

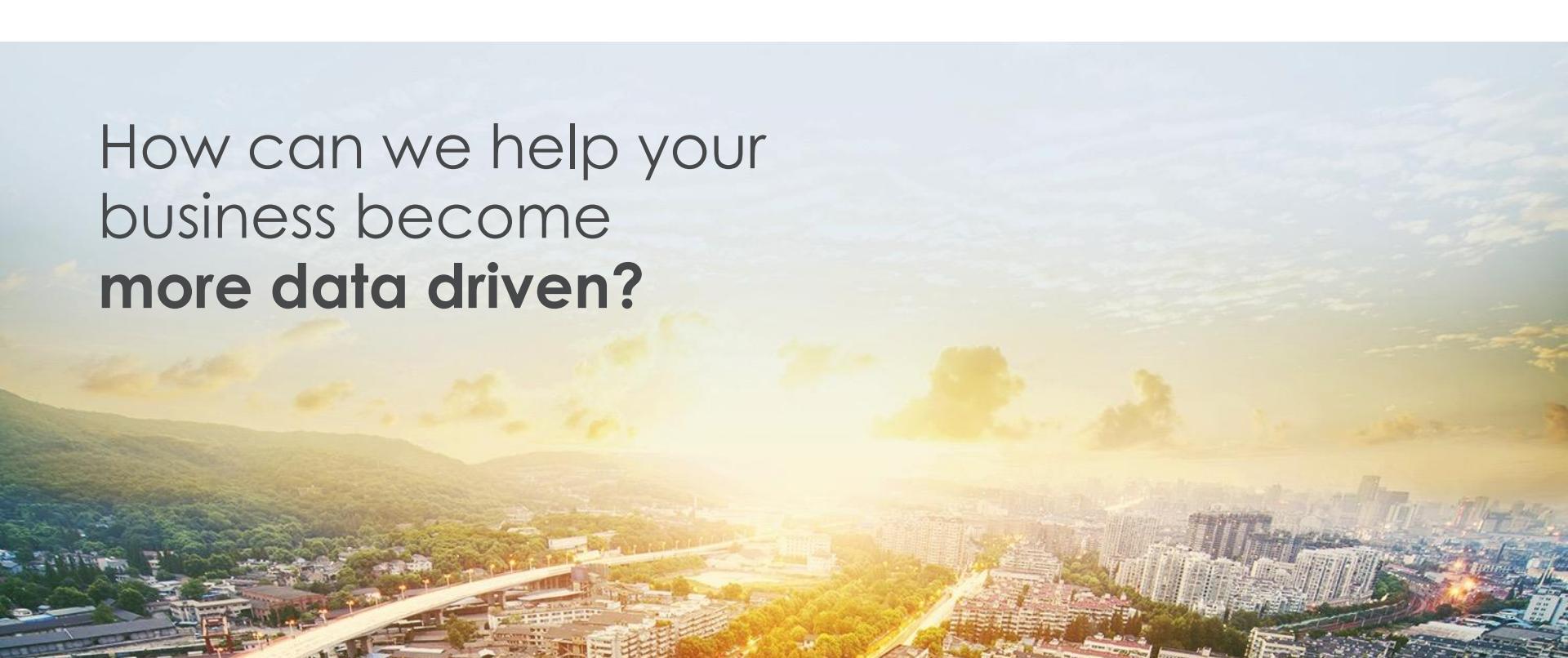
## Big Data: Big Health?

# Realität und Zukunftsvisionen von Big Data im Gesundheitssystem

Peter Grolimund, Senior Industry Consultant

3. März 2016 Bern SAMW





# How can we help your business become **more data driven?**

## **Corporate Vision**

Enabling data-driven  
business

## **Mission**

Providing the world's best analytic data solutions  
to drive competitive advantage for our customers



# US based Healthcare Customers



Anthem



Health Care Service Corporation,  
a Mutual Legal Reserve Company



blue of California



BlueCross BlueShield of Kansas City



Independence  
Blue Cross



GroupHealth

GEISINGER



Horizon

Horizon Blue Cross Blue Shield of New Jersey



TRUVEN  
HEALTH ANALYTICS

More Than Data.  
Answers.

BlueCross  
BlueShield  
Minnesota

Wake Forest University Baptist  
MEDICAL CENTER

TERADATA

Eidgenössischer Datenschützer

## Die Wahl, die keine ist

### Die wichtigsten Herausforderungen beim Datenschutz

dgy. · Daten sind Gold wert – für die Wirtschaft, für die öffentliche Hand, für die Polizei, im Gesundheitsbereich. Das sind die grossen Herausforderungen für den künftigen Datenschützer:

► Big Data: Daten werden zum Rohstoff. Sie können in Zukunft nicht mehr nur zweckbestimmt gespeichert, sondern in grossen Mengen aus verschiedenen Quellen verknüpft und ausgewertet werden – zum Beispiel für einen effizienten und sicheren Strassenverkehr. Big Data stellt dabei ein grosses Wirtschaftspotenzial dar, bedroht aber die Privatsphäre, wenn Informationen aus verschiedenen Lebensbereichen systematisch strukturiert, gesammelt und ausgewertet werden.

► Open Data: Die Behörden verfügen über wertvolle Datenbestände, die der Öffentlichkeit zugänglich gemacht werden sollen. Persönliche Daten gehören zwar typischerweise nicht dazu, doch wenn Angaben beispielsweise über die Nutzung von ÖV-Angeboten verknüpft werden, kann unter Umständen ein Problem entstehen.

Datenschutz und Krebsregistergesetz

### Krebserkrankungen wirksamer behandeln

Um aussagekräftige Vergleiche zwischen den Kliniken machen zu können, braucht es auch einzelne Auswertungen in nichtanonymisierter Form.

Gastkommentar | von Pascal Strupler, BAG | 2.3.2016, 05:30 Uhr

Diese Zustimmungsregelung kritisiert Peter Schumacher in seinem Gastkommentar in der NZZ (25. 2. 16) mit dem Verweis auf das im Krankenversicherungsgesetz verankerte Gebot der regelmässigen Überprüfung der Wirksamkeit, Wirtschaftlichkeit und Zweckmässigkeit (WZW) von medizinischen Leistungen. Dabei geht er fälschlicherweise davon aus, dass das Einverständnis der behandelnden Institutionen für jede Art von Datenerhebung und Datenauswertung nötig ist. Dem ist aber nicht so: Es braucht diese Einwilligung nur, wenn die Daten bezüglich Leistungserbringer in nicht anonymisierter Form ausgewertet werden sollen – und das ist, wie oben beschrieben, nur in Ausnahmefällen nötig.

NEWS

## Big Data Project Takes Aim At Pediatric Cancer

by JOSEPH GOEDERT  
MAR 1, 2016 6:30am ET

information  
management

Over time, the platform will support research on other rare diseases and bring in data from other sources, such as electronic health records and medical imaging. The plan is to not only make data available to researchers worldwide, but to physicians, and give patients access to their own data, Resnick says.

# A BIG DATA CURE?

**US Healthcare Costs = \$2.9 Trillion**

**17.4% of GDP**

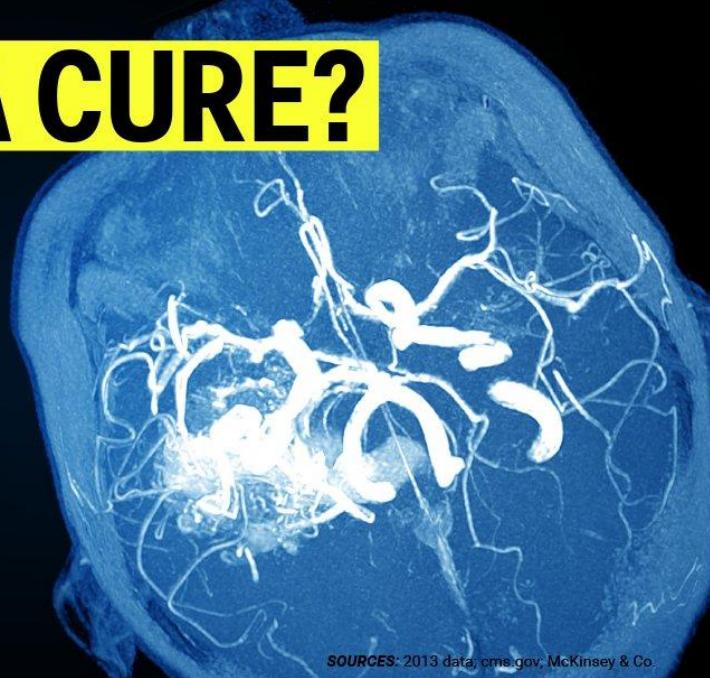
**\$9,255 per person**

## PROMISES

Savings of up to \$1B a year.  
Shift toward "evidence-based" treatment.

