TECHNOLOGIES IN ACHIEVING POPULATION HEALTH

Tan Yen Ling
Tan Zhi Rong

Contents

1

Overview of the way technologies can be applied in healthcare

- Breakdown of Connected Care Technologies
- Breaking down AI into specific analytics
- 4 Applications of Big Data Analytics in Healthcare

2

Understanding Healthcare Data

- Types of Data
- Identifiable VS non-identifiable data
- Data Security & Privacy
- Processing & integration of data
- Visualization of Data Flow

3

Case Studies on the use of big data & AI in population health

- CareSkore's Chronic Care Management Platform
- UK BioBank
- Population Health Metrics & Analytics (PHMA)
- Identifying undiagnosed Type 2 Diabetes cases
- Predictive Analytics Model
- CancerLinQ Data Lake

4

Moving forward – how Singapore can be better equipped to leverage on these technologies

- Singapore's Healthcare Sustainability Challenges
- Singapore's Health IT Masterplan
- Review of Singapore's current approach
- Key challenges in population health management
- For consideration

1

Overview of the way technologies can be applied in healthcare

(i) Breakdown of Connected Care Technologies

What are connected care technologies?

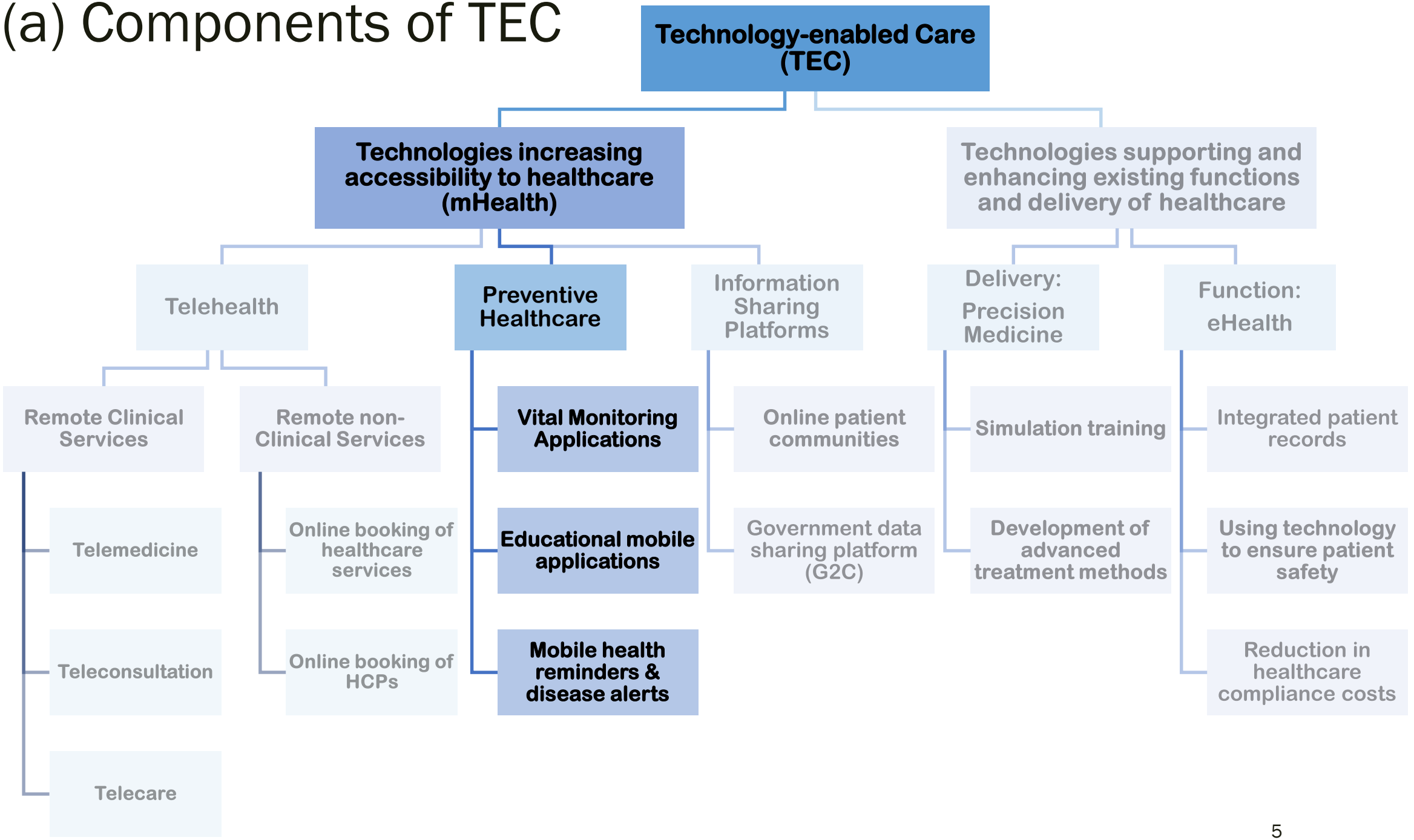
- Technology-enabled care (TEC), also known as connected care, is a convenient, cost-effective and sustainable way of meeting the rising demand for healthcare services through the integration of technology and healthcare.

- We will be zooming in on Preventive Healthcare (Population Health Management) which falls under TEC.

