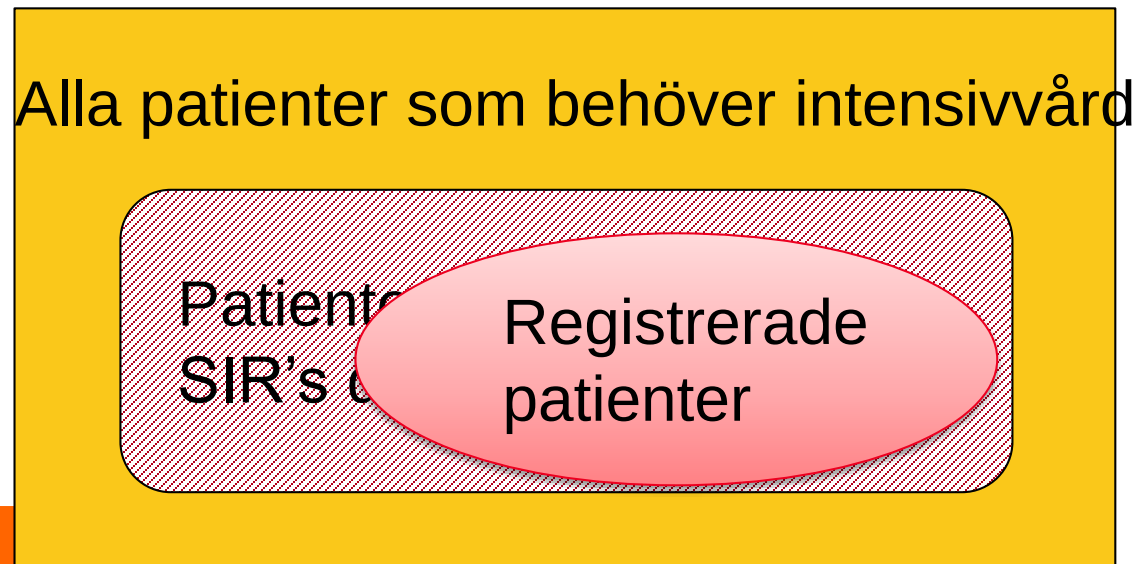
- Rampopulation/registerpopulation (frame/registry population)

Alla patienter som behöver intensivvård

Patienter som passar in i
SIR's definition.

