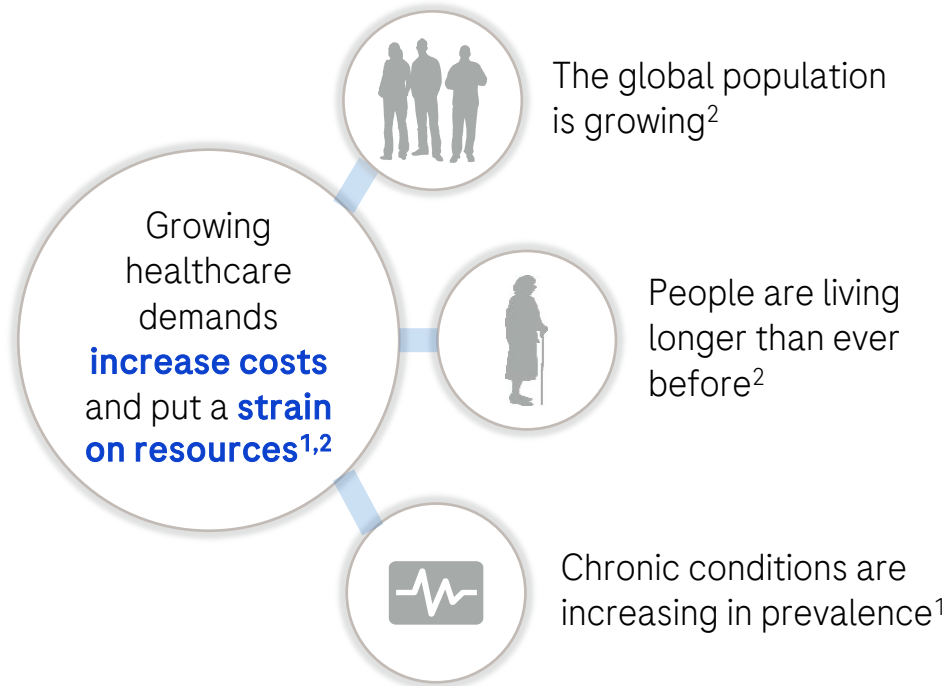# Multimodality: applying advanced analytics in the pharmaceutical and diagnostics industries to improve patient outcomes and enable personalised care

Gunther Jansen
Personalised Healthcare Centre of Excellence, Roche

Applied Machine Learning Days 2022

The current course of healthcare is unsustainable

Growing healthcare demands **increase costs** and put a **strain on resources**[1,2]

The global population is growing[2]

People are living longer than ever before[2]

Chronic conditions are increasing in prevalence[1]

Profound transformations are needed to meet healthcare demands and improve patient outcomes

1. Deloitte. 2019 Global Health Care Outlook. Link (Accessed 12 August 2021); 2. United Nations. World population prospects 2019. Link (Accessed 12 August 2021).

The convergence of science and technology provides an exceptional opportunity to transform healthcare

Data

Provide a deeper understanding of the biology of each patient

Technology

Offers new ways of collecting and analysing data and enhancing care

Machine learning

Provides new insights from data and technology
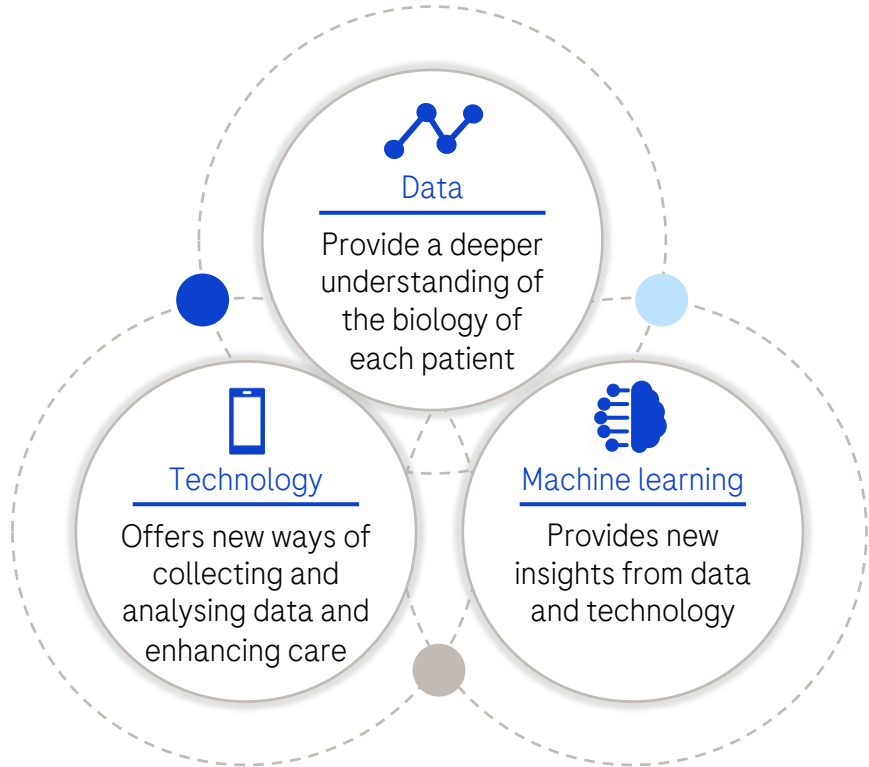
Data, technology and machine learning analytics provide opportunities to address healthcare challenges, accelerate R&D and improve patient outcomes