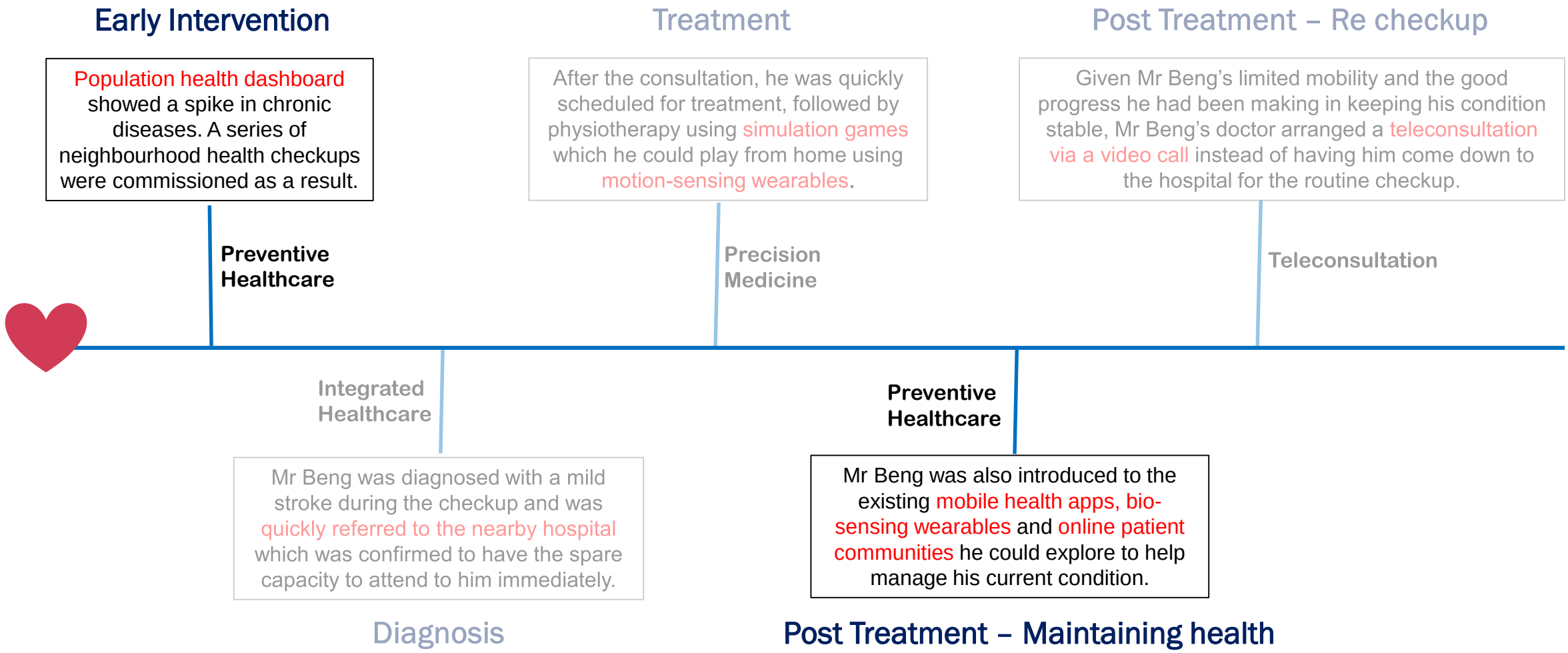
- We analyse TEC in 3 ways:

a) In terms of its components
b) In terms of a timeline for a patient
c) In terms of the technological tools

(a) Components of TEC



(b) TEC visualized as a timeline of a patient (Mr Beng)



(c) TEC as Technological Tools

Internet-of-Things (IoT)	• E.g. Apple watches for fitness tracking & preventive healthcare • IoT sensors in wards to monitor patients through motion detection
Artificial Intelligence	• E.g. AI algorithms that mine medical records to design treatment plans, alert doctors to rare diseases, or create drugs faster than current pace
Virtual & Augmented Reality	• E.g. AR to improve visualization of tumors via 3D reconstructions • Simulation trainings for HCPs
Robotics	• E.g. surgical robots, exoskeletons, pharmabotics • Robot social companions
Nanotechnology	• E.g. Nanomedicine • Patient monitoring device
3D Printing	• E.g. Tissue Engineering – 3D-printed skin for burnt victims • Prototyping for healthcare products

