Health data drives innovation

Learning Health Community

Scaling up trustworthy RWD for clinical research

The momentum in Europe

Professor Dipak Kalra
President

The European Institute for Innovation through Health Data
HD is a neutral, not-for-profit, European institute uniting stakeholders.

To co-create solutions for:
- the capture and sharing of better quality health data
- its trustworthy use for smarter health care and efficient research

- Citizen and patient associations
- Clinical and biomedical research companies
- Health data aggregators and analytics companies
- Multi-national decision makers
- Scientific centres, Reference Networks
- ICT companies, standards developers
- Health system funders, care commissioners
- Healthcare providers and provider organisations
IMI – Europe’s partnership for health

**IMI1: 2008-2013**
€2 bn budget
59 projects

**IMI2: 2014-2024**
€3.3 bn budget
More ambitious
More open
Greater scope

Slide courtesy of Pierre Meulien, Executive Director of IMI / IHI
Patient recruitment a major cause of trial delays

- Identifying and recruiting suitable patients and trial sites are principal causes of trial delays

The percentage of studies that complete enrolment on time:
- 18% in Europe,
- 7% in the US

Almost half of all trial delays caused by patient recruitment problems

Each day a drug is delayed from market, sponsors lose up to $8m

50% of today’s clinical trials fail to achieve the target recruitment rate

---

The EHR4CR project

- EHR4CR – Electronic Health Records for Clinical Research
  - 4+1 year project (2011-2016), 35 partners, budget >17M€
- Objectives & Scope
  - Provide a platform for **trustworthy re-use of EHR data** to support innovation in clinical research and healthcare operations.
  - Unlocking **Health data** for optimising clinical trials
  - 7 pilot sites across Europe
- Status
  - Extended into 2016 for making the transition to a sustainable platform.
  - Initiated a **Champion Programme**, connecting hospitals to an operational platform, building up experience with pharma
  - The **European Institute for Innovation through Health Data** – an independent governance body

For more information: [http://www.ehr4cr.eu/]
### Feasibility

Enabling **protocol testing with real world data** in potential trial sites rather than with guestimates.

### Recruitment

Speeding up recruitment by making EHR data **searchable for investigators** and establishing a **unified communication path** between sponsors and sites.

### Study conduct

Facilitating **EHR data extraction** for applications used during trial execution (e.g. pre-filling of CRFs and of SAE reports).
InSite – Technical Overview, for Protocol feasibility

InSite provides expertise and tools to support local sites with mappings.

Only aggregated data (patient counts) leave the hospital, only on approval.

Local InSite applications to support recruitment.

Protected by Privacy Enhancing Techniques e.g. suppression of small counts.

Full audit trail inside hospital.

External governance by i~HD.

Secure access for researchers.
InSite – Protocol feasibility query

Patient results have been obfuscated for sites MCW. Approximated results are indicated by an *-icon.

Patient Reach for Baseline query

- **58** patients
- **34** in Netherlands
- **34** in MCW

Site & Country Scores

- **16,62** patient matches
- **24** patient matches
- **[Netherlands]**
- **[United Kingdom]**
- **[MCW]**
- **[EHFT]**
Clinical research platform scale up

Custodix merger with TriNetX

Growth within 5 years

InSite

The Largest European Live Clinical Data Network

11 Countries in which InSite is active
31 Million patient lives in the InSite network
50 Hospitals in the InSite network
>100 Additional actively discussing hospitals

Key hospitals include:
- Vall d’Hebron Hospital
- UZ Brussels
- Hospital Universitario 12 de Octubre
- M4H
- Universitetssjukhuset Örnsköldsvik

Network in Numbers

- 157M+ Patients
- 141 Healthcare Organizations
- 29 Industry Clients

- 40B+ Clinical Facts
- 16,260+ Studies Analyzed
- 317+ Sponsored Clinical Trials

100% Industry Retention
Why some eligibility criteria cannot be converted into EHR queries

- Criteria that could not be formalised, such as
  - conditional events (e.g. medication used for another reason)
  - medical response, treatment response
  - measurements usually taken at home
  - toxicity grades
  - "symptomatic", "may or may not", "treatment naive"
  - "more than one medication" but not specified which ones
  - "must have recovered from all side effects"
  - "uncontrolled"

- Likely willingness of the patient to provide informed consent or to comply with abstinence
- Criteria that should apply at a particular visit or at screening
- Participation in another trial
- Investigator's opinion
Redundant data entry

