

Workshop 4: Data Revolution

(Room Louis Armand)



Moderator:
Remi Salomon
PhD, Vice-Chairman, AP-HP (FR)
[@APHP](#)



Roland Eils
Data Management and
Genomics Computing,
Cancer Research
Institute DKFZ (G)
[@CaptainSysBio](#)



Dominique Polton
Chairwoman, IDS (FR)



Eva Nourney
Scientific
communication at the
Federal Ministry of
Education and
Research (BMBF) (G)

Organised by



Big Data in Cancer Research and Patient Care

- Roland Eils

We all are similar, but (slightly) different



Evolution of Large-scale Genome Analysis

- 2000: Human genome working drafts
 - All data freely released
- Project took about 10 years and cost about \$3 billion
- 2008: Major genome centers can sequence the same number of base pairs as were produced for the HGP

Every 16 hours

~~Every day~~

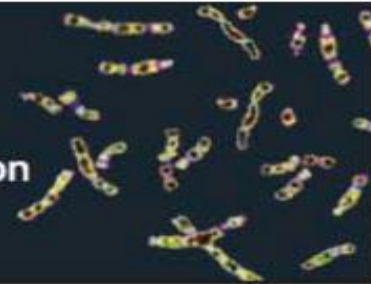
~~Every 2.5 days~~

- ~~Every 4 days~~



1000 Genomes

A Deep Catalog of Human Genetic Variation



Big Data in Genomics: eilslabs vs. Facebook



600 Terabytes per day

(Source: Vagata, P., & Wilfong, K. (2014). Scaling the Facebook data warehouse to 300 PB.

<https://code.facebook.com/posts/229861827208629/>)



12 Terabytes per day

(Source: Zhao, L., Sakr, S., Liu, A., & Bouguettaya, A. (2014). Cloud Data Management, Springer)



11 Terabytes per day

NATIONAL CENTER FOR TUMOR DISEASE

NCT HEIDELBERG

- FIRST COMPREHENSIVE CANCER CENTER IN GERMANY
- JOINT VENTURE OF DKFZ AND UNIVERSITY HOSPITAL
- 10.000 CANCER PATIENTS ANNUALLY
- FOR A THIRD OF THEM CANCER GENOME SEQUENCING MIGHT BE AN OPTION

dkfz.