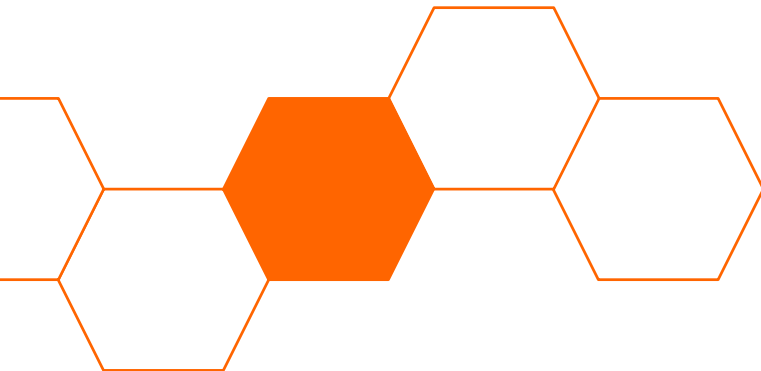
Registerpopulation

- Täckningsgrad



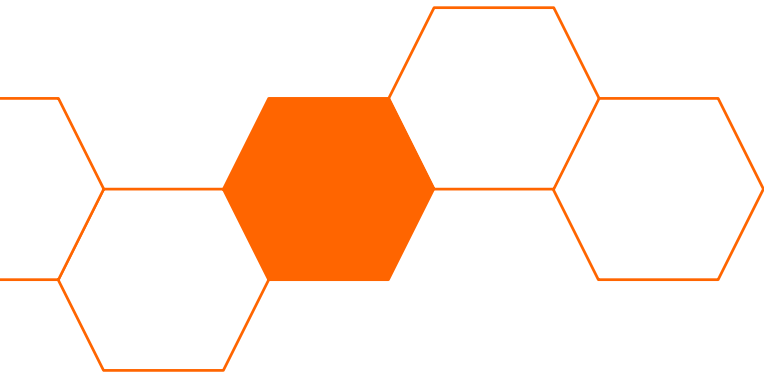
Registerpopulation

Anslutningsgrad



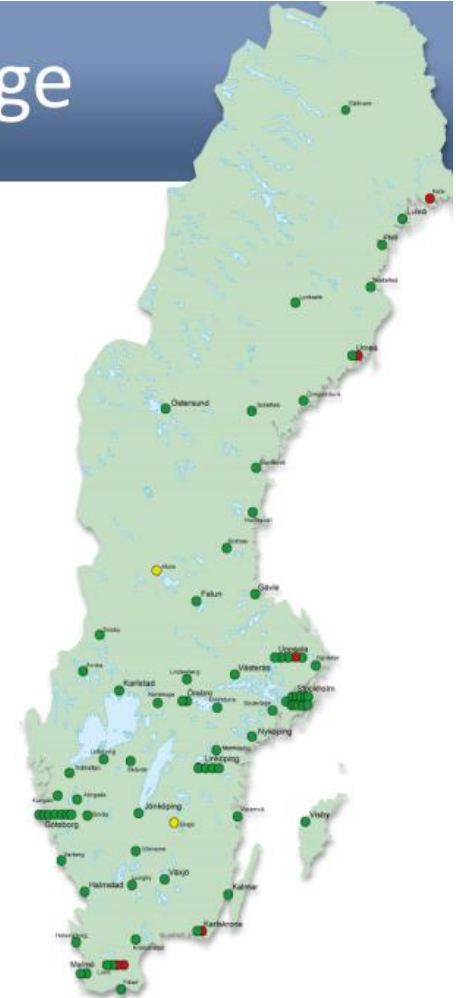
Registerpopulation

Hög anslutningsgrad är en förutsättning för att uppnå god datakvalitet men säger lite om datakvalitet i registret.



Intensivvård i Sverige

- 84 intensivvårdsavdelningar (78 medlemmar)
 - 27 finns på länsdelsjukhus
 - 25 finns på länssjukhus
 - 32 finns på universitetssjukhus
- Struktur
 - 64 allmän IVA
 - 8 thoraxintensivvård
 - 5 neurointensivvård
 - 4 barnintensivvård
 - 2 brännskadeintensivvård
 - 1 ECMO- center (intensivvård)



2014-11-27

Svenska Intensivvårdsregistret

Consequences of Statistical Analysis with Missing Data

March 28, 2014 11:38 am

Big data: are we making a big mistake?

By Tim Harford

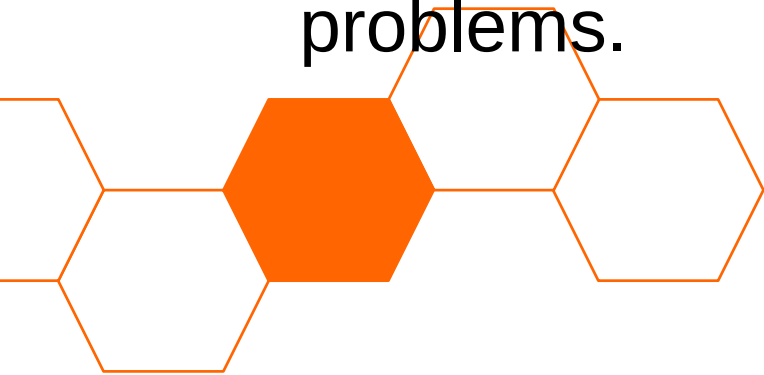
Big data is a vague term for a massive phenomenon that has rapidly become an obsession with entrepreneurs, scientists, governments and the media

“There are a lot of small data problems that occur in big data,” says Spiegelhalter. “They don’t disappear because you’ve got lots of the stuff. They get worse.”



Consequences of Statistical Analysis with Missing Data

- It depends on why the data is missing:
 - missing at completely at random: in general it is problematic from a power perspective but it doesn't bias the results.
 - missing not at random: bias + power problems.





