

The data collection process established for the German DRG system over the last six years with focus on data quality

Mathias Rusert
Department IT & Statistics,
Institute for the Hospital Financing System (InEK),
Siegburg, Germany

DRG-forum Norway
Oslo, March 11, 2009

Agenda

1. Introduction
2. Methods
3. Results
4. Conclusions

Data quality

In general: Quality of data is the basis to obtain quality results.
Data should be: correct, adequate, valid, reliable.

Goal: Development of a DRG system

Basis: High quality data

Means: Combination of expert knowledge and technical routines

- knowhow: medical / economical / technical / analytical / ...
- experience in med.-econ. problems
- not only using the data but understanding how it comes to it
=> communication with hospitals
- plausibility checks / data validation
- reliable software tools to control, verify and analyse

Data collection process

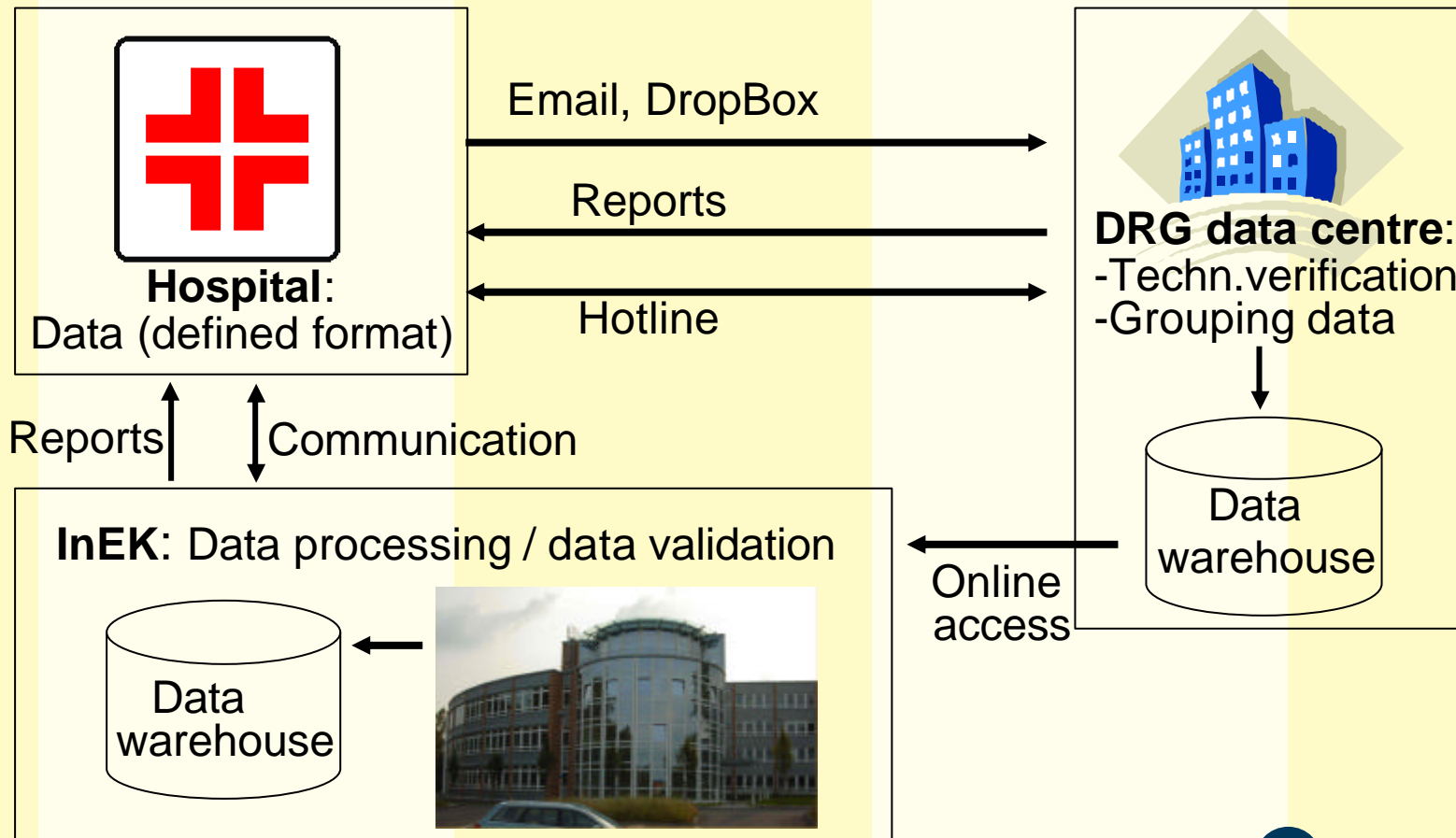
By law nearly all German hospitals provide obligatory

