CHALLENGE IN CONVERTING HIGH VOLUME DATA TO RELEVANT CLINICAL DATA FROM REGULATORY PERSPECTIVE

Kaushik Maity

October 13, 2018

FSPx

Flexible.Scalable.Proactive

COVANCE
SOLUTIONS MADE REAL®

Agenda

- What is Big data
- Big data in Clinical Trial
- Application of Big Data in Pharma from Regulatory Perspective
- Data Networks for Analysis of Big Data in Healthcare
- Regulatory Challenges in using Big Data
- Conclusion

Big/High Volume Data

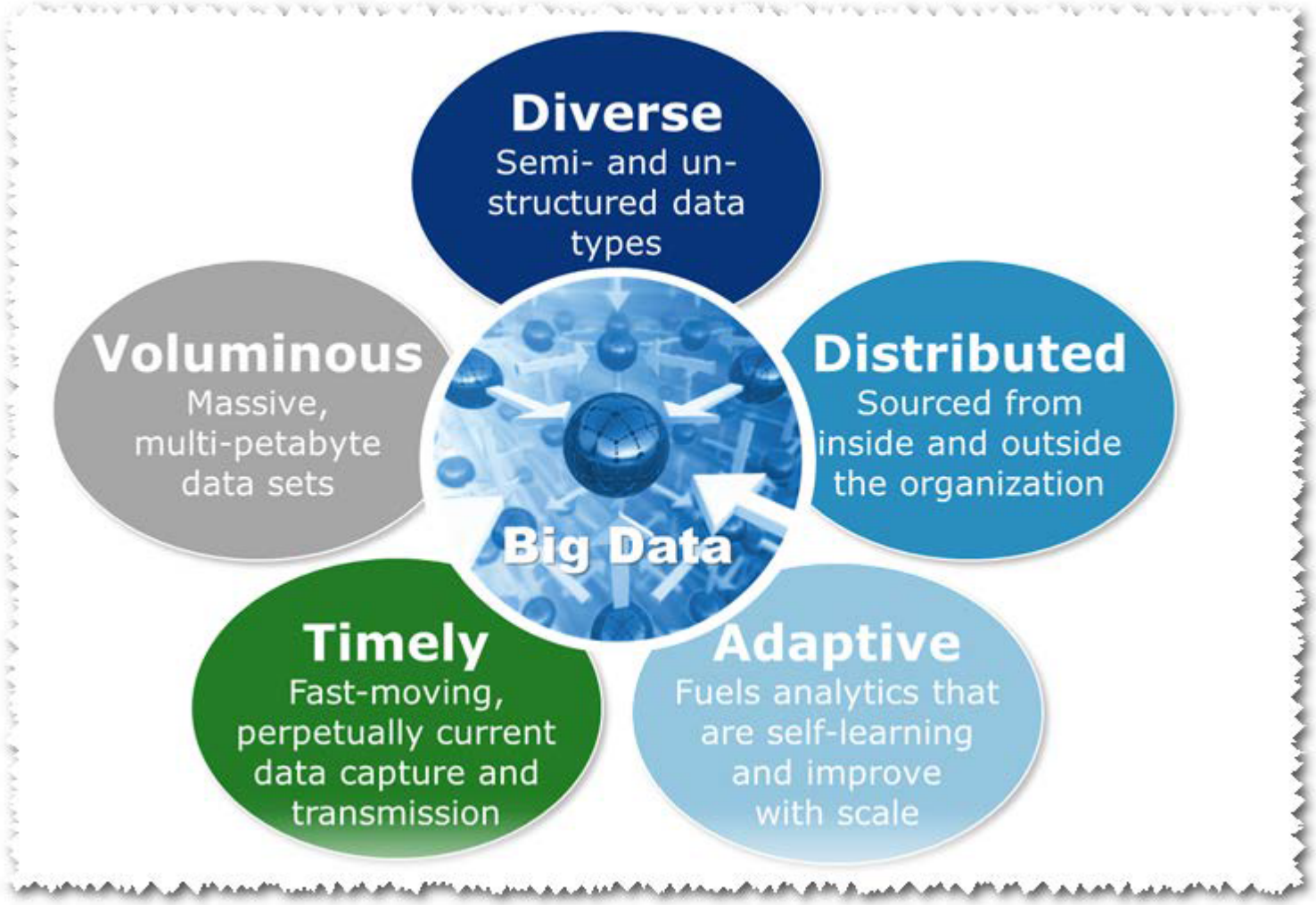
‘**Big data**’ has been defined as 'high-volume, high-velocity and/or high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight, decision making, and process optimization'.

