Paolo Vineis
Imperial College London

Oomics, new biomarkers and “meet-in-the-middle” design in environmental epidemiology

Lyon Canceropole  22 March, 2013
Some environmental exposures can be studied by epidemiology with confidence, i.e. measurement error is relatively low and has little impact on estimates. Advancement in exposure assessment due e.g. to GIS techniques for air pollution.

When measurement error is too high we may need biomarkers (e.g. number of sexual partners, OR for cervical cancer around 2; HPV strains, OR around 100-500).

Some discoveries that support the original model of molecular epidemiology

<table>
<thead>
<tr>
<th>Marker linked to exposure</th>
<th>Exposure</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Internal dose</strong></td>
<td></td>
</tr>
<tr>
<td>Urinary metabolites (NNK, NNN)</td>
<td>Nitrosocompounds in tobacco</td>
</tr>
<tr>
<td><strong>Biologically effective dose</strong></td>
<td></td>
</tr>
<tr>
<td>DNA adducts</td>
<td>PAHs, aromatic compounds</td>
</tr>
<tr>
<td>Albumin adducts</td>
<td>AFB 1</td>
</tr>
<tr>
<td>Hemoglobin adducts</td>
<td>Acrylamide, Styrene, 1,3-Butadiene</td>
</tr>
</tbody>
</table>

Vineis and Perera, 2007
**Exposome - definition**

The exposome concept refers to the totality of environmental exposures from conception onwards. The *internal exposome* is based on measurements in biological material of complete sets of biomarkers of exposure, using *repeated biological samples* especially *during critical life stages*.

Biomarkers which can be measured in this context cover a wide range of molecules, ranging from xenobiotics and their metabolites in blood (*metabolomics*) to covalent complexes with DNA and proteins (*adductomics*).

The term *omics* generally refers to the rigorous study of a complete set of biological and non-biological molecules with high-throughput techniques (Rappaport and Smith 2010).
Serum exposome

DATA-DRIVEN DISCOVERY (EWAS)

KOHNLEDGE-DRIVEN APPLICATIONS

Molecular epidemiology

Exposure biology

Systems biology

Drug development

Causality and prevention

Diagnosis, prognosis and treatment

S. Rappaport, Biomarkers, 2012, 17(6), 48: 3-9
Omics in cohort studies: examples from our research
Manhattan plots for EWAS results for smoking status in two case-control studies (Shenket et al, 2012).
<table>
<thead>
<tr>
<th>Marker</th>
<th>Genomic locus</th>
<th>Smoking status</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>AUC never vs former (95% CI)</th>
<th>PPV</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Never</td>
<td>Former</td>
<td>Current</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Test set, n</strong></td>
<td></td>
<td>33</td>
<td>30</td>
<td>18</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AHRR_p1</td>
<td>Chr5:373,299</td>
<td>79.64</td>
<td>73.33</td>
<td>59.63</td>
<td>0.65</td>
<td>0.67</td>
<td>0.71 (0.56 – 0.86)</td>
</tr>
<tr>
<td>2q37_p1</td>
<td>Chr2.233,284,11</td>
<td>18.96</td>
<td>15.88</td>
<td>18.00</td>
<td>0.67</td>
<td>0.77</td>
<td>0.68 (0.56 – 0.85)</td>
</tr>
<tr>
<td>2q37_p3</td>
<td>Chr2.233,284,66</td>
<td>59.56</td>
<td>53.94</td>
<td>43.45</td>
<td>0.61</td>
<td>0.64</td>
<td>0.66 (0.46 – 0.81)</td>
</tr>
<tr>
<td>6p21.33</td>
<td>Chr6:30,720,080</td>
<td>73.03</td>
<td>66.09</td>
<td>58.45</td>
<td>0.63</td>
<td>0.65</td>
<td>0.63 (0.45 – 0.85)</td>
</tr>
<tr>
<td>Cotinine</td>
<td>n/a</td>
<td>4.96</td>
<td>3.77</td>
<td>980.74</td>
<td>0.04</td>
<td>0.90</td>
<td>0.47 (0.32 – 0.63)</td>
</tr>
<tr>
<td>MI</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>0.69</td>
<td>0.90</td>
<td>0.82 (0.64-0.99)</td>
</tr>
<tr>
<td><strong>Validation set, n</strong></td>
<td></td>
<td>102</td>
<td>45</td>
<td>33</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MI</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>0.71</td>
<td>0.80</td>
<td>0.83 (0.70-0.96)</td>
</tr>
</tbody>
</table>

PPV, positive predictive value using optimum thresholds; 95% CI, 95% confidence intervals; MI, methylation index.
In the last years, research on SES has expanded with the aim of identifying the biological mechanisms through which socioeconomic status is embedded and eventually “gets under the skin”

In humans, low socioeconomic status across the lifecourse has been associated with greater diurnal cortisol production, increased inflammatory activity, higher circulating antibodies for several pathogens (suggesting dampened cell-mediated immune response), reduction in prefrontal cortical grey matter and greater amygdale reactivity to threat, among others. Social and financial adversities over the entire lifespan (and more critically in early life) would program a “defensive” phenotype.