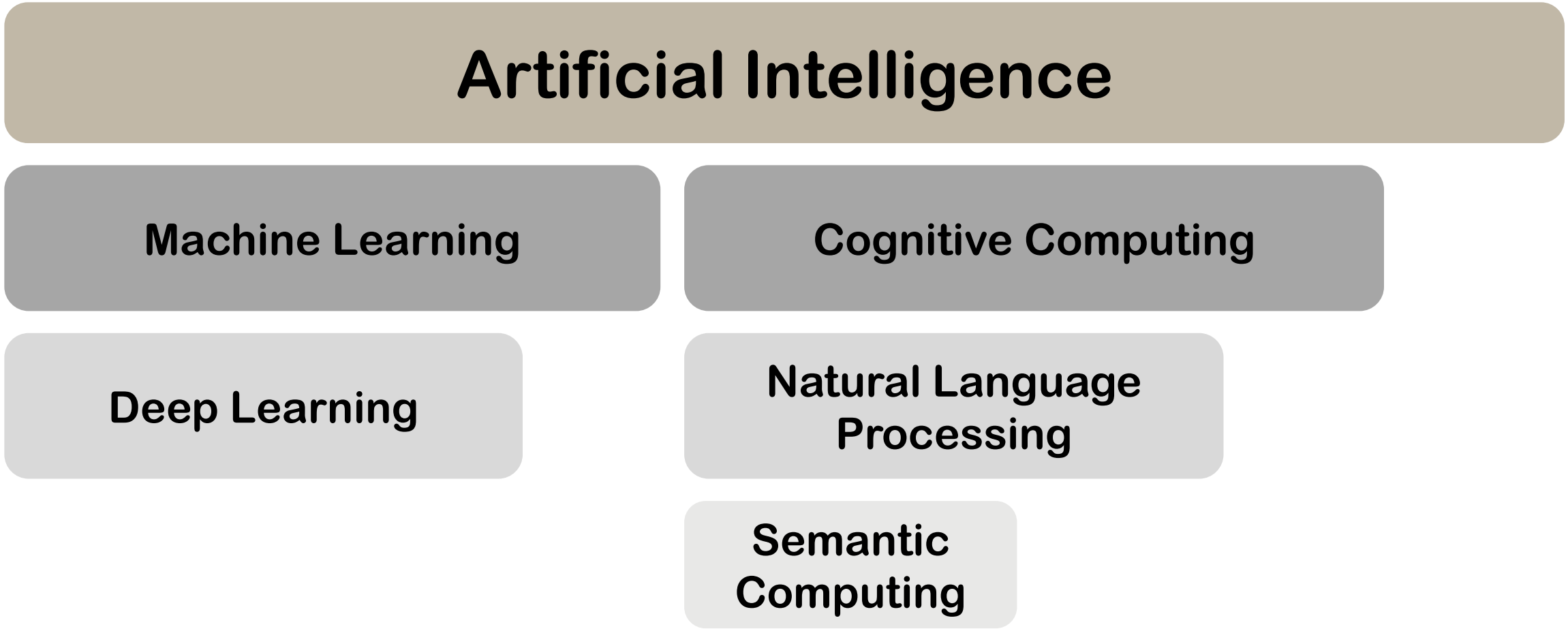
Big Data Analytics

1

Overview of the way technologies can be applied in healthcare

(ii) Breaking down AI into specific analytics

Breaking down AI into specific analytics



1

Overview of the way technologies can be applied in healthcare

(iii) 4 Applications of Big Data Analytics in Healthcare

4 Applications of Big Data Analytics

1. Population health management and patient management
   → Early intervention, post-treatment
2. Revenue cycle, administrative, and operational analytics
   → Healthcare resource optimization
3. Clinical analytics and decision support
   → Diagnosis, treatment
4. Quality reporting and provider benchmarking
   → Checks and balances

