Validity of data sources in pharmacoepidemiology

Jesper Hallas MD DrMedSc
Dept of clinical pharmacology
University of Southern Denmark, Odense
jhallas@health.sdu.dk
Disposition

1. Defining validity
2. Internal vs external validity
3. Describing the process from prescribing to ingestion
4. Areas of major and minor uncertainty in the prescribing process
   - Prescribed vs purchased medication
   - Purchased vs ingested medication
   - Ingested medication vs verbal account of medication
5. Some examples of validity studies of data sources.
6. How to read a paper on validity
7. Does the choice of data source make a difference in QI?
8. Some conclusions
Validity definition

- Of Latin verb *valere*: being strong

- The validity of a measurement can be defined as the degree with which the measured value reflects the characteristic it is intended to measure.

  Roger J. Lewis,

- implies the comparison of a measurement against a superior representation (or measurement) of its object.
Internal and external validity

- Internal: pertains to the study sample itself
  - Are the data correct?
    - Technical problems in data transfer?
    - Which data source is most accurate? (for what purpose?!)  
- External: pertains to the reality outside the study sample
  - Is the studied population representative?
Describing the process from prescribing to ingestion
Describing the process from prescribing to ingestion

Event

Proxy

Physician administrative system, e.g. GPRD in UK

Pharmacy-based prescription database, e.g. MEDICAID, MEMO, OPED

Interview
Primary non-compliance; what prescriptions are not redeemed?

20,921 prescriptions issued in a large rural practice retrieved in a pharmacy-based research database (MEMO, Scotland).

"Observational studies of drug exposure can be more accurately estimated from dispensing rather than prescriber data"

Validity problems; can patients’ account be trusted?

- 38 patients
- Pilocarpin eyedrop t.i.d. for 40 days
- Compliance estimate by structured interview
- Compliance measured by electronic device

Comparison of interview and electronic monitoring device: findings

- No statistical correlation between estimated and actual compliance
- Gross non-compliance occurred among subjects with an estimated perfect compliance
- Patient were never more compliant than claimed

Data on drug exposure
Which is best?

Physician administrative system, e.g. GPRD

Pharmacy-based prescription database

Interview

Area of major uncertainty
Validity studies: grid of data flow

**Event**

- **Level I**
  - GP
  - →
  - GP database

- **Level II**
  - Pharmacy
  - →
  - Pharmacy database

- **Level III**
  - Patient
  - →
  - Patient source
Measuring internal validity: taxonomy

- Indicate levels of comparison (level I-III)
- Indicate patient group
- Indicate gold standard for comparison
- Indicate tools for data acquisition (level I-III)
- Indicate possible limitations/exemptions of data transfer (level I-III)
Validity grid: Beardon study

Event

GP

Assumed without error

Proxy

GP database

Pharmacy database

Assumed without error

Pharmacy

Patient source

Patient

Study object

Measuring point

Gold standard

Measuring point
Validity grid: Norell study

Event | Proxy
---|---
GP | GP database
Pharmacy | Pharmacy database
Patient | Structured interview
| Monitoring device output
Comparison between interview and register data

- Participants in a questionnaire-based national survey of pregnant women (N=2041) were retrieved in a prescription database (PDNJ) by use of a 120 day window

<table>
<thead>
<tr>
<th>Drug</th>
<th>Reported by interview</th>
<th>Retrieved in database</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beta-blockers</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>Insulin</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>Thyroid drugs</td>
<td>12</td>
<td>9</td>
</tr>
<tr>
<td>Diuretics</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Antidepressants</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>Analgesics</td>
<td>828</td>
<td>17</td>
</tr>
<tr>
<td>Topic corticosteroid</td>
<td>25</td>
<td>55</td>
</tr>
<tr>
<td>Antibiotics</td>
<td>204</td>
<td>310</td>
</tr>
</tbody>
</table>

Olesen C et al, J Clin Epidemiol 2002
Olesen C et al, Epidemiology 2001, in press
Validity grid, Olesen study

<table>
<thead>
<tr>
<th>Event</th>
<th>Proxy</th>
</tr>
</thead>
<tbody>
<tr>
<td>GP</td>
<td>GP database</td>
</tr>
<tr>
<td>Pharmacy</td>
<td>Pharmacy database</td>
</tr>
<tr>
<td></td>
<td>OTC-analgesics exempt from registration</td>
</tr>
<tr>
<td>Patient</td>
<td>Patient source</td>
</tr>
<tr>
<td></td>
<td>Study object</td>
</tr>
<tr>
<td></td>
<td>Posted Questionnaire</td>
</tr>
</tbody>
</table>

Measuring point:
- Gold standard
- Measuring point
Does the choice of data source make a difference in QI?
Stepwise Approach to Asthma Therapy - Adults

Outcome: Asthma Control

Controller:
- Daily inhaled corticosteroid
- Daily long-acting inhaled β2-agonist

Reliever:
- Rapid-acting inhaled β2-agonist prn

Outcome: Best Possible Results

Controller:
- Daily inhaled corticosteroid
- Daily long-acting inhaled β2-agonist
- plus (if needed)
- Theophylline-SR
- Leukotriene
- Long-acting inhaled β2-agonist
- Oral corticosteroid