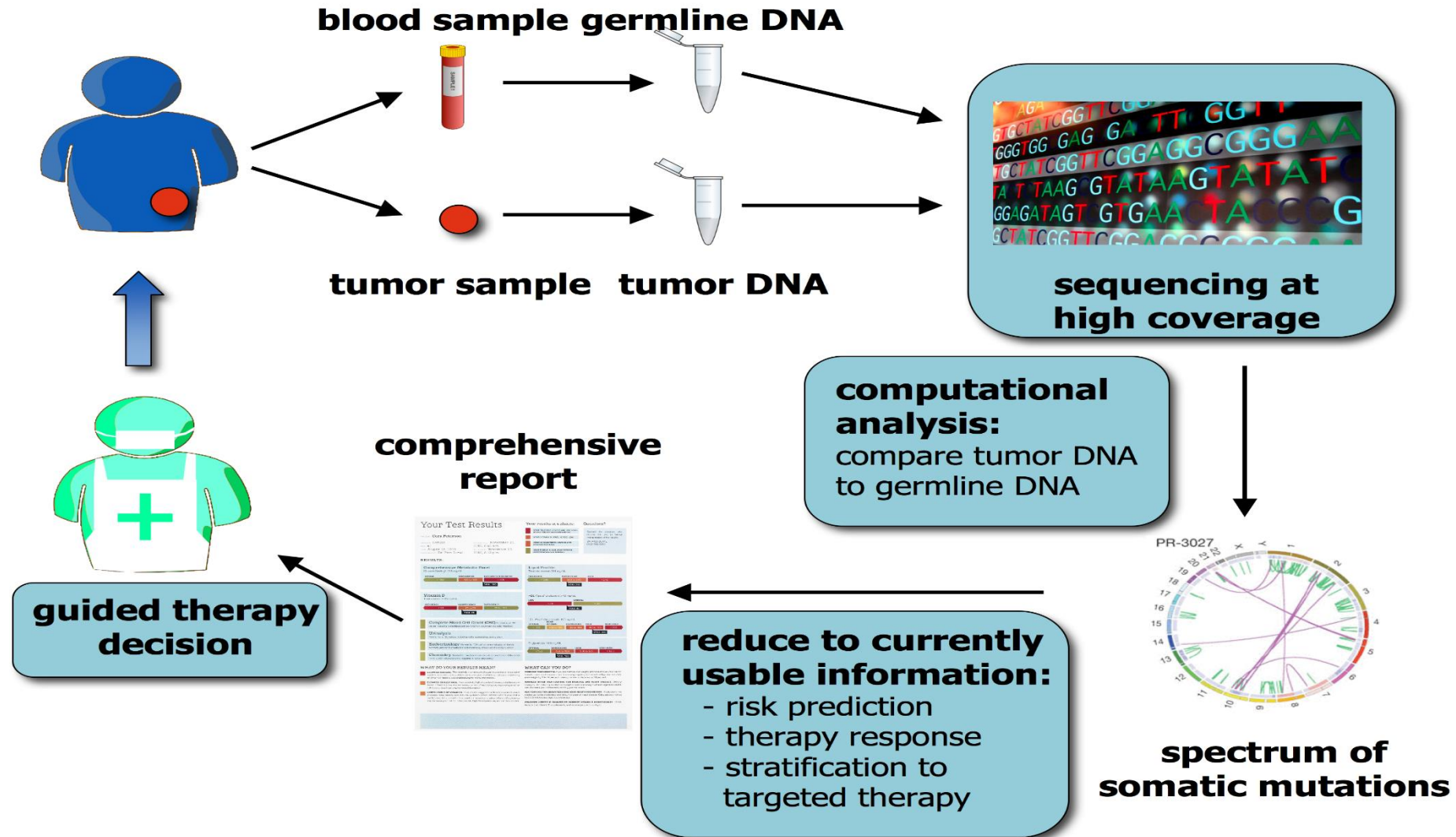


UniversitätsKlinikum Heidelberg

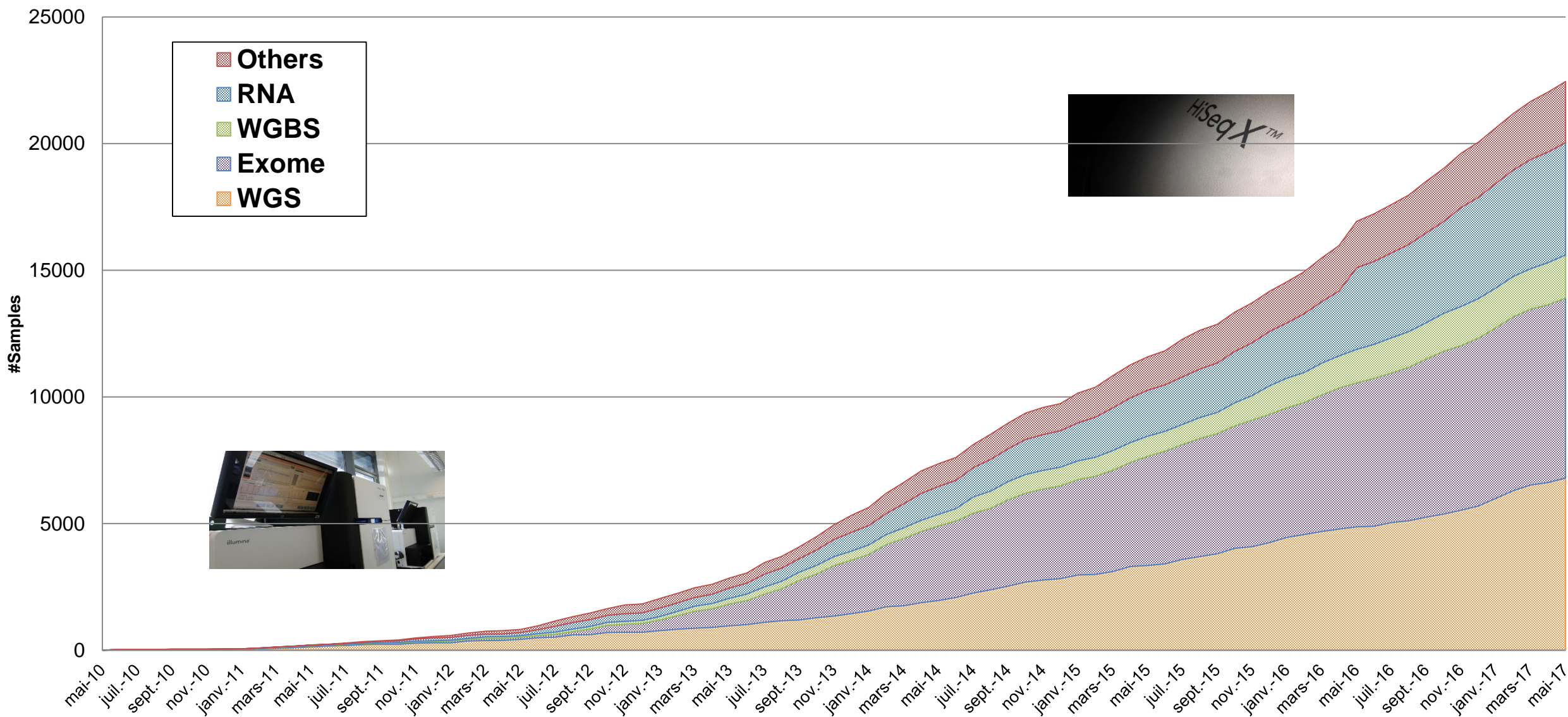
50 Jahre – Forschen für
ein Leben ohne Krebs



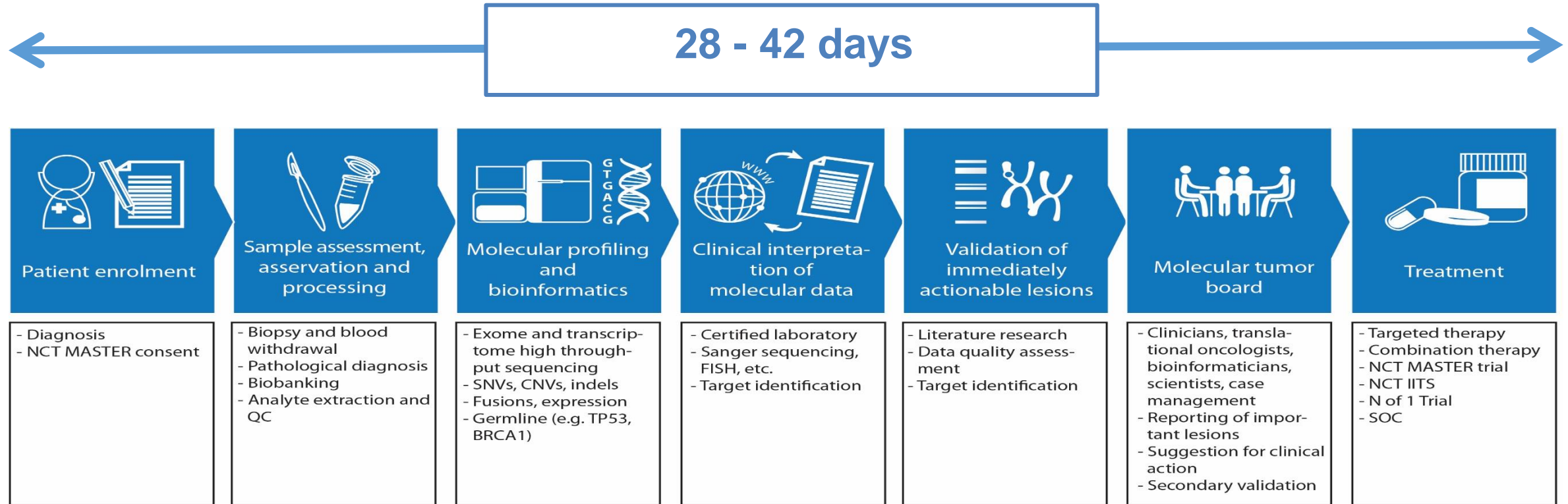
NCT CLINICAL SEQUENCING PROGRAM



Number of Samples (approx. 25,000) sequenced in Heidelberg (clinical) sequencing program