Pharmaceutical industry is emerging as one of the industry where a large volume of data is generated from sources like patients, retailers, and R&D processes.

The high volume/big data allow manufacturers to understand their complex business processes much better, which in turn results in better risk management by pharma companies along with increasing patient safety, and enhanced collaboration between the different stakeholders of the pharma ecosystem.

5 “V”s in Big Data

- **Volume**
- **Variety**
- **Veracity**
- **Velocity**
- **Value**



Impact of Big Data to Pharma Industry

Research

- In silico target screening
- Genomic diagnostics
- Toxicity prediction

Development

- Trial simulation
- Patient recruitment
- Trial design
- Asset prioritization
- Competitive insights
- Unmet need
- Reimbursable dossier development

Market Access

- Formulary/ protocol negotiation
- Value-based pricing
- Payor collaboration (e.g., patient selection, adherence)

Medical

- Safety monitoring
- Targeted physician/ patient education

Operations

- Quality analytics
- End-to-end supply chain forecasting/ planning
- Externalization
- Distribution channel strategy

Few Facts about Big Data in Pharma Industry

- High volume of healthcare data are continually being generated, offering huge opportunities but making it impossible for human to keep pace with all the information.
- Harnessing of the potential of high volume / big data by researchers and regulators is hindered by the fact that it is often unstructured, noisy and inaccessible.
- Storing and processing disparate data sets is challenging and may be facilitated using cloud solutions and machine learning technologies.
- Big data enables us to generate a lot of conclusions; we have to be able to discriminate whether they represent causal relationships or spurious coincidence.
- Big data informatics allows for networks-driven systems pharmacodynamics whereby drug information can be coupled to cellular- and organ-level physiology for determining whole-body outcomes.
- FAIR (Findable, Accessible, Interoperable and Reusable) is not a standard, but a set of principles to ensure data can be used, not just by human but also by computers.
- Pharmaceutical companies are participating in the Innovative Medicines Initiative (IMI), the world's largest public-private partnership for health research, which runs several projects involving big data, including "Big data for better outcomes".

Few Facts about Big Data in Pharma Industry (Contd.)

- Experiences with incorporation of pharmacogenomics information into medicine labels till date have revealed issues of clinical utility, test variability and inconsistencies in what information is presented in the label and when tests are recommended.

- Big data analysis can provide knowledge that may not be available from traditional studies such as clinical trials, but evidence should be validated so that it is accepted by regulators and healthcare professionals.

- To perform precision medicine, organizations need to have various types of data stored in appropriate infrastructures. Advancement towards such systems will be propelled by the growing movement of health insurers towards outcome-based reimbursement.

