Public Health e-Labs for a Global Digital Economy

Link to publication record in Manchester Research Explorer

Citation for published version (APA):

Citing this paper
Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

General rights
Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Takedown policy
If you believe that this document breaches copyright please refer to the University of Manchester's Takedown Procedures [http://man.ac.uk/04Y6Bo] or contact uml.scholarlycommunications@manchester.ac.uk providing relevant details, so we can investigate your claim.
Public Health ‘e-Labs’
for a Global Digital Economy

WHO-PHI 2008 (Asia-Pacific), Delhi, 3-4 Nov 2008

Prof. Iain Buchan
University of Manchester
PHI Targets

• Digital Deserts
  Building e-readiness for the public’s health

• Digital Dust
  Turning digital commodities into actions for the public’s health
Situational Awareness of Rising Child-BMI: Example Wirral 3-yr-olds from 1988 to 2004

Month of measurement by Health Visitor

SDS = standard deviation score from 1990 British Growth Reference charts – adjusts for age and sex of the child
Secular trend to increasing BMI is much greater in taller children

Source: Buchan et al. 2006
Health data-silo anthropology

‘data-tombs’...
Digital Dust (data deposit > use)
Public Health Info-economics

Problems with Public Health Information

– Too little
– Too late
– Can’t find it
– Can’t reproduce it
– Consumes more resource than it needs to
– Benefits invisible to healthcare providers
– Cost savings not measured
Cloud of millions of messages in the local health economy

Structured Data

Organise

Transform & Examine

Structured Data & Metadata
Patient-driven information into records

Person-identifiable and sensitive information removed

Real-time

Data Repository in PCT

24-hourly updates

Anonymised Data Repository in PCT

Trusted person poses question(s)

FIREWALL

Optometrist
Eye screening
Community nurses
Podiatry

Link on NHS number

Deaths, Demographics etc.

Outputs

Biomics Data
What is an e-Lab?

...an information system bringing together data, analytical methods and people for timely, high-quality decision-making
Clinical audit question: “is diabetes care picking up enough treatable anaemia in patients with mild kidney impairment?”

Answer: No

Care pathway improvements

Next similar e-Lab query made easier

Deeper research...
Anaemia at lower levels of kidney impairment than commonly thought

Clinical (audit) questions leading to scientific findings: supporting sustainable healthcare-academic partnership

Anaemia at lower levels of kidney impairment than commonly thought
Dataset ➔ Digital Commodity

Serving health communities with high-quality health intelligence requires **metadata** from **local uses**...
Excellent research by-products of excellent service development

Local NHS → e-Lab → Research Networks → Research

Service development

Federation of e-Labs → scalable & sustainable
Summarising care quality

Care improvement or case-mix change?

Total Cholesterol

Systolic BP

1993 1995 1997 1999 2001 2003 2005 2007

136 138 140 142 144 146 148 150

136 138 140 142 144 146 148 150
Developing models and software to make complex scenarios easy to explore in real time → democratise commissioning?

Outputs: Population-based incidence, prevalence; Deaths prevented; Life-Years; Life expectancy; Costs; Cost-effectiveness ratios
Increasing Expectation of Models

• Research
  – Multi-level stochastic
  – Machine-learning
    • Omics
    • Image analysis

• Service-development
  – Graphical models & discrete event simulations

• Clinical & self-care decision support?
Crude Pan-Genome Scans

for( i = 1 to #random permutations)
{
    for( j = 1 to #SNPs)
    {
        for( k = 1 to #patients)
        {
            disease status vs. locus status $\chi^2$
        }
    }
}

Given a typical 5k patients, 0.5m SNPs and 10k permutations:
20k $\chi^2$ calcs per sec on modern single core $\Rightarrow$ 70 hrs single SNPs;
$\Rightarrow \approx 1,980$ years for $[n*(n-1)]/2$ SNP pairs
Computational free-thinking, for insights from richly-observed health & environments.
...the e-Research Digital Economy
Obesity Attributable Cancers

• What is & will be the obesity-attributable cancer burden?

• Setting: 30 countries

• Inputs needed:
  – site- and sex-specific cancer risk data
  – standardised risk estimation by site
  – sex- and age-specific risk exposure data (present & past)
  – up-to-date cancer incidence
  – trends in cancer numbers & population demographics