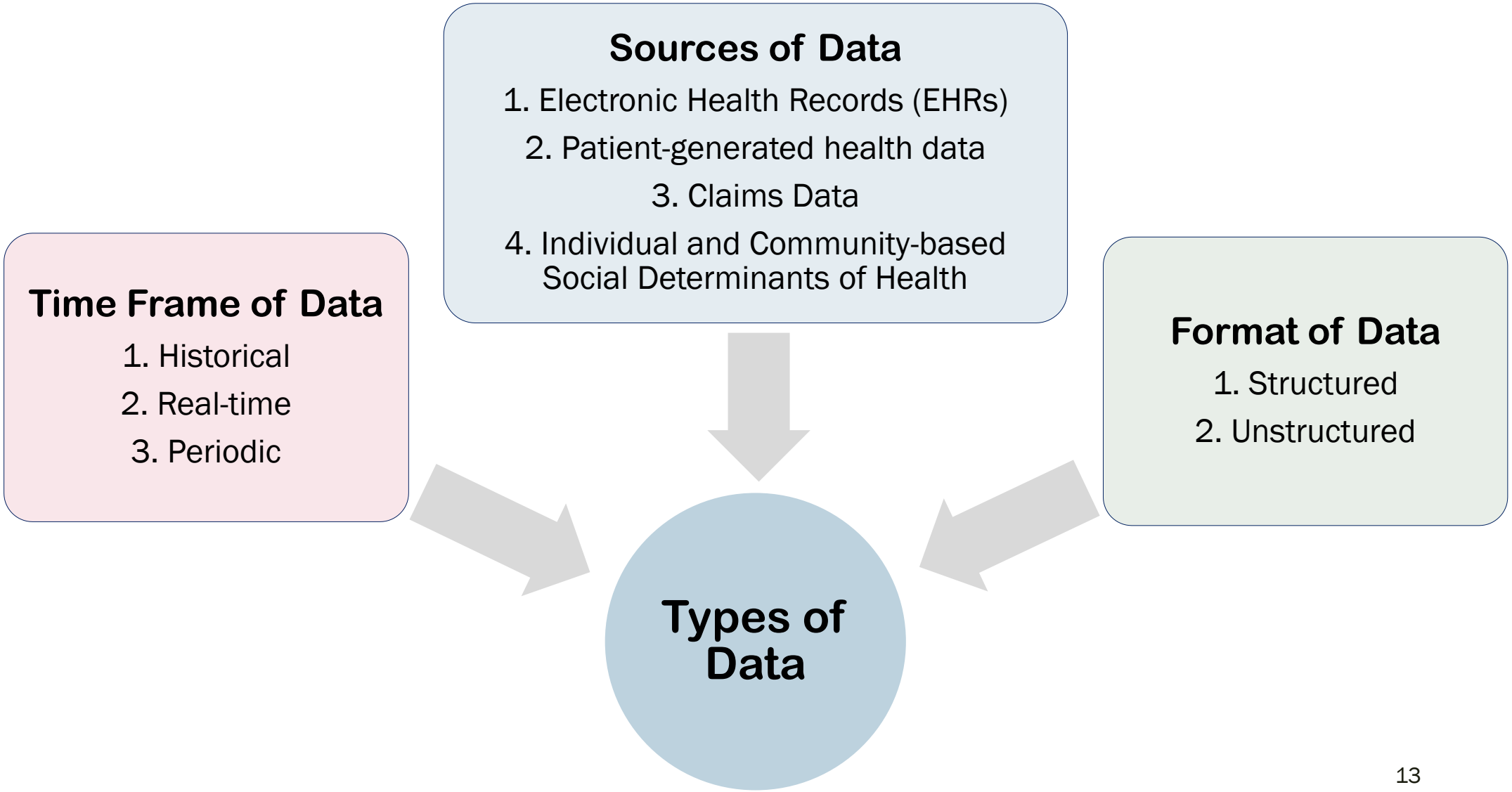
Area of Focus

2

Understanding Healthcare Data

(i) Types of Data

Types of Data



2

Understanding Healthcare Data

(ii) Identifiable VS Non-identifiable Data

Stages of identifiability

- Green data involves the average values for large and aggregated groups of patients. An example is the average time it took for patients to have an operation in hospital after diagnosis. This data is public.
- Red data is personal, confidential data that includes one's postcode, date of birth, identification number, that can be used to easily identify a patient. Such information may also be used as an **identifier**, to connect healthcare data across different sources for an individual. This red data may be made available in exceptional circumstances, such as a public health emergency.
- Amber data is ”anonymised", and each patient's identifiers have been stripped out and replaced with a pseudonym. The amber data can be used for tracking how individuals interact with different parts of the healthcare ecosystem over time.

Data Anonymization

Rationale for anonymizing data

- To ensure data privacy and security
- Different analytics require a different nature of data

Non-identifiable data

Predictive Analytics

Determines patterns and predicts future outcomes and trends with a certain probability

Identifiable data

Prescriptive Analytics

Informs a user what course of action would produce the highest likelihood of maximum benefit when a predicted event does occur.

Potential cause of concern in terms of data privacy

2

Understanding Healthcare Data

(iii) Data Security & Privacy

Data Security & Privacy – Issues & Innovations

Data security and privacy remain the top prioritization in all data-driven endeavours. Issues include:

1. Policy safeguards for storing and sharing protected health information
2. Technical safeguards: proper transmission and authentication protocols, controls over access, integrity, and auditing
3. Transparency & coherent communication regarding data usage and transfer to all stakeholders

Innovations in data security

- **Google DeepMind** is using Block Chain's **Distributed Ledger System** to enhance data security and auditing protocols
- Termed “Verifiable Data Audit”, the aim is to create a special digital ledger that automatically records every interaction with patient data in a cryptographically verifiable manner, through a method called Merkle Tree

Data Security & Privacy – NHS Digital's Care.data (2014)

The programme was poorly received, and eventually cancelled in 2016

Concern 1: Ambiguous Data Usage

- Lack of clarity on how data would be used
- Citizens are worried data may be sold to 3rd parties, especially of insurance companies getting hold of the data, re-identifying it and using it to influence insurance premiums

Concern 2: Poor Communications

- NHS did not focus on the importance and potential benefits of the programme
- Stakeholders, including doctors, were neither consulted nor informed before the public announcement
- Previous dubious NHS data-sharing incidents (NHS-DeepMind collaboration) added to the negative sentiment

2

Understanding Healthcare Data

(iv) Processing and Integration of Data

Processing of Data

In integrating data of multiple types across multiple sources, there are often differing formats, semantics, and data granularities

- Data processing also involves transforming data into the right format, and ensures that datasets are relevant, accurate, and not corrupted.

- Unstructured data (which includes text, images, audio) extracted from its source systems have to be reworked into a readable form that is consistent across stakeholders.

2 approaches to structuring unstructured data are:

1. Reducing the amount of unstructured data via a widely accepted protocol for representing certain types of data
2. Using Natural language processing to transcribe the data into a structural form

Integration of Data – Innovations in Data Warehouses

Transition from data storage in traditional data warehouses to semantic data lakes will enable HCPs to integrate vastly different datasets more easily

- Semantic data lakes enable more efficient use of data without the need for deliberate structuring (avoids forming data siloes)

- The use of cognitive computing techniques enables the database to draw links between seemingly disparate datasets e.g. it is easier to identify the cause of a patient's sudden complication

- Allows for meta-analysis as new information can be added in easily

Integration of Data – Comparison between traditional data warehouses & semantic data lakes