- Clinical trial data are manually entered into dedicated electronic clinical trial systems (EDC) and the same information is often also entered into EHR systems
  - Cumbersome and slow processes
  - Transcription inconsistencies

- Over 40% of clinical trial data are entered into the patient’s health record, the clinical trial EDC system, and, possibly, a third paper copy¹

- 70% of data are duplicated between EHR and clinical trial systems²

---

¹. Integrating Electronic Health Records and Clinical Trials: An Examination of Pragmatic Issues, Michael Kahn, University of Colorado.
EHR2EDC Data Flow

Hospital control

Patient data

Research patients

Study consented patients only

Privacy by design and by default

InSite CDW

Original copy, audited transfer

Investigator chooses what to import

Investigator control

Standard operating rules for all personnel and for the ICT

Sponsor control

Data transfer when confirmed by the investigator

Audited annotation, additional study data

Sponsor EDC system

GDPR compliant Data Protection Impact Assessment

GCP compliance

Compliance with regulatory guidance on the use of EHR data in studies

Study investigator control

EHR

Clinical and Research patients

Study consented patients only

Privacy by design and by default

InSite EHR2EDC database

Original copy, audited transfer

Investigator chooses what to import

Investigator control

Sponsor EDC system
Mapping eSource EHR Data Standards

FHIR@Hospital
- Hospital A Data Model
- Hospital B Data Model
- Hospital C Data Model

FHIR2CDISC
- EHR2EDC FHIR based Common Information Model
- CDISC - CDASH-SDTM

CDISC@Sponsor
- Pharma A Data Model
- Pharma B Data Model
- Pharma C Data Model

EHR eSource

Slide courtesy of Christel Daniel @ AP-HP
The EHR2EDC results in a nutshell

- Six protocols from three pharma companies (AstraZeneca, Janssen, Sanofi)
- Four categories of health data (demographics, vital signs, laboratory and medication)
- Hospitals from three countries: AP-HP (France), 12 Octubre (Spain), IRST (Italy)
- >11000 data points automatically transferred
- Covering 37% of the patient data needed for these studies
- Several companies now developing commercial solutions
Strategic alliance between the public and private sectors to:

- Transform the way clinical trials are conducted
- Improve and accelerate drug development processes
- Place the patient at the center (co-designed by patients)

by developing a common framework for platform clinical trials/Integrated Research Platforms (IRPs)
**Platform**
To test multiple drugs for a single disease in a continuous manner, with drugs allowed to enter and leave the platform on the basis of a decision algorithm.

*EU-PEARL focuses on PLATFORM TRIALS*

**Basket**
To test one drug for multiple diseases or disease subtypes.

**Umbrella**
To test multiple drugs for a single disease.

---

**EU-PEARL**
EU PATIENT-CENTRIC CLINICAL TRIAL PLATFORMS

---

**BENEFITS OF PLATFORM TRIALS**

- Relevant for complex and/or rare diseases with high unmet needs.
- Drugs are tested in parallel so treatments can be developed faster.
- Shared infrastructure can result in trials becoming more efficient.
- “Plug and play” system allows for potential drugs to enter and exit the trial according to the results observed.
- Less strain in patient recruitment as only one control group is needed to test several drugs.
- Outcomes and learnings are shared amongst different companies, researchers, etc., thus advancing science faster.

---

- Shared master protocol and methodology.
- Scientific, legal, regulatory and ethical requirements.
- Network of hospitals, clinicians and researchers.
- Data governance policies and procedures.
- Regulated access to patient electronic health records and patient cohorts.
- Pathway for patients’ participation in trials design.
Value to hospitals = value to all health stakeholders

- Better data access, and tools, to analyse their own data
- Efficient capability to conduct research: income and reputation
- Ability to measure health outcomes and improve care
- Stronger drive to improve data quality
Example data quality issues from hospitals

Leeds Hospital's 'Own Data' Stopped Surgery

The NHS chief who halted children's heart surgery at Leeds General Hospital says the hospital's faulty data was to blame.

© Tuesday 9 April 2013 10:33, UK

Twice as many babies and children seemed to be dying at the unit compared with specialist facilities elsewhere in the country.