## OBSTACLES

Privacy concerns.  
Resistance from patients and doctors.



SOURCES: 2013 data; cms.gov; McKinsey & Co.

The machines we do have..

The first nuggets as well..

But what now?



**MINING**



There's a gold rush on and it's all about the data. Data miners are scrambling to extra value from mountains of data. But like the 1849 gold rush in California, the

# Operational system vs. an analytical platform



Patient Dossier

Visit

Diagnosis

Treatment

## Main purpose of operational systems

Efficient support of the process from admission, diagnostic, therapy until the release or admission of the patient (handling the electronic health records over the patient journey such the effort gets minimized)



Not in scope  
for TD



Patient Dossiers

Epidemiology

Finance data

Biomarker

Reference data

Scientific Data

Public data



## Scope of an analytical Platform

Analytics of data for

- Research
- Decision support
- Safety / Quality
- Process tracking and optimization
- Performance analytics (costs, etc.)
- Outcome

Sweet spot  
for TD

## Real World Data Analytics

# Internet of everything....



## Wearable Motion Detector: Wrist Device Helps Patients & Doctors Track Symptoms of Parkinson's Disease

Posted August 25, 2015 - Add Comment



## FDA approves use of electronic chips in medications

August 1, 2012 by Bob Yirka

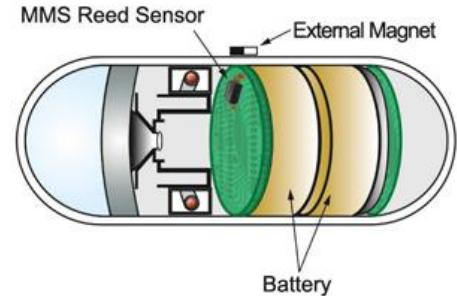


A novel method to measure tremor in daily life

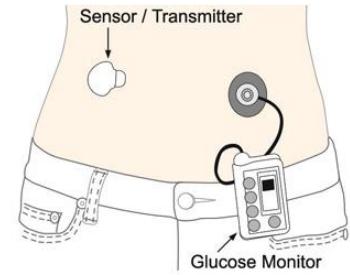
The Skin Analytics website features a blue header with the company logo and a sub-header: "Knowing when a mole has changed can save your life". Below this, a sub-sub-header reads: "Skin Analytics lets you track your moles for change online". Two buttons are present: "Buy Personal \$12.50/quarter" and "Buy Family \$20/quarter". A section titled "What do we do?" includes three items: "Baseline you and your family", "Track changes over time", and "Share your history with a GP". Logos for BBC, The Telegraph, SMART, London Business School, Innovations, and the Jacobs Institute of Innovation and Entrepreneurship are displayed. A close-up image shows a small, thin, flexible sensor device being held between fingers.

## Micro-machining technology: a primer

Mon, 10/17/2011 - 1:17pm by John Beigel



**Figure 6.** An ingestible pill that videos the entry into the stomach and the small intestine as it moves through the digestive tract



**Figure 5.** A glucose monitoring system worn at the belt level



TERADATA

# Open data

## ClinicalTrials.gov

A service of the U.S. National Institutes of Health

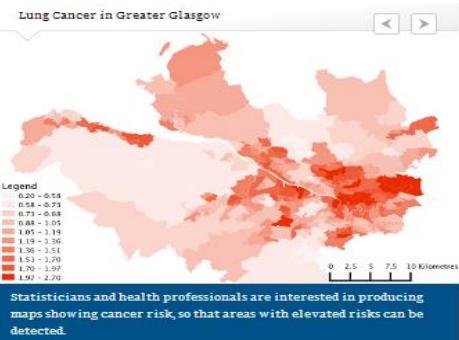
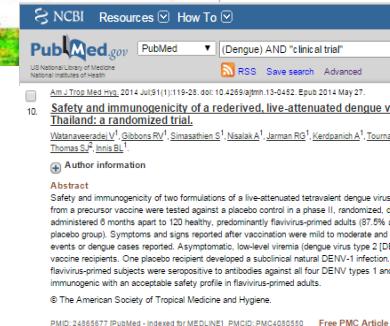
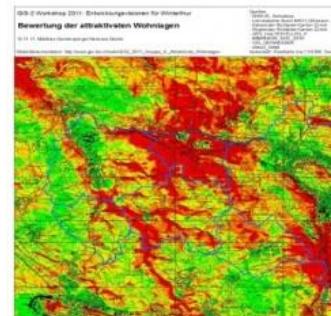
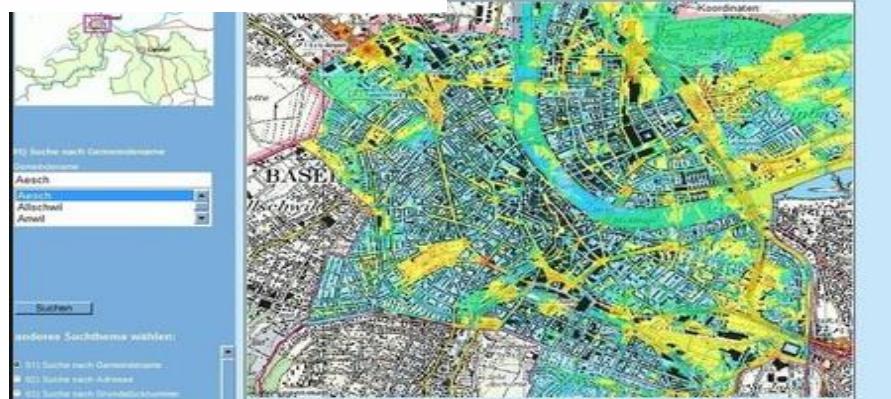


Schweizerische Eidgenossenschaft  
Confédération suisse  
Confederazione Svizzera  
Confederaziun svizra



MORE DIMENSIONS TO DATA

Bundesamt für Meteorologie und  
Klimatologie MeteoSchweiz



eurostat  
Ihr Schlüssel zur europäischen Statistik

INFORMATIONEN ZU EINEM BESTIMMTEN THEMA

Allgemeine und Regionalstatistiken

Wirtschaft und Finanzen

Industrie, Handel und  
Dienstleistungen

Landwirtschaft und Fischerei

Verkehr

Umwelt und Energie

Bevölkerung und soziale  
Bedingungen

Internationaler Handel

Wissenschaft und Technologie

Stadt Zürich  
Wasserversorgung

Qualitätsüberwachung  
Hausberg 9  
Postfach 1179  
CH-8021 Zürich

Telefon +41 (0)44 415 21 11  
Fax +41 (0)44 415 25 57  
[www.zuercherwasser.ch](http://www.zuercherwasser.ch)