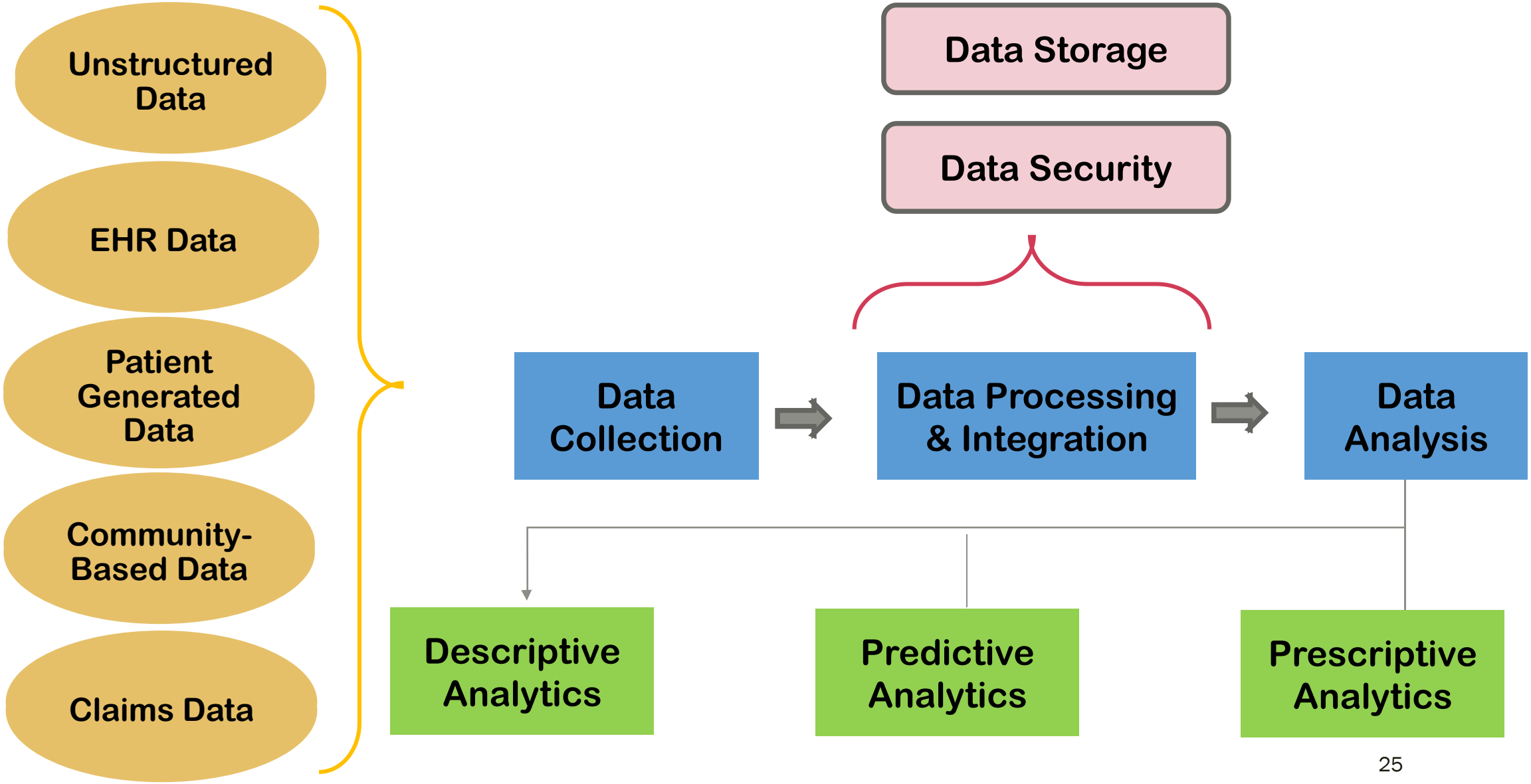
Traditional Data Warehouses (Relational databases)	Semantic Data Lakes (Graph databases)
Has to be targeted to a specific use case – Requires a fine structure thus problems have to framed in a specific way. Information is accepted, compared and retrieved based on a relatively narrow set of parameters.	Does not have to be targeted to a specific use case thus no structure has to be defined - many different types of information can be stored in their original formats in a single repository, allowing for unprecedented flexibility.
User has to predict all future-use cases as altering the problem statement results in data silos since different problems require different architectures.	Uses cognitive computing techniques (natural language processing) to draw connections across vastly different datasets, be it in size, detail, or scope.
	Each element is given a standardized, unique identifier, which allows the database to link separate concepts and generate complex insights the way a human brain does, using natural language queries.
	New data elements can be added and curated to enrich existing datasets or create more capabilities, allowing an organization to include previously unavailable metadata or load fresh data sources into the system to generate more sophisticated insights.

2

Understanding Healthcare Data

(v) Visualization of Data Flow

Visualization of Data Flow



3

Case Studies on the use of big data & AI in population health (focusing on early intervention & post-treatment care)

Case Study – CareSkore's Chronic Care Management Platform

- Who is leading this?

CareSkore, implemented by the Methodist Hospital of Chicago

- What problem are they trying to solve?

Aims to provide a holistic view of a patient by digesting all the factors at play in a patient's care e.g. clinical factors, financial factors, claims data, doctor's notes, socioeconomic factors, and social differences

- What type of data is required?

Patients' medical records including clinical and claims data, as well as patient data and updates in near real-time

- How are they using the data to solve the problem?

CareSkore's Chronic Care Management Platform applies artificial intelligence, machine learning to the data collected and applies its models to do predictive analytics on every individual patient. This allows CareSkore to inform healthcare organizations of a patient's risk of readmission, risk of hospitalization, risk of sepsis, risk of falls and even risk of mortality.

