Passion for Innovation.
Compassion for Patients.™

Daiichi-Sankyo

Opportunities and Challenges in the Use of Real-World Evidence and AI/ML in the Pharmaceutical Product Development

Hoa V. Le, MD, PhD

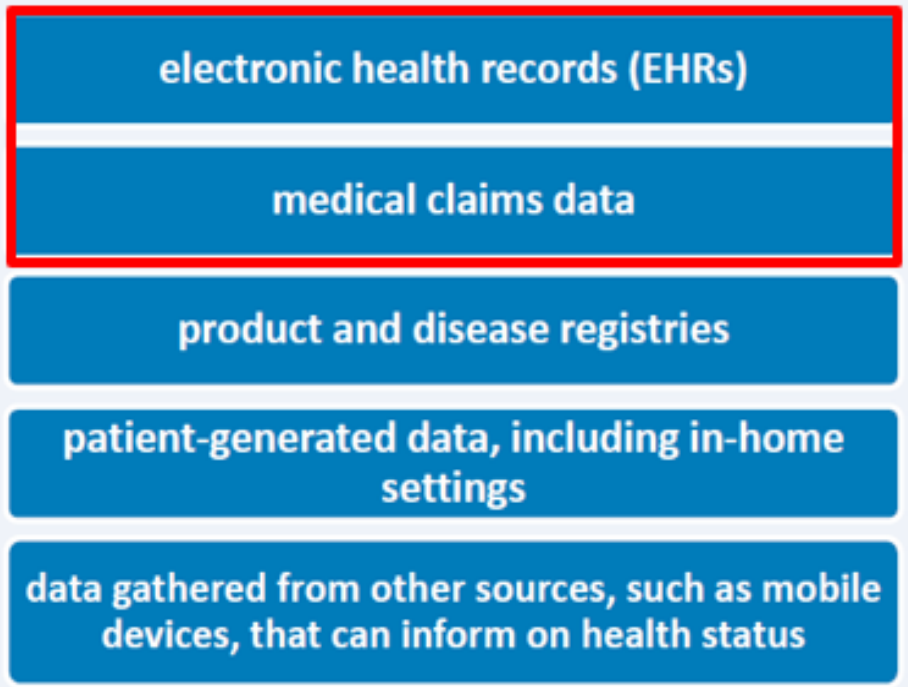
31 May 2023
Tokyo Institute of Technology

OUTLINE

- **Real-World Data (RWD) and Real-World Evidence (RWE)**
- **RWE in the product development**
- **AI/ML in the product development**
- **Case Examples**
  - External Control Arm (ECA) to support regulatory decision-making
  - Using Machine Learning algorithm to predict medication initiations and discontinuations
- **Opportunities, Challenges and Take-home messages**

‘REAL-WORLD’ DEFINITIONS

Real World Data (RWD) are data relating to patient health status and/or the delivery of health care routinely collected from a variety of sources



Real-world evidence (RWE) is the clinical evidence regarding the usage and the potential benefits or risks of a medical product derived from the analysis of RWD.

Source: Framework for FDA’s Real-World Evidence Program, Dec 2018

RWD SOURCES

Data Sources

Medical record
(e.g. EMR, paper chart)

Claims systems

Provider reported measures
(e.g. surveys, registries)

Patient reported
(e.g. PRO, social media)

Wearable device

Self-monitoring device

Evidence Types*

Patient / Caregiver Data

- Demographics
- Health status
- Lifestyle risk factors
- Treatment history
- Laboratory and radiology results
- Comorbidities
- Consumer statistics

Healthcare Provider Data

- Demographics
- Practice profile
- Treatment patterns
- Quality metrics
- Reimbursement type
- Referral patterns

Plan / Formulary Data

- Plan profile
- Benefit details
- Coverage
- Formulary/Tiers/Co-pay/Co-insurance

Treatment Data

- Diagnostic tests
- Line of therapy
- Treatment cost
- Prescription written /filled
- Treatment patterns

Outcomes Data

- Clinical
- Humanistic
- Economic
- Quality

* Non-exhaustive, not mutually exclusive categories

*Source: *IQVIA*

DATABASE – US EXAMPLES

Database Name	1. Patient Counts[1]	2. Oncology Indications[2]	3. Biomarker/ Genetic Data	4. Oncology Feature Depth[3]	5. Dataset Types, incl. Linked[4]	6. Geographic Coverage
ConcertAI Definitive Oncology	4.5M oncology patients	Broad coverage	Available	Enriched	EMR, Claims, Genomic	United States
Flatiron Health	2.4M oncology patients	Broad coverage	Available	Enriched	EMR, Genomic	United States
Guardant Health InformDB	147K oncology patients	Broad coverage	Available	Enriched	Clinical, Genomic	United States
Truven Health (IBM) MarketScan	263M patients	Available	Not available	Not enriched	EMR, Claims, Hospital	United States
IQVIA Charge Detail Master (CDM)	16.7M patients	Available	Not available	Not enriched	Hospital claims	United States
Optum EHR	102M patients	Broad coverage	Not available	Enriched	EMR, Claims	United States
Premier Healthcare Database	231M patients	Available	Not available	Not enriched	Hospital	United States
SEER Registry	160M patients	Broad coverage	Available	Enriched	Registry	United States
Medicare Current Beneficiary Survey	16,000 beneficiaries	Not available	Not available	Not enriched	Survey	United States

[1] M = million, K = thousand.