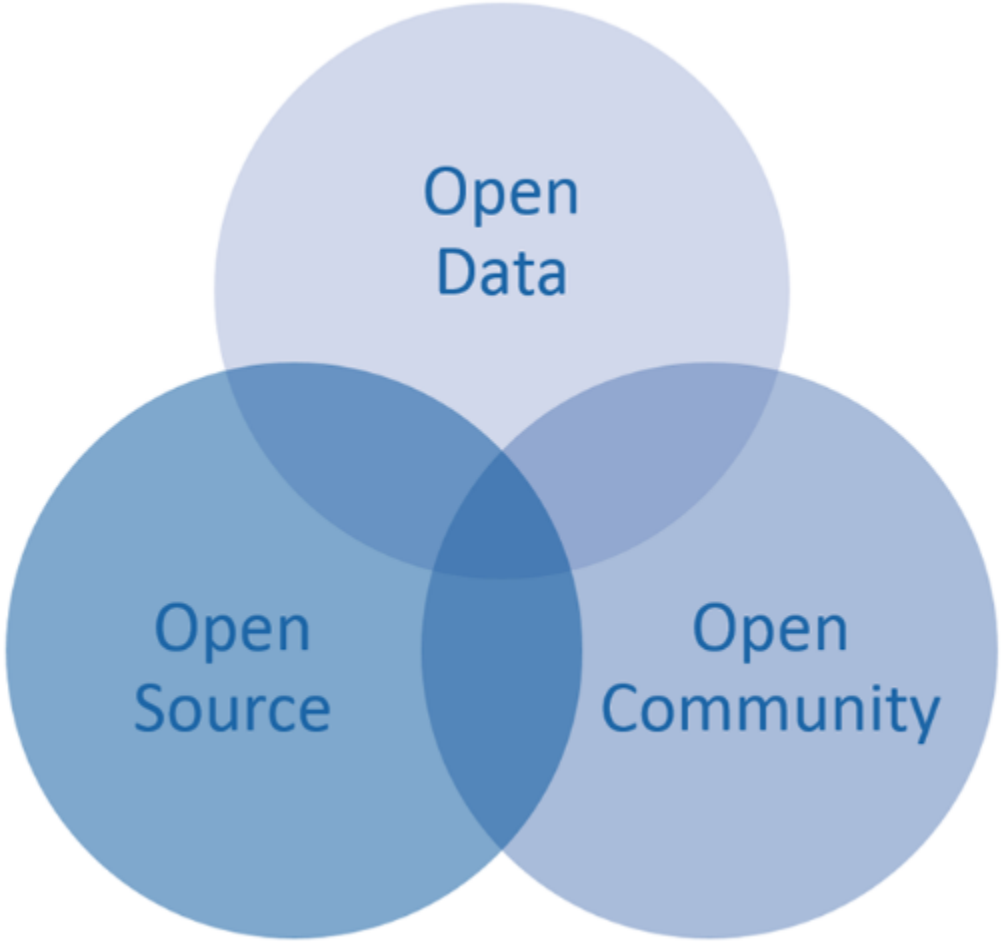
- Serialization of medicines, where each medicine pack will have a unique barcode, will be implemented soon to prevent medicines falsification; however it offers possibilities for use in pharmacovigilance and pharmacoepidemiology.

FDA perspective about Big/High Volume data

- **Data** are raw measurements
- **Information** obtained from data is combined with critical context about what is being measured
- **Evidence** is derived from the analysis of information
- FDA is working on an innovative technology environment that can handle vast amount of data and provide powerful tools to identify and extract the information we need to collect, store and analyze.
- A classic example of leveraging high volume data is cloud computing followed by FDA.
- “Cloud Computing” is computing on demand.

openFDA
open.fda.gov

OpenFDA is an Elasticsearch-based API that serves public FDA data about nouns like drugs, devices, and foods.



COVANCE
SOLUTIONS MADE REAL®

OpenFDA (contd.)

OpenFDA promotes data sharing, data access, and transparency in our regulatory and safety processes, and spurs innovative ideas for mining various types of data and promoting the public health.

Big data is important to the way we carry out regulatory science, which is the science of developing new tools and approaches to assess the safety, efficacy, quality, and performance of FDA-regulated products. Through innovative methods such as cloud computing, we are taking advantage of this flood tide of new information to continue to protect and promote the public health.

OpenFDA API (Application Program Interface)

NDC: National Drug Code

SPL: Structured Product Labelling

UNII: Unique Ingredient Identifier

RxNorm: Normalized naming system

EMA (European Medicines Agency) perspective about Big/High Volume data

- The GDPR (General Data Protection Regulation) comes into force in 2018 and aims to facilitate free flow of data while protecting privacy.
- The regulation covers consent, data collection principles and security requirements and introduces the requirement of accountability for data controllers and how to handle ethical problems.
- Three Principles:
  - Necessity
  - Proportionality
  - Subsidiarity
- Privacy, security, ethics and transparency measures should be implemented from the outset of projects to maintain trust of participants.
- Patients can provide valuable data and insights to regulators and can drive innovation and data capture.

Data Networks for Analysis of Big Data in Healthcare

- To increase reliability and realize the potential of observational evidence; best practices and open-source tools for analytics should be established.
- Collaboration across networks in the observational research community could enable generation of knowledge for entire research areas to meaningfully inform decision-making of medicines.
- Big data is used throughout the medicines life cycle, especially facilitating discovery. Real world data is increasingly important at authorization and during the post-authorization phase.
- Digitalization and social media are now ubiquitous and data can be analyzed using machine learning.

Data Networks for Analysis of Big Data in Healthcare (contd.)