NCT CLINICAL CANCER PROGRAM: MOLECULAR SEQUENCING DIAGNOSTICS



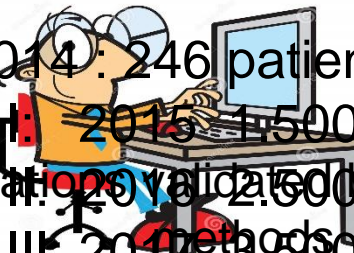
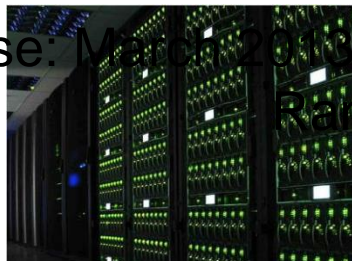
Pilotphase: March 2014 – December 2014: 246 patients, Success rate* 56%

Ramp-up phase I: 2015: 1.500 patients p.a. (120X)

* actionable mutations validated by certified diagnostic methods

phase II: 2016: 2.500 patients p.a. (120X)

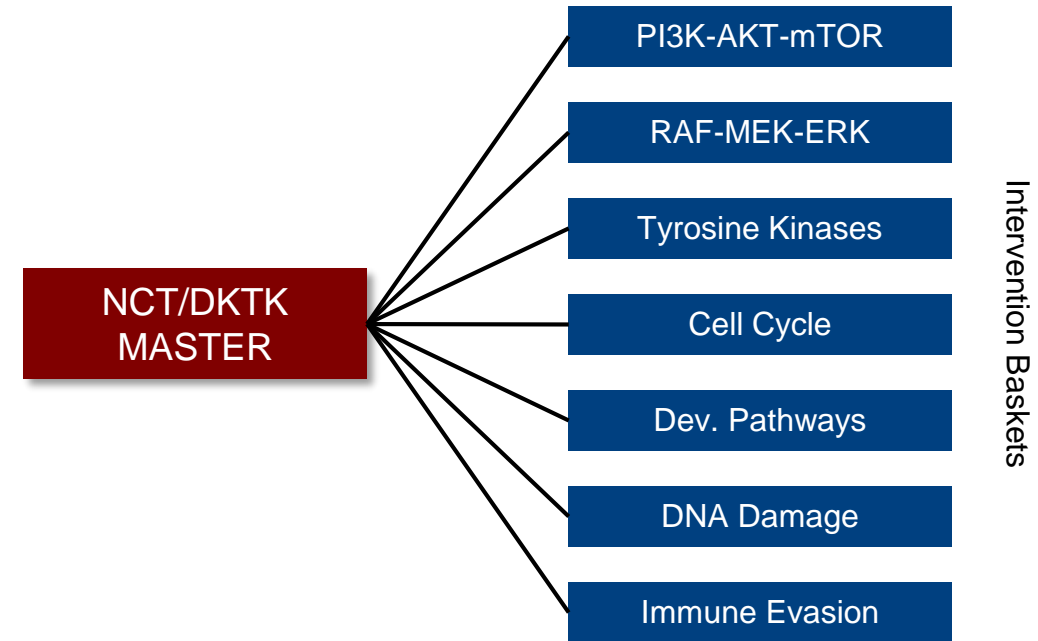
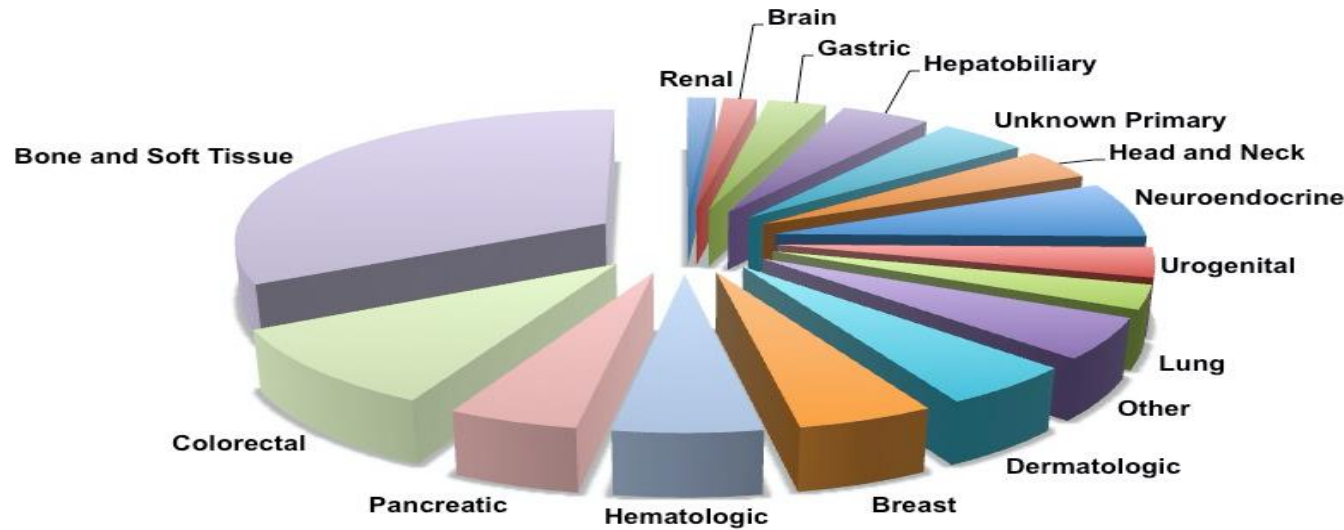
phase III: 2017: 3.500 patients p.a. (120X)



5 hours

NCT Master Match: basket trials based on molecular profiles

(Stefan Fröhling, Benedikt Brors, Hanno Glimm)



October 25, 2017

- **Molecular Tumor Board: 793 patients**
- **Reevaluation of clinical diagnosis: 5%**
- **Management recommendation (Level 1-4): 75% of cases (up from 05/2016: ~60%)**
- **Genomics-guided clinical management: 30% of cases (05/2016: ~25%)**
- **Response or disease stabilization: 45% of cases**