- Were they successful?

Methodist Hospital's orthopedists were able to **reduce readmission rates from 12% to 4%** as the CareSkore platform enabled them to identify primarily, social determinants of reasons why patients would have a higher readmission rate – patients required verbal reminders and follow ups

Case Study – UK BioBank

- Who is leading this?

Established by the Wellcome Trust medical charity, Medical Research Council, Department of Health, Scottish Government and the Northwest Regional Development Agency (UK)

- What problem are they trying to solve?

Aim to improve the prevention, diagnosis and treatment of a wide range of serious and life-threatening illnesses – including cancer, heart diseases, stroke, diabetes, arthritis, osteoporosis, eye disorders, depression and forms of dementia.

- What type of data is required?

Extensive phenotypic and genotypic detail about its participants, including data from questionnaires, physical measures, sample assays, accelerometry, multimodal imaging, genome-wide genotyping and longitudinal follow-up for a wide range of health-related outcomes.

- How are they using the data to solve the problem?

UK Biobank is linking to a wide range of electronic health records (cancer, death, hospital episodes, general practice), and is developing algorithms to accurately identify diseases and their sub-sets. Blood biochemistry is being analysed (such as hormones & cholesterol) and genotyping has been undertaken on all 500,000 participants and these data are being used in health research.

Case Study – Population Health Metrics & Analytics (PHMA)

- Who is leading this?

Saw Swee Hock School of Public Health (Singapore)

- What problem are they trying to solve?

Prevention of chronic diseases by providing risk prediction to individuals, possibly through behavioural interventions or suitable reminders at the right time.

- What type of data is required?

A combination of demographic data, risk factor data (e.g. smoking, obesity), real-time data (e.g. food consumption, physical activity) etc. at the population level

- How are they using the data to solve the problem?

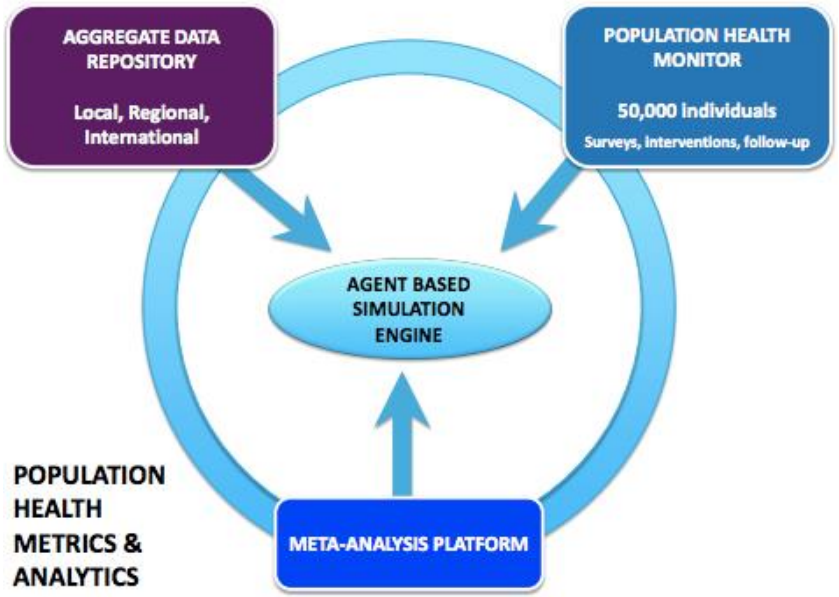
The Population Health Metrics and Analytics (PHMA) consists of 4 major components:

• The **Modelling and Simulation Core (DEMOS++)** makes predictions from agent-based simulations

• The **Population Health Monitor** continuously follows up with a large cohort of Singapore's population to understand their health risks and health-seeking behaviours through questionnaires, interviews, health checks and laboratory tests

• The **Aggregate Data Repository** creates a comprehensive database that enables a researcher to locate and utilise data in an efficient manner

• The **Meta-analyses Platform** analyses and organises epidemiological data from studies around the world



Case Study –
Identifying people with undiagnosed Type 2 Diabetes

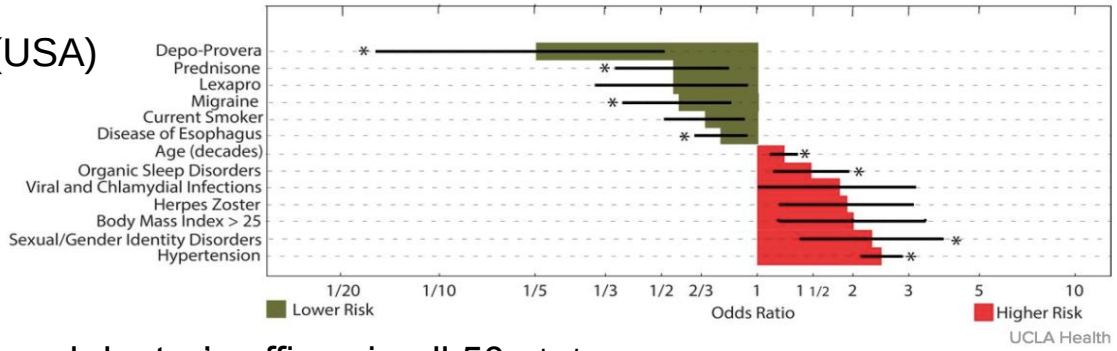
- Who is leading this?