- hospital specific structure data and
- case related service data (patient data).

In partial census about 270 of these hospitals („**sample hospitals**“)
provide additionally and voluntarily

- detailed patient cost data
- additional information concerning hospital structure and approach of calculation
- additional patient data (blood products, drugs, etc.)

Data collection process: Data Flow



Agenda

1. Introduction
2. Methods
3. Results
4. Conclusions

Hospital specific structure data

- Adress, number of beds, email adress
- Unique hospital ID (char 9)
- Operator (public, NPO, private)
- Name of hospital information system / DRG-grouper software
- if applicable: hospital IDs of merged hospitals

Only sample hospitals:

- Details about transferred cost data (cost centre, cost type, definition of cost elements)
- Further information (hospital structure and approach of calculation):
 - hospital structure (e.g. number of beds intensive care)
 - information about basics of cost calculation (e.g. total costs)
 - „clinical allocation models“
 - details about time of operation & anaesthesia

Patient data

Data set defined by self administration / InEK.

- Socio-demographic data of patient: age, sex, zip code, ...
- Medical data: reason and date of admission/ discharge/ relocation, hours of mechanical ventilation (hmv), ...
- Diagnosis: up to 50 in total, one main diagnosis
- Procedure codes: up to 100, with location and operation date
- Reimbursement: DRG, additional charges, ...
- Departmental data: date of admission / discharge for each hospital department

In 2007: about 19 million in-patient (discharges 2007) case data from about 1,700 hospitals

Patient cost data (sample hospitals)

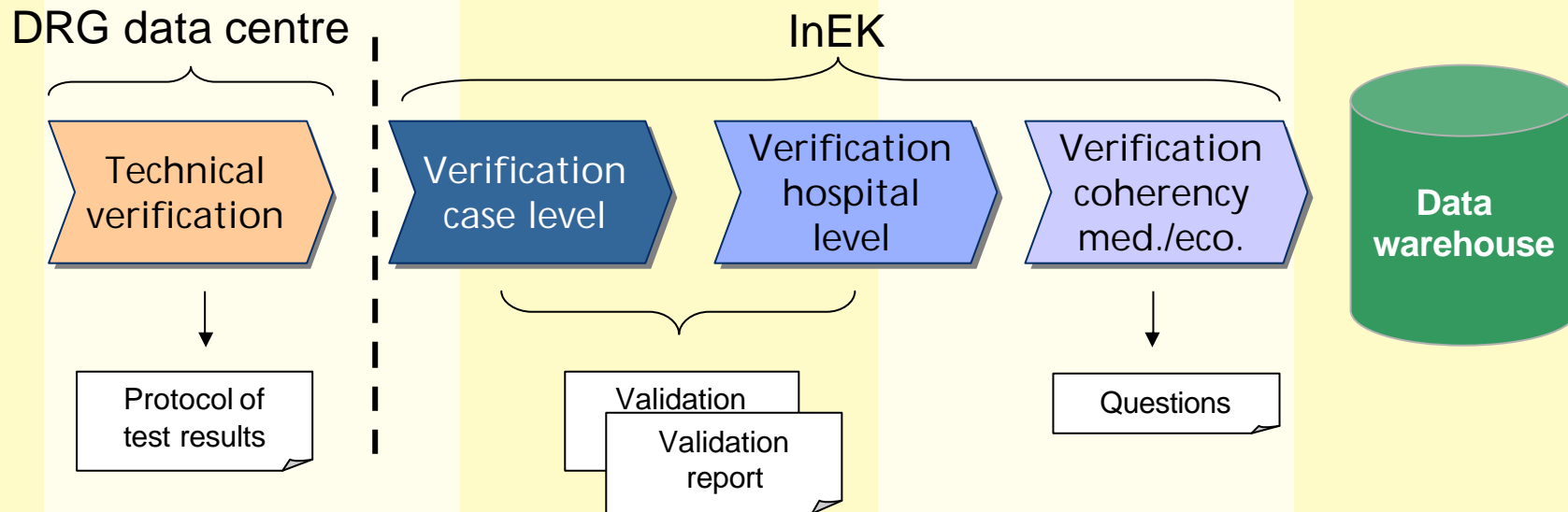
Data is based on accounting data of hospitals.

