Medical data analysis infrastructure
- Attempt at the Kochi Medical School -

Hiromi KATAOKA, Yutaka HATAKEYAMA,
Yoshiyasu OKUHARA, Shu KURAMOTO
Kochi Medical School
Background

Electronic Health Record

Penetration: 21.8% (<400 beds)
68% (>400 beds)

Japanese Association of Healthcare Information Systems industry

Data analysis
- Quality enhancement of medical treatment
- Secondary use for EBM

USA
AMGA (American Medical Group Association)

Health, Labour and Welfare Ministry Online by receipt

Kochi Medical School Hospital
IMIS (Integrated Medical Information System)
Purpose

- Data anonymizing
  - Privacy protection
- Large variety of data type
  - Number, text, waveform data, image
- Standardization of test data in different hospitals
  - Existence of test error
Fig1. Need for Analytical Database

- **Used for hospital management**
- **Used for medical care**
- **Used for medical research**

**Medical Records Database**
- Feedback
- Restructuring of system
- Operations Improvement

**Analytical Database**
- Evaluation of system
- Evaluation of the quality of medical care
- Quality indicator
- EBM (Evidence-Based Medicine)
- Cross-section study
- Longitudinal study

**Research outcome**
- Hypothesis

**Planning of prospective study**
- Retrospective study

**Operations Improvement**
- Feedback
Fig. 2 DWH system in Kochi Medical School Hospital
Process for anonymous data

• Anonymizing for Patient ID
  – reversible mapping
• Change of date information (shifting)
  – Definition of changing date
• Data removal for few disease

<table>
<thead>
<tr>
<th>Patient ID</th>
<th>Anonymous ID</th>
<th>Shifting date</th>
</tr>
</thead>
<tbody>
<tr>
<td>0000001</td>
<td>1234567890</td>
<td>5</td>
</tr>
<tr>
<td>0000002</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>0000003</td>
<td>0987654321</td>
<td>-3</td>
</tr>
<tr>
<td></td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td></td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
Ratio of patient (HbA1c < 7.0mg/dl) over 3 months diabetes care

Fig.3. Ratio of control patient for diabetes care
Distribution of lactase dehydrogenase (LD) in 28 years (distribution by year)

Fig. 4. Standardization of test data in measurement change

Change of measurement 2001/10/1
Definition of coefficient of transformation based on data distribution

Test data in long period

Data by year based on
Old process for measurement

BoxCox transform → Truncation → coefficient of transformation

Mean value

Data for evaluation

Coefficient of transformation

Regression coefficient

Evaluation

Little data in other hospitals!!!

Concurrent measurement values
In inspection department

Fig. 5. Proposal of normalization process
Seamless analysis of longitudinal data with considering accuracy of assay

Fig.6. Sample of standardization result
5. Conclusion

• Infrastructure development for medical records
  • Kochi medical school hospital from 1981
• Secondary use
  • medical economics, estimation, ...
• Data integration process with other hospitals
  • standardization
• Right management system of each medical data
  • Patients, hospitals, ...