Our Personalised Healthcare approach

Deeper insights into disease

Technologies

Insights

Early & accurate diagnosis

Tailored care plan

Faster, more efficient clinical studies

Study data

Faster, more efficient R&D

Real-world data

Improved patient outcomes

Insights

Better, earlier go/no-go decisions

Machine learning

Continuous & remote patient management

Full access to optimal care intervention
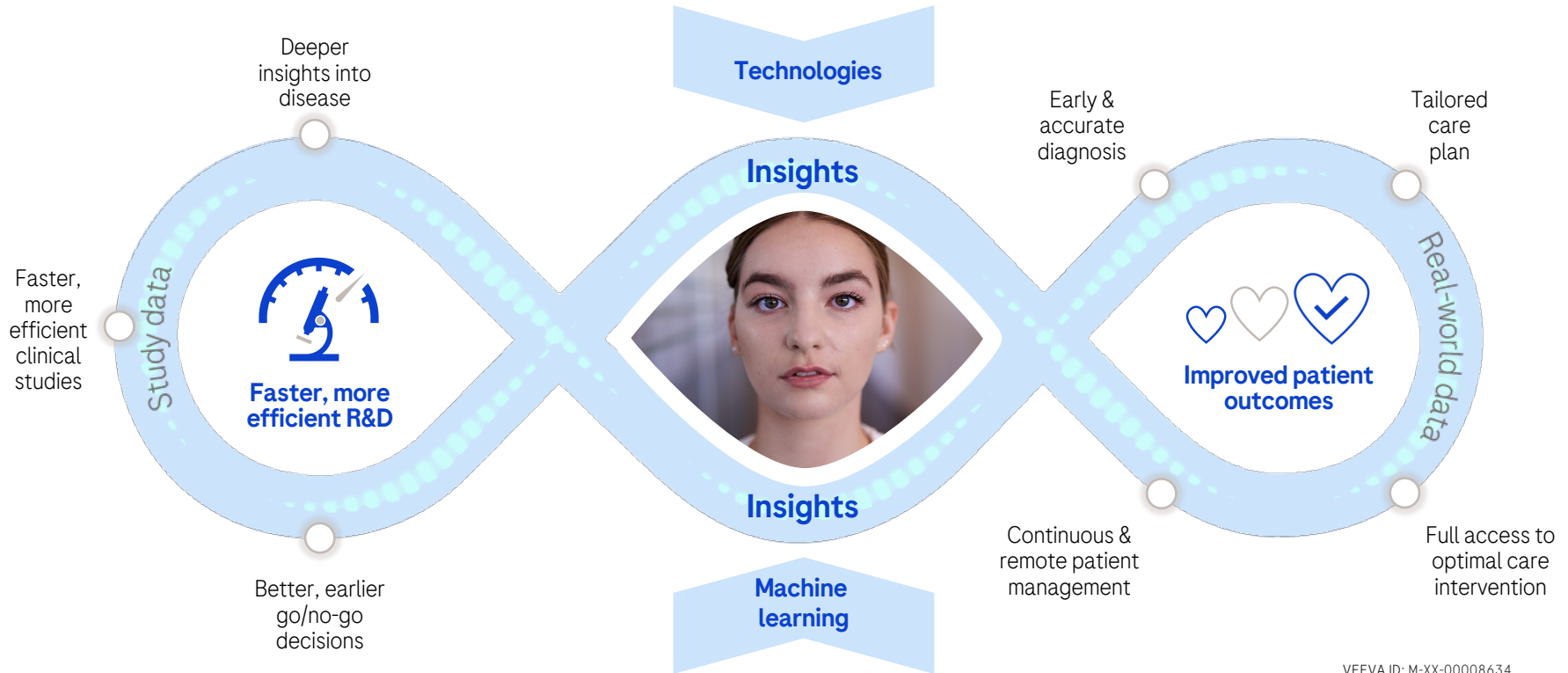
Multimodal data enable a deeper representation of the patient

Age? Sex? Smoker?

Tumour mutations?

Family history?

White-grey lung mass on X-ray?

Abnormal tissue detected on biopsy?

Cough? Night sweats?

Healthcare data volume is projected to increase by 36% per year[1]

~10 500 exabytes[3]

153 exabytes[2]

1218 exabytes[1]

2013

2018

2025

Volume of healthcare data
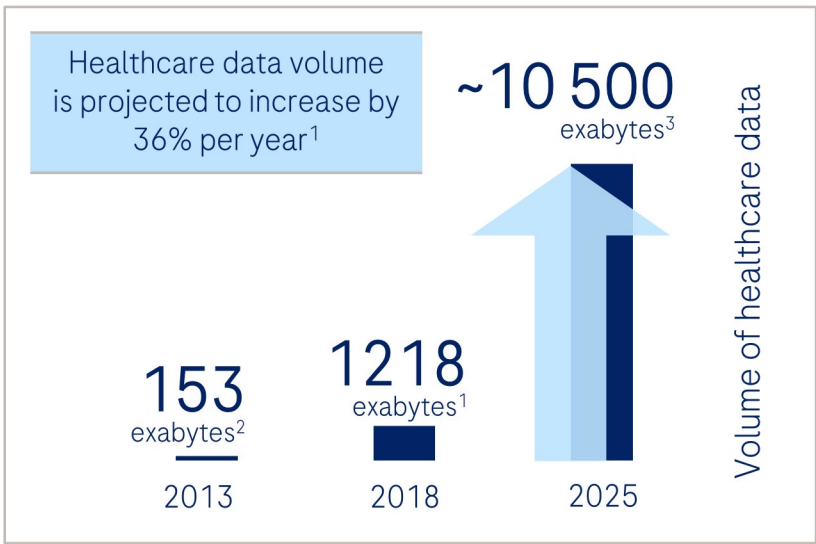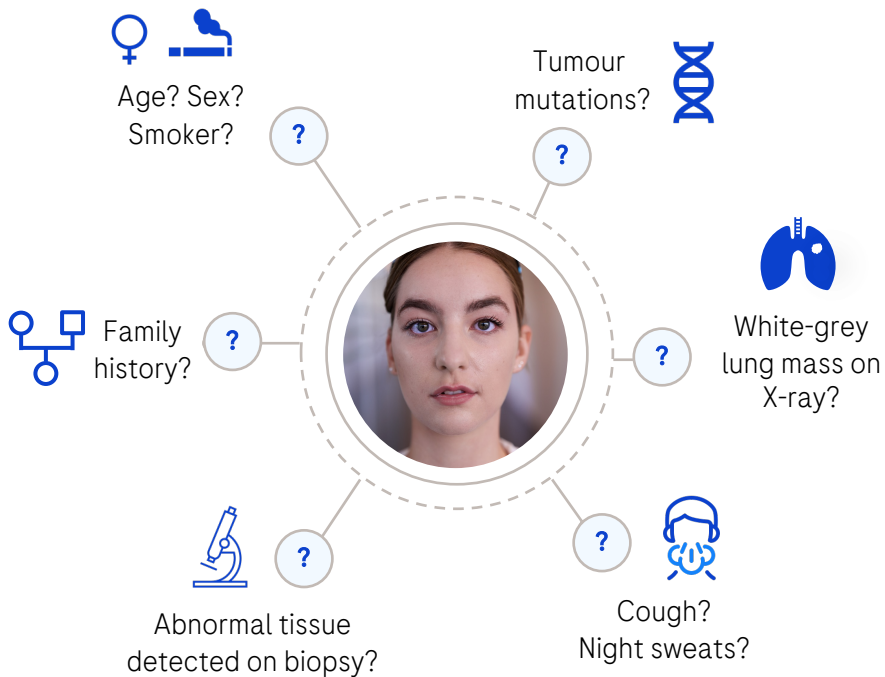
1. IDC White paper: Data Age 2025, 2018. Link (Accessed 10 March 2022); 2. EMC Digital Universe with Research and Analysis by IDC, 2014. Link (Accessed 10 March 2022); 3. IDC White paper. DATCON Level 3. 2018. Link (Accessed 10 March 2022)