- Methodology for cost calculation set out in calculation handbook
- Goal: cost allocation based on causality (use of resources)
- Structure of cost data: 11*10 matrix of cost centres vs. cost types.

In 2007: about 4.6 million in-patient case and cost data

		cost type								
		personnel			material				overhead	
		doctors	nurses	other personnel	drugs (general costs)	drugs (specific costs)	implants	medical products (general costs)	medical products (specific costs)	medical overhead costs
cost centre	wards									
	intensive care unit									
	dialysis									
	operating room									
	anaesthesia									
	delivery room									
	cardiology									
	endoscopy									
	radiology									
	laboratories									
	other (US, EEG...)									

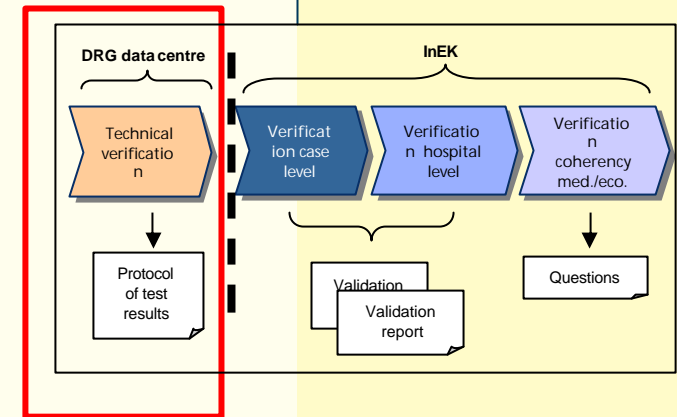
Data verification



Technical verification

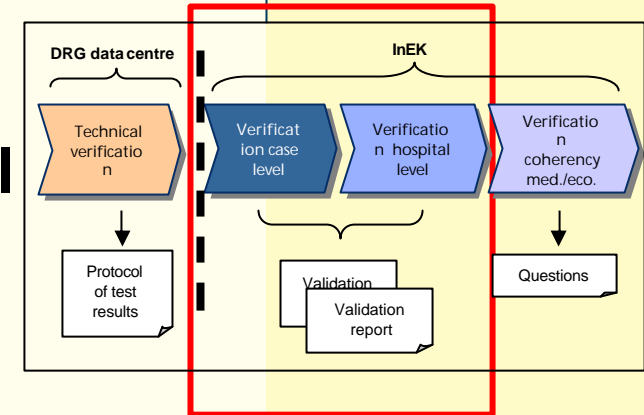
External audit in DRG data centre

- Fundamental plausibility checks
 - all files transmitted completely?
 - format of specific data correct?
 - correct version of diagnosis and procedure classification (ICD-10-GM, OPS)
 - information given plausible / valid (e.g. age, length of stay, ...)?
 - ...
- Report about errors and unaccepted data to hospital
- Hospital has chance to resend corrected data



Verification of data content

Verification on case & hospital level



➤ Economic verification

Costs are checked for the presence and absence of values

- admissibility of values
- infringement of fixed parameter criteria

➤ Medical verification

Compliance of coding guidelines, use of correct codes, ...

