ON BIOMARKER DATA QUALITY

Paolo Verderio

Unit of Bioinformatic and Biostatistics

Fondazione IRCCS Istituto Nazionale dei Tumori di Milano

CHALLENGES IN DATA-DRIVEN GENOMIC COMPUTING

Organized by the Advanced ERC Project 693174 “GeCo” – data-driven Genomic Computing

Workshop in Como, Villa del Grumello – March 6-8 2019
Outline

I. Background
   ✔ Biomarker definition
   ✔ Biomarker state of the art

II. Biomarker development process
   ✔ Biomarker Discovery
   ✔ Biomarker Validation
   ✔ Clinical Utility

III. Concluding remarks
I. Background
What is a biomarker?

- A defined characteristic that is measured as an indicator of normal biological processes, pathogenic processes, or responses to an exposure or intervention, including therapeutic interventions.

- Molecular characteristic
- Histological characteristic
- Radiographical characteristic
- Physiological characteristic
- ....

https://www.ncbi.nlm.nih.gov/books/NBK338448/#IX-B
Hayes DF. Biomarker validation and testing. Molecular Oncology. 2015(9): 960-966
What is a tumor biomarker?

might be a:

- **molecular change** (i.e. nucleic acid, protein, or metabolite) OR
- **process change** (i.e. alteration in tissue or image appearance)

Can be detected in:

- ✓ tissue
- ✓ blood
- ✓ secretions (i.e. urine, stool, sputum, or breast nipple aspirates, CSF, etc)
What is a biomarker test?

✓ a test used to identify or measure the perturbation reflected by the tumor biomarker

✓ must be applicable in clinical routine and reproducible and accurately translate the biomarker into a measurement parameter

Hayes DF. Biomarker validation and testing. Molecular Oncology. 2015(9): 960-966
Kisser A. Issues and challenges in the systematic evaluation of biomarker tests. Precision in Medicine. 2018
Type of biomarker (1)

- Diagnostic biomarker
- Pharmacodynamic response biomarker
- Prognostic biomarker
- Predictive biomarker
- Susceptibility risk biomarker
- Monitoring biomarker
- Safety biomarker

Type of biomarker (2)

**Diagnostic**: “A biomarker used to detect or confirm presence of a disease or condition of interest or to identify individuals with a subtype of the disease.”

**Predictive**: “A biomarker used to identify individuals who are more likely than similar individuals without the biomarker to experience a favorable or unfavorable effect from exposure to a medical product or an environmental agent.”

**Prognostic**: “A biomarker used to identify likelihood of a clinical event, disease recurrence or progression in patients who have the disease or medical condition of interest.”
huge number of research studies focused on biomarkers in many different field and specimen
~30% are focused “oncological” biomarkers
However....

- a limited number of biomarkers is really translated and fully applied in the real clinical practice
- very few discovered biomarkers are used in the routine clinical setting

Why?

Poste G. Nature 2011;469:156-157
Challenges in biomarker development (1)

- Analytical factors
- Pre-analytical factors
- Statistical issue

Biomarker determination
## Challenges in biomarker development (2)

### Pre-analytical factors

- ✔ starting material
- ✔ sample collection and processing
- ✔ sample storage conditions
- ✔ type of extraction method
- ✔ ...

### Analytical factors

- ✔ type of detection method
- ✔ type of throughput
- ✔ type of commercial platform
- ✔ ...

### Statistical analysis

- ✔ data normalization
- ✔ level and type of analysis
- ✔ type of predictor
- ✔ methods for multivariate model building
- ✔ ...


Challenges in biomarker development (3)

- Molecular features measurable from **body fluids** (i.e. Liquid biopsies)
  - a great opportunity to develop non invasive biomarkers

- Availability of **high-throughput technologies**
  - new diagnostic and therapeutic perspective especially in the oncological scenario
Challenges in biomarker development (3)

II. Biomarker development process
Comprehensive evaluation of biomarker development process
1. **identification of the context of use [study planning]**
   - the fundamental use of the biomarker
   - the target population (with inclusion and exclusion criteria)
   - specific clinical context
   - …

2. **setting-up of the conditions for experiment running [data generation]**
   - sample specimen and collection (e.g. collection tubes, timing of collection)
   - sample processing (e.g. timing of processing, temperature of storage)
   - sample analysis (e.g. type of extraction, platform)
   - …

3. **biomarker identification/selection [data analysis]**
   - pre-processing and quality controls
   - definition of data analysis procedure
   - identification of candidate biomarker
   - development of omics-based signatures (e.g. statistical model)
   - …
Biomarker discovery (2)

omics-based signature generation

Signature discovery

• setting-up of the conditions for experiment running
• pre-processing and quality controls
• data normalization
• definition of data analysis procedure

assay optimization

• signature confirmation

technical validation

Signature validation

• Identification of candidate biomarkers
• preliminary signature(s) building

Identification of validated signature(s)