STEP 1: Intermittent
STEP 2: Mild Persistent
STEP 3: Moderate Persistent
STEP 4: Severe Persistent
STEP Down

When asthma is controlled, reduce therapy
Monitor

Alternative controller and reliever medications may be considered (see text).
Validity of quality indicators for assessing the quality of asthma prescribing

- Dutch GP based prescription database (N=30,486) in 1997
- 146 patients aged 18-49 with asthma
- Subjected to database analysis
- Subjected to interview and diagnostic work-up (gold standard)

## Sensitivity and positive predictive values for database-derived quality indicators in asthma

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Problem</th>
<th>Sensitivity</th>
<th>Positive predictive value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1</td>
<td>No inhaled beta-agonist</td>
<td>0.86</td>
<td>0.52</td>
</tr>
<tr>
<td>Step 2A</td>
<td>No inhaled corticosteroid, +/- inhaled betaagonist</td>
<td>0.74</td>
<td>0.46</td>
</tr>
<tr>
<td>Step 2B</td>
<td>Daily use of an inhaled beta-agonist, no inhaled corticosteroid</td>
<td>0.37</td>
<td>0.71</td>
</tr>
<tr>
<td>Step 3</td>
<td>Inhaled beta-agonist &gt; once daily but low dose inhaled corticosteroid</td>
<td>0.07</td>
<td>0.20</td>
</tr>
<tr>
<td>Step 4</td>
<td>Inhaled beta-agonist &gt;once daily, adequate dose inhaled corticosteroid, no LABA</td>
<td>Too few patients to validate</td>
<td></td>
</tr>
</tbody>
</table>

Validity grid: Pont study

Event

GP

↓

Pharmacy

↓

Patient

? → GP database

↓

Pharmacy database

↓

Patient source

Interview + Clinical work-up

Proxy

Measuring point

Gold standard
Lack of validity for database derived asthma QI, why

Fig. 2. Distribution of ACQ scores according to individual annual use of SABA in DDD. Area of circles proportional to the number of subjects for each data point. ACQ = Asthma Control Questionnaire. SABA = short-acting beta-2-agonists. DDD = defined daily dose.

Davidsen et al. Pulm Pharmacol Ther 2011
# Studies of agreement between prescription data sources

<table>
<thead>
<tr>
<th>Reference</th>
<th>Drugs</th>
<th>Comparators</th>
<th>Finding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Norell 1981</td>
<td>Eye drops</td>
<td>Interview vs EMD III vs III</td>
<td>Poor correlation between estimated and actual compliance</td>
</tr>
<tr>
<td>Pont LG 2004</td>
<td>Asthma medication</td>
<td>Interview + work-up vs GP. I vs III</td>
<td>Poor performance of GP derived quality indicators</td>
</tr>
<tr>
<td>Barat I 2001</td>
<td>All, elderly patients</td>
<td>Interview vs GP I vs. III</td>
<td>20-70% non-adherence, depending on dose regime, knowledge about drug, toxicity etc</td>
</tr>
<tr>
<td>Olesen C 2002</td>
<td>All, pregnant women</td>
<td>Questionnaire vs pharmacy. II vs III</td>
<td>Good agreement for chronic treatments, poor for antibiotics and analgesics</td>
</tr>
<tr>
<td>Enlund H 1981</td>
<td>Users of antihypertensives</td>
<td>Interview vs pharmacy. II vs III</td>
<td>94% agreement. Patient interview gave an overestimation of compliance</td>
</tr>
<tr>
<td>de Jong-van den Berg LT 1993</td>
<td>All, pregnant women</td>
<td>Interview vs pharmacy. II vs III</td>
<td>The two modes are complementary. Performance dependent on study object. Pharmacy record preferable in long recall or complicated regimes</td>
</tr>
<tr>
<td>Van den Brandt PA 1991</td>
<td>All prescribed</td>
<td>Questionnaire vs pharmacy. II vs III</td>
<td>Sensitivity 61%, specificity 100%</td>
</tr>
<tr>
<td>Lau HS 1997</td>
<td>All prescribed</td>
<td>Medical inventory vs pharmacy II vs III</td>
<td>Sensitivity improved with increased observation window</td>
</tr>
<tr>
<td>West SL 1995</td>
<td>NSAID and estrogen up to 12 years in the past</td>
<td>Interview vs pharmacy. II vs III</td>
<td>Sensitivity 41-78% depending on number of prescriptions. Specificity 92-100%</td>
</tr>
</tbody>
</table>
Studies of internal data validity, statements

- True gold standards do not exist. They are all alloys.
- Technical data errors are rarely reported
- Non-compliance is substantial and is difficult to ascertain
- Failure to redeem prescription is a minor problem for most drug classes
- Databases may have important limitations, e.g. absence of non-subsidized prescription medication
- Recall errors of omission are much more common than errors of commission. Rather forget than make something up.
- Recall is dependent on chronicity
Data on drug exposure
An example; antibiotics and myocardial infarction

Background: Antibiotics with effect against Chlamydia Pneumonia might protect against AMI

Exposure: any antibiotic 3 years before the index date.

Cases: 3315 persons with first-time MI
Controls: 13,139 without MI.

Question

- Given that the source of data has a strong impact on QI, should we endeavour to use
  - GP source?
  - Patient level source?
  - Something else?

When assessing the quality of the prescribing?