➤ Medical-economic verification

coherence / plausibility between medical and economic case data

Development of plausibility checks

- Use the whole data set (sample hospitals) of the previous year (codes: „map forward“ to next years data situation) as data base.
- Evaluation of (number of) hits per check procedure. Adjustment of limit parameters based on the experiences of the previous year(s) or update of comparison values if necessary. Mapping / adjustment of code lists if necessary.
- Use of hospitals comments of the previous year(s) for certain plausibility issues.
- Differentiation of check procedures if adequate to validation problem.
- Technically: executed automatically via Stored Procedures (flexibility during the verification process - e.g. table of certain check values belongs to flexibel number of plausibility checks)

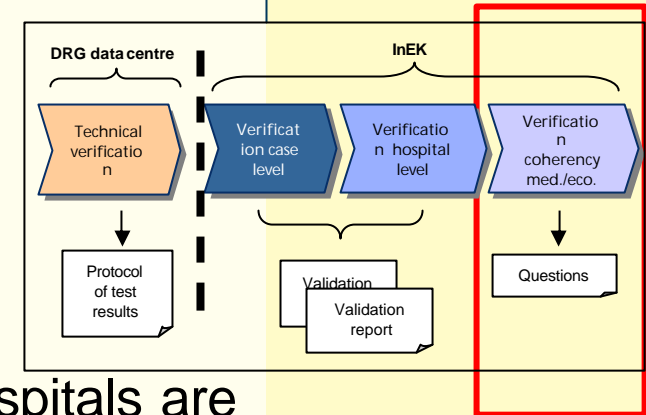
Communication with hospitals

Verification of coherency

Each hospital: **fixed contact expert in InEK.**

During the phase of plausibility checks the hospitals are in contact with InEK:

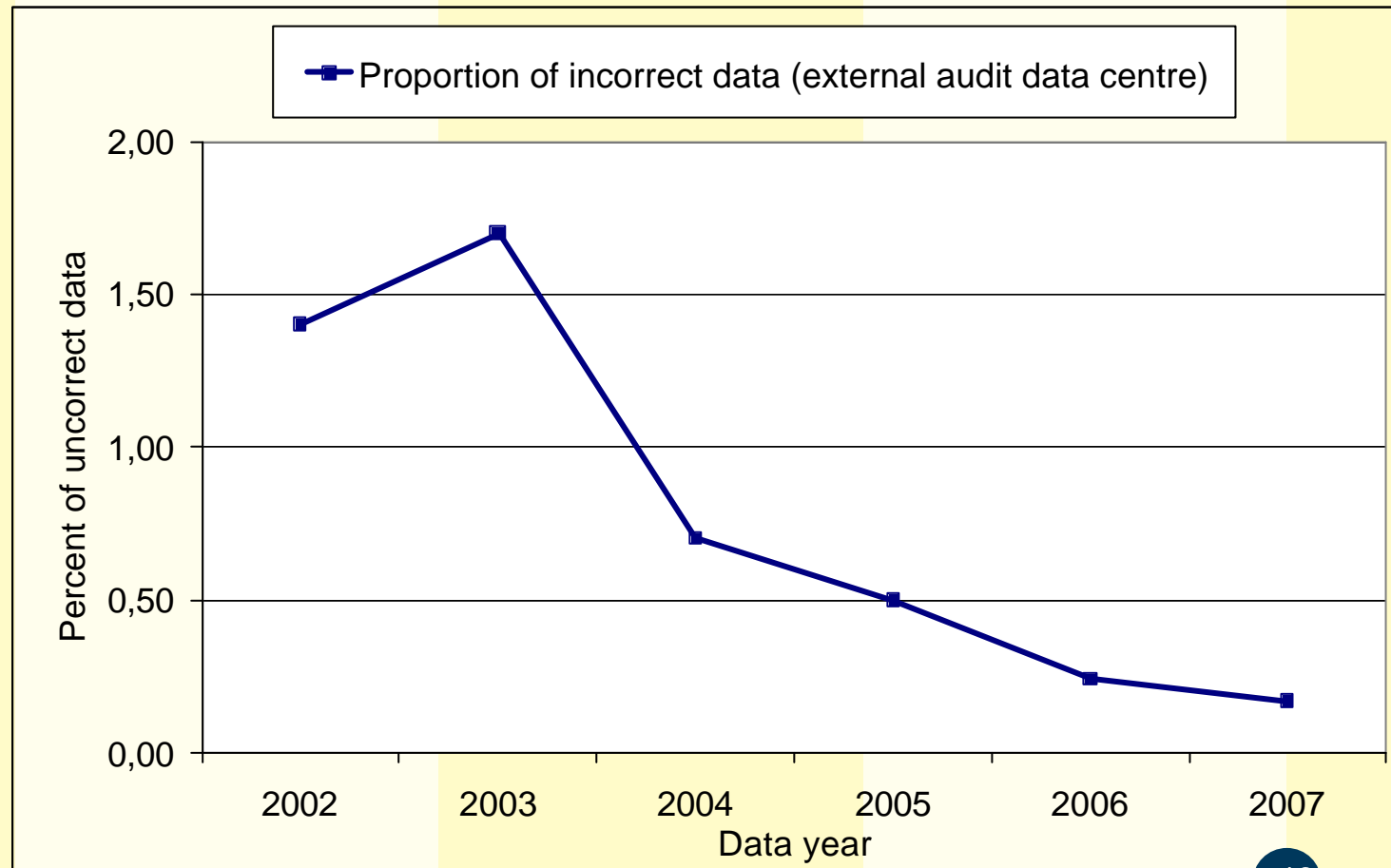
- questions concerning certain plausibility checks
- explanation / commentary concerning marked cases not to be plausible
- information if there are serious errors on hospital level or systematic errors on case level
- the sooner the hospitals know how to improve data the better: **iterative process of comprehension and improvement**
- Chance to send data several times