HiGHmed: Medical Informatics Consortium

Das Konsortium — HiGHmed

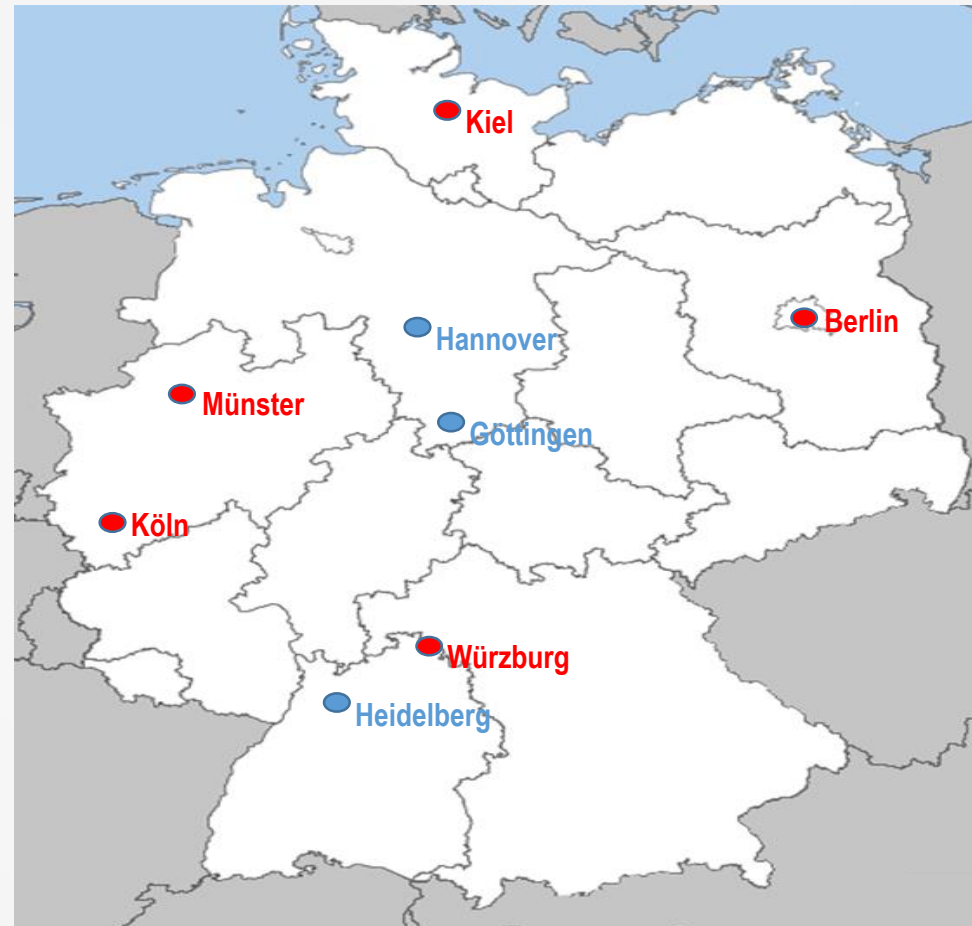
Göttingen
Universitätsmedizin Göttingen

Hannover
Medizinische Hochschule Hannover

Heidelberg
Ruprecht-Karls-Universität Heidelberg
Universitätsklinikum Heidelberg
Deutsches Krebsforschungszentrum (DKFZ)

New Partner sites:
University Hospitals

- Köln
- Würzburg
- Münster
- Kiel
- Berlin



Berlin
Robert Koch-Institut

Berlin
Ada Health GmbH

Braunschweig
Helmholtz-Zentrum für Infektionsforschung GmbH

Braunschweig
Technische Universität Braunschweig

Darmstadt
Technische Universität Darmstadt

Erlangen
Siemens Healthcare GmbH

Hannover
Hochschule Hannover

Heidelberg
NEC Europe Laboratories

Heilbronn
Hochschule Heilbronn

Hildesheim
Hochschule für angewandte Wissenschaft und Kunst (HAWK)

Ismaning
Sana Kliniken AG

Potsdam
Hasso-Plattner-Institut für Softwaresystemtechnik GmbH

Walldorf
InterComponentWare AG

Walldorf
SAP SE

Use of health data in France: opportunities and obstacles.