34% of weight errors led to medication-dosing errors
48% of these patients required additional monitoring, examination or treatment

- Most health is captured by busy junior staff, using various EHR systems
- Staff have no access to training in data quality
- Patients also have no training! (but their data is increasingly important)
Methods and dimensions of electronic health record data quality assessment: enabling reuse for clinical research
Nicole Gray Weiskopf, Chunhua Weng

A Harmonized Data Quality Assessment Terminology and Framework for the Secondary Use of Electronic Health Record Data
Michael G. Kalin

Secondary Use of EHR: Data Quality Issues and Informatics Opportunities
Taslarchis Botsis, Gunmar Hartvigsen, Fei Chen, Chunhua Weng

A practical framework for data management processes registries
M. Sariyar, A. Borg, O. Fiedinger & K. Pomerening

A Pragmatic Framework for Single-site and Multisite Data Quality Assessment in Electronic Health Record-based Clinical Research
Michael G. Kalin, MD, PhD.†,‡, Marsha J. Rachel, PharmD.†,‡, Jason M. Glanz, PhD, MS.†,‡, Karen Riedlinger, MPH, MT (ASCP),‡, and John F. Steinman, MD, MPH.‡

A Data Quality Assessment Guideline for Electronic Health Record Data Reuse
Nicole G. Weiskopf, PhD, Suzanne Bakken, RN, PhD,‡,§, George Hripcak, MD, MS,‡, Chunhua Weng, PhD

Applying probabilistic temporal and multisite data quality control methods to a public health mortality registry in Spain: a systematic approach to quality control of repositories
Carles Sánchez,1,2, Marí­a Zuriaga1,3, Jordi Perea-Penalosa3, Imane Mezchor3, Moisés Ballesteros1 and Juan M. Sarda-García1,2
<table>
<thead>
<tr>
<th>Name</th>
<th>Definition</th>
<th>Name</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Completeness</td>
<td>Data values are present</td>
<td>Timeliness</td>
<td>Data is up-to-date to their real world state for the task at hand</td>
</tr>
<tr>
<td>Consistency</td>
<td>Data satisfy constraints (format, allowable ranges and values, domain rules, relations)</td>
<td>Stability</td>
<td>Data inherent concepts and statistics are comparable among sources (hospitals, professionals, etc) and over time</td>
</tr>
<tr>
<td>Correctness</td>
<td>Values are true and unbiased with respect to their real-world state</td>
<td>Relevance</td>
<td>Data are useful for their task</td>
</tr>
<tr>
<td>Uniqueness</td>
<td>Records representing a single patient are not replicated</td>
<td>Contextualization</td>
<td>Data are annotated with the acquisition context, their meaning and semantics</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Trustworthiness</td>
<td>Data can be trusted based on the reputation of the stakeholders involved in their acquisition</td>
</tr>
</tbody>
</table>
Data quality rules, assessment tools, improvement support

Data Quality Monitoring Dashboard

Completeness 75%
Consistency 77%
Correctness 82%
Uniqueness 96%
Stability 73%

Recorded past medical conditions over time

Age at death over time

Violations of decision rules in last 30 days

Length of stays in 2017

Drug usage over time

Uniqueness of data records

Distribution of height vs. weight

The European Institute for Innovation through Health Data
Standards and regulations on EHR systems and eSource

• ISO 18308: Requirements for an electronic health record reference architecture, 2011
• CDISC: Leveraging the CDISC Standards to Facilitate the use of Electronic Source Data within Clinical Trials, Version 1.0, November 2006
• EMA Reflection paper on expectations for electronic source data and data transcribed to electronic data collection tools in clinical trials (General-EMA/INS/GCP/454280/2010)
• FDA Guidance for Industry Electronic Source Data in Clinical Investigations, September 2013
• MHRA Position Statement and Guidance Electronic Health Records, Version 1.0, 16 September 15
• ICH Guideline for good clinical practice E6(R2)
• FDA Guidance for industry on the Use of Electronic Health Record Data in Clinical Investigations, July 2018
Growing a Hospital Network of Excellence

Assessment criteria reflect:
• ISO EHR standards
• Regulatory eSource guidelines
• Industry quality frameworks

Two level assessment process
• 1: Online self-assessment
• 2: Optional on site assessment, Certification

Maturity Compass
• EHR Information Governance
• eSource readiness
• Data Quality
Growing a Hospital Network of Excellence

Assessment criteria reflect:
- ISO EHR standards
- Regulatory eSource guidelines
- Industry quality frameworks

Two level assessment process
- 1: Online self-assessment
- 2: Optional on site assessment, Certification