Thanks to: Andrew Renehan
Localising Evidence Needs PHI

<table>
<thead>
<tr>
<th>Cancer site and type</th>
<th>Number of studies</th>
<th>RR (95% CI)</th>
<th>p</th>
<th>p*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oesophagus and stomach</td>
<td>5</td>
<td>1.67 (1.32-2.07)</td>
<td>&lt;0.0001</td>
<td>24%</td>
</tr>
<tr>
<td>Thyroid</td>
<td>4</td>
<td>1.33 (0.96-1.86)</td>
<td>0.02</td>
<td>77%</td>
</tr>
<tr>
<td>Colon</td>
<td>22</td>
<td>1.24 (0.80-2.02)</td>
<td>&lt;0.0001</td>
<td>22%</td>
</tr>
<tr>
<td>Breast</td>
<td>11</td>
<td>1.24 (0.95-1.62)</td>
<td>&lt;0.0001</td>
<td>32%</td>
</tr>
<tr>
<td>Liver</td>
<td>4</td>
<td>1.24 (0.95-1.62)</td>
<td>0.32</td>
<td>85%</td>
</tr>
<tr>
<td>Male genital carcinomas</td>
<td>6</td>
<td>1.37 (1.05-1.80)</td>
<td>0.004</td>
<td>48%</td>
</tr>
<tr>
<td>Multiple myeloma</td>
<td>7</td>
<td>1.33 (1.04-1.71)</td>
<td>&lt;0.0001</td>
<td>7%</td>
</tr>
<tr>
<td>Rectum</td>
<td>18</td>
<td>1.09 (0.76-1.12)</td>
<td>&lt;0.0001</td>
<td>3%</td>
</tr>
<tr>
<td>Gallbladder</td>
<td>4</td>
<td>1.09 (0.54-2.12)</td>
<td>0.72</td>
<td>0%</td>
</tr>
<tr>
<td>Leukaemia</td>
<td>7</td>
<td>1.08 (0.76-1.40)</td>
<td>0.009</td>
<td>0%</td>
</tr>
<tr>
<td>Prostate</td>
<td>12</td>
<td>1.09 (0.99-1.20)</td>
<td>0.05</td>
<td>70%</td>
</tr>
<tr>
<td>Non-Hodgkin lymphoma</td>
<td>6</td>
<td>1.06 (0.96-1.17)</td>
<td>0.002</td>
<td>9%</td>
</tr>
<tr>
<td>Prostate</td>
<td>27</td>
<td>1.00 (0.99-1.00)</td>
<td>0.11</td>
<td>71%</td>
</tr>
<tr>
<td>Gastrointestinal</td>
<td>8</td>
<td>0.90 (0.80-1.00)</td>
<td>0.49</td>
<td>35%</td>
</tr>
<tr>
<td>Lung</td>
<td>11</td>
<td>0.76 (0.60-0.98)</td>
<td>&lt;0.0001</td>
<td>63%</td>
</tr>
<tr>
<td>Oesophagus and stomach</td>
<td>5</td>
<td>0.71 (0.56-0.85)</td>
<td>&lt;0.0001</td>
<td>42%</td>
</tr>
</tbody>
</table>

Figure 3: Summary risk estimates by cancer sites in men

WHO Infobase
GloboCan

Future Population
Impact Numbers

Current Population
Impact Numbers

Risk exposure trends
Tumour registries

Interpretation & Report
Meta-analysis
Systematic review
Protocol

Rising complexity & computational cost
Safety Blind-spot: Tamoxifen

• Question: Is there a substantial burden of recurrent breast cancer due to interaction of tamoxifen with anti-depressants?
  – Plausible CyP450-2D6 competition (tam → end-oxifen)

• Blind-spots (missing from registers)
  – Recurrent cancers
  – Adjuvant therapies
  – Concurrent therapies
1. User logs on and submits query
2. Access control module authorizes request
3. Broker performs distributed query; 4. generate pseudonym keys
5. Per request keyed pseudonymisation
6. Data integration
7. Anonymisation and inference control
8. Storage
9. Data analysis and visualization

Salford PCT e-Lab

NWeH e-Lab Federation

NWCIS e-Lab

E-Lab Repository

3 & 4 & 6 Broker

5. Per request keyed pseudonymisation

NWCIS Registry

5. Per request keyed pseudonymisation

8 Research Object Repository

Access Control
e-Lab Anatomy is Simple

e-Lab = community + work objects + methods for building work objects

A research object is a story about an investigation.

A decision object is a critical mass of evidence to support a decision.
e-Lab Activity at Manchester

• >100 person years of activity planned for next 3-5 years
  – Healthcare and Public Health
    • North West e-Health: 19 fte to 2012
    • Care Pathway Simulators: 6 fte to 2013
    • Obesity e-Lab: 3 fte to 2011
  – Biology, Chemistry, Social Science, other...
    • Taverna, myGrid & myExperiment: 16 fte 2012

• Ethos
  – Use open-standards, service oriented arch., simple APIs
  – All software freely available in open source
  – Contribute to & learn from global family of innovation
Open Source Projects
Sustained by the Value they Add
through Crowd-Wisdom
+ Cloud Resources Shared

Care ↔ Service Development ↔ Research

e-Lab: Sense-Making Layer

Standards-based Health Information Systems

Powerful Models
Agile Communities
Conclusion

Vision: Global Network of e-Health e-Labs

– Sharing data, expertise & computational resources
– Free, open-source sense-making layer built on top of standards-based healthcare IT

– Innovation is local
– Inspiration is global
– Let’s keep talking