Roche

Secondary multimodal data is the key to unlock Personalised Healthcare

Omics

Diagnostic laboratory values

Imaging

Histology

EMRs

Insurance claims data

Clinical trial data

Registry data

Patient-reported data (apps and wearables)
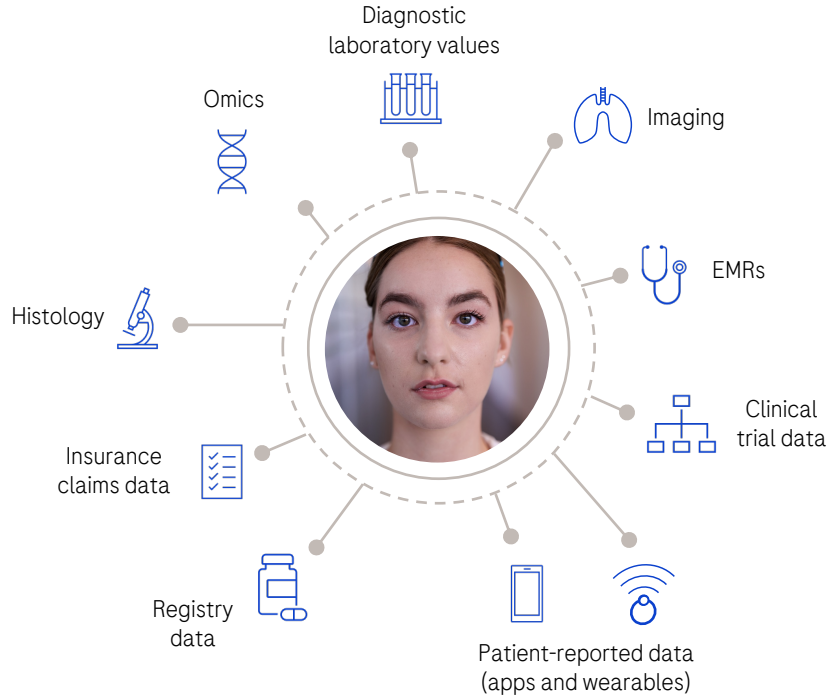
Integration challenges mean that multimodal data are rarely utilised to their full potential to benefit clinical research and clinical practice

Heterogeneity...