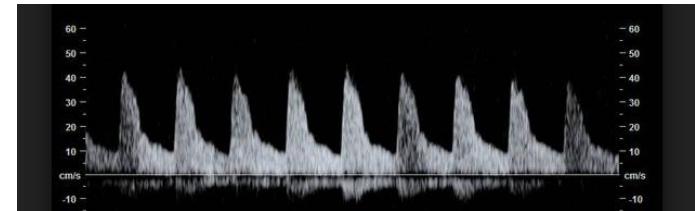
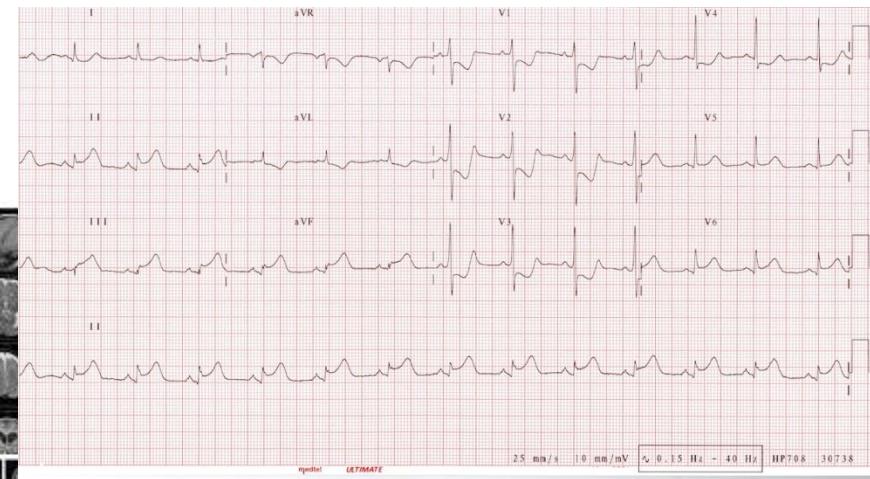
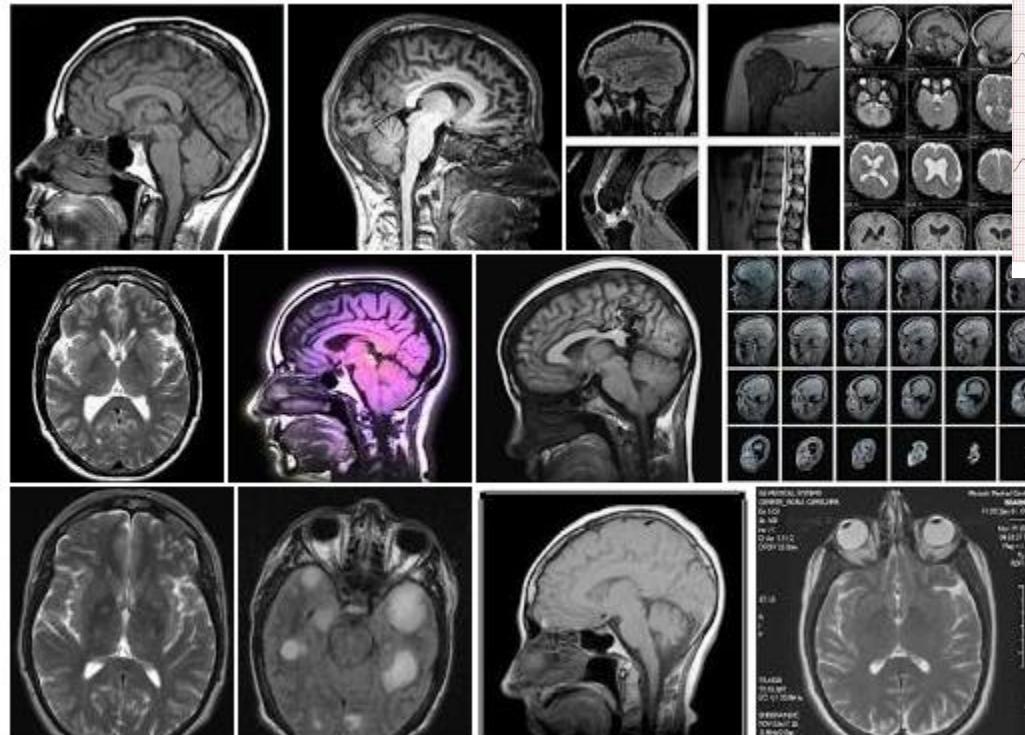
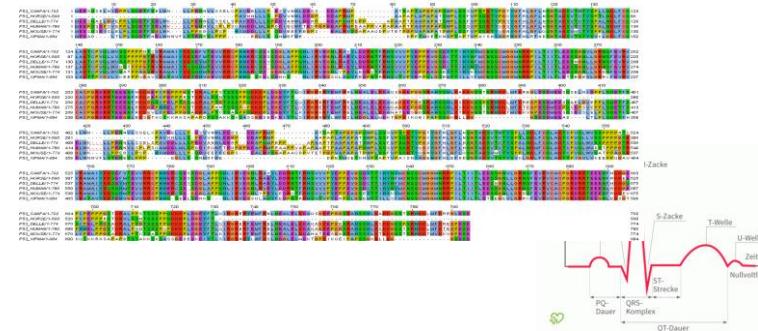
Die Qualität des Zürcher Trinkwassers 2012

1. Allgemeine und anorganische Parameter

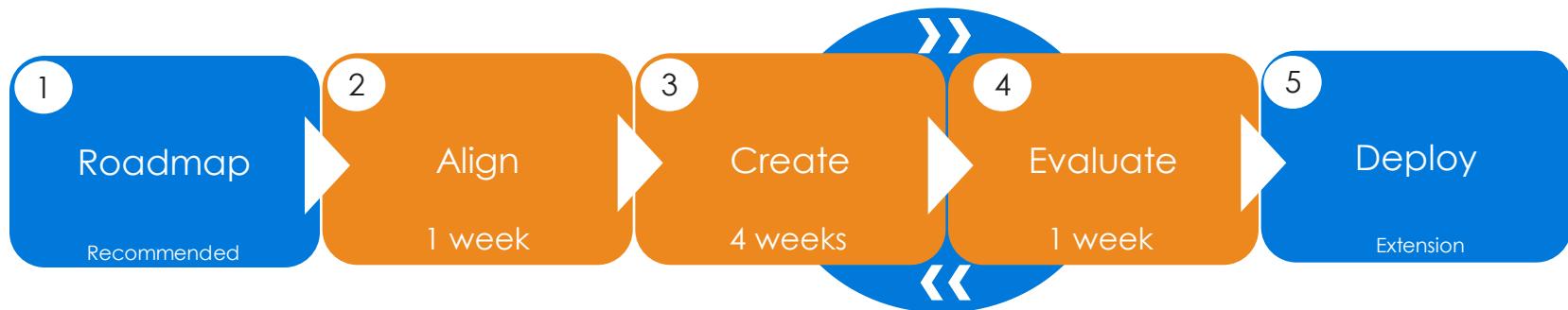
| Parameter               | Einheit | Mittel | Min  | Max  | EW      | TW | GW |
|-------------------------|---------|--------|------|------|---------|----|----|
| Temperatur, °C          | NTU     | 8.9    | 4.2  | 17.0 | 8-15    |    |    |
| Redox-Potential         | mV      | 578    | 362  | 513  |         |    |    |
| pH                      |         | 7.01   | 7.75 | 8.29 | 6-8.2   |    |    |
| Gleichgewichts-pH       |         | 7.73   | 7.60 | 7.90 |         |    |    |
| Leitfähigkeit bei 20 °C | µS/cm   | 297    | 232  | 369  | 200-400 |    |    |
| Eisen, Mangan, Kupfer   | µg/l    | 2.0    | 0.78 | 40.3 | 0-31    |    |    |