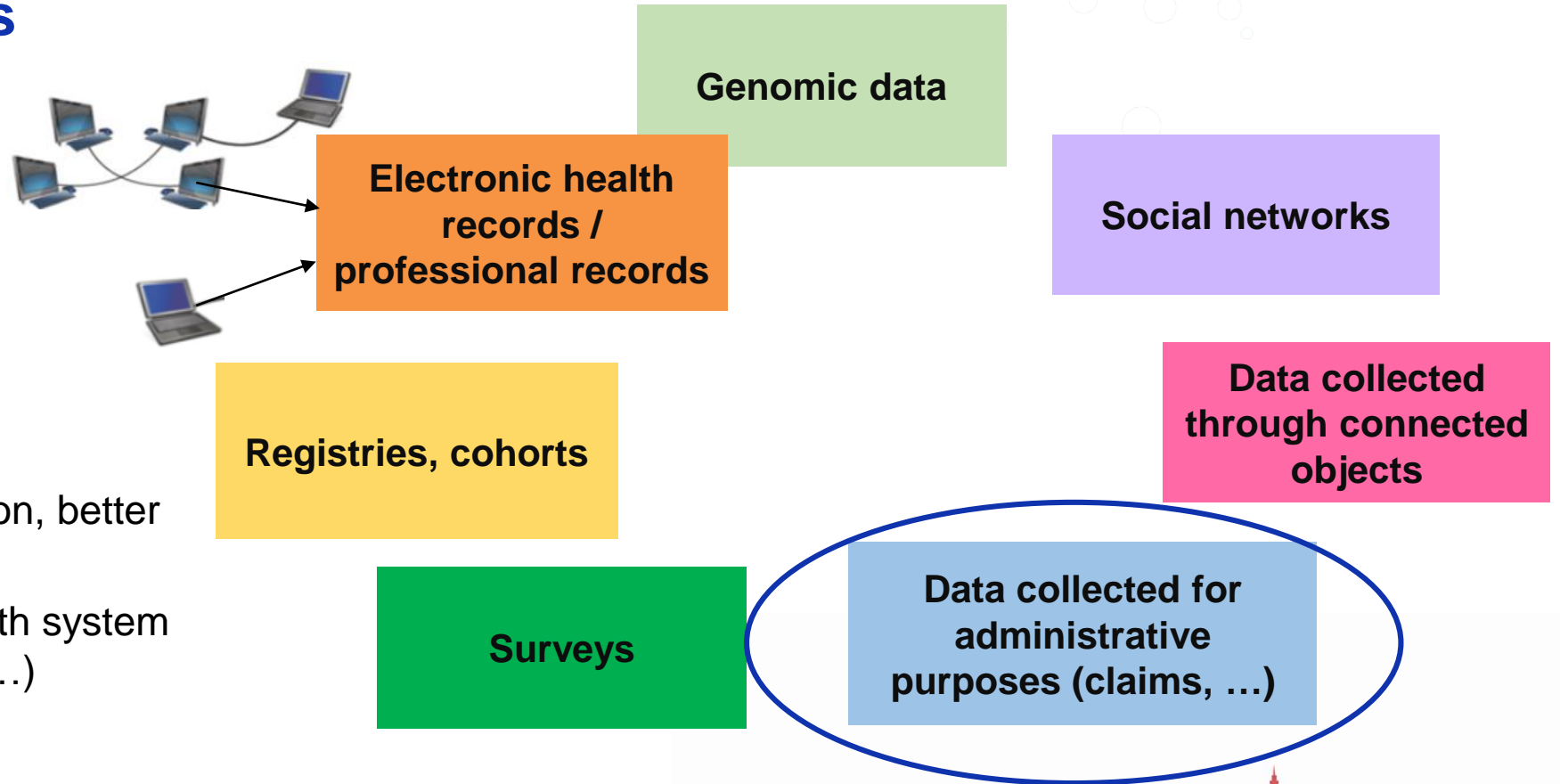
PARIS
DEC, 13th 2017

Dominique POLTON
Chairwoman, National
Institute for health data

Sources of health data

Multiple sources

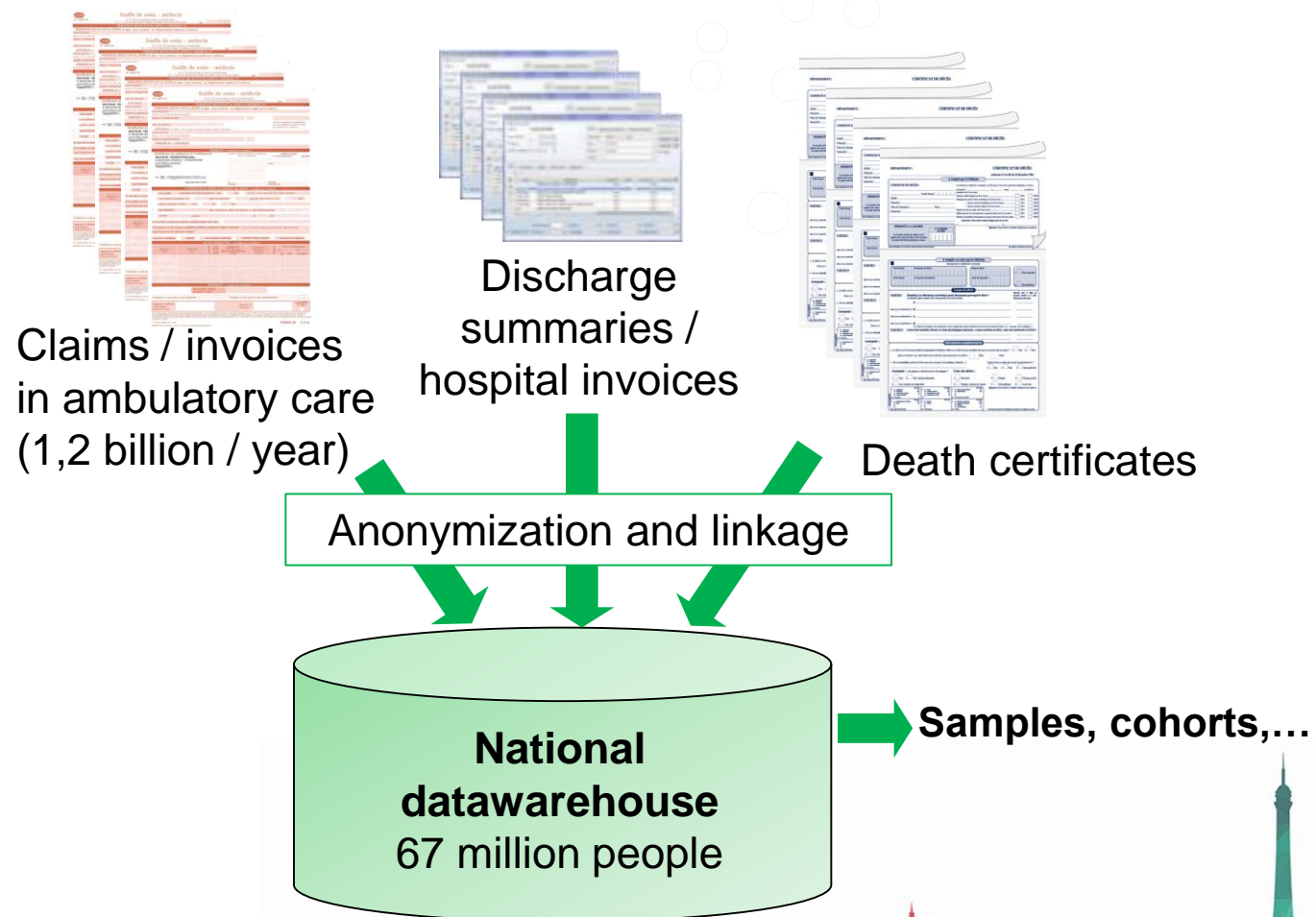
- Enormous amounts of data produced by health care systems (and also outside health care systems)
- **Individual use** (prevention, better / personalized care)
- and **collective use** (health system monitoring, drug safety, ...)