...noise...
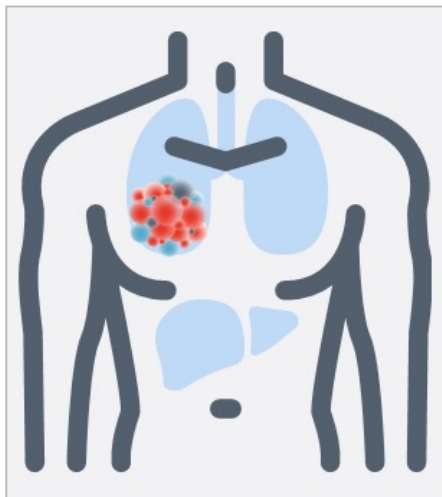
...lack of standardisation

EMR, electronic medical records

Case study
Multimodal patient representation in cancer

Secondary multimodal data are the key to unlock Personalised Healthcare

Faster, more efficient R&D

Improved patient outcomes

Case study

Multimodal patient representation in cancer

Our vision

Data of **differing** dimensionality and magnitude

Transformations

Data of **similar** dimensionality and magnitude

Task-agnostic, holistic patient representation

Systematically integrating data of different modalities from different sources

Information sharing and **uncertainty propagation**

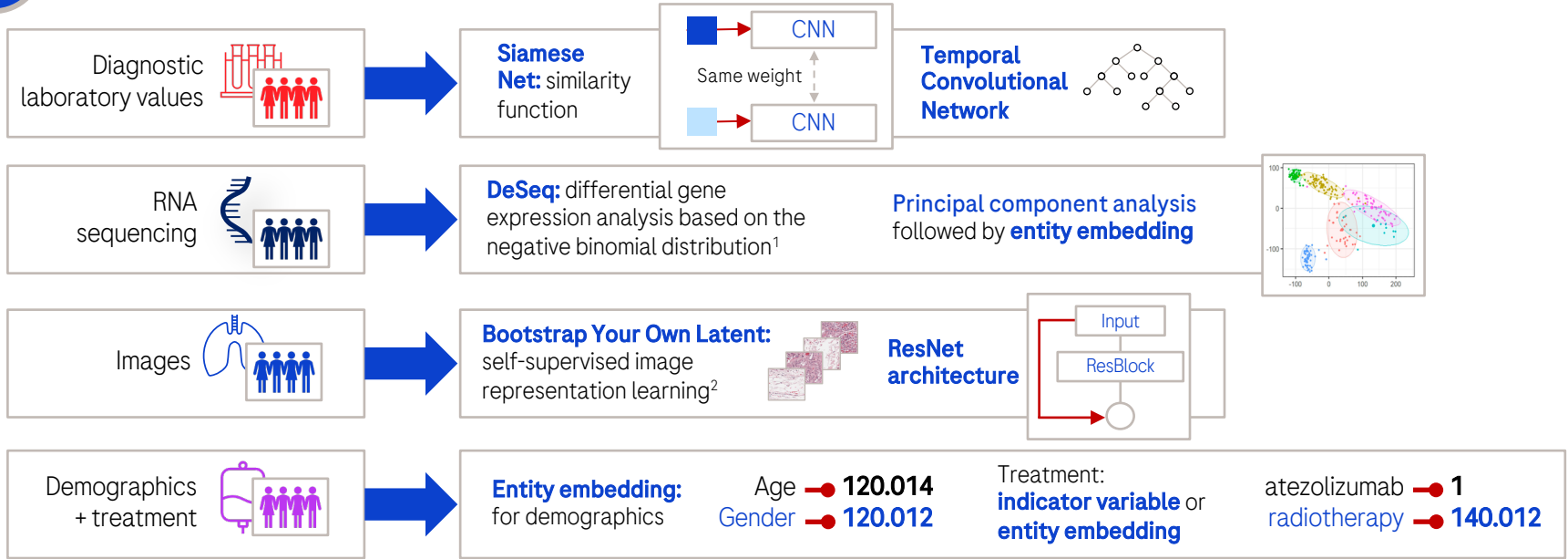**Combining independent and robust representations** to use as inputs for downstream models

8
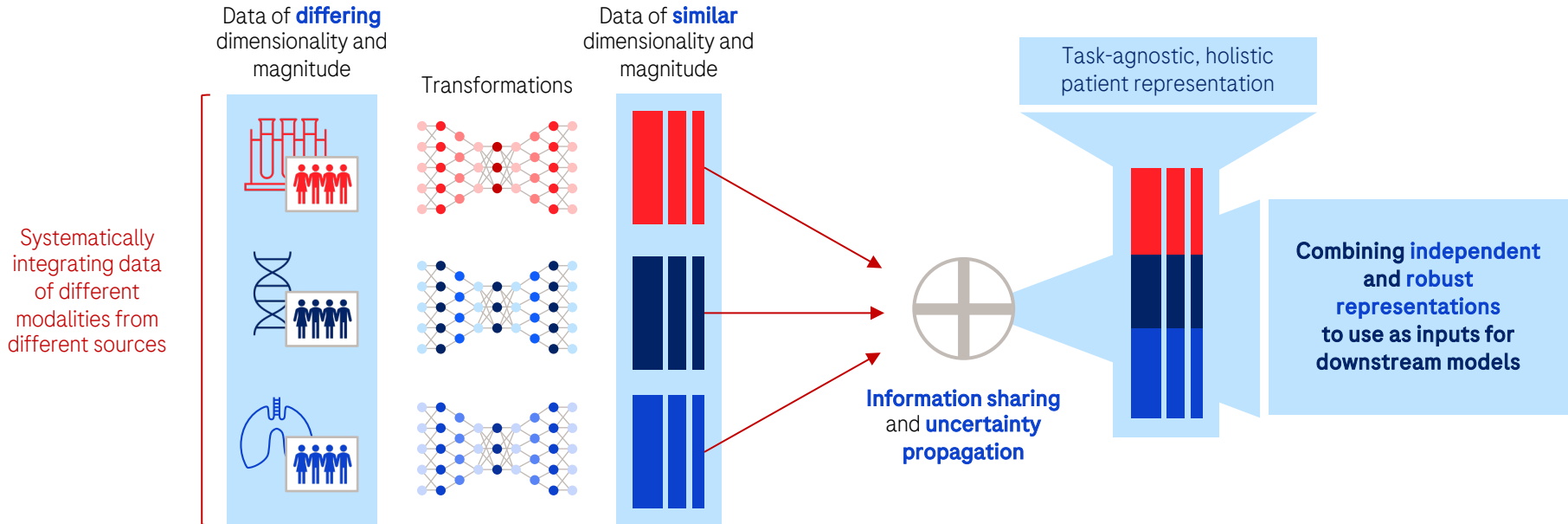
Case study

Multimodal patient representation in cancer

We are reusing our vast repository of clinical trial data and applying different representation learning methods for each data type
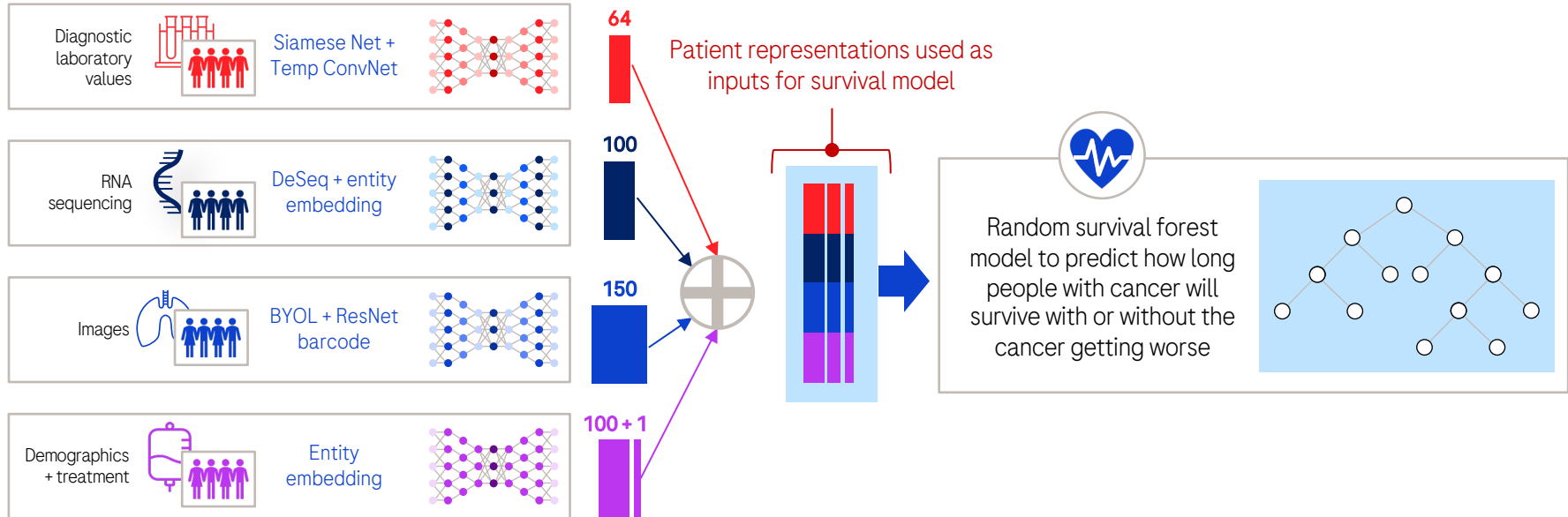
Diagnostic laboratory values → **Siamese Net:** similarity function