Sentinel Initiative

Sentinel is a network of distributed data approach which allows the FDA to rapidly and securely access information via a CDM from large amounts of electronic healthcare data, such as EHRs, insurance claims data and registries. Pilot project delivers access to 99 million patient lives, 2,9 billion drug prescriptions and 38 million acute hospital stays

CNODES
CANADIAN NETWORK FOR OBSERVATIONAL DRUG EFFECT STUDIES

The CNODES network delivers access to the health and prescription records of over 40 million people and a widely distributed network of academic and data analytics experts to rapidly evaluate the risk:benefit profiles of medicines

OHDSI
OBSERVATIONAL HEALTH DATA SCIENCES AND INFORMATICS

OHDSI is a multi-stakeholder, interdisciplinary collaborative to bring out the value of health data through large-scale analytics. All the solutions are open-source. Currently the community has converted >50 databases covering >660 milliion patients

Sentinel Common Data Model

- Sentinel common data model includes Claims, Electronic Health Record (EHR) and registry data.
- It can incorporate other data domains (eg. free text), and is extensible to any data source.
- It supports any type of analysis because the data are stored at the most granular level available.
- It was designed to meet FDA needs for analytic flexibility, transparency, and control.
- Sentinel distributed querying approach allows automated query execution and response.
- Sentinel approach gives FDA maximum control of the network, data, and tools.
- Sentinel provides information about following groups:

    Public health surveillance

    Health services research

    Clinical trial planning and enrollment

    Patient level prediction modeling

Canadian Network for Observational Drug Effect Studies

Canadian Institute of Health Research (CIHR) established the **Drug Safety and Effectiveness Network** (**DSEN**) in collaboration with **Health Canada** and other stakeholders.

DSEN created Canadian Network for Observational Drug Effect Studies (**CNODES**), to expand the skills to study specific drug safety and effectiveness questions using multiple healthcare databases. **CNODES** also delivers access to the health and prescription records of over approx. 40 million people and a widely distributed network of academic and data analytics experts to rapidly evaluate the risk:benefit profiles of medicines.

The goals of CNODES are as follows:

Organize sufficient financial and human resources

Co-ordinate responses to such safety signals

Standardize methodological approaches

Obtain rapid access to datasets that are large enough to give precise estimates of medication risks and benefits.

(Observational Health Data Sciences and Informatics)

OHDSI is a multi-stakeholder, interdisciplinary collaborative between researchers and observational health databases to bring out the value of health data through large-scale analytics.

All the solutions are open-source.

Currently the community has converted >50 databases covering >660 million patients.

Common Data Model (CDM) is one of the process under OHDSI to standardize the structure and content of healthcare data in preparation for observational analyses.