The National Health Data System (1/2)

- **A national data warehouse**

- Linking data from several sources
 - Diagnoses (ICD10 codes for hospitalizations & for patients with chronic/severe diseases)
 - Individuals encounters with health professionals in ambulatory care (identification of professionals)
 - Hospital stays (acute care, mental health, rehabilitation, hospital at home)
 - Coding of procedures (diagnostic and therapeutic) in ambulatory care and hospital care, drugs, appliances
- Vision of the patient journey / care pathway



The National Health Data System (2/2)



• Strengths

Entire population (67 million people)

No lost to follow-up

Homogeneous coding (for a large part of the data)

Historical perspective (10 years follow-up now)

...and weaknesses

Very limited clinical data

No diagnostic tests results, no medical history or risk factors

Limited socio-economic data

Grey areas (but getting covered progressively)



Use of SNDS data: a new legal context



• Until 2016

Limited number of users

- Large use by the sickness funds for their regulatory missions (feedback to health professionals & academic detailing, benchmarking, assessment of the quality of care processes, analysis of expenditures by disease,...)
- Drug Agency (drug safety), Public Health Agency (epidemiology)
- Public research teams (growing use)
- Limited possibilities of linkage with other sources (restricted use of the national identifier)

• New access rules (2016 Act)

Objective = increase the uses of available data + facilitate the linkage with other sources

- Open data for anonymous data (no risk of re-identification)
- Access for all stakeholders (subject to the compliance to the requirements in terms of data protection and purposes of the data processing)
- Deadlines for the procedures
- Easier linkages

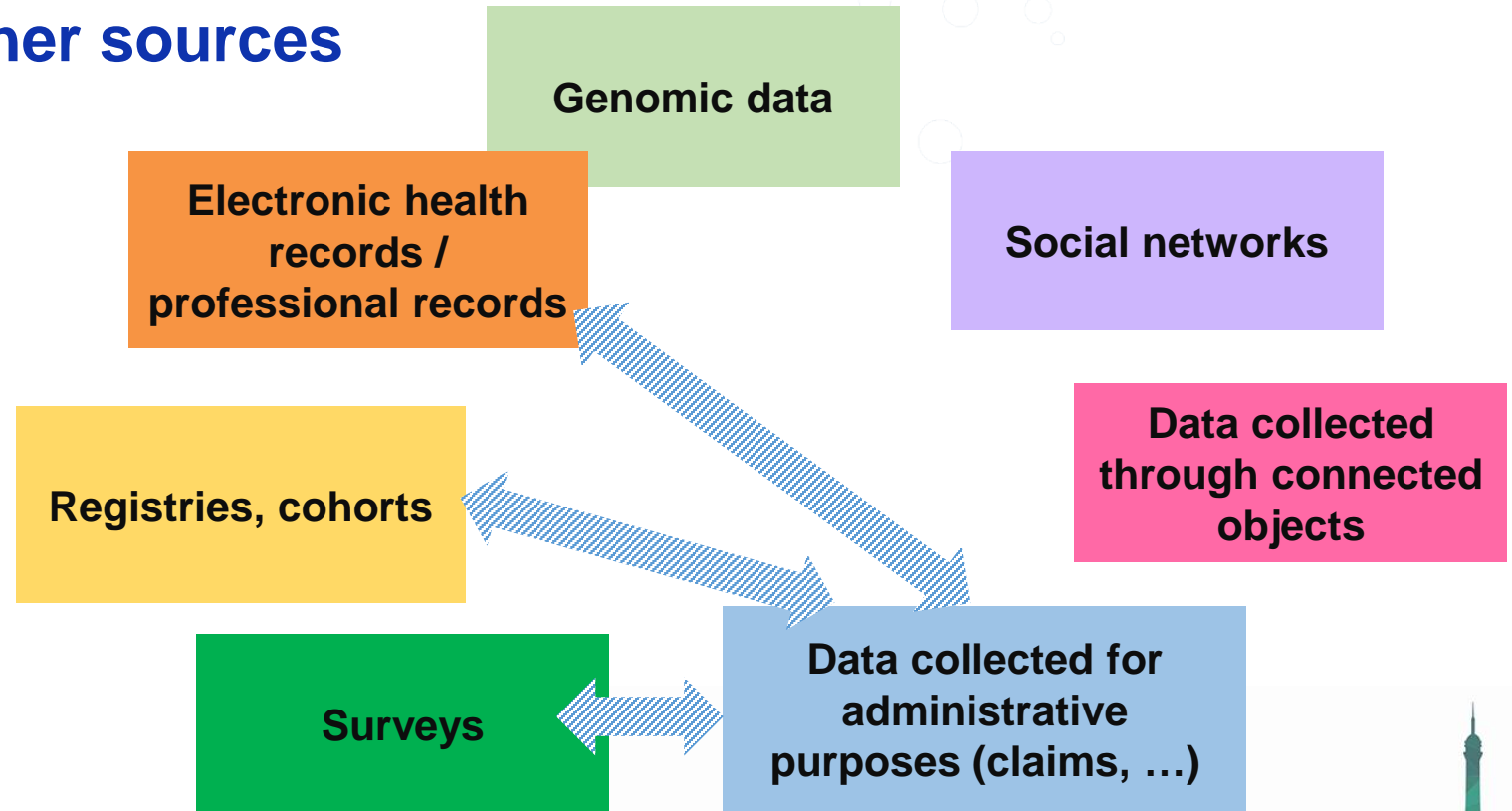


Further developments

- **Linkages with data from other sources**

Examples of projects

- Linkage with registries (e.g. the Transcatheter Aortic Valve Implantation (TAVI) registry, cancer registries...)
- Linkage with clinical data extracted from electronic medical records of GPs (sample)
- Linkage with hospital warehouses ?