Missade VTS-pass



Fråga: När det idag saknas vissa data godkänner inte valideringsprogrammet ett inskick exempelvis missade registreringar av VTS-pass. Det gör att man idag efterkonstruerar värden på dylika variabler vilket riskerar det statistiska underlaget. Finns det någon tanke från SIR om att låta ett framtida valideringsprogram godkänna missade data istället vilket skulle göra att underlaget blir mer korrekt om än något mindre.

2014-11-27

Svenska Intensivvårdsregistret

30

http://www.icuregswe.org/Documents/Education/Adminkurs/Utbildningsdagen_del1_2014-11-27.pdf

Example: Spanish National Acute Coronary Syndrome Registry

Circulation
Cardiovascular Quality and Outcomes



Patient Registries of Acute Coronary Syndrome: Assessing or Biasing the Clinical Real World Data?

Ignacio Ferreira-González, Josep R. Marsal, Francesca Mitjavila, Antoni Parada, Aida Ribera, Purificación Cascant, Núria Soriano, Pedro L. Sánchez, Fernando Arós, Magda Heras, Héctor Bueno, Jaume Marrugat, José Cuñat, Emilia Civeira and Gaietà Permanyer-Miralda

Circ Cardiovasc Qual Outcomes. 2009;2:540-547; originally published online September 22, 2009;

doi: 10.1161/CIRCOUTCOMES.108.844399

Circulation: Cardiovascular Quality and Outcomes is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231

Copyright © 2009 American Heart Association, Inc. All rights reserved.

Print ISSN: 1941-7705. Online ISSN: 1941-7713

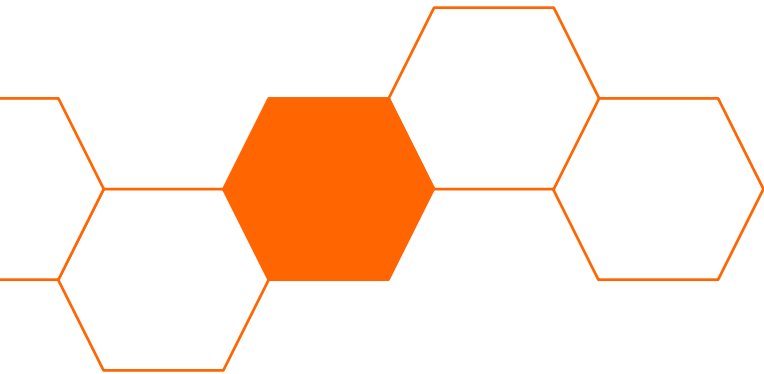
Example: Spanish National Acute Coronary Syndrome Registry

- Audit of the Spanish national acute coronary syndrome registry [Ferreira-Gonzalez et al, Circulation: Cardiovascular Quality and Outcomes, 2009; 2: 540-547].
- They compared enrolled patients with those that were not enrolled for some participating hospitals (17 of 50).
- Missed patients were of higher risk and received less recommended therapies than the included patients. In-hospital mortality was almost 3 times higher in the missed population.