[2] Coverage of oncology indications. “Broad coverage”: full list of oncology indications and patient counts are confirmed by vendor. “Available”: full list of oncology indications and patient counts are unknown, though oncology patients are known to be included in this dataset.

[3] Indicates if additional feature enrichment (ex. chart review, natural language processing) was performed to supplement completeness of oncology features

[4] Dataset types available from this vendor and for linking

Abbreviations: EMR – electronic medical record.
SEER: Surveillance, Epidemiology, and End Results
.

RWD – EU EXAMPLES

	PHARMO (NL)	CPRD (UK)	AARHUS (DK)	Caserta (IT)	Palermo (IT)	ARS (IT)	GePaRD (DE)	IMS (ES)	IMS (BE)
Affiliation	PHARMO Institute for Drug Outcomes Research		Aarhus University, Department of Clinical Epidemiology	University of Messina	University of Messina	Agenzia regionale de sanità della Toscana	BIPS		IQVIA RWES
Name	PHARMO Database Network	CPRD		Caserta claims database	Palermo claims database	ARS	GePaRD	LPD (Longitudinal Patient Database)	Hospital Disease Database
Geographic coverage	~25% of the Dutch population	United Kingdom	National	Caserta	Palermo	Tuscany	All German regions	Spain	North of country, Brussels (no South)
Source population, N	More than 9 mn for an average of 12 yrs	>35 mn	5.7 mn	1.17mn	1.3 mnin	3.6 mn	Members of 4 German SHI providers, > 25 mn	800.000 total	1.1 mn distinct patients
Start of data collection	01/01/1998	1988	01/01/1968	01/01/2002	01/01/2005	Jan 2003	01/01/2004	Start of panel 2008 with maturity in 2014	01/01/2013
Gaps in data collection	No	No	No	No	Yes	No	No	Yes	No
if yes, please specify					Outpatient drug dispensing available until May 2018				
Frequency of database update / data collection	From once per month to once per year	Monthly	6-24 months depending	Annually	Annually	Every 2 months	Annually	Daily	Not updated since 2014 due to GDPR alignment
Database update lag time	12 months		6-24 months depending on the database	2 months	6-10 months	3-4 months	~ 18 months	1 month (data validation)	Not applicable

RWD – JAPAN EXAMPLES

	Hospital-based			Insurance-based				Pharmacy-based	
Database Name	**RWD database**	**Medical Information Database Network (MID-NET)**	**EBM Provider®**	**National Clinical Database**	**National Database of Health Insurance Claims and Specific Health Checkups of Japan (NDB)**	**JMDC Payer-based database**	**JammNet database**	**IQVIA NPA data**	**NIHON CHOUZAI Pharmacy Claims DB**
Organizer	Real World Data, Co., Ltd.	Pharmaceuticals and Medical Devices Agency	Medical Data Vision Co., Ltd.	NCD	Ministry of Health, Labour and Welfare	JMDC Inc.	JammNet	IQVIA Solutions Japan K.K.	Japan Medical Research Institute Co.,Ltd.
Start of data collection	Before 2000	2009	2008	2011	2008	2005	2008	2015	2001
Number of unique Identifiers, 10K	2300	530	3669	1500	12000	1200	216	7434	1760
Annual number of unique identifiers, 10K	151	-	1105	200	-	850	194	3012	310
% coverage of population	1.2%	-	・9% ・about 26% of total number for acute Hospitals	More than 95% of the surgeries performed in Japan	almost equal to total population in Japan	・7% ・About 28% of the people belong to health insurance association	approx2% of insurance programs	approx. 19.8% of total extramural dispense Claims	1.7% Of annual prescription
Age distribution: 65-74 years old [%]	12.8	-	16.4	-	-	4	17	15.1	17.5
Age distribution: 75 years old or older[%]	27	-	18.7	-	-	0	6	18.3	21.4
Latest data	1 month	No lag time	2 months	2 months	-	5 months	3 months	1 month	Previous day

Source: Website of Japanese Society for Pharmacoepidemiology https://www.jspe.jp/committee/020/0210/, as of August 2021)

DATABASE SUMMARY

- Each data source should be evaluated to determine whether the available information is **appropriate** for addressing a specific **research questions**