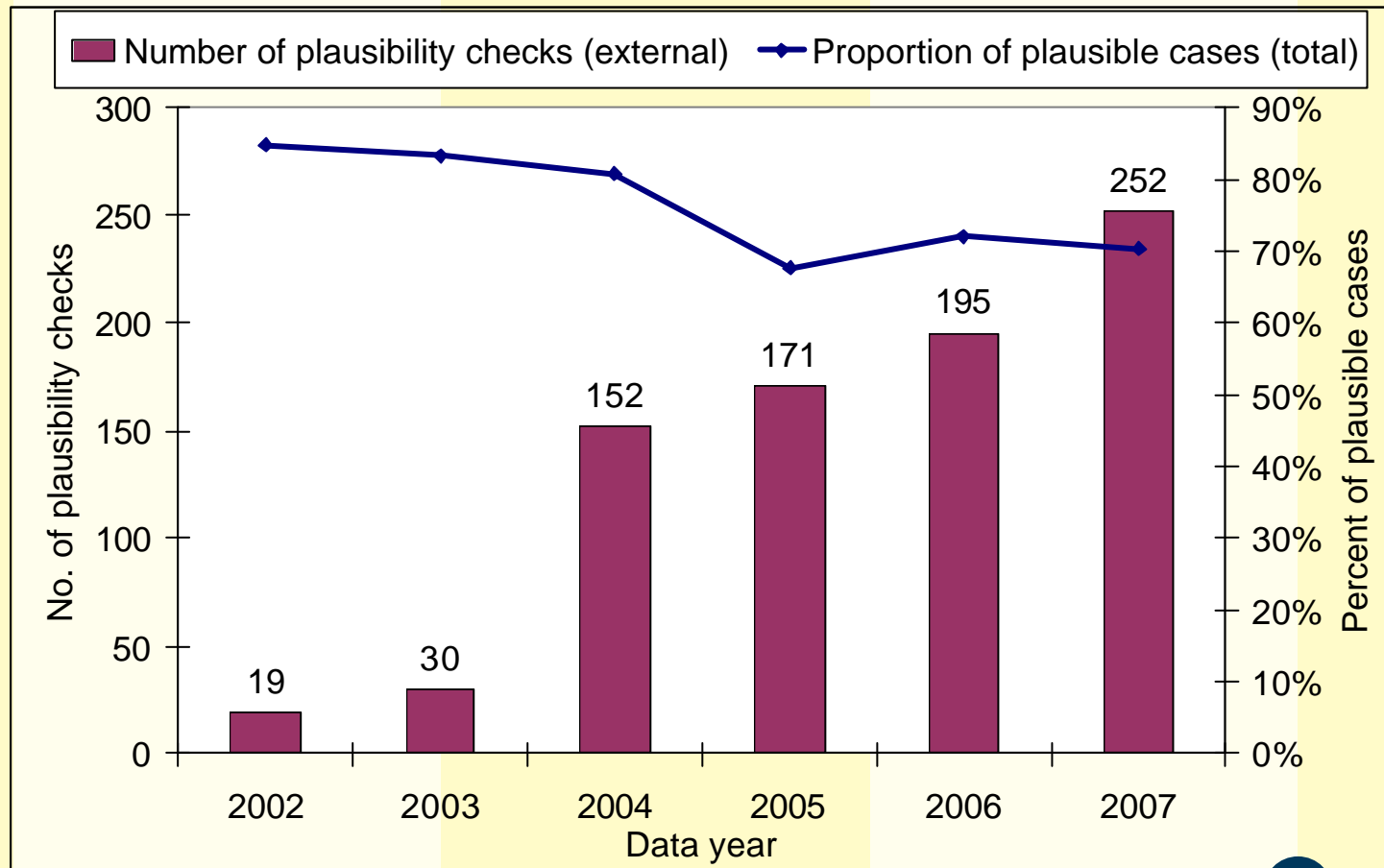
Agenda

1. Introduction
2. Methods
3. Results
4. Conclusions

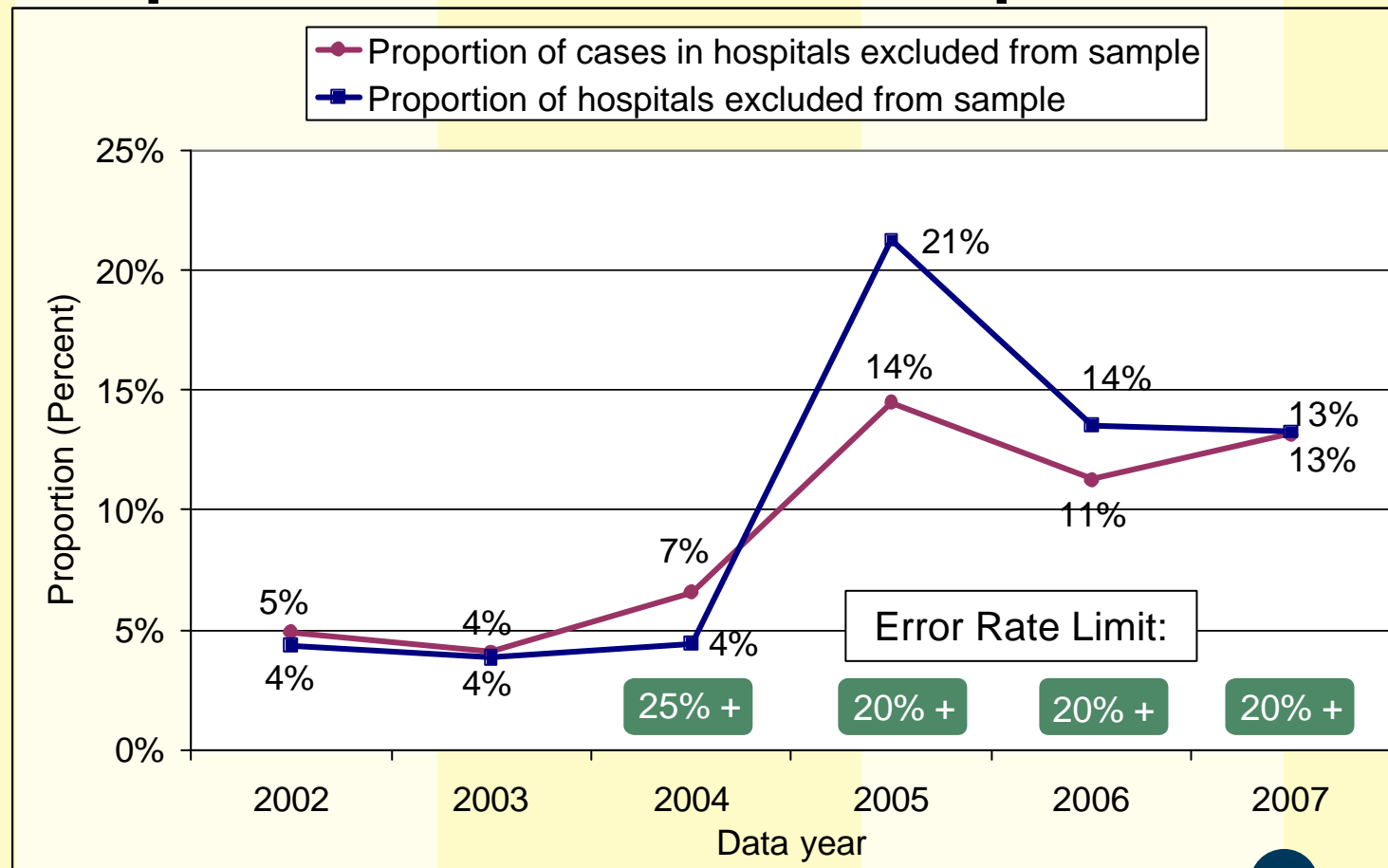
External audit data centre



No. of plaus. checks vs. Prop. of plausible cases



Hospitals excluded from sample



Data application and publication

- Data is basis for further development of the G-DRG-system
- Further applications: Data transmission and publication (annually) according to legal specifications, e.g.
 - hospital planning (Federal level „Bundesländer“),
 - hospital statistics (Federal Statistical Office),
- Public admission to aggregated data: www.g-drg.de
 - e.g. data browser (e.g. Calculation data, Secondary Research)

Data application

Data is basis for further development of the G-DRG-system:

- Every DRG-simulation is based on the plausible cases of sample hospitals. Proprietary software and control processes are installed to execute the simulations and calculations resp.
- Results are then applied to the total data set to analyse the effect of the simulated system to the German hospital service spectrum. Proprietary reporting tools are used to support these analyses.
- Data is basis to calculate the annual baserate (underlying the case weights of the DRGs).

Data application: Expert knowledge

Even if you have a reasonable number of high quality data:

➤ there are still aspects unanswered concerning the allocation of services to remuneration.

(Easy to produce arbitrary data nonsense automatically ...!)

Therefore:

➤ Expert analyses beyond standard techniques are essential, together with the knowledge to use the software efficiently (ask proper questions to get appropriate answers for your problem).

Agenda

1. Introduction
2. Methods
3. Results
4. Conclusions

Conclusion

- Data quality is essential for a data-driven DRG system.
- To achieve a high level of data quality
 - intensive communication with the participating hospitals is needed
 - experience from previous years should be integrated into the development of data validation
- Even with a large amount of high quality data there are still questions unanswered concerning the allocation of services to remuneration => need expert analyses beyond standard techniques.
- Publication of data, methods and results leads to transparency and acceptance.

Thank you for your attention!