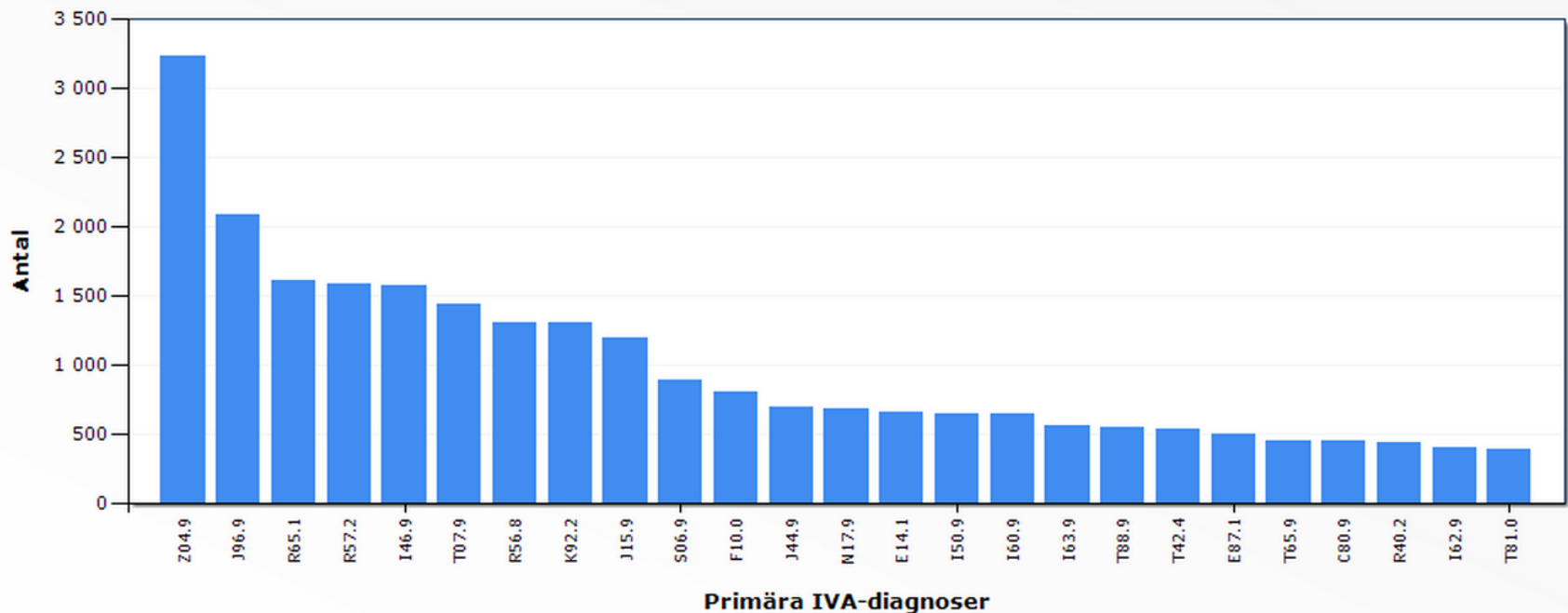
Consequences of Statistical Analysis with Misclassification

- It depends:
 - non-differential: typically underestimation.
 - differential: under- and overestimation.



Diagnoser

25 i topp – primära IVA-diagnoser (vårdtillfälle), inskrivningsperiod 2014-02-26 - 2015-02-26



Detta är en originalrapport från SIR

Svenska Intensivvårdsregistret 2015-02-26

<http://portal.icuregswe.org/Rapport.aspx>

Exempel: Diagnoser

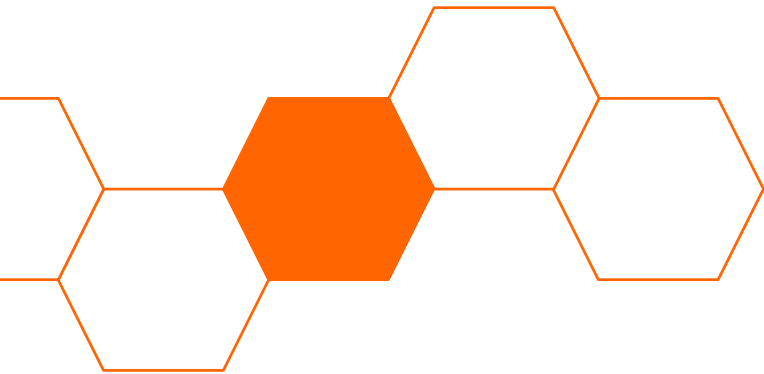
Diagnoser	Antal
Z04.9	3200
J96.9	2000
I46.9	1500
R57.2	1500
R65.1	1500
T07.9	1400
R56.8	1400
K92.2	1200
J15.9	1200

Andel felaktiga diagnoser som blev "Z04.9"	Andel korrekta ranking
2 %	4 % (2 % - 10 %)
5 %	5 % (2 % - 12 %)