UCLA's Semel Institute for Neuroscience and Human Behavior (USA)

- What problem are they trying to solve?

Identify people who have undiagnosed Type 2 diabetes

- What type of data is required?



The team used Electronic Health Records from hospitals, clinics and doctor's offices in all 50 states which were not identifiable but contained vital signs, prescription medications and reported ailments, categorized according to the International Classification of Diseases diagnostic codes.

- How are they using the data to solve the problem?

The team examined electronic records for 9,948 people. The researchers used half of the records to refine an algorithm that allowed them to predict the likelihood of an individual having diabetes, and then tested this pre-screening tool on the other half.

- Were they successful?

The pre-screening tool based on the entirety of a patient's electronic health record proved **2.5 percent better** at identifying people with diabetes than the standard approach, and **14 percent better** at identifying those who do not have it. The researchers calculated that if the new method were used nationally, it would **identify 400,000 people** who have not yet been diagnosed with the disease.

Case Study – Predictive Analytics Model based on text-mining

- Who is leading this?

University of Arizona (USA)

- What problem are they trying to solve?

Predict ED use of people with asthma

- What type of data is required?

The researchers collected Twitter data from October through December 2013, identifying more than 1.3 million asthma-related messages, or more than 15,000 worldwide tweets each day mentioning key terms related to the disease.

- How are they using the data to solve the problem?

The team used machine learning to extract a relevant dataset narrowed to English-language messages containing specific keywords and narrowed down the tweets by geographic areas of interest. They then correlated the Twitter data to known emergency department use information, environmental trigger data, and Google searches to produce a predictive analytics model that could identify increased social media activity with higher levels of ED visits.

- Were they successful?

The predictive analytics model can suggest patterns of asthma-related ED use with 75 percent accuracy.

Case Study – CancerLinQ Data Lake

- Who is leading this?

SAP & American Society of Clinical Oncology (USA)

- What problem are they trying to solve?

Improve quality of care of oncologists

What type of data is required?

The data lake comprises of anonymized electronic health data from one million cancer patients as well as oncologists' experiences in treating their patients.

- How are they using the data to solve the problem?

The system provides real-time feedback for oncologists by using the big data repository to uncover patterns in outcomes, benchmark performance against national guidelines, and provide clinical decision support.

Gleaning Insights from the Case Studies

Although none of the case studies demonstrated true population-based health management, we can extract a few key insights from the examples

- EHRs are not the only data source which can identify risk factors & incidence of diseases
- Data of different velocities and different sources have to be integrated (& thus interoperable) in order to support AI competencies
- Elements of artificial intelligence are necessary to identify patterns and correlations from large, unstructured datasets, as well as to provide predictions and prescriptive analytics

4

Moving forward – How Singapore can be better equipped to leverage on these technologies

(i) Singapore's Healthcare Sustainability Challenges

Singapore's situation:
Our Healthcare Sustainability Challenges

Increased demand for Healthcare Services

- Rapidly ageing population → likely to double in the next 13 years
- Increasing prevalence of chronic diseases e.g. diabetes
- 30,000 more healthcare workers will be needed by 2020 to meet the rising demand

Burgeoning Healthcare Expenditure

- Projected to rise to $12 billion in 2020 from $4 billion in 2011

4

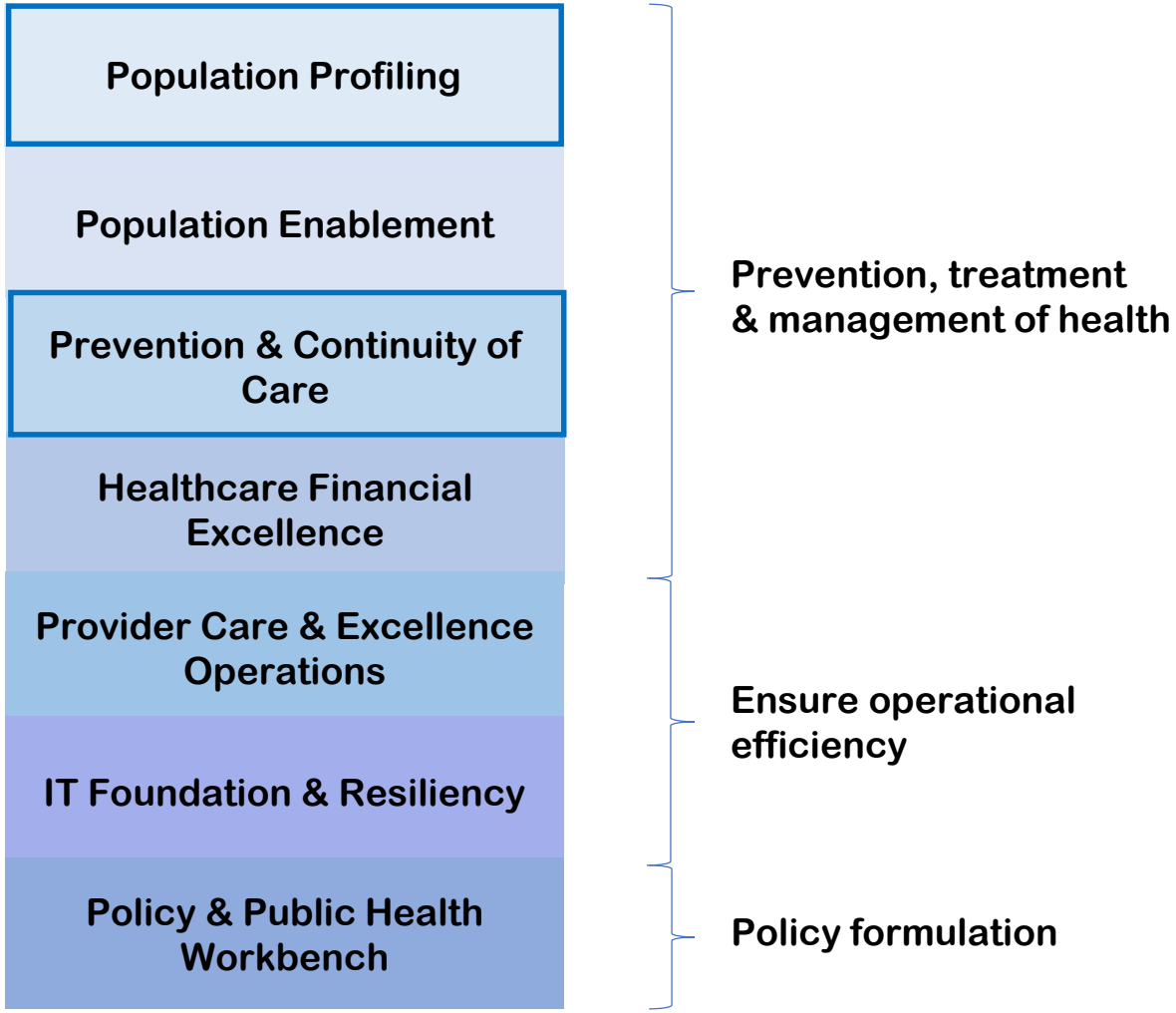
Moving forward – How Singapore can be better equipped to leverage on these technologies

