Data Linkage of Clinical Trial Data to Research Biomaterial Repositories

Prof. Ruth Ladenstein (Vienna, AT)
St. Anna Children’s Hospital and Cancer Research Institute
The Case of Neuroblastoma
– A Rare Cancer of Childhood!

- Population-based data from Cancer Registries participating in RARECARE:
  Gemma Gatta: European Journal of Cancer (2012) 48, 1435 ff
  - About 2000 new embryonal cancers every year in EU27
  - Annual incidence rate of 4 per million
    (1.8 neuroblastoma, 1.4 nephroblastoma, and 0.5 retinoblastoma);
  - 91% of cases in patients under 15 years

- Cancer of the sympathetic nervous system
- Adrenal glands, but also in nerve tissues in neck, chest, abdomen, pelvis
- 50% before the age of 2 years
- 50% wide spread dissemination at diagnosis
- It is a disease exhibiting extreme heterogeneity – Biology is key!
  - Low-risk disease most common in infants and good outcomes are common with observation only or surgery
  - High-risk disease is difficult to treat successfully even with intensive multi-modal therapies.
A Case for International Collaboration!
International Neuroblastoma Risk Group (INRG) Data

2004: INRG Task Force established
(52 investigators from US, Europe, Japan, Australia) to develop a consensus approach to pre-treatment risk stratification

Methods:
- “Double Pseudonymisation” of Clinical Trials and Research Data Sets (via a honest broker = trusted third party)
- Data collected on 8,800 unique patients diagnosed between 1990-2002 and treated on studies from COG, SIOPEN, GPOH, JANB and JINCS with follow-up to 2004
  - Demographics
  - 36 prognostic markers (Genetic markers: 1p, 11q, MYCN, ploidy)
  - Treatment
  - Outcome (EFS, OS)

Factors prognostic of event-free survival were identified using survival tree regression

Will we need to go back to every single patient/parent for „specific“ and „explicit“ consent in the future?
Secondary Use of Data to built the “The INRG Classification System”
Survival Tree Regression: Top Level – New Insights!

Overall
n=8,800
EFS 63%±1%
OS 70%±1%

INSS Stage 1,2,3,4S
n=5,131
EFS 83%±1%
OS 91%±1%

INSS Stage 4
n=3,425
EFS 35%±1%
OS 42%±1%

GN, maturing
GNB, intermixed
n=162
EFS 97%±2%
OS 98%±2%

NB & GNB, nodular
n=4,970
EFS 83%±1%
OS 90%±1%

MYCN NON-AMP
n=3,926
EFS 87%±1%
OS 95%±1%

MYCN AMP
n=349
EFS 46%±4%
OS 53%±4%

EFS: >85%
EFS: >75-<85%
EFS: >50-<75%
EFS: <50%

Terminal Node

Non 4S
Metastatic Disease

Histological Category

MYCN Status
Benefits of Secondary Use of Data
“The INRG Classification System”

7 factors identified that were highly statistically significant and also considered clinically relevant

- Non 4S Metastatic Disease
- New Age Cut Point: < 18 months vs. ≥18 months]
- Histological Category – Ganglioneuroma, ganglioneuroblastoma – intermixed vs. neuroblastoma, ganglioneuroblastoma – nodular
- Grade of Tumour Differentiation differentiating vs. undifferentiated or poorly differentiated
- 3 Biological Factors
  - MYCN status
  - Presence/absence of 11q aberration
  - Ploidy (≤ 1.0 versus >1.0)

Such efforts rely on a „broad“ **One-Time Only Consent**!
- Trying to trace back patients absorbs enormous time and resources
- Likely to result in loss of data or abandoned research
Benefits of Secondary Use of Data

The INRG Classification System

- Ensures that children diagnosed with neuroblastoma in any country are stratified into homogenous pre-treatment groups.

- Facilitates the comparison of risk-based clinical trials conducted in different regions of the world.

- Enhances our ability to develop international collaborative studies.
The Issue:
Need for Secondary Use of Data

Collaborative and shared research

- INRG data are available for investigator-initiated data mining studies
- Approximately 30 research projects completed or still ongoing
- Analysis conducted by INRG statisticians
- Published in high profile journals
  - Bagatell et al., J Clin Oncol, 2009
  - Taggart et al., J Clin Oncol, 2011
  - Baruchel et al., Eur J Cancer, 2011
  - London et al., J Clin Oncol, 2011
  - Schleiermacher et al., Br. J. Cancer 2012
  - And more...
One Example of many…
Achievements of the INRG Biology Committee

- Development of precise definitions
- Standardisation of techniques
- Proposition of standard operating procedures for the determination of genetic markers used for treatment stratification (MYCN)
Continued Need for Secondary Use of Data and Follow UP!

Limitations of original INRG Data

- **Original INRG Data Base outdated!**
  - Consists of prognostic factors identified > 30 years ago
  - More recent whole genome data generated by labs around the world are not included in the database (GWAS, array cGH, omic signatures, NGS)

- **GOAL**
  - Transform the originally flat-field application housing the INRG data
  - Use new technology facilitating links with other databases (i.e. biobank data, genomic data, …)
  - Create an Interactive INRG database (iINRGdb)

The future potential of biomarker and mode of actions discovery rely on Data Linkage and Patient Traceability!

Does not work with anonymised data sets!
Evolution of Techniques

New datasets, using new technologies, have been generated!

<table>
<thead>
<tr>
<th>Technique</th>
<th>Nº of features</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLPA</td>
<td>100 loci (Ambros et al 2011)</td>
</tr>
<tr>
<td>aCGH</td>
<td>4k – 1000k</td>
</tr>
<tr>
<td>SNP</td>
<td>&gt; 1 Mio</td>
</tr>
<tr>
<td>SNP (SNP6, cytoscanR)</td>
<td></td>
</tr>
</tbody>
</table>

- Sequencing data
- Whole exome
- Whole genome
Next Steps: Expansion Phase

New datasets, using new technologies, have been generated!

- DNA copy number profiles (array CGH, SNParrays)
- Somatic mutations (NGS techniques)
- Coding gene expression profiles
- miRNA and non coding gene expression profiles
- Methylation and other epigenetic profiles

- Genomics of peripheral samples (ctDNA)
- Germline genomics

- **Currently updating outcome data and expanding data fields on existing patients** (race, ethnicity, sex, second malignancies, etc)

- **Adding data on new patients after approval from Cooperative Group Chairs**
The Need: Large Scale Data Integration in Rare Diseases

i.e. “An Interactive iINRGdb” – under construction

- Fostering research in Biomarker Discovery & Mode of Actions
- Basis for Innovative Drug Development
- Basis for “Personalized Medicine” approaches in Rare Diseases

Impact of DPR

NEED TO MAINTAIN RESEARCH IN RARE DISEASES
- One-time “broad” consent
- Pseudonymisation
- Safe-Guard Measures
Acknowledgments
INRG Task Force

- **Co-Chairs** - Andrew D. J. Pearson, U.K. and Susan L. Cohn, USA
- Investigators: pediatric oncologists, biologists, statisticians, pathologists, surgeons, radiologists, and young investigators
- Investigators were assigned to chair one of 4 committees:
  - Surgery (Tom Monclair)
  - Statistics (Wendy London)
  - Biology (Peter Ambros)
  - Metastatic Disease (Kate Matthay)

- INRG investigators
- The Forbeck Foundation – Sponsor of the 2005 INRG Conference
- Little Heroes Pediatric Cancer Research Foundation
  (Friends For Steven Pediatric Cancer Research Foundation)