Maturity Compass
- EHR Information Governance
- eSource readiness
- Data Quality

Improvement programmes for these domains
- Compass, dashboards
- Benchmarking intra & inter hospitals of excellence
- Peer learning events
  - Webinars
  - Tutorials
  - Physical workshops
- Sponsored in partnership with industry

Sustainable on-line catalogue
- A searchable & accessible online tool of hospitals of excellence
▪ Quality educational programs to support all stakeholders
  ▪ Trustworthy use of high-quality health data to continuously improve care and accelerate research

▪ “By the i~HD community for the i~HD community”

- Data Quality
- Trustworthy health ICT systems
- Data interoperability standards
- GDPR and information governance

Videos
Narrated presentations
Assessment tools e.g. quiz
# The spectrum of data use: from care to research

## Individual level health data
- EHR systems, apps, sensors, genomics, Clinical Decision Support, AI

**Used for:**
- Health status monitoring
- Continuity of care (including the patient and caregivers)
- Care pathway tracking, clinical workflow management
- Real-time feedback and guidance to patients and clinicians
- Personalised medicine
- Disease interception, prevention and wellness
- Healthcare provider reimbursement

## Population level health data
- EHR systems, regional & national eHealth infrastructures

**Reused for:**
- Healthcare provider performance and planning
- Quality and safety, care pathway optimisation
- Medical device and algorithm refinement
- Pharmacovigilance
- Public health surveillance
- Public health strategy
- Health services and resource planning

## Big health data
- National & international research infrastructures, federated query platforms + cross-sectoral services

**Reused for:**
- Epidemiology
- Digital innovation: devices, sensors, apps
- AI development
- Personalised medicine and biomarker research
- Diagnostics development
- Drug development
- Disease understanding and stratification
Some research findings from “big data”

- >700 million patient records
- 4.9 million hypertensive patients
- 8,000 leukaemia outcomes
- 174,000 prescriptions

- new cancer risk predictors
- best drug to prevent complications
- stronger case for treating elderly
- quality of respiratory treatments
Benefits of federated networks

- Data remains under the control of the data owner
- Locally required legal and ethical approvals apply
- No patient level data leaves the owner’s site, only aggregated counts, thereby ensuring patient privacy
- GDPR – ‘Privacy by Design’
- Analysis is “brought to the data” rather than creating central data repository
- Use of common data model allows for efficient search / analysis across multiple data sets
- Requires close collaboration with data owners which builds trust
The OMOP Common Data Model

- **Person**
  - Observation_period
  - Specimen
  - Death
  - Visit_occurrence
  - Procedure_occurrence
  - Drug_exposure
  - Device_exposure
  - Condition_occurrence
  - Measurement
  - Note
  - Observation
  - Fact_relationship

- **Standardized health system data**
  - Location
  - Care_site
  - Provider
  - Payer_plan_period
  - Cost

- **Standardized meta-data**
  - CDM_source
  - Concept
  - Vocabulary
  - Domain
  - Concept_class
  - Concept_relationship
  - Relationship
  - Concept_synonym
  - Concept_ancestor
  - Source_to_concept_map
  - Drug_strength
  - Cohort_definition
  - Attribute_definition
Data harmonisation

**Local concepts**

- Cohort 1
- Cohort 2
- Cohort n

**Global concepts**

- C1
- C2
- \( \ldots \)
- Cn

**Data custodians**

- Identify local concepts
- Specify mappings
- Define security

**Community**

- Specify global and derived concepts
- Define research groups

Source: Rudi Verbeeck and Michel Van Speybroeck - Janssen
The OHDSI community, working with EHDEN

The Observational Health Data Sciences and Informatics (OHDSI) programme. https://ohdsi.org
Why we need big health data

EHDEN Supported Study on Low Neurological Risk with COVID-19 Vaccines published in British Medical Journal

17th March 2022

BMJ Press Release:

Study finds no increased risk of rare neurological events after COVID vaccination

8 330 497 people who received at least one dose of covid-19 vaccines

735 870 unvaccinated individuals with a first positive reverse transcription polymerase chain reaction test result

14 330 080 participants from the general population (control group)
Big health data sharing initiatives

- Myriad of initiatives to share health data across jurisdictional, institutional and domain borders:
  - Sharing data for cross-border care or for research
- Emerging paradigm for analysing personally-identifiable health data:
  - Federated infrastructure model: network of repositories with an overarching governance and interoperability layer
Possible concept model for the European Health Data Space