Biomarker discovery (3)

omics-based signature generation

Signature discovery → assay optimization → Signature validation

Validated omics-signature(s) = biomarkers

Clinical Cancer Research

Plasma microRNA levels for predicting therapeutic response to neoadjuvant treatment in HER2-positive breast cancer: Results from the NeoALTTO trial

Serena Di Cosmo, Valentina Appietto, Sara Pizzamiglio, Paola Tiberio, Marilena Viasio, Florentine Hober, Evandro de Azambuja, Lorena de la Pena, Miguel Ángel Izquierdo, Jens Huober, José Baselga, Martine Piccart, Filippo G. De Braud, Giovanni Apollone, Paolo Verderio, and Maria G Daidone

DOI: 10.1158/0007-4194.CCR-18-2507
From biomarker discovery ... to biomarker validation

omics-based signature generation

Signature discovery → assay optimization → Signature validation → assay development → clinical application

Biomarker discovery

Clinical utility

Biomarker validation
Biomarker validation (1)

Signature discovery → assay optimization → Signature validation → assay development → clinical application

omics-based signature generation

Subjects classification

- assay performance
- clinical interpretability

Biomarker validation (2)

✓ **Analytical validation**: demonstrates the **robustness** (accuracy, precision, reproducibility) of the test

  *how well does the test measure what it claims to measure?*

✓ **Clinical Validation**: demonstrates the **effectiveness** of the test

  *how relevant is the test measurement to the clinical condition?*

  *is the biomarker related to the clinical endpoint?*
The conclusion that a given use of a medical product will lead to a net improvement in health outcome or provide useful information about diagnosis, treatment, management, or prevention of a disease.

Does use of the biomarker result in patient benefit?
Focus on the analytical validation

Biomarker discovery → Biomarker validation → Clinical utility

Analytical validation:
- Operating procedures setting-up
- Operating procedures standardization
- Internal Quality Control (IQC)
- External Quality Assessment (EQA)

- Definition of operating procedures (OPs) for biomarker determination
- Validation of the Ops in terms of precision and accuracy according to the standards
- Evaluation of the validated standards within laboratory
- Between laboratories comparison and assessment of their accuracy

Verderio P et al. New Biotechnology 2019 (submitted)
External Quality Assessment (EQA)

Purpose of EQAs:
(a) evaluation of laboratory performance for specific tests and its continuous monitoring
(b) identification of inter-laboratory differences
(c) evaluation of method/diagnostic system performances
(d) degree of comparability between methods/diagnostic systems
(e) monitoring of the success of harmonization/standardization efforts for improving results comparability

✓ can be organized in various biomedical research field
✓ can be focused on the pre-analytical, analytical or post-analytical phase of biomarker determination
Dynamicity of the EQA schemes

- State of the art (survey)
- EQA implementation
- Corrective actions
- Identification of critical factors
- Evaluation of the proposed corrective actions
## EQA: scheme design

<table>
<thead>
<tr>
<th>MAIN STEPS</th>
<th>ACTIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>COORDINATING AND MANAGEMENT ACTIVITIES</td>
<td>• principal investigators meeting</td>
</tr>
<tr>
<td>PROTOCOL WRITING</td>
<td>• definition of the: study rationale, operating procedures for testing sample analysis, participants actions, evaluation sheets, statistical analysis plan (SAP)</td>
</tr>
<tr>
<td>SELECTION OF PARTICIPANTS</td>
<td>• definition of participants inclusion/exclusion criteria</td>
</tr>
<tr>
<td>SELECTION AND PRODUCTION OF TESTING SAMPLES</td>
<td>• definition of sample’ selection criteria</td>
</tr>
<tr>
<td></td>
<td>• preparation of testing samples</td>
</tr>
<tr>
<td>CIRCULATION OF TESTING SAMPLES</td>
<td>• management procedures and monitoring</td>
</tr>
<tr>
<td>EVALUATION OF TESTING SAMPLES</td>
<td>• definition of reference value(s)</td>
</tr>
<tr>
<td></td>
<td>• actions by each participants according to the protocol</td>
</tr>
<tr>
<td>DATA COLLECTION</td>
<td>• data management</td>
</tr>
<tr>
<td></td>
<td>• database implementation</td>
</tr>
<tr>
<td>DATA ANALYSIS</td>
<td>• statistical analysis according to the SAP</td>
</tr>
<tr>
<td>DATA REPORTING</td>
<td>• participant’ report</td>
</tr>
<tr>
<td></td>
<td>• summary report</td>
</tr>
<tr>
<td>DISSEMINATION AND TRAINING</td>
<td>• dissemination of the overall results</td>
</tr>
<tr>
<td></td>
<td>• guidelines writing</td>
</tr>
<tr>
<td></td>
<td>• training activity and corrective actions</td>
</tr>
</tbody>
</table>
Our experience in EQAs

“Preanalytical errors still account for nearly 60%-70% of all problems occurring in laboratory diagnostics, most of them attributable to mishandling procedures during collection, handling, preparing or storing the specimens”.