Dominique Polton
dominique.polton@indsante.fr

PARIS
DEC, 13th 2017



German Medical Informatics Initiative

PARIS
DEC, 13th 2017

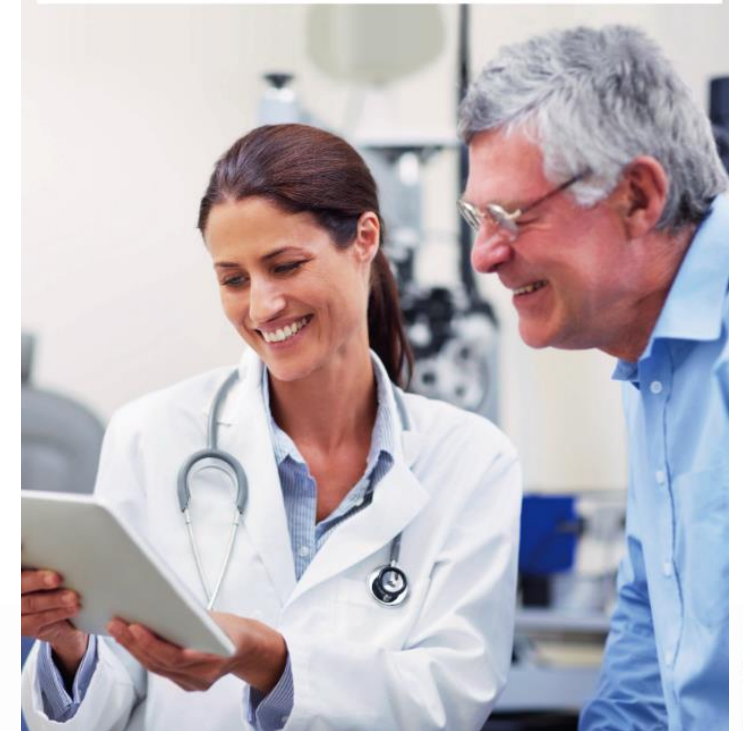
Eva Nourney
German Federal Ministry of
Education and Research
(BMBF)

- **The larger picture**

- High priority for Education and Research initiatives
- 2017 BMBF budget: 17,6 Billion Euro (more than doubled compared to 2008)
- **High Tech Strategy** – „Digital economy and society“; „Innovative working environment“; „Research for healthy living“
- **National Digital Agenda**
- **Health Research Framework Program**
- **Medical Informatics Initiative**



- **Aims of the funding scheme**
 - **Improve patient care and research opportunities** using innovative IT solutions
 - **By facilitating the exchange and use of data** from health care, clinical and biomedical research across institutions and sites
 - **Position medical informatics as a progressive field** in research, teaching and continuing education
- Funded by the Federal Ministry for Education and Research (BMBF): **150 Million €** for the years **2018-2021**



- **Who receives funding?**

Consortia intending to share and exchange patient-related research and health care data

consisting of

- **at least two university hospitals**
- **further partners** (e.g. research institutions, higher education institutions, businesses, private clinics, other health care providers)



- **Focus points of the funding scheme**

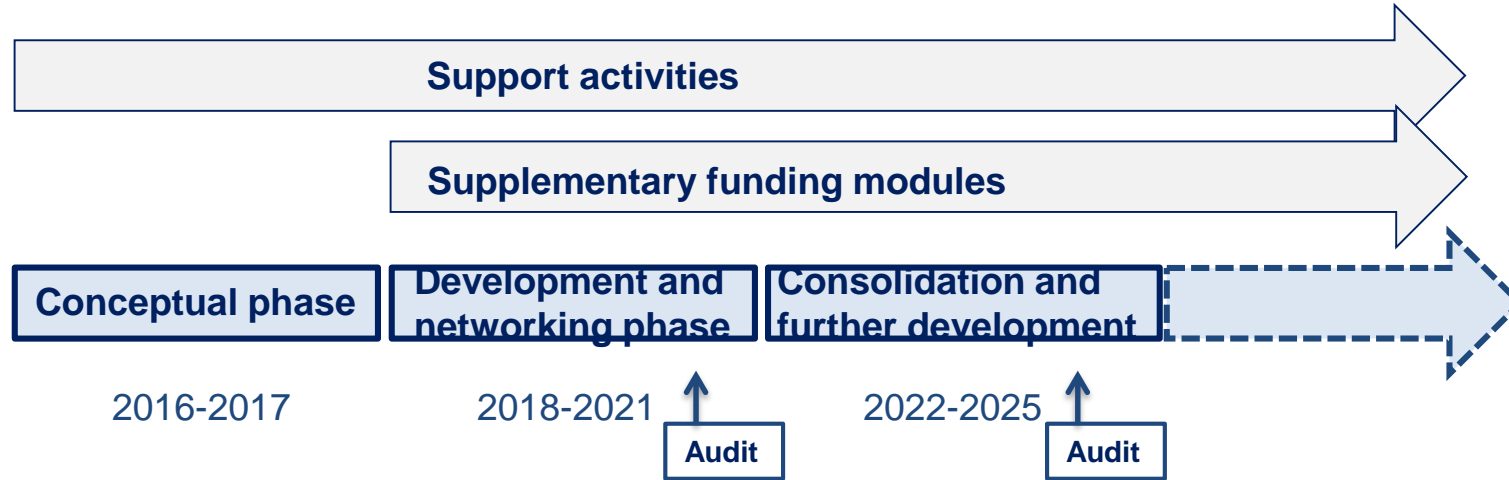
- **Data integration centres**
- **„Use Cases“**
- **Junior Research Groups**
- **National Steering Committee**
- **Modular approach**