TERADATA

# Raw data



# Process

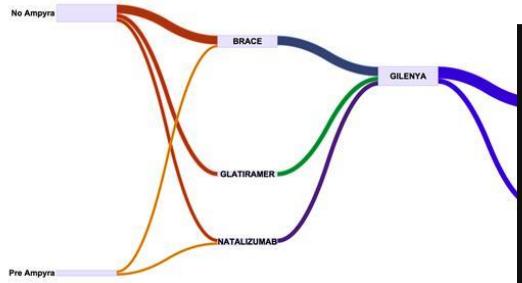


Today

Tomorrow

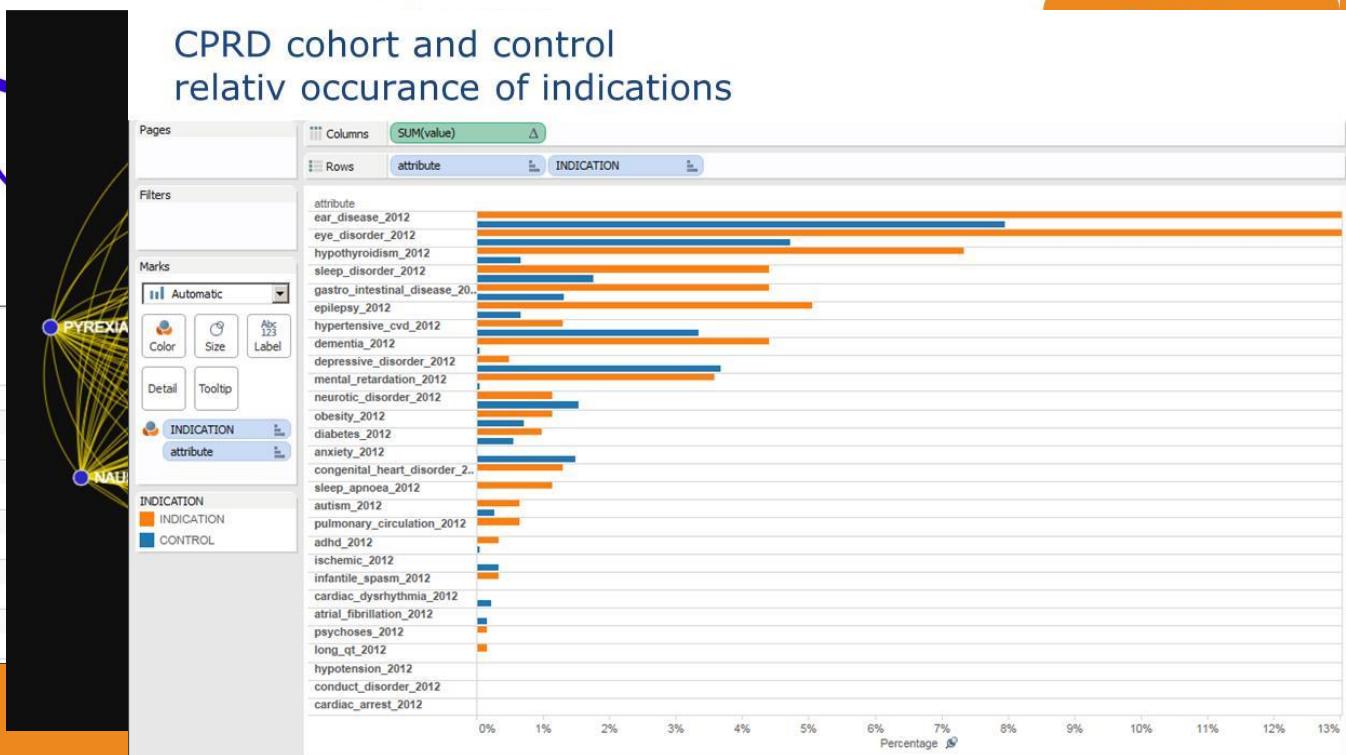
After Tomorrow

# Today



## Comparing number of pre-index relapses with number of post-index relapses by age and gender

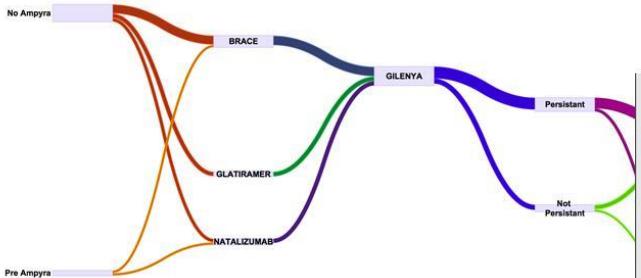
| age (bin) | der_sex_dec.. |
|-----------|---------------|
| 10        | female        |
| 20        | female        |
|           | male          |
| 30        | female        |
|           | male          |
| 40        | female        |
|           | male          |
| 50        | female        |
|           | male          |
| 60        | female        |
|           | male          |



# Real World Evidence

- Modeling clinical trials for minimizing number of patients and feasibility check
  - Pricing calculation and justification
  - HTA assessment and continues monitoring
  - Epidemiology
  - Co-medications and disease know-how

## CPRD cohort and control relativ occurrence of indications



Comparing number of pre-index relapses with number of post-index relapses by age and gender

| age (bin) | der_sex_dec.. | Fewer |          |
|-----------|---------------|-------|----------|
|           |               | Count | % of Tot |
| 10        | female        | 1.0   | 100.00   |
| 20        | female        | 23.0  | 82.14    |
|           | male          | 6.0   | 85.71    |
| 30        | female        | 75.0  | 88.24    |
|           | male          | 13.0  | 86.67    |
| 40        | female        | 116.0 | 87.88    |
|           | male          | 22.0  | 70.97    |
| 50        | female        | 79.0  | 91.86    |
|           | male          | 39.0  | 88.64    |
| 60        | female        | 15.0  | 75.00    |
|           | male          | 6.0   | 85.71    |

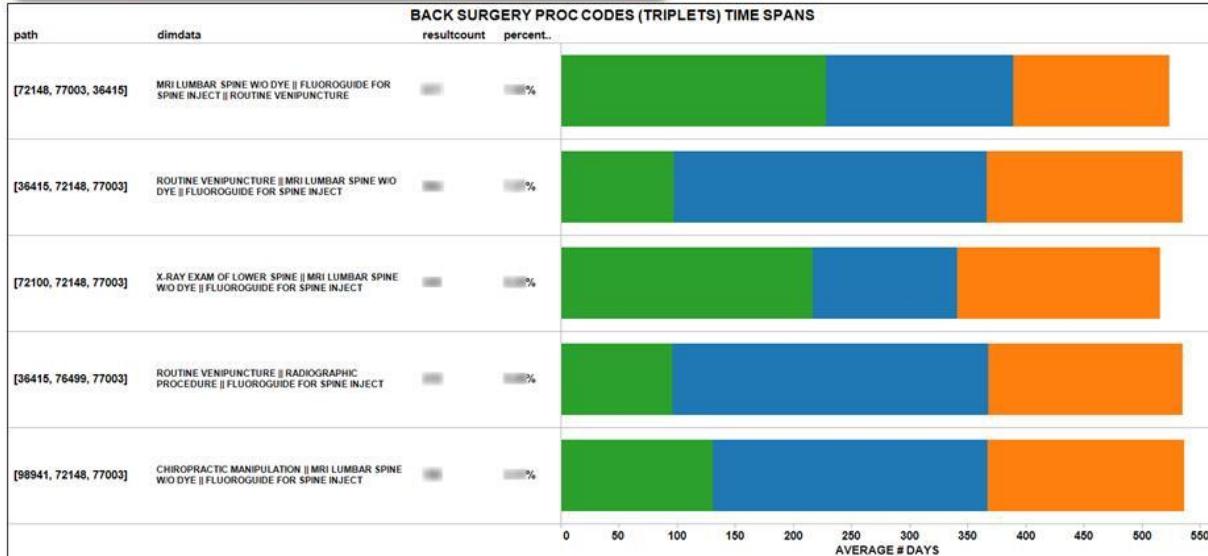


## Real World Evidence

- Modeling clinical trials for minimizing number of patients and feasibility check
- Pricing calculation and justification
- HTA assessment and continues monitoring
- Epidemiology
- Co-medications and disease know-how

# Today

- STEP 1: FIND ALL POSSIBLE PATHS
- STEP 2: GENERATE PATH STATISTICS
- STEP 3: REVIEW AND VISUALIZE THE RESULTS
- STEP 4: USE RESULTS TO GENERATE TARGET LIST