<http://portal.icuregswe.org/Rapport.aspx>

Error Sources in National Quality Registries

1. Before registration
2. During registration
3. After registration



Error Sources in National Quality Registries

1. Before registration:

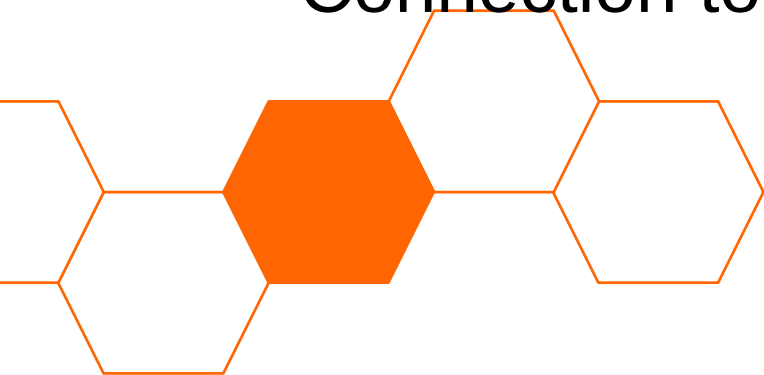
- wrong data is registered in the journal.
- the patient is not enrolled in the registry.
- It is very difficult to estimate. One way is to use the registries from the National Board of Health and Welfare.



Error Sources in National Quality Registries

2. During registration:

- misinterpretation and inaccurate typing: wrong value, wrong calculation, wrong alternative.
- incomplete data.
- Connection to different lab systems



Radera vårdtillfälle

- Om man raderar vårdtillfälle i sitt lokala system som redan rapporteras till SIR så måste ni meddela SIR detta
- Raderade vårdtillfälle raderas inte med automatik hos SIR om ny fil för samma period skickas.

Error Sources in National Quality Registries

3. After registration:

- programming errors.
- Communication problems between different databases.

- They are very difficult to find and have severe consequences.

Dutch National Intensive Care Evaluation Registry (NICE)

NICE contains data from patients who have been admitted to Dutch intensive care units and provides insight into the effectiveness and efficiency of Dutch intensive care.

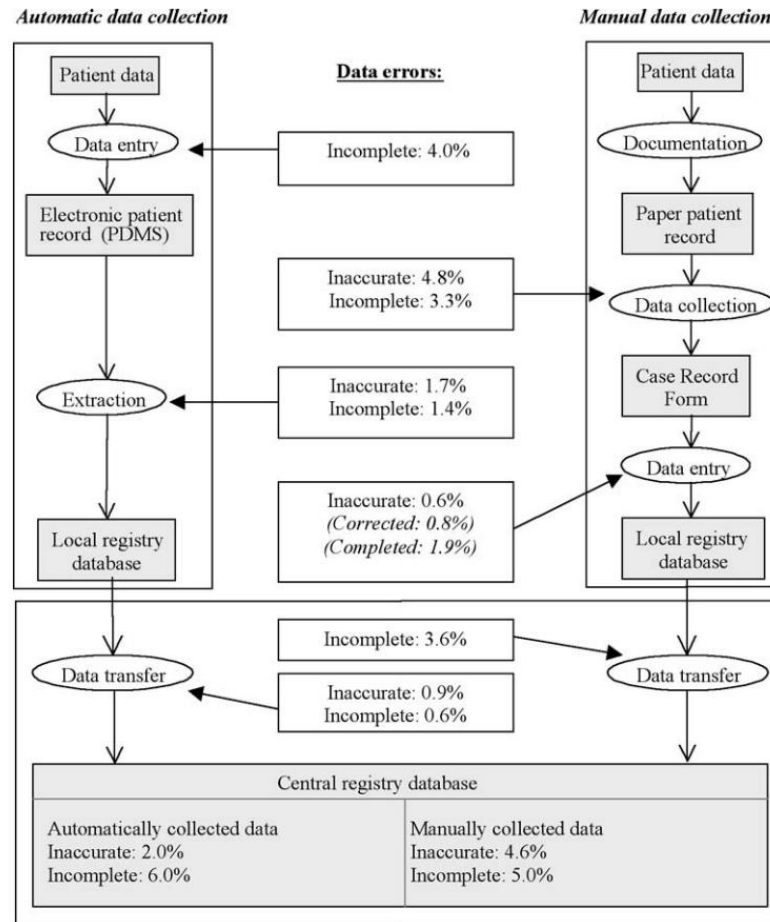


Figure 1 Results from the case study: types and percentages of newly occurring data errors at the different steps in the data collection process.