- When evaluating a study design and a specific data source, important to consider also:
  - Relevance of data source
  - Patient population
  - Data capture
  - Information content and missing data considerations
  - Validation

Source: FDA website 2021 Guidance for Industry (fda.gov)

Passion for Innovation.
Compassion for Patients.™

Daiichi-Sankyo

RWE in the Product Development

RWE POTENTIAL FOR PHARMACEUTICAL COMPANIES

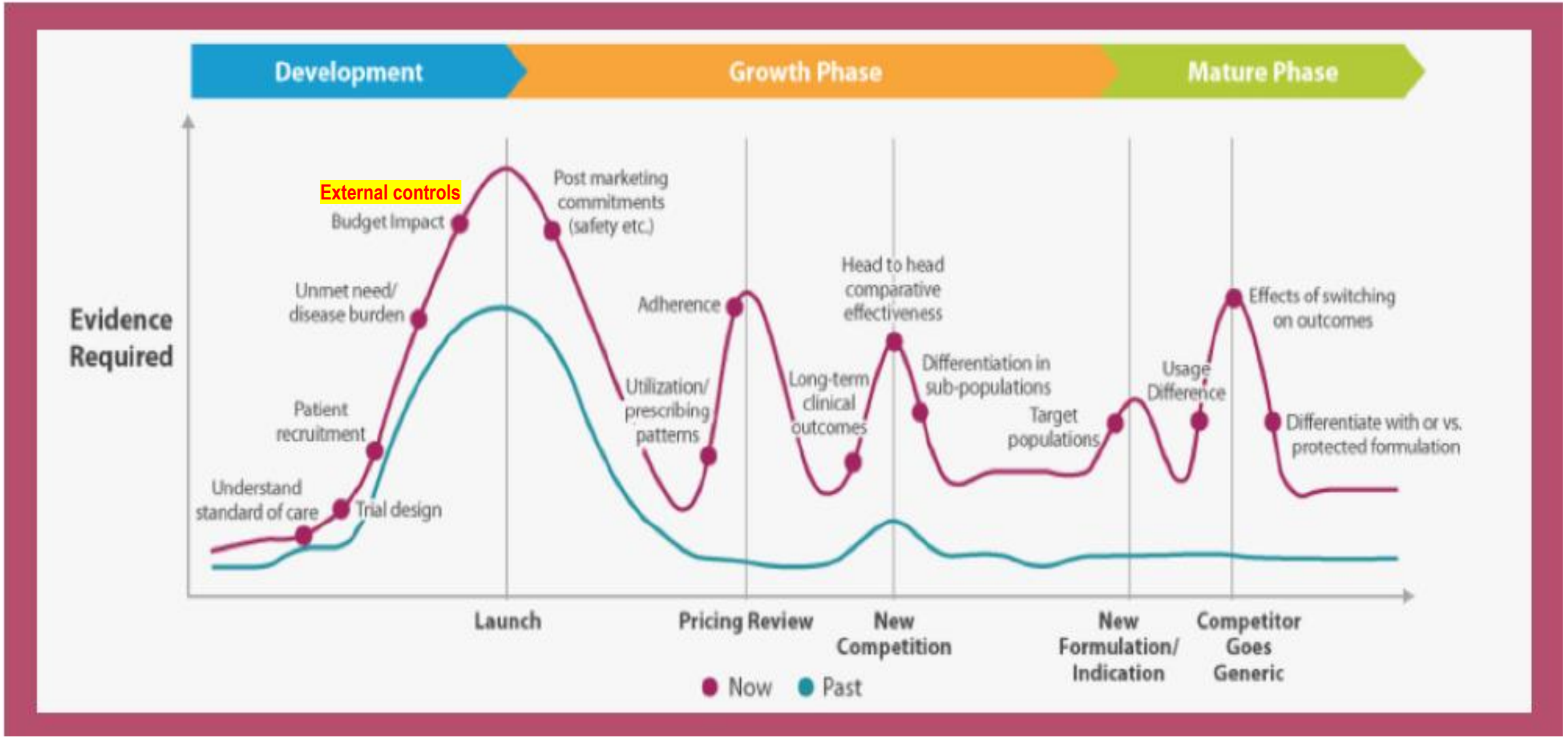
More companies are realizing the value of real-world evidence (RWE) – the clinical evidence regarding the potential benefits or risks of a therapy generated from data relating to patient health or the delivery of health care

RWE can add at every stage of a therapy's lifecycle:

- inform innovation
- trial design
- regulatory filings
- product's safe and appropriate use in a post-approval setting
- personalized medicine

Source: PinkSheet Feb 07 2022

RWE APPLICATIONS ACROSS THE PRODUCT LIFECYCLE



Adapted from McKinsey Practice Perspectives on RWD.