## SUMMARY

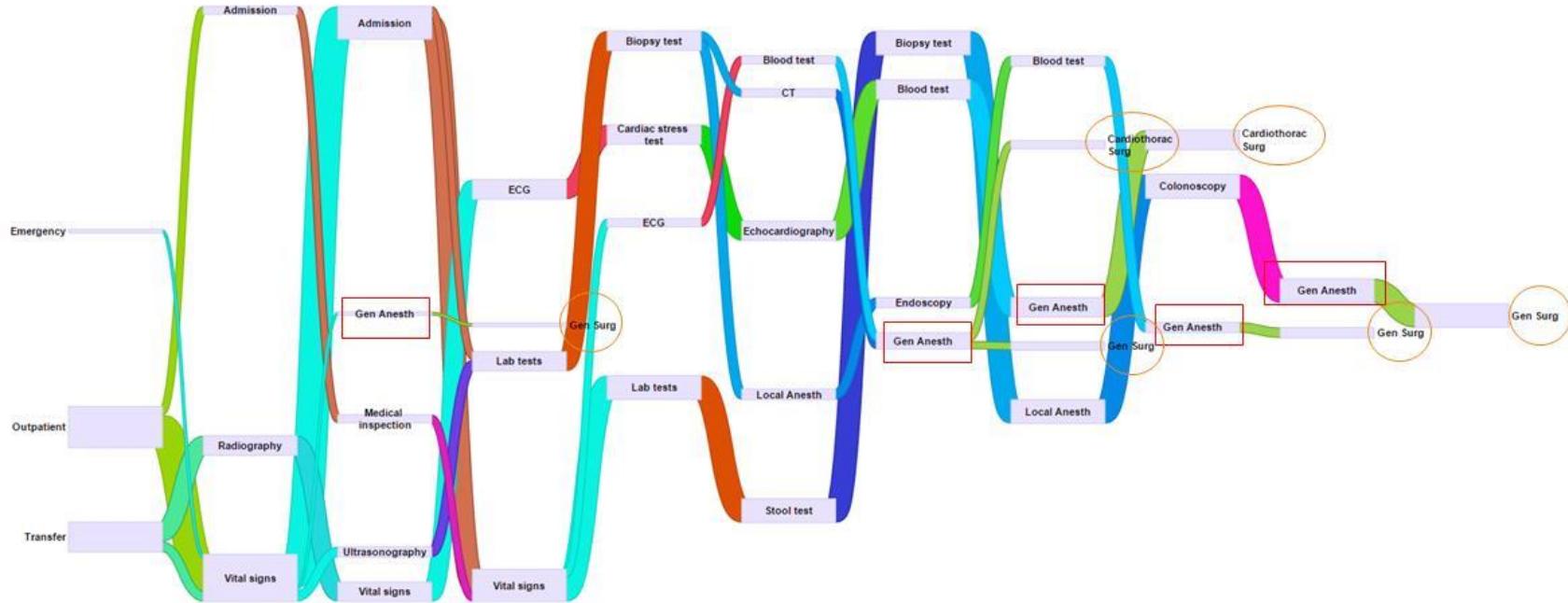
- In the visualization above, the **GREEN** represents the average number of days from the first recorded visit to the beginning of the pattern, the **BLUE** represents the average number of days from the beginning of the pattern to the end of the pattern and the **ORANGE** represents the average number of days from the end of the pattern to the date of the surgical procedure.

## Optimization of treatments

- Process related overview (path to surgery)
- Non-surgical approach (path to avoid surgery)

## PATH TO SURGERY

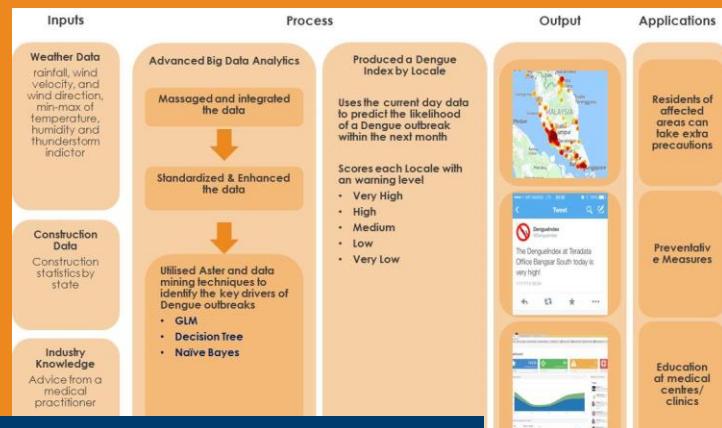
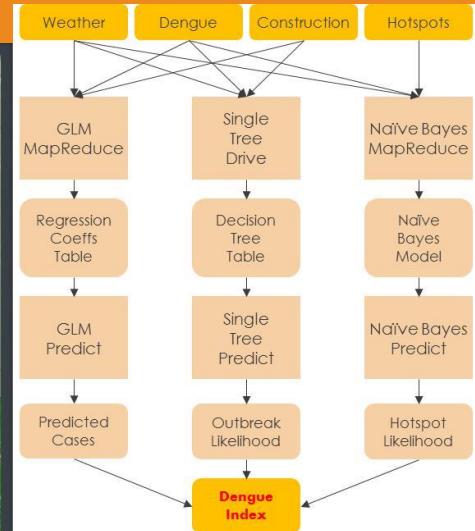
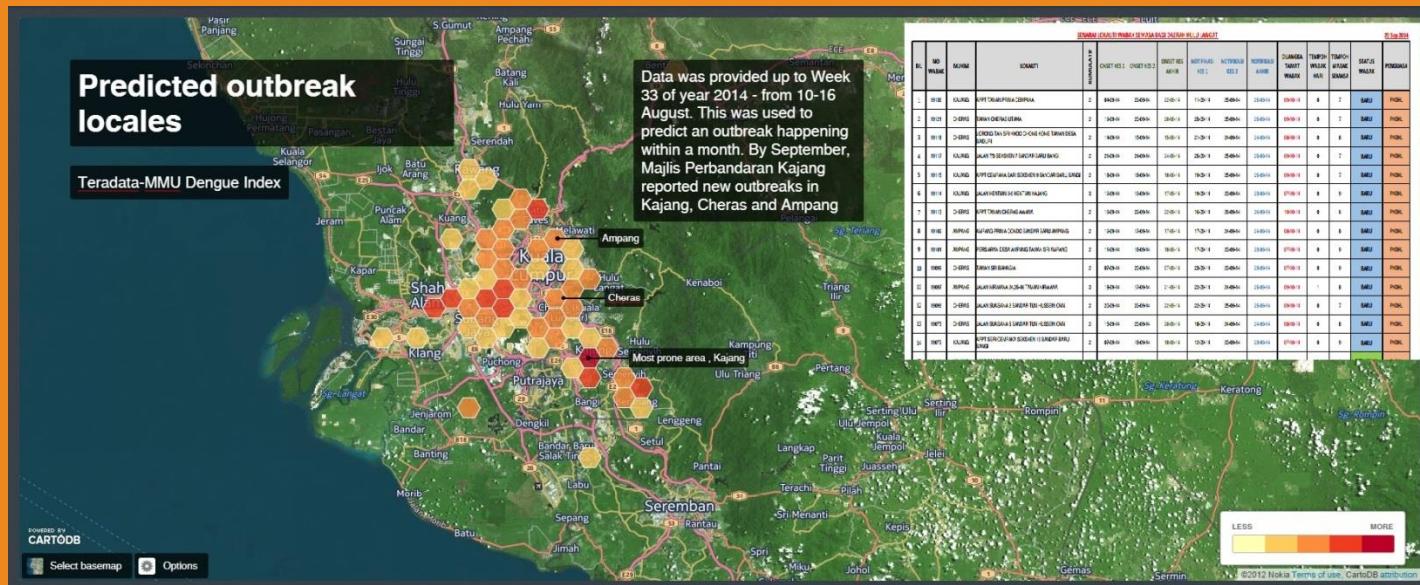
- Path to surgery Sankey Graph generated by Aster nPath Functions