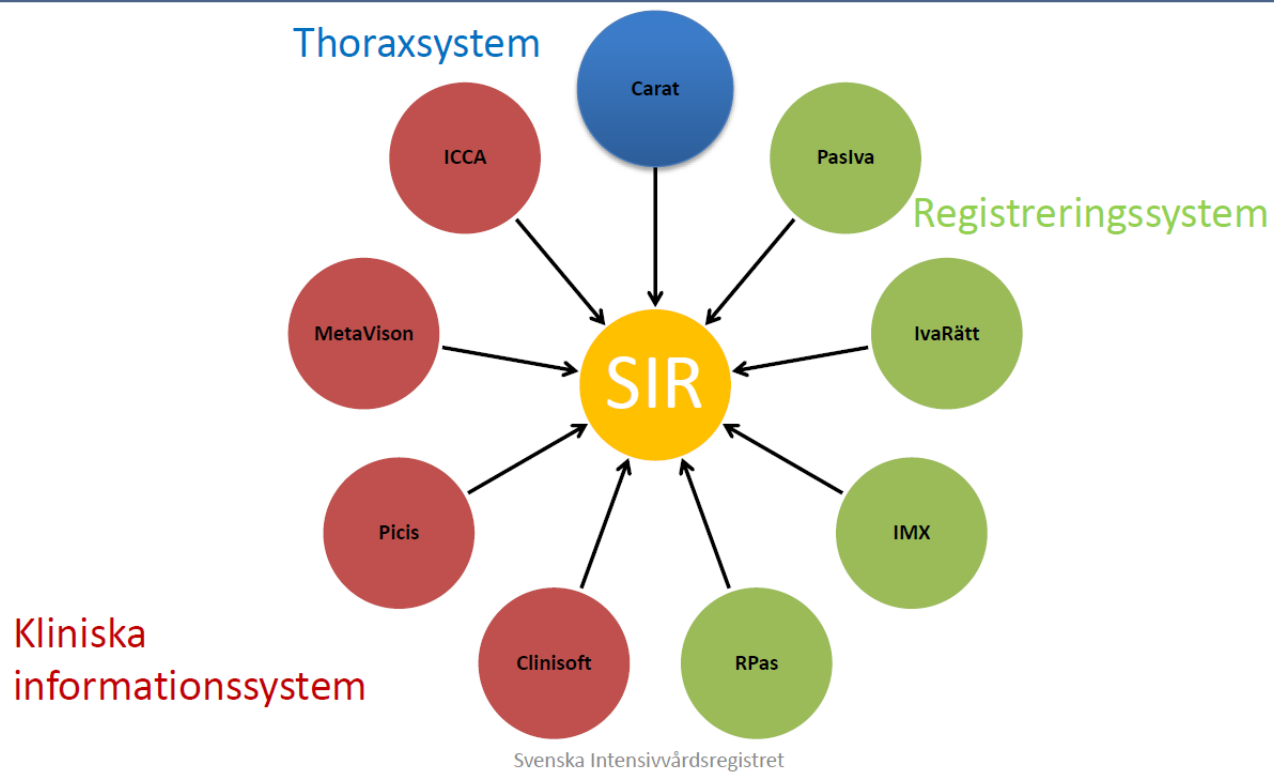
How to Minimize Errors in the Registries

1. Before registration:

- clear objectives with the registry
- clear definitions
- adequate training
- adjudication