Common EHDS

- Data Governance
- Data Interoperability and Quality
- Infrastructural Building Blocks

Non-Federated

Public health crisis solution creators

Industry research and innovation

Patients & citizens

Cross sectorial EDS

Federated

Research

Networks

Innovation

Continuity of care

Regulation

Possible concept model for the European Health Data Space
EMA DARWIN initiative

• EMA’s planned network for Real World Evidence generation

• Focus:
  • Drug development – disease epidemiology, unmet need, historical controls, planning
  • Authorisation – contribution to BR, controls, extrapolation to general & special populations
  • On market – benefit risk monitoring, extension of indication

Source: Peter Arlett, EMA, 2021
Building an ecosystem for better monitoring and communicating safety of medicines use in pregnancy and breastfeeding: validated and regulatory endorsed workflows for fast, optimised evidence generation

Co-ordinated by Novartis and UMC Utrecht
EMA is a partner
Some areas of RWE generation methodology are being prepared for EMA ITF submission. One or two areas are candidates for EMA Qualification Advice.
Medicines information for patients

• Equip and empower citizens with digital information tools, the **Gravitate Lens (G-Lens)**

• offering trustworthy, up-to-date and personalised medicines information sourced from the ePI

• targeting
  • safe use of medicines
  • confident, active, and responsive in their patient journey
  • better health outcomes and quality of life
Observatories will initially collect outcomes for three diseases: Diabetes · IBD · Cancer

Source: Stamm et al. (2021). NEJM Catalyst
How do we reach societal acceptability?

- Data protection regulations prioritise the rights of the individual to privacy
- Clinical research can bring important benefits to society
- Many surveys indicate patients are in favour of their data being re-used for research
- The public need greater transparency about why and how health data are used, safeguarded, and the benefits of that use

We need to find the right balance between the rights of the individual and the benefits for society
The challenge with gaining public acceptance of health data reuse

**Individual level health data**

- EHR systems, apps, sensors, genomics, Clinical Decision Support, AI guidance

**Used for:**
- Health status monitoring
- Continuity of care (including the patient and caregivers)
- Care pathway tracking, clinical workflow management
- Real-time feedback and guidance to patients and clinicians
- Personalised medicine
- Disease interception, prevention and wellness
- Healthcare provider reimbursement

**Population level health data**

- EHR systems, regional & national eHealth infrastructures

**Used for:**
- Healthcare provider performance and planning
- Quality and safety, care pathway optimisation
- Medical device and algorithm refinement
- Pharmacovigilance
- Public health surveillance
- Public health strategy
- Health services and resource planning

**Big health data**

- National & international research infrastructures, federated query research platforms + cross-sectoral infrastructures & services

**Reused for:**
- Epidemiology
- Digital innovation: devices, sensors, apps
- AI development
- Personalised medicine and bio-marker research
- Diagnostics development
- Drug development
- Disease understanding and stratification

---

**Decreasing public understanding of why and how data are used**

**Increasingly unfamiliar data users**

**Increasing time from data use to demonstrated value**

**Increasing distance of data results from the patient**

**Perceived lessening choice and greater cybersecurity risk = harder to trust**
Example ecosystem events and reports

Joining the Dots synergies conference Nov 2019

DigitalHealth Europe project:
EHDS: Policy White Paper, industry consultation, patient and citizen consultation 2020

eHealth Stakeholder Group member

Multi-stakeholder consensus events and reports 2020-21, joint with DHS, sponsored by MS, J&J and MSD
Data protection and data altruism

Data Protection
- GDPR
- HIPAA Privacy Rule

Data Altruism
- Data Governance Act
  (Health Record Banks)

- Public preferences respected
- Societally acceptable data uses
- Agreed minimum safeguards
- Codes of data conduct widely upheld
- Accountability for use
- Public involvement and transparency
- Visible societal benefits
Personalised Health and Learning Health Systems…

- Clinical understanding of disease
- Molecular understanding of disease
- Evidence Based Medicine
  - Predictive modelling
  - Innovative medicines
  - Targeted therapy
- Person centred care
- New value based care models
- Citizen empowerment (promoting wellness)
- Health system capability and capacity
- Patient preferences, goals, lifestyle

…need all stakeholders to collaborate, to maximise the insights we can gain from health data