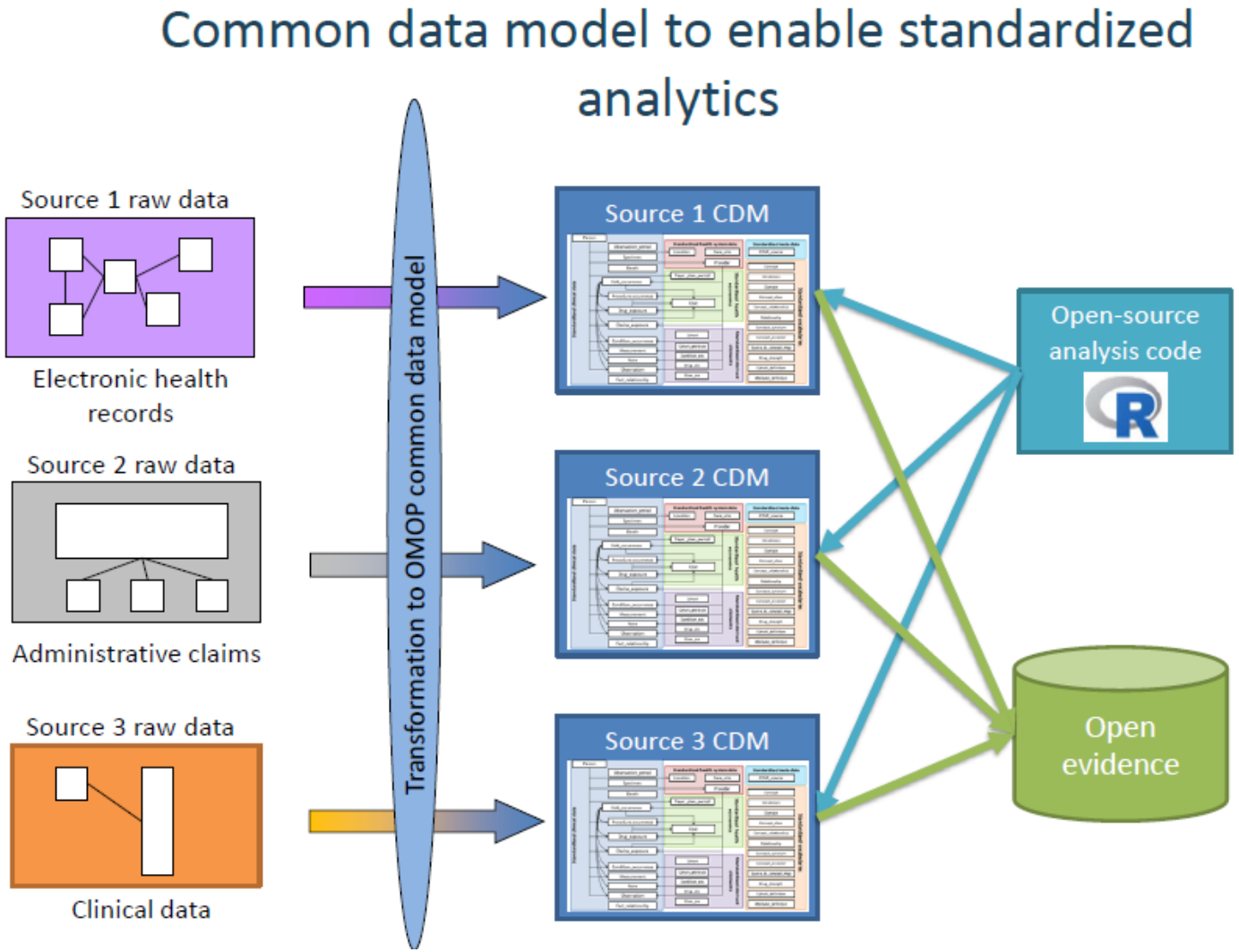
CDM specific standardized **Specification** and **Vocabulary** are available to integrate healthcare data with latest version V5.0.

OMOP (Observational Medical Outcomes Partnership) CDM Principles

- Many networks are using a common data model to enable use of multiple databases across organizations. Data control stays at the local site but data can be accessed as needed in a common format for analysis to respond to questions posed centrally.
- OMOP model is an information model blended between Vocabulary (Conceptual) and Data Model
    - Domain-oriented concepts
    - Patient centric
    - Accommodates data from various sources
- Preserves data provenance
- Extendable
- Evolving

OHDSI (Observational Health Data Sciences and Informatics)

Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM) V5.3.1 14June2018



Different Tools for Analysis of Healthcare data in OHDSI

ATLAS: It is free, publicly available, web based, open source software tool for researchers to conduct scientific analyses on standardized observational data converted to the OMOP Common Data Model V5.

ACHILLES: Automated Characterization of Health Information at Large-scale Longitudinal Evidence Systems – descriptive statistics about an OMOP CDM v4 and V5 database.

White Rabbit: It is a software tool to help prepare for ETLs (Extraction, Transformation, Loading) of longitudinal healthcare databases into the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM).

USAGI: It is used to help in the process of mapping codes from a source system into the standard terminologies stored in the Observational Medical Outcomes Partnership (OMOP) Vocabulary.

WebAPI: It is a service layer which provides access to data and method execution. It offers documentation on the available endpoints and expected value formats.

Regulatory Challenges of Big Data

- An increasing number of medicines with genomic mechanism of action and/or genomic biomarkers enabling smaller, focused Randomized Clinical Trials is one of the regulatory challenges.
- Innovative medicines and personalized prescribing creates regulatory challenges.
- Rare diseases may be associated with more limited information at authorization level.
- Unknown generalizability of RCT results to normal clinical practice: need for new approaches to gather complementary evidence.
- Increasing interest in combination therapies to treat complex diseases creates regulatory challenges.

Regulatory Challenges of Big Data (contd.)

