Italian Cystic Fibrosis Registry (ICFR) data quality assessment

Natalia Cirilli

2nd International Workshop
RARE DISEASE AND ORPHAN DRUG REGISTRIES

(Rome, 22nd October 2013)
Disease Registry
Data Quality Assessment

This hard work is essential and preliminary to any epidemiological study based on these data.

Analysis of registry events and outcomes requires an assessment of data quality.

That means evaluate if:
- all important variables are collected
- the data are complete
- missing data are handled correctly
- data are accurate
ICFR project

Data input

Quality Control Working Group

Data OK?

YES

Data analysis

Italian CF centres

Data OK?

NO

European CF registry

Web platform

Annual Report

Specific queries
ICFR Quality Control Working Group:
Natalia Cirilli, Gianluca Ferrari

Project partners:

ICFR Scientific Committee:
• Serena Quattrucci, Carla Colombo, Vincenzina Lucidi on behalf of Italian CF Society
• Domenica Taruscio, Antonia Stazi, Virgilia Toccaceli on behalf of Italian National Institute of Health

Lega Italiana Fibrosi Cistica:
Gianna Puppo Fornaro

European CF Registry
ICFR QA web platform
- Data transfer procedures
- ICFR list of variables (definitions and input rules)
- Minutes from meetings
- European CF Registry QA reports
- Contacts (QA WG and CF centres)
- ICFR QA reports
- Next deadlines
ICFR Data Quality Assessment

• Start date: 1st August 2012
• End date: 30th April 2013
• ICFR sections:

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>Pregnancy</th>
<th>Microbiology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sweat test</td>
<td>Paternity</td>
<td>Therapies</td>
</tr>
<tr>
<td>Genetics</td>
<td>Transplants</td>
<td>Complications</td>
</tr>
<tr>
<td>Lung function</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Growth</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
RIFC: missing data

- 25 CF centres
- CF patients (2010 year) = 3388

![Graph showing missing data across different CF centres for PERT therapy, Sweat test, Mut1, and Mut2.](image)
RIFC: missing data

- 25 CF centres
- CF patients (2010 year) = 3388
CF diagnosis

- 25 CF centres
- CF patients (2010 year) = 3388
Lung function

- 25 CF centres
- CF patients (2010 year) = 3388

- Error type A: FEV1, FVC volumes in millilitres and not in litres. 4/25 (16%) CF centres are affected by this type of error (range: 0.6% - 87.3%)

- Error type B: FEV1, FVC (% predicted) = 0. 5/25 (20%) CF centres are affected by this type of error (range: 0.6-70.1%)

- Error type C: incomplete FEV1, FVC: 6/25 (24%) CF centres are affected by this type of error (range: 1.0% - 34.5%)

- Error type D: non-homogeneous pulmonary function values (0, 1, 2 decimal numbers): 5/25 (20%) CF centres are affected by this type of error (range: 0.5% - 23.5%)

- Error type E: FEV1 and FVC volume (L) and FEV1, FVC (% predicted) inverted: only 1/25 (4%) CF centres is affected by this type of error
Growth

- 25 CF centres
- CF patients (2010 year) = 3388

🎉 error type A: anthropometric and pulmonary function parameters are not collected on the same day. 24/25 CF centres are affected by this type of error (range: 1.5-100%)

😢 error type B: wrong weight/height; weight and height reversed. 8/25 CF centres are affected by this type of error (range: 0.3-2.6%)

😢 error type C: weight without decimal numbers: 25/25 CF centres are affected by this type of error (range: 3.8-100%)

😢 error type D: incomplete data: 7/25 centres are affected by this type of error (range: 0.5-11.7%)

😢 error type E: height/weight = 0: 2/25 centres are affected by this type of error (range: 2.3%-31.2%)
Next steps

– constant monitoring of the data with a team of experts
– active collaboration between QC working group and CF centres
– train health care professionals to input data correctly