Optimization of treatments

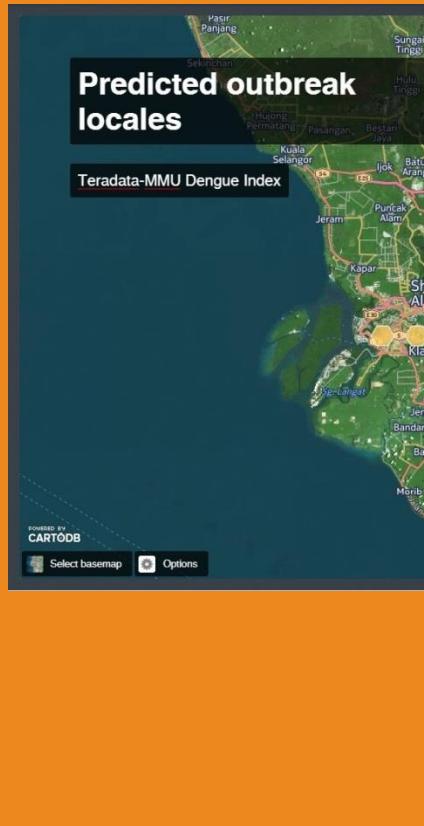
- Process related overview (path to surgery)
- Non-surgical approach (path to avoid surgery)

Today



# Dengue Fever and Flu

- Forecast and actions in Malaysia
  - Planned extension in Pakistan
  - Pharma-customer to drive logistics of new drug
  - Tracking of mutations and the source



## H1N1 PANDEMIC

### PATIENT ZERO ANALYSIS

H1N1 DNA - Patient Zero

AGTCTCTGAAG

Aster IdentityMatch Function

H1N1 DNA - Patient 1

AGTCTCTGAAG

H1N1 DNA - Patient 2

AGTCTCTGAAG

H1N1 DNA - Patient 3

AGTCTCTGAAG

H1N1 DNA - Patient 4

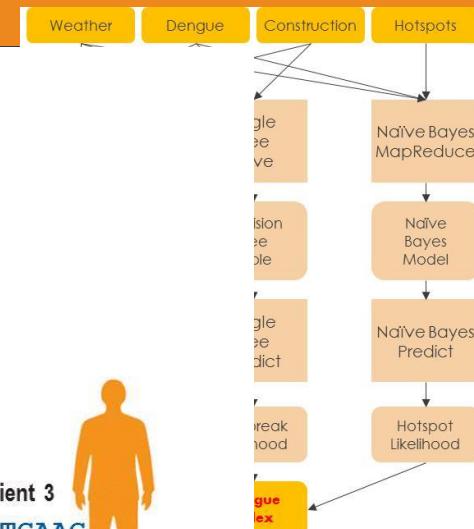
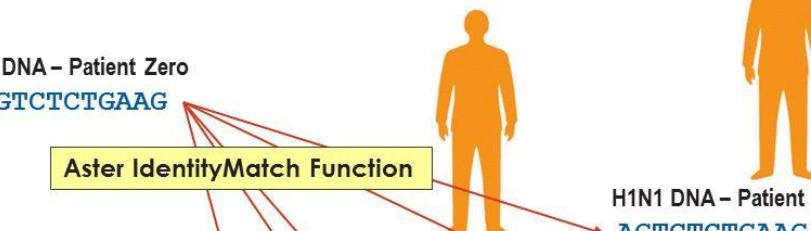
AGTCTCTGAAG

H1N1 DNA - Patient *n*

AGTCTCTGAAG

Patient Zero

The first person to  
Have contracted the virus



## Dengue Fever and Flu

- Forecast and actions in Malaysia
- Planned extension in Pakistan
- Pharma-customer to drive logistics of new drug
- Tracking of mutations and the source

Construction Data

Construction statistics by state

Industry Knowledge

Advice from a medical practitioner

Utilised Aster and data mining techniques to identify the key drivers of Dengue outbreaks

- GLM
- Decision Tree
- Naïve Bayes

High

Medium

Low

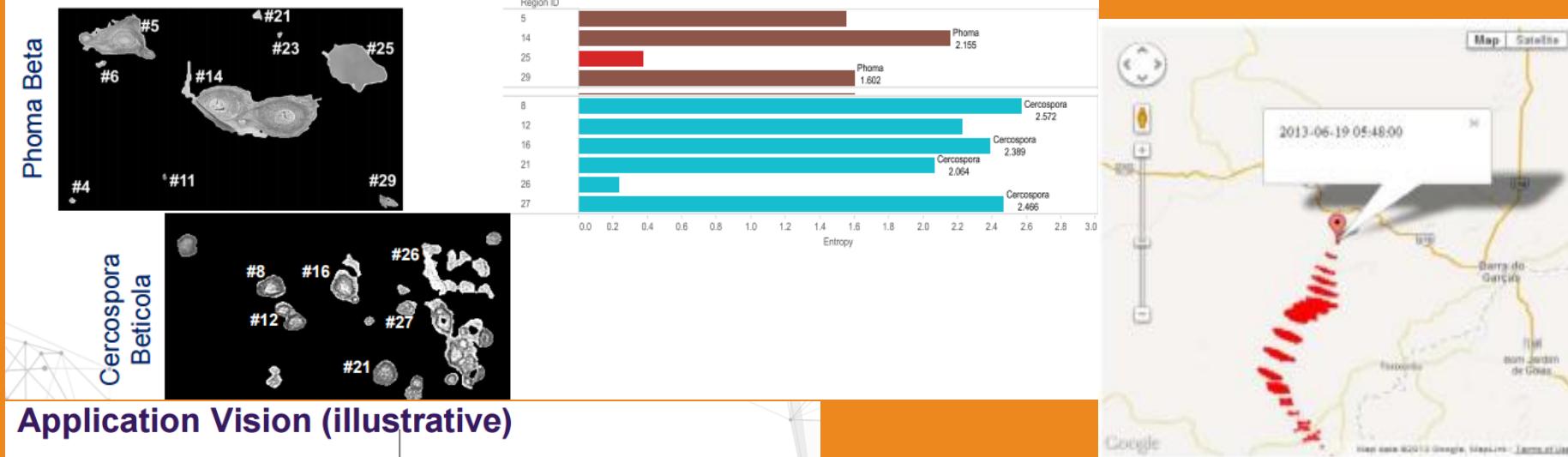
Very Low

Residents of affected areas can take extra precautions

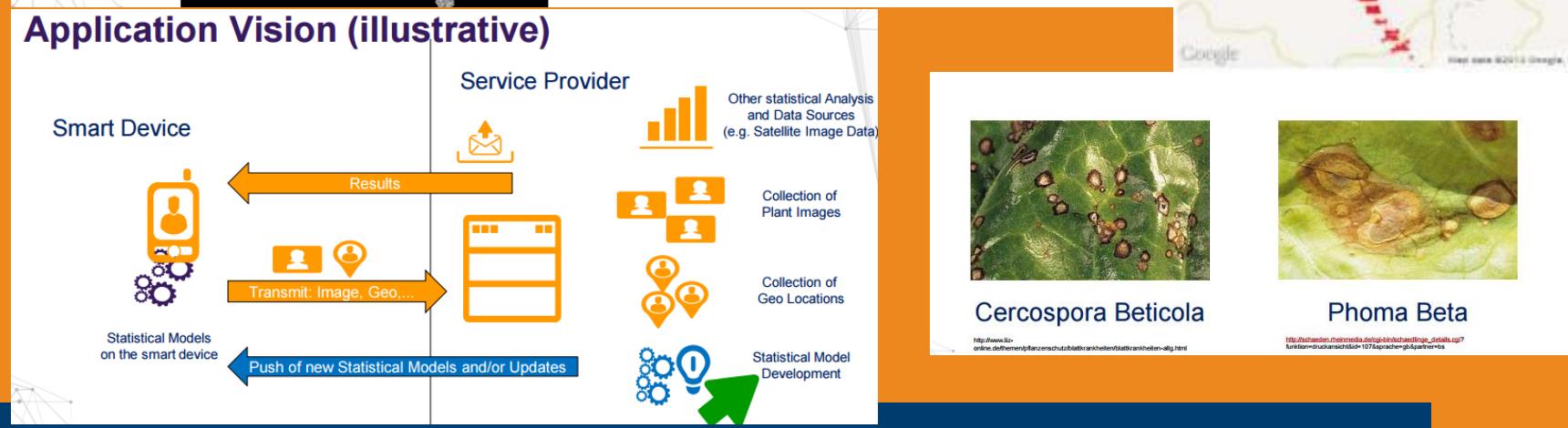
Preventive Measures

Education at medical centres/clinics

- The simple entropy based approach



## Application Vision (illustrative)



## Precision Farming

- Image analytics from mobile device
- Diagnosis and
- Decision support

# Tomorrow



## Welcome to OHDSI!

The Observational Health Data Sciences and Informatics (or OHDSI, pronounced "Odyssey") program is a multi-stakeholder, interdisciplinary collaborative to bring out the value of health data through large-scale analytics. All our solutions are open-source.