Human and animal studies have shown that socioeconomic status influences DNA Methylation and gene expression, in particular across genome regions regulating the immune function.
Association of household’s highest occupational position with DNA methylation.

Unpublished data from the EPIC-Torino cohort (n=830)(S Stringhini et al)
From Illumina 450K results 16 candidate genes have been selected as being implicated in psycho-social stress. Results after Bonferroni correction.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Probe</th>
<th>N</th>
<th>MODEL 1 + lifestyle factors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>β</td>
</tr>
<tr>
<td>NFATC1</td>
<td>cg22532194</td>
<td>822</td>
<td>-0.78</td>
</tr>
<tr>
<td>NFATC1</td>
<td>cg02763290</td>
<td>822</td>
<td>-0.79</td>
</tr>
<tr>
<td>NFATC1</td>
<td>cg13580107</td>
<td>822</td>
<td>-1.19</td>
</tr>
<tr>
<td>AHRR</td>
<td>cg24688690</td>
<td>822</td>
<td>-1.09</td>
</tr>
<tr>
<td>GPR132</td>
<td>cg20968821</td>
<td>822</td>
<td>-0.99</td>
</tr>
</tbody>
</table>
Dominance rank and expression level of pro-inflammatory genes (macaques)

Incorporation of omics into epidemiological research: challenges in design

1. precious biobanked material, not easily released by PIs

2. ethical issues

3. single (spot) biological samples

4. no cohorts allow life-course epidemiology

5. in-depth exposure assessment is limited by feasibility (for cancer you need large sample sizes)

6. lab measurements and omics have the same limitations related to sample size and feasibility
Experimental Short Term Studies (STS)

**New measurements**
- Oxford Street 2
- TAPAS

**Existing resources**
- Oxford Street 1
- RAPTES

**Outline**
- Volunteers experience contrasting levels of air pollution (high/low) during exposure periods
- PEM measurements made during exposure periods
- Blood samples obtained for omics before and after each exposure period

---

**Diagram**

- PEM
- START Condition 1
- Blood Sampling

- Exposure Analysis

- END Condition 2
- Omics Analysis
Mother-Child Cohort Studies (MCO) and Adult Long-Term Studies (ALTS)

**New measurements**
- INMA
- EPIC-ESCAPE and East Anglia
- SAPALDIA

**Existing resources**
- INMA
- EPIC-ESCAPE and East Anglia
- Rhea
- ALSPAC
- Piccoli+

**Outline**
- Participants in existing cohort studies with stored blood samples will be selected
- PEM measurements will be made during three periods at five sites

![Diagram showing PEM measurements over three periods at five sites](image)
- “Dedicated” smart phone in a pouch on sensor pack to enable user input/output (i.e. we do not intend to “leverage” the user’s personal phone, at this stage).
- Rechargeable Li battery pack supplies power to instruments and smartphone via USB hub for 36 hr autonomy.
- Each Sensor and Smartphone log data independently (synched in time during initial setup).
The “meet-in-the-middle” approach

Schematic representation of the implementation of the ‘meet-in-the-middle’ approach (Chadeau-Hyam et al, Biomarkers 2011).
Study on metabolomics in breast and colon cancer nested in EPIC-Italy

No markers found in association with breast cancer, 8 signals found in association with colon cancer (Chadeau-Hyam et al., 2010)

**Dietary fibers** intake was found to be associated to four putative markers out of 235 (with corresponding p-values ranging from 0.003 to 0.02).

One marker indicates a possible link with gut microbial fermentation of plant phenolics in the colon (Nicholson et al., 2005, Phipps et al., 1998, Aura, 2007), a process also plausibly linked to higher dietary fibers exposure and lower colon cancer risk.
Human metabolic phenotype diversity and its association with diet and blood pressure

Elaine Holmes¹*, Ruey Leng Loo¹,²*, Jeremiah Stamler³, Magda Bictash¹,², Ivan K. S. Yap¹,², Queenie Chan², Tim Ebbels¹, Maria De Iorio², Ian J. Brown², Kirill A. Veselkov¹, Martha L. Daviglus⁵, Hugo Kesteloot⁴, Hirotsugu Ueshima⁵, Liancheng Zhao⁶, Jeremy K. Nicholson¹ & Paul Elliott²

Supervised analysis defined urinary profiles robustly associated with known dietary factors

(Nature 2008)
The end