CNN

Same weight

CNN

Temporal Convolutional Network

RNA sequencing → **DeSeq:** differential gene expression analysis based on the negative binomial distribution[1]

Principal component analysis followed by **entity embedding**

Images → **Bootstrap Your Own Latent:** self-supervised image representation learning[2]

ResNet architecture

Input

ResBlock

Demographics + treatment → **Entity embedding:** for demographics

Age ●— **120.014**
Gender ●— **120.012**

Treatment: **indicator variable** or **entity embedding**

atezolizumab ●— **1**
radiotherapy ●— **140.012**

1. Love MI *et al. Genome Biol* 2014;15:550; Link 2. Grill J-B *et al. Adv Neural Inf Process Syst* 2020;33:21271-21284 Link. CNN, convolutional neural network; DP, digital pathology; LSTM, long short term memory; ResNet, residual network

Case study

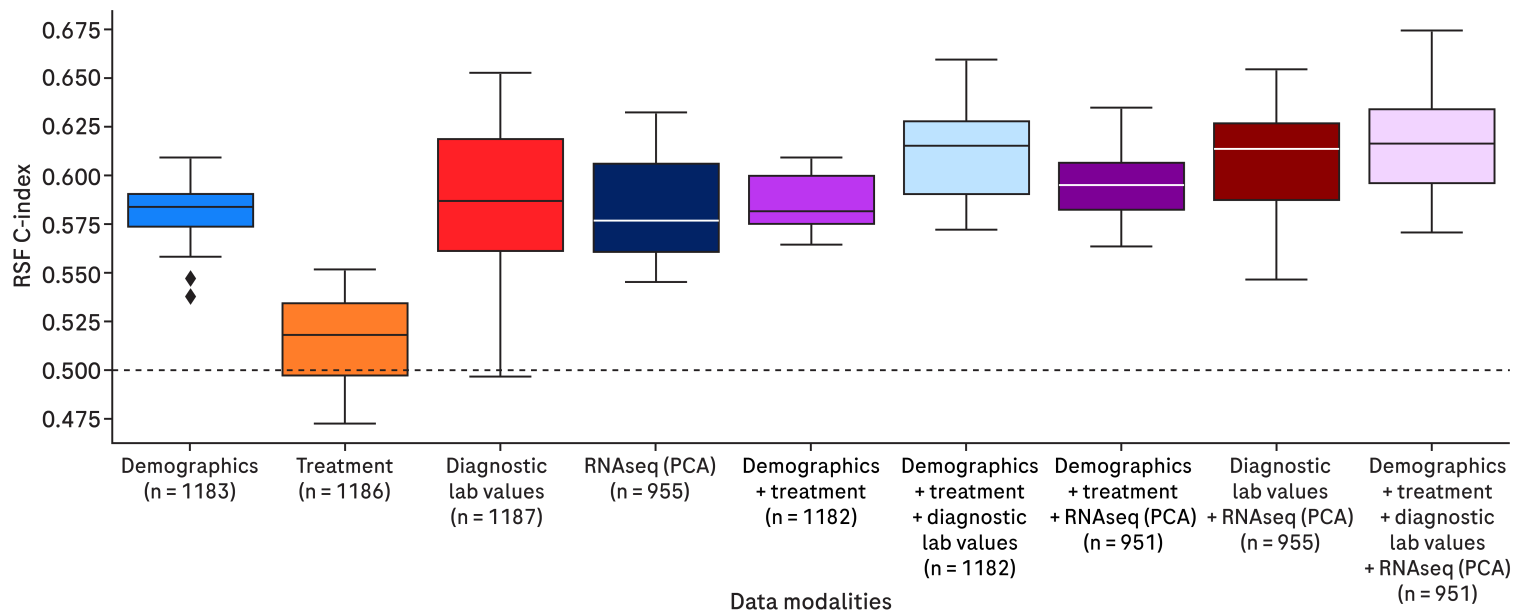## Multimodal patient representation in cancer

Status of the current pipeline

Diagnostic laboratory values		Siamese Net + Temp ConvNet
RNA sequencing		DeSeq + entity embedding
Images		BYOL + ResNet barcode
Demographics + treatment		Entity embedding

64

100

150

100 + 1

Patient representations used as inputs for survival model

Random survival forest model to predict how long people with cancer will survive with or without the cancer getting worse

BYOL, Bootstrap Your Own Latent; ResNet, Residual Neural Network; Temp ConvNet, Temporal Convolutional Network

Case study

Multimodal patient representation in cancer

Overall survival: predictive power increases with the addition of multiple data types