NAs: Nucleic Acids;
CEN/TS: Comité Européen de Normalisation/ Technical Specification
ISO/IS: International Organization for Standardization/ Internation Standrds
Pre-analytical workflow

Sample collection

Sample shipment

Sample arrival

DNA/RNA extraction

Profiling analysis

Type of sample collection

Type of sample shipped

Type of sample storage

Type of extraction method

Type of platform

Adapted from Verderio P. J Clin Oncol. 2012;30:1912-5
Our experience in EQAs

“Preanalytical errors still account for nearly 60%-70% of all problems occurring in laboratory diagnostics, most of them attributable to mishandling procedures during collection, handling, preparing or storing the specimens”.


NAs: Nucleic Acids;
CEN/TS: Comité Européen de Normalisation/ Technical Specification
ISO/IS: International Organization for Standardization/ Internation Stanrds
The EU SPIDIA and SPIDIA4P projects

- **SPIDIA** *(Standardisation and improvement of generic pre-analytical tools and procedures for in-vitro diagnostics)* was a 4.5-year project funded by the EU-FP7 programme (GA n. 222916, 2008-2013)

- **SPIDIA4P** *(Standardisation and improvement of generic pre-analytical tools and procedures for in-vitro diagnostics for personalize medicine)* is a 4-year project funded by the H2020 programme (GA n. 73312, 2017-2020)
  - 19 partners coordinating by QIAGEN
  - built upon SPIDIA’s results
SPIDIA blood EQAs: participant laboratories

2 runs of EQAs for blood-DNA/DNA-plas and blood-RNA

<table>
<thead>
<tr>
<th></th>
<th>SPIDIA – DNA/DNAplas</th>
<th>SPIDIA - RNA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>applicants</td>
<td>effective</td>
</tr>
<tr>
<td>1st EQA</td>
<td>197(a)</td>
<td>183</td>
</tr>
<tr>
<td>2nd EQA</td>
<td>188(b)</td>
<td>174</td>
</tr>
</tbody>
</table>

(a) 130 labs took part to the SPIDIA-DNA, only; 67 labs to both SPIDIA-DNA and SPIDIA-DNAplas EQAs
(b) 127 labs took part to the SPIDIA-DNA, only; 61 labs to both SPIDIA-DNA and SPIDIA-DNAplas EQAs
SPIDIA blood EQAs: some results

Participants were asked to extract RNA from Tube C (RNA C) immediately upon arrival of the tubes and from Tube D (RNA D) 24 h after Tube C. PAXgene Tube D was stored at RT while EDTA Tube D was stored either at RT or at 4°C according to a randomized scheme.

Overall distribution of IL8 according to blood collection tube (panel A) and to storage temperature/collection tube (panel B). The box horizontal sides identify the 25th and 75th centile, the horizontal line inside the box the median, the two whiskers correspond to minimum and maximum, and the dashed line indicates the T0 value zero.
SPIDIA: the achieved goals

Development of **9 CEN Technical Specifications (CEN/TS) for pre-analytical workflows** in Europe, within the CEN/Technical Committee 140 for “In vitro medical devices”:

- CEN/TS 16826-1, snap frozen tissue – Part 1: Isolated RNA
- CEN/TS 16826-2, snap frozen tissue – Part 2: Isolated proteins
- CEN/TS 16827-1, FFPE tissue – Part 1: Isolated RNA
- CEN/TS 16827-2, FFPE tissue – Part 2: Isolated proteins
- CEN/TS 16827-3, FFPE tissue – Part 3: Isolated DNA
- CEN/TS 16835-1, venous whole blood – Part 1: Isolated cellular RNA
- CEN/TS 16835-2, venous whole blood – Part 2: Isolated genomic DNA
- CEN/TS 16835-3, venous whole blood – Part 3: Isolated circulated cell free DNA from plasma
- CEN/TS 16945, metabolomics in urine, serum and plasma

http://www.spidia.eu/publications/technical-specifications/
III. Concluding remarks
It is important in the –omics era to have:

• **reliable** biomarkers easily **translatable** in the clinic

• **guidelines** for the **whole** biomarker determination **process** from sample collection to biomarker detection and analysis

SPIDIA4P aim to implement comprehensive portfolio of 21 pre-analytical CEN/TS to be developed into global ISO/IS documents
THANK YOU!