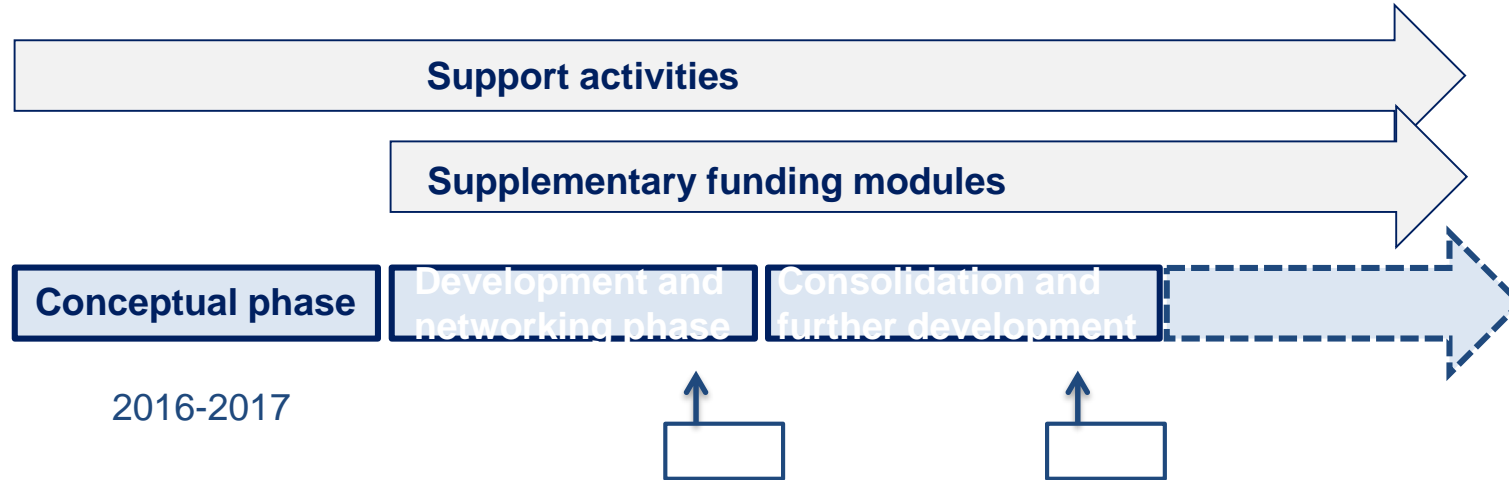
German Medical Informatics Initiative



- **Modular approach - Overview**



- **Conceptual phase**



In the **conceptual phase** 7 consortia developed strategic plans for

- data integration centres
- use cases
- data exchange and sharing
- change management to incorporate data exchange und usage in daily clinical life



- **National Steering Committee**

- **High level of commitment**
- **Agreement / working groups on:**
 - Patient consent
 - Core data sets
 - Interoperability
 - Use and access
 - Mission paper: research-compatible electronic health records

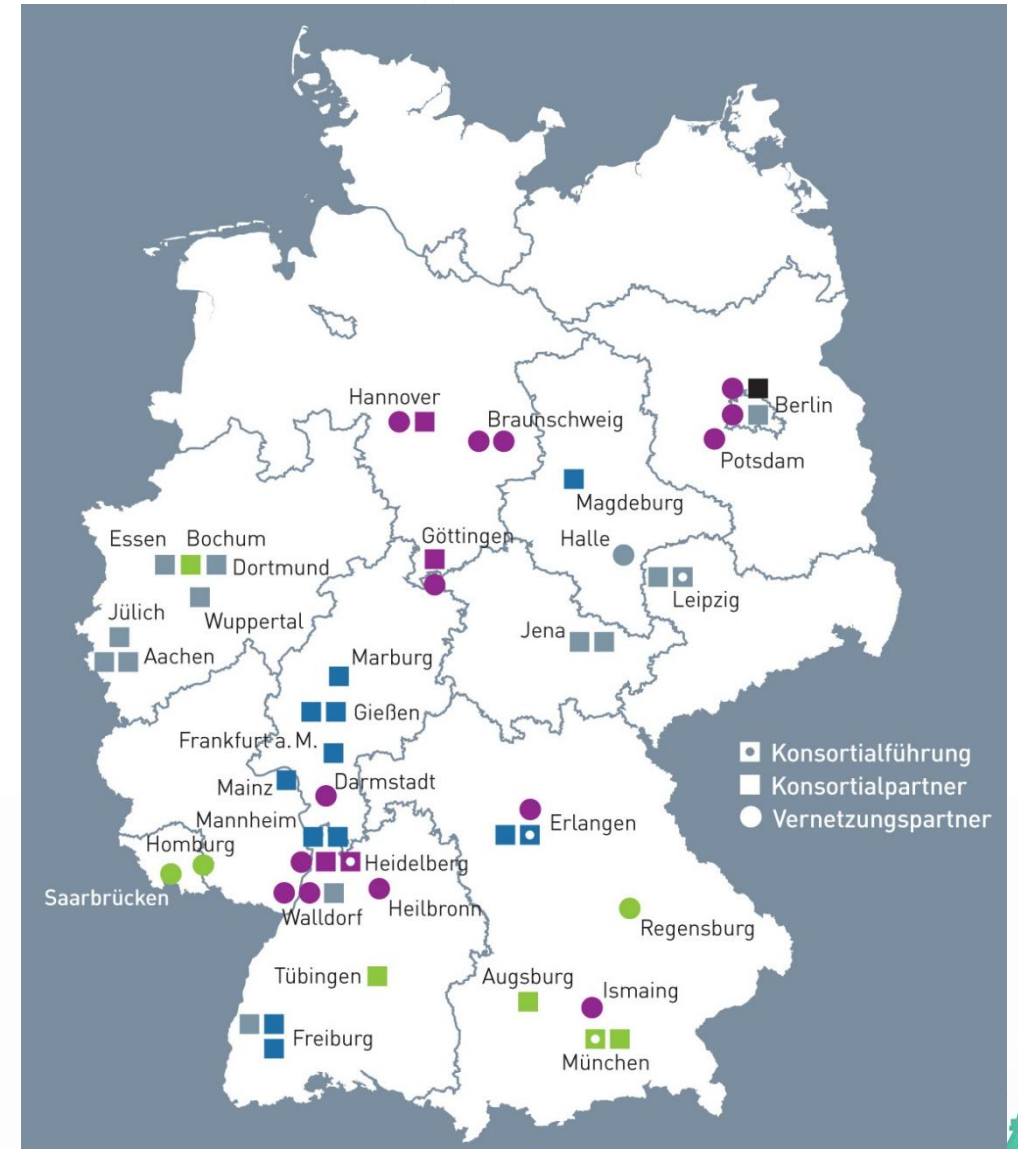


German Medical Informatics Initiative



- **Development and networking phase**

- ✓ **4 consortia with 17 university hospitals and other partners from academia and industry**
 - **DIFUTURE**
 - **HIGHmed**
 - **MIRACUM**
 - **SMITH**
- ✓ **further university hospitals from the conceptual phase will join these 4 consortia in 2018/2019**



Eva Nourney
Eva.Nourney@bmbf.bund.de

PARIS
DEC, 13th 2017