Observational  
Medical  
Outcomes  
Partnership



SITE SEARCH

HOME TUTORIALS DATABASES PRICING MEDIA ABOUT

[Home](#) » Database Profile Listing

## Database Profile Listing



Your connection to healthcare databases worldwide®

[www.bridgetodata.org](http://www.bridgetodata.org)

B.R.I.D.G.E. TO DATA® has the following 282 database and registry profiles as of February 26, 2016:

- Aarhus University Prescription Database (aka "Prescription Databases of the Central Denmark Region and the North Denmark Region" or "Pharmacoepidemiologic Prescription Database of North Jutland") (Denmark)
- Achmea Health Database (AHD) (formerly Agis Health Database) (Netherlands)
- Alaska Cancer Registry (USA)
- AnalytCare Long Term Care (LTC) Data (USA)
- Anionix Active Patient Network (AAPN) (USA)
- Antihypertensive Drug Use-Results Survey Database (Japan)
- Antihypertensive Drug Use-Results Survey Database (Japan)

## Open source / Crowdsourcing

- Open source as OMOP, OHDSI, transMART might accelerate analytics
- Public provided aligned open data (connectors and infrastructure)
- Collaborative models (e.g. Pharma-Science-Insurance)
- Patient open source services

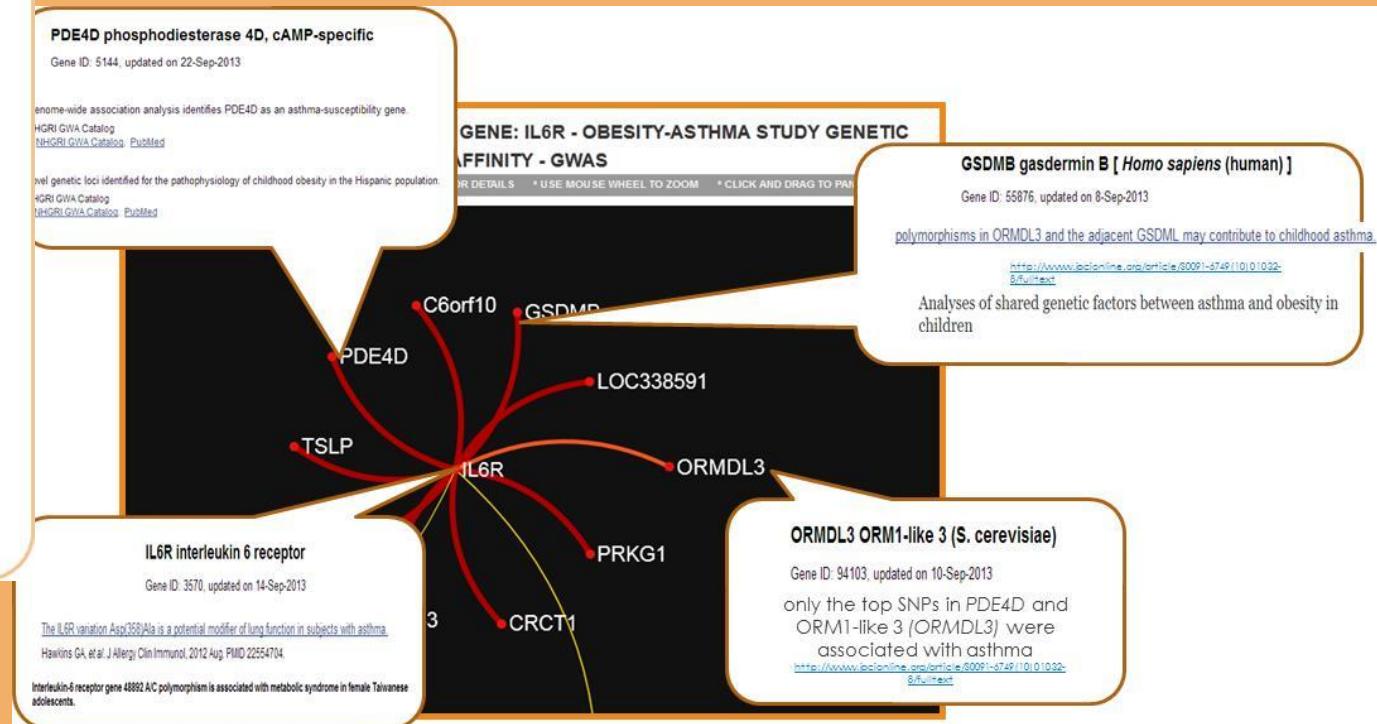
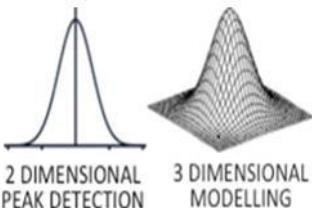
# Tomorrow



Combining LIMS metadata with raw Mass Spec data

Identify and quantitate protein(s) in experiments measuring the impact of drugs and environmental conditions on human cells.

Analyse multi-dimensional protein properties, including chemical structures, modification sites, localisation etc.



## -OMICS and generally Biomarker

- All kind of -omics data e.g. Proteomics, Genomics,
- New biomarkers (e.g. history or path of the patient and not single values)

# Tomorrow

**Arztzeugnis detailliert**

**Arztzeugnis detailliert ohne IV-Entscheid**

**betreffend (Angaben der unterstützten Person, Name, Vorname)**

Gemäss gesetzlicher Vorschrift müssen Personen, die Sozial Arbeitsunfähigkeit mit einem ärztlichen Zeugnis nachweisen, wobei ein Untersuchung vorbehalten wird. Das Formular ist vollständig aus Sozialhilfebehörde Musterdorf zu senden.

Die oben genannte Person ist seit \_\_\_\_\_ bei mir in  Behar

Die oben genannte Person war/ist wegen  Krankheit  Schw<sup>a</sup>  
 Unfall  Spital;

|     |     |    |
|-----|-----|----|
| vom | bis | zu |
| vom | bis | zu |

Er/Sie ist ab \_\_\_\_\_ wieder voll arbeitsfähig.

Falls keine volle Arbeitsfähigkeit resultiert, bitte folgende Angaben mai

- Die oben genannte Person ist ab/seit \_\_\_\_\_ in ih  
als \_\_\_\_\_ dauernd zu
- Welche alternativen Tätigkeiten kann die oben genannte Person  
Einschränkungen? Welches Entwicklungspotential ist vorhanden?  
 \_\_\_\_\_  
 \_\_\_\_\_  
 \_\_\_\_\_  
 \_\_\_\_\_  
 \_\_\_\_\_

• Diese soeben aufgeführten, alternativen Tätigkeiten sind der oben genannten Person  
ab/seit \_\_\_\_\_ im Umfang von \_\_\_\_\_ % zumutbar.