(ii) Singapore's Health IT Masterplan

Singapore's Health IT Masterplan

- Singapore's Health IT Master Plan (HITMAP) laid out 7 transformation programmes which guides the development of ICT initiatives to advance the Ministry of Health's (MOH) three key shifts in the Healthcare 2020 Master Plan to move:

1. Beyond quality to value
2. Beyond hospital to community
3. Beyond healthcare to health



4

Moving forward – How Singapore can be better equipped to leverage on these technologies

(iii) Review of Singapore's current approach to Connected Care

Review of Singapore's current approach to Connected Care

Successful

- HCPs have access to patient's records, preventing patients from telling the same story repeatedly to different HCPs
- Integrated healthcare system with 3 regional clusters that are able to deliver more comprehensive and coordinated care

1. Central region - NHG
2. Eastern region - SingHealth
3. Western region - NUHS

Can be improved

- Patients don't have complete access to their Electronic Health Records (EHRs).
- Lack of interoperability between patient records and mobile health applications
- Perception gap – citizens are unaware of the existing connected care technologies
- Current financial claims do not factor in claims for remote clinical services e.g. teleconsultation & telemedicine
- Healthcare campaigns generally focus on healthy lifestyle & dietary tips rather than promoting the use of IoT health devices capable of monitoring vital signs.

4

Moving forward – How Singapore can be better equipped to leverage on these technologies

(iv) Key Challenges specific to Population Health Management

Key Challenges specific to Population Health Management

- Achieving interoperability of EHRs (or in Singapore's case, NEHR) with key sources of data e.g. patient generated data
- Integrating data of different velocities in order to make useful inferences i.e. EHRs, patient-generated data, claims data and environmental data
- Data privacy and security issues concerning public-private collaborations in achieving population health

4

Moving forward – How Singapore can be better equipped to leverage on these technologies

(v) For Consideration

For Consideration

Achieving population-wide interoperability of EHRs

- Electronic sharing of healthcare data depends on the existence of a functioning technological infrastructure, and the interoperability of separate EHR systems.
- Without set standards, EHR vendors might develop systems that are not interoperable

USA's HITECH ACT

- Under USA's Health Information Technology for Economic and Clinical Health (HITECH) Act, the Office of the National Coordinator for Health Information Technology (ONC) is given the authority to endorse technical standards.
- Because HITECH authorizes ONC to review and endorse technical standards for EHR systems, ONC can guide different EHR vendors on how to develop interoperable systems.

For Consideration

Data privacy and security issues concerning public-private collaborations

- **UK National Data Guardian** Fiona Caldicott recommended an **eight-point consent model** for patient data sharing, which involves:

    1. Explaining to patients about data usage, security and transmission
    2. Four different options for asking the public if they wish to opt out of personal data being used for purposes other than direct care.

        →Single opt-out

        →Opt-out distinguishing between using their data to run services, and using it for research

- Ms. Caldicott also recommend that the government considers introducing stronger sanctions to protect anonymized data, including criminal penalties “for deliberate and negligent re-identification of individuals”.

For Consideration

Transition towards semantic data lakes

- Allows for easier storage of different types of data as well as finding linkages between different datasets hence making it easier to identify common patterns and trends that may not be obvious on the surface

Encourage the use of IoT health devices

- As the case studies have shown, EHRs are not the only vital source of information - patient generated data (collected through IoT health devices) are also as important in identifying risk factors.

THE END