Lokala IT-system



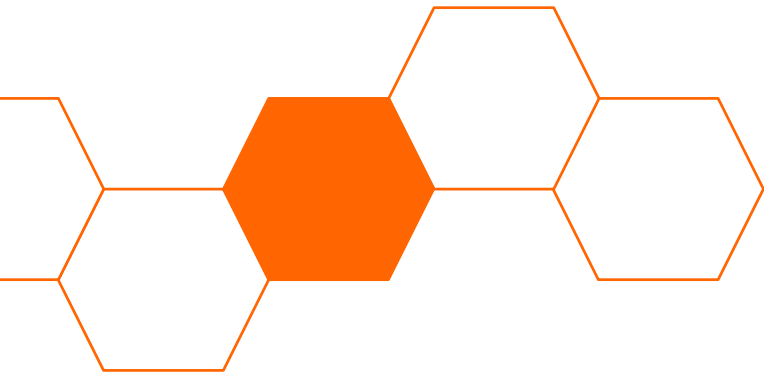
2014-11-27

12

How to Minimize Errors in the Registries

2. During registration:

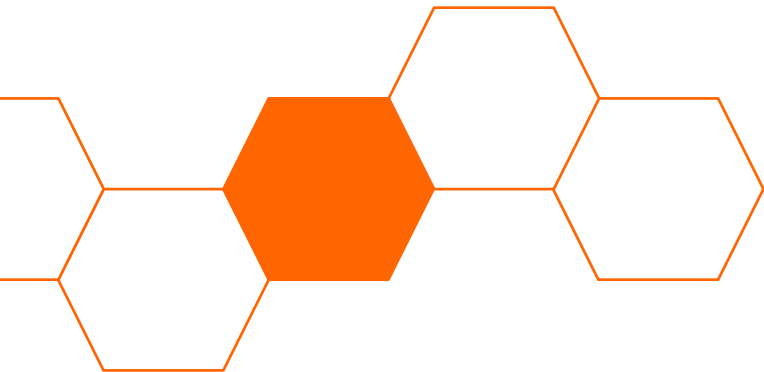
- easy-to-use registry platform
- automatized data control
- mandatory variables



How to Minimize Errors in the Registries

3. After registration:

- data validation using the journals
- data quality reports
- frequent control of data



Some Issues...

- Bias concerns?
- Information about the data quality and the missing data in the publications and online websites directed to patients and politicians?
- Ethical and legal viewpoints



Regulations for the Swedish Quality Registries

- Law about use of personal information (PUL)*.
- **Grundläggande krav på behandlingen av personuppgifter**
- **9 §** Den personuppgiftsansvarige skall se till att:
 - b) personuppgifter alltid behandlas på ett korrekt sätt och i enlighet med god sed,
 - e) de personuppgifter som behandlas är adekvata och relevanta i förhållande till ändamålen med behandlingen,
 - g) de personuppgifter som behandlas är riktiga och, om det är nödvändigt, aktuella,
 - h) **alla rimliga åtgärder vidtas för att rätta, blockera eller utplåna sådana personuppgifter som är felaktiga eller ofullständiga med hänsyn till ändamålen med behandlingen, ...**

Regulations for the Swedish Quality Registries

- **7 kap. Nationella och regionala kvalitetsregister***
- **Inledande bestämmelse**
- **1 §** Med kvalitetsregister avses en automatiserad och strukturerad samling av personuppgifter som inrättats särskilt för ändamålet att systematiskt och fortlöpande utveckla och säkra vårdens kvalitet. **Kvalitetsregistren ska möjliggöra jämförelse inom hälso- och sjukvården på nationell eller regional nivå.** Bestämmelserna i detta kapitel gäller för nationella och regionala kvalitetsregister i vilka personuppgifter samlas in från flera vårdgivare.

Validation Using Sampling Theory

Principle: we only sampled some patients and compare their journals to the registered data. We extend this information to all the patients in the registry.

References:

- “Sampling of populations”, P. Levy, 2008, Willey.
- “Sampling Essentials: Practical Guidelines for Making Sampling Choices”, J. Daniel, 2012, SAGE Publications.



Simple Random Sampling

1. Select randomly some patients.
2. Estimate the proportion of incorrect data.
3. Extrapolate to the registry.



Simple Random Sampling

- Usually without replacement.
- It may give a sample that doesn't represent the registry very well.
- “It may require more patients than other sampling techniques”.
- Expensive (transportation cost).



Summary

- There is always going to be errors and missing values in the database but we need to characterize them to be able to trust the results from the analysis.
- The results from the validation give information about the data stored in the registry.
- Not all missing data is MCR and not all misclassification is non-differential.
- Sampling theory helps us to design a sample that represents the registry.
- It doesn't have to be very expensive to validate a registry.