**SILØAH**

Akutklinik, Pflege und Rehabilitation  
 Worbstrasse 316 | 3073 Gämigen  
 Tel +41 31 958 11 11 | Fax +41 31 958 10 90  
 info@siløah.ch | www.siløah.ch

**Arztzeugnis**

An die ärztliche Leitung

|                 |   |  |
|-----------------|---|--|
| Name / Vorname: | Name des Spitals, des Heimes, der Institution |  |
| Geburtsdatum:   | Tel.-Nr.:                                     |  |
| Strasse:        | PLZ/Weinort:                                  |  |
| PLZ / Wohnort:  |   |  |

Definitiver Aufenthalt  ja  nein

Kurzaufenthalt  ja  nein

Weitere Heimammeldungen  ja  nein

Wenn ja, wo?

**Diagnosen**

**FORUM**

**Das Arztzeugnis zum Discountpreis**

Ein Aufruf an gleichgesinnte Kolleginnen und Kollegen

**Therapie** (Medizin)

**Diagnosen**

**Hauptdiagnosen**

**Letzter Spitalaufenthalt**

**Soziale Situation**

Die korrekte und seriöse Erstellung eines Arztzeugnisses beziehungsweise eines Arbeits-unfähigkeitszeugnisses ist nicht etwa eine Sache von zwei Minuten, wie allgemein angenommen wird, und mit den gegenwärtig empfohlenen 10 Franken auch nicht vermeintlich grosszügig abgegolten.

**ROBERT KURMANN**

Der behandelnde Arzt ist verpflichtet, die Arbeitsunfähigkeit im Rechtmäße zuhanden von Versicherungen und/oder Arbeitgebern zu beurteilen und festzuhalten. In einigen Fällen, zum Beispiel wenn es um juristisch ertragte Beratungen geht, aber auch sonst, kann dem Arbeitsunfähigkeitszeugnis auch ein geringerer Wert beigemessen werden.

Die korrekte Beurteilung solcher Situationen erfordert bereits eine ausführliche Krankheits-, Arbeits- und Sozialanamnese. Natürlich sollen wir uns als Haustiere hinter den Patienten und suchen die für ihn bestmögliche Lösung, allein aber seine Aussage, er könnte mit seinen Beschwerden so nicht arbeiten, bedarf nicht automatisch zum Arzt einer Arbeitsunfähigkeit. Vielmehr muss das Problem genau expliziert werden:

- Welches Ausmaß haben die objektivierbaren Funktions einschränkungen?
- Wie wirken sich diese auf den Arbeitsplatz aus?
- Worin genau besteht die Artik des Patienten?
- Worin besteht die subjektive Einschränkung?
- Bestehender Druck im Arbeitsumfeld?
- Bestehen persönliche Probleme oder Probleme im familiären Umfeld?
- Wie ist das Anstellungsverhältnis: arbeitslos, ausgesteuert, in Kündigung oder in der Sparstift?
- Welches sind die Versicherungsleistungen im Krankheitsfall, vor allem im längeren Krankheitsfall?
- Wurden bereits Sozialleistungen angemeldet oder werden solche schon bezogen?

Die korrekte Erhebung einer solchen Anamnese kann rasch

## General Practitioners and Testimonial

- Fill forms by machine learning mechanism
- Minimize efforts and focus on patient value time

# Tomorrow

The four major subject areas (types of data) that are hosted on this server are:

- A. → Human genome sequence data—This data, sometimes called WGS (Whole Genome Sequence) data, is stored on the server both in its raw form from the sequencing device in a format called FASTQ, and a more compressed format called BAM. The genetic sequencing devices used are made by Illumina, and the sequencing is performed by BGI (Beijing Genomics Institute).
- B. → Text files—These documents, like .pdf files, are journal articles and research papers that are used for reference and meta-research by the investigators/researchers.
- C. → Electronic Medical Record (EMR) data—The EPIC EMR (sometimes referred to as EHR) has been implemented in several large hospitals in Copenhagen. On a nightly basis, EPIC data is updated on the supercomputer.
- D. → Miscellaneous MySQL databases & Excel files—Reference data is also stored in various formats.

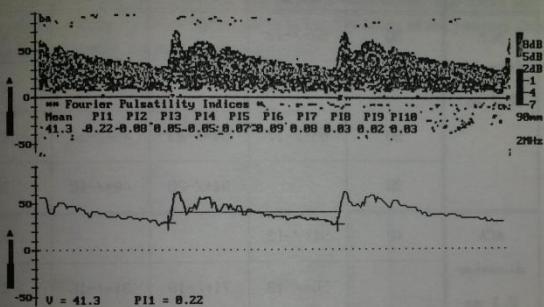


Fig.5.8 d): The vertebral and basilar arteries were examined from position 4 on the neck. Both supply the brain and show a similar flow pattern as the one observed in the other brain supplying vessels. For the vertebral artery it is almost impossible to get an approximation of the angle of insonation due to colling which may be present in the region of the atlas. However, for the basilar artery a small angle (below 30° degrees) can be assumed.

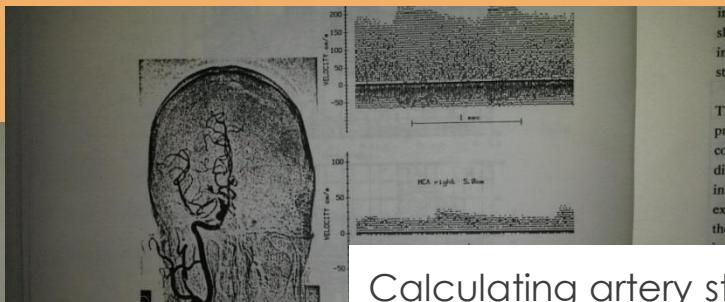


Fig.6.3: Patient suffering from TIA due to a severe i...  
The stenosis could be confirmed by the angiogram and velocity results. The pulse wave form is damped (n...  
lesion.

Calculating artery stiffness out of Transcranial Ultrasound Doppler recordings and link to other data to forecast Arteriosclerosis or even improvements by therapies

## Broader analytics of clinical information

- In depth analytics of all clinical data in specific indications
- Including Raw data comments etc.

DRG's are gone as re-imbursements is fully driven by case, outcome, complexity in relation to all other patients

Because of personalized medicine prescription information is personalized on the most recent data and matching my profile

As a patient I pay more when I'm not compliant and get fully reimbursed when I'm compliant

No registries any more

Half of radiologists

As a patient I can fully analyze my records against others and crosscheck treatments and predict it

Doctors will have the full experience of healthcare and decision support available on their fingertips and are able to focus on their patients

Payers will pay pharmaceutical companies based on outcomes



Life | Tue Jun 30, 2015 9:11am EDT

David Epstein said he was talking to several healthcare customers about a system under which they would get the drug at a discount but then pay Novartis more if, as expected, it successfully reduces the need for costly hospital visits.

"We are beginning to share the risk," he said in an interview.



Life | Thu Nov 12, 2015 8:40am EST

Instead Jimenez and Severin Schwan, CEO of cross-town rival Roche, have been forced to concede that insurance companies, especially in the United States, are not yet ready for such "outcome-based" pricing models.

A key hurdle, Jimenez and Schwan said in recent interviews, is that electronic medical record systems aren't capable of accurately tracking a drug's role in reducing hospital stays or preventing further trips to the emergency room.

**Personalisierte Medizin** ist ohne Big-Data Ansätze nicht möglich

**Ethisch betrachtet** müssen alle Daten in der Medizin genutzt werden

**Daten generieren** keine Fragestellung sondern die Nutzung der richtigen Daten für eine Fragestellung ist zentral

**Die Nutzung von Big Data** ist nicht primär eine Frage der Technologie

**Big Data** ist nur eine weitere Fazette einer Daten- und Analysestrategie

**Fähigkeiten** zur Nutzung / Analyse der Daten gepaart mit Fachwissen ist nötig

**Big Data Strategien** sollten mit schrittweisen Ansätzen umgesetzt werden

Thanks for your attention and looking forward to discuss



**Beobachtung von Mustern** in den Daten kann helfen Vorhersagen betreffend Outcome zu machen



**Unerwartete Muster** in den Daten können wichtige Erkenntnisse generieren

TERADATA®