90%

Of the world's data has been created in the past 2-4 years

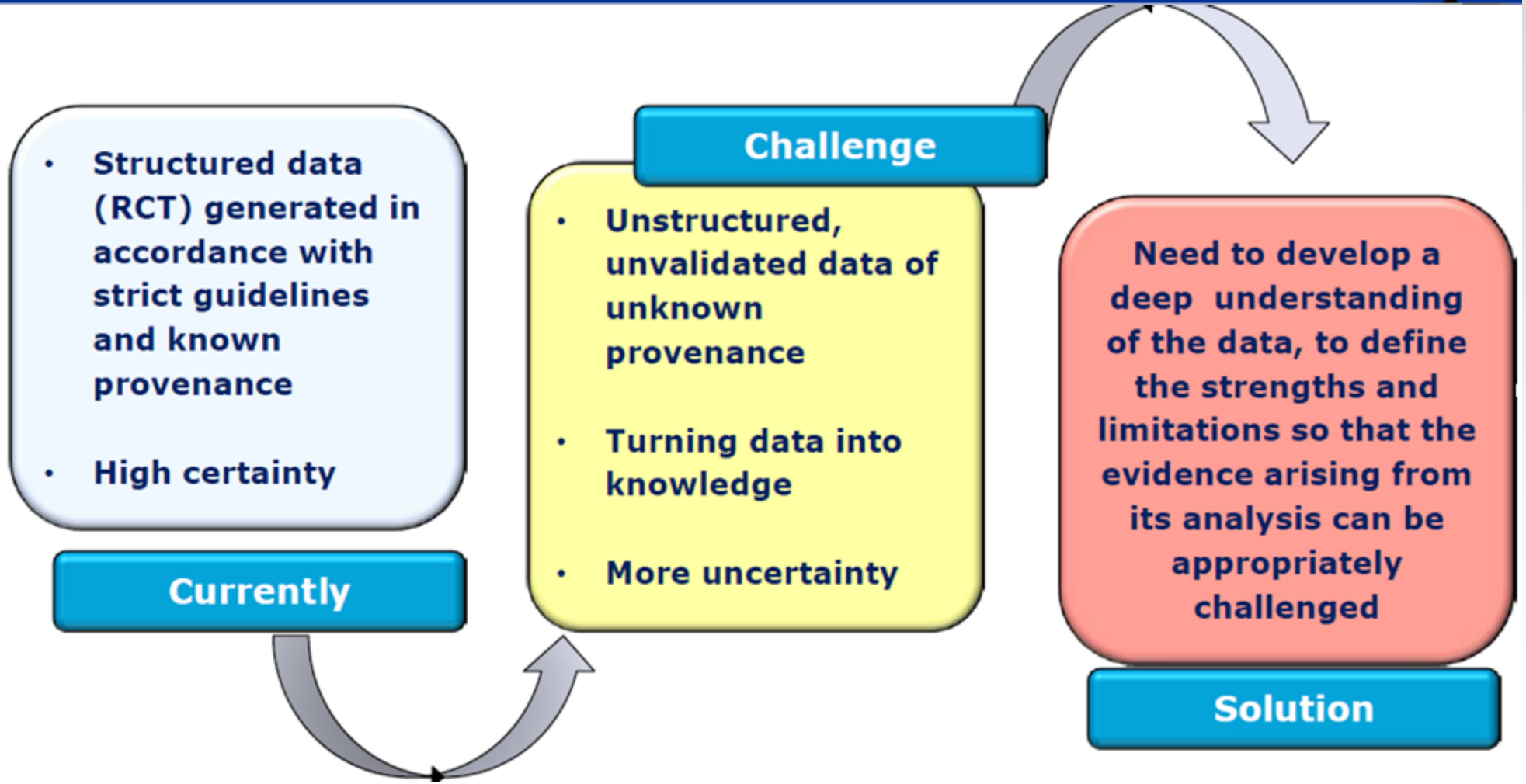
24 months

Frequency at which electronic healthcare data doubles

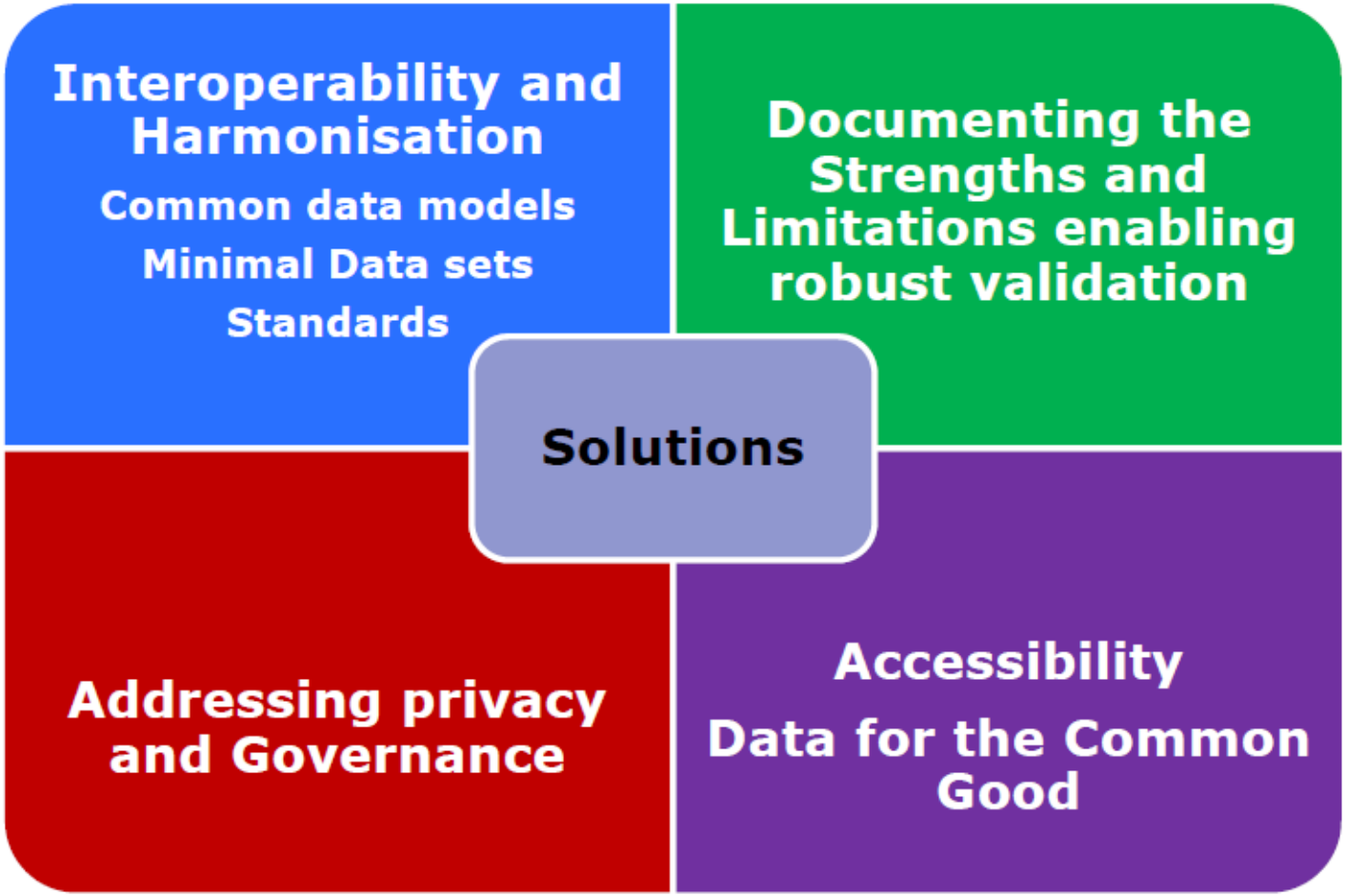
75%+

Percentage of patients expected to use digital health services in future

Changes in the Traditional Regulatory Paradigm



Conclusion



Sources

- https://blogs.fda.gov/fdavoice/index.php/2014/06/fda-leverages-big-data-via-cloud-computing/
- http://www.ema.europa.eu/docs/en_GB/document_library/Presentation/2018/01/WC500241036.pdf
- http://www.ema.europa.eu/docs/en_GB/document_library/Presentation/2018/01/WC500241032.pdf
- https://open.fda.gov/about/introduction/
- http://www.ema.europa.eu/docs/en_GB/document_library/Presentation/2018/01/WC500241031.pdf

Questions

COVANCE
SOLUTIONS MADE REAL®