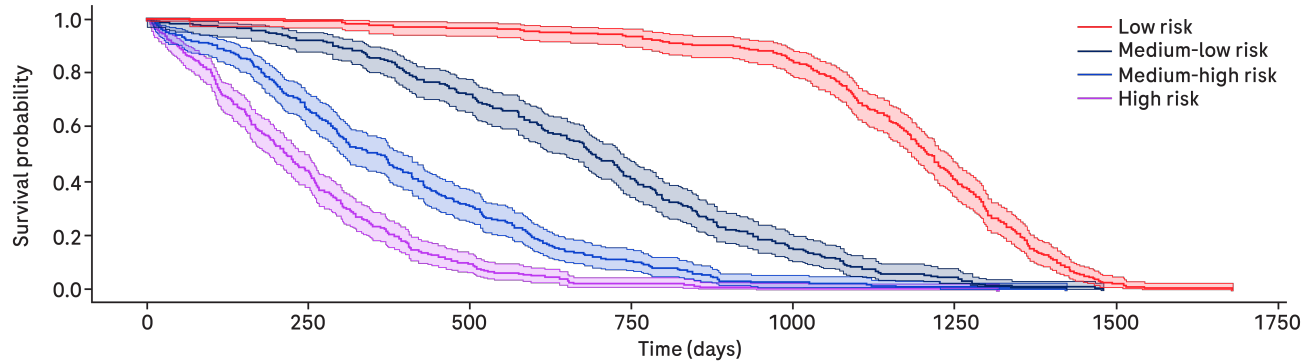PCA, principal component analysis; RSF C-index, random survival forest concordance-index

Case study

Multimodal patient representation in cancer

**The multimodal approach enables the prediction of overall survival and risk stratification of patients**

		0	250	500	750	1000	1250	1500	1750
Low risk	At risk	247	245	239	231	208	101	5	0
	Censored	0	0	0	0	0	0	0	0
	Events	0	2	8	16	39	146	242	247
Medium-low risk	At risk	230	212	166	94	34	10	0	0
	Censored	0	0	0	0	0	0	0	0
	Events	0	18	64	136	196	220	230	230
Medium-high risk	At risk	228	151	70	24	6	1	0	0
	Censored	0	0	0	0	0	0	0	0
	Events	0	77	158	204	222	227	228	228
High risk	At risk	246	108	23	5	1	1	0	0
	Censored	0	0	0	0	0	0	0	0
	Events	0	138	223	241	245	245	246	246

Case study

Multimodal patient representation in cancer

Delivering benefits to clinical research and clinical practice

...clinical research

...clinical practice
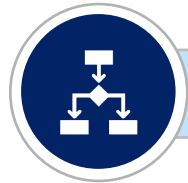
...and patients

Enabling smaller, more efficient clinical trials with enriched patient populations

Deeper, individualised, data-driven treatment decisions

Evaluating the relevance of clinical study endpoints

Deep similarity matching

Deeper insights into treatment responses
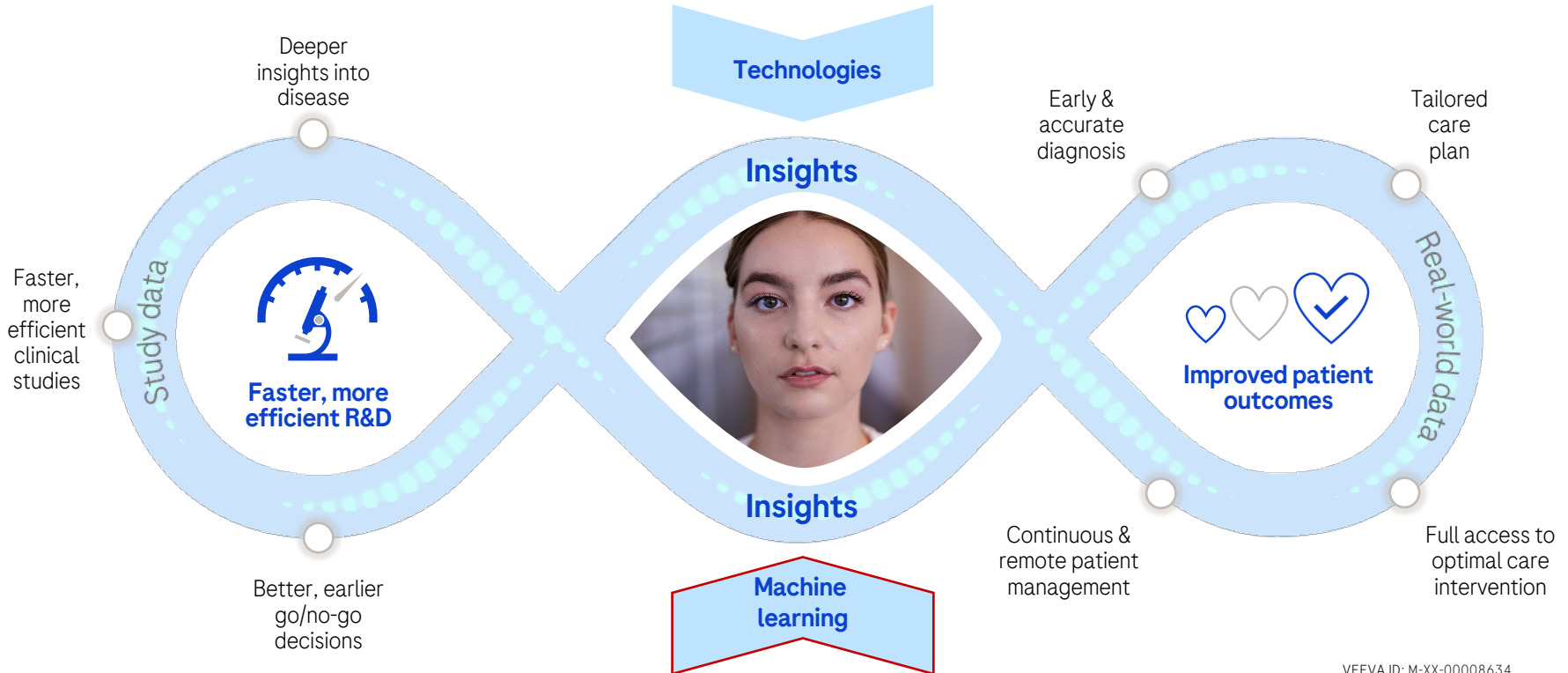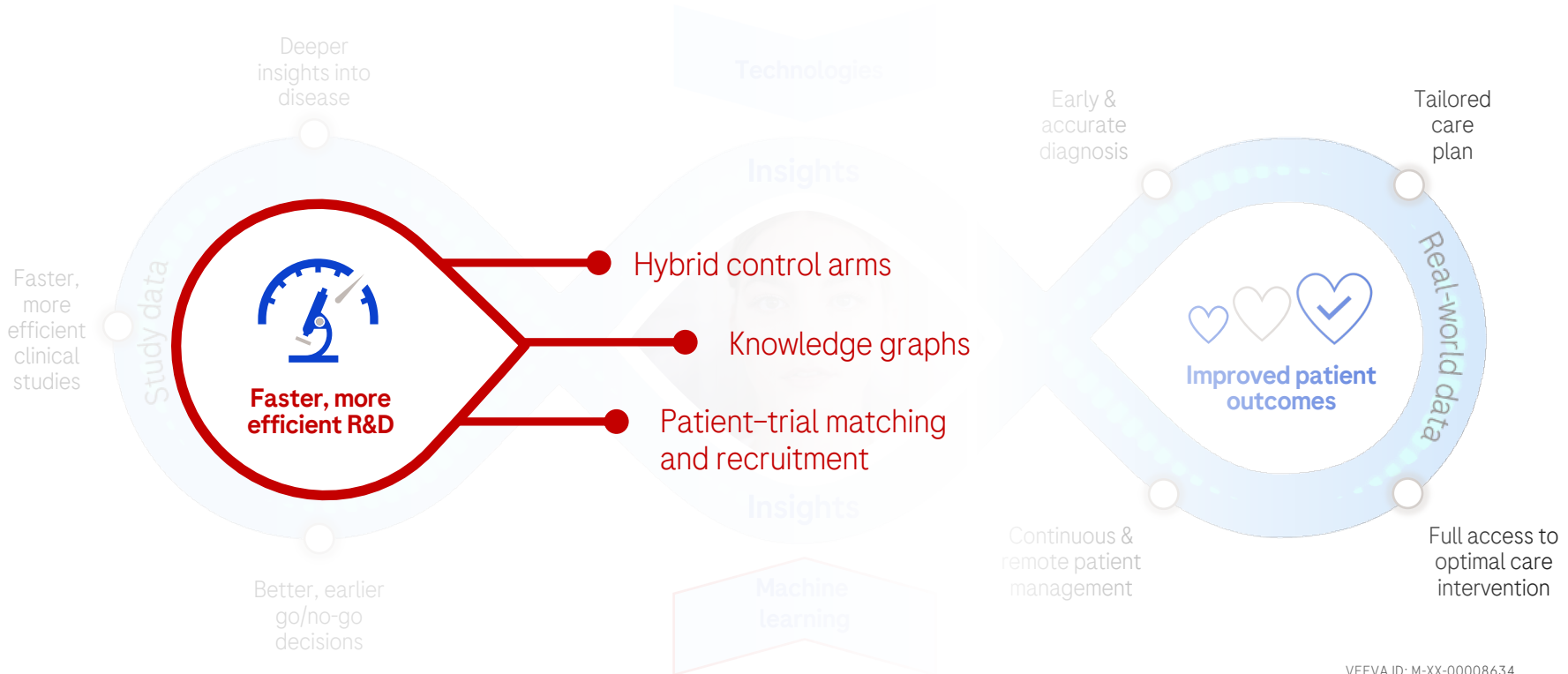
Improving the standard of care in oncology enabling patients to access the right treatment at the right time

Identifying candidate predictive biomarkers

AI is fundamental to our Personalised Healthcare approach

Roche

Technologies

Insights

Insights

Machine learning

Deeper insights into disease

Faster, more efficient clinical studies

Study data

Faster, more efficient R&D

Better, earlier go/no-go decisions

Early & accurate diagnosis

Tailored care plan

Real-world data

Improved patient outcomes

Continuous & remote patient management

Full access to optimal care intervention

AI, artificial intelligence

AI is fundamental to our Personalised Healthcare approach

Deeper insights into disease

Technologies

Insights

Faster, more efficient clinical studies

Study data

Early & accurate diagnosis

Tailored care plan

Real-world data

Faster, more efficient R&D

Hybrid control arms

Knowledge graphs

Improved patient outcomes

Patient-trial matching and recruitment

Insights

Continuous & remote patient management

Full access to optimal care intervention

Better, earlier go/no-go decisions

Machine learning

AI, artificial intelligence

AI is fundamental to our Personalised Healthcare approach

Deeper insights into disease

Technologies

Insights

Early & accurate diagnosis

Tailored care plan

Faster, more efficient clinical studies

Study data

Faster, more efficient R&D

Predictive and prognostic disease algorithms

Supporting more informed clinical decisions

Improved patient outcomes

Real-world data

Better, earlier go/no-go decisions

Insights

Machine learning

Continuous & remote patient management

Full access to optimal care intervention

AI, artificial intelligence

The Roche Advanced Analytics Network

R AAN

A global community of more than **1400 data scientists**
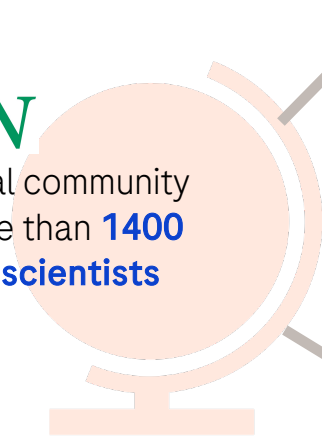
Connecting and empowering the Roche Advanced Analytics community

Fostering knowledge, sharing and developing Advanced Analytics expertise

Building foundational Advanced Analytics expertise in emerging business areas
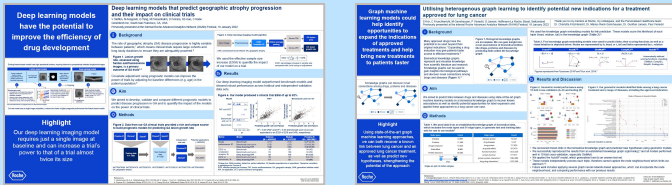
data challenges
White papers   **networking**
Advisory boards
webinars   Virtual events
hackathons
training   internships
expertise sharing

Impacting our research, business and patients by creating key insights from data

Roche at AMLD 2022

Roche Advanced Analytics Festival

Visit the Roche booth to speak to two of our poster presenters from the RAAN festival

AI & Healthcare track

Co-organised by Roche
- 09:00–15:00, Tuesday 29 March

Elif Ozkirimli, Roche

Opportunities and challenges of using Natural Language Processing for biomedical text

- AI & Industry track
- 15:23–15:43, Tuesday 29 March

AI, artificial intelligence; AMLD, Applied Machine Learning Days; RAAN, Roche Advanced Analytics Network

Acknowledgement and thanks

All those involved in Roche clinical studies

The patients and their families

The investigators and clinical study sites

Roche

Contributors

- Fei Yang
- Ben Torben-Nielsen
- Jason Lai
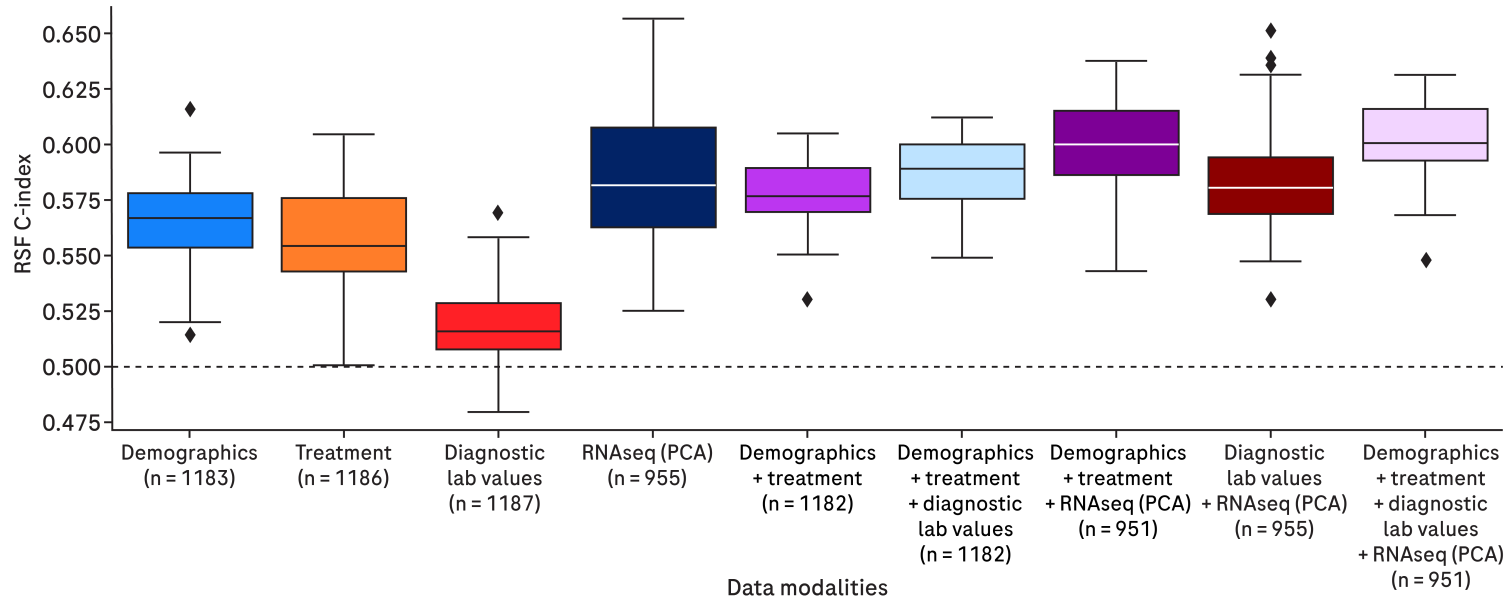- Otto Fajardo
- Phil Arnold
- Marius Garmhausen

Backup

Case study

Multimodal patient representation in cancer

PCA, principal component analysis; RSF C-index, random survival forest concordance-index

Case study
Multimodal patient representation in cancer

The multimodal approach enables the prediction of progression-free survival and risk stratification of patients

Legend: Low risk, Medium-low risk, Medium-high risk, High risk
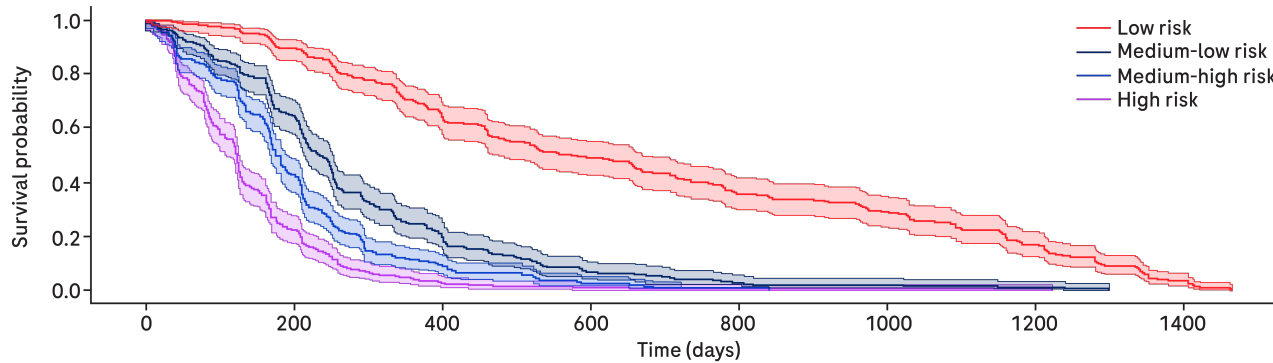
Low risk	At risk	239	214	153	117	85	69	40	8
	Censored	0	0	0	0	0	0	0	0
	Events	0	25	86	122	154	170	199	231
Medium-low risk	At risk	230	148	45	15	6	4	2	0
	Censored	0	0	0	0	0	0	0	0
	Events	0	82	185	215	224	226	228	230
Medium-high risk	At risk	242	103	21	6	1	0	0	0
	Censored	0	0	0	0	0	0	0	0
	Events	0	139	221	236	241	242	242	242
High risk	At risk	240	53	6	2	1	1	1	0
	Censored	0	0	0	0	0	0	0	0
	Events	0	187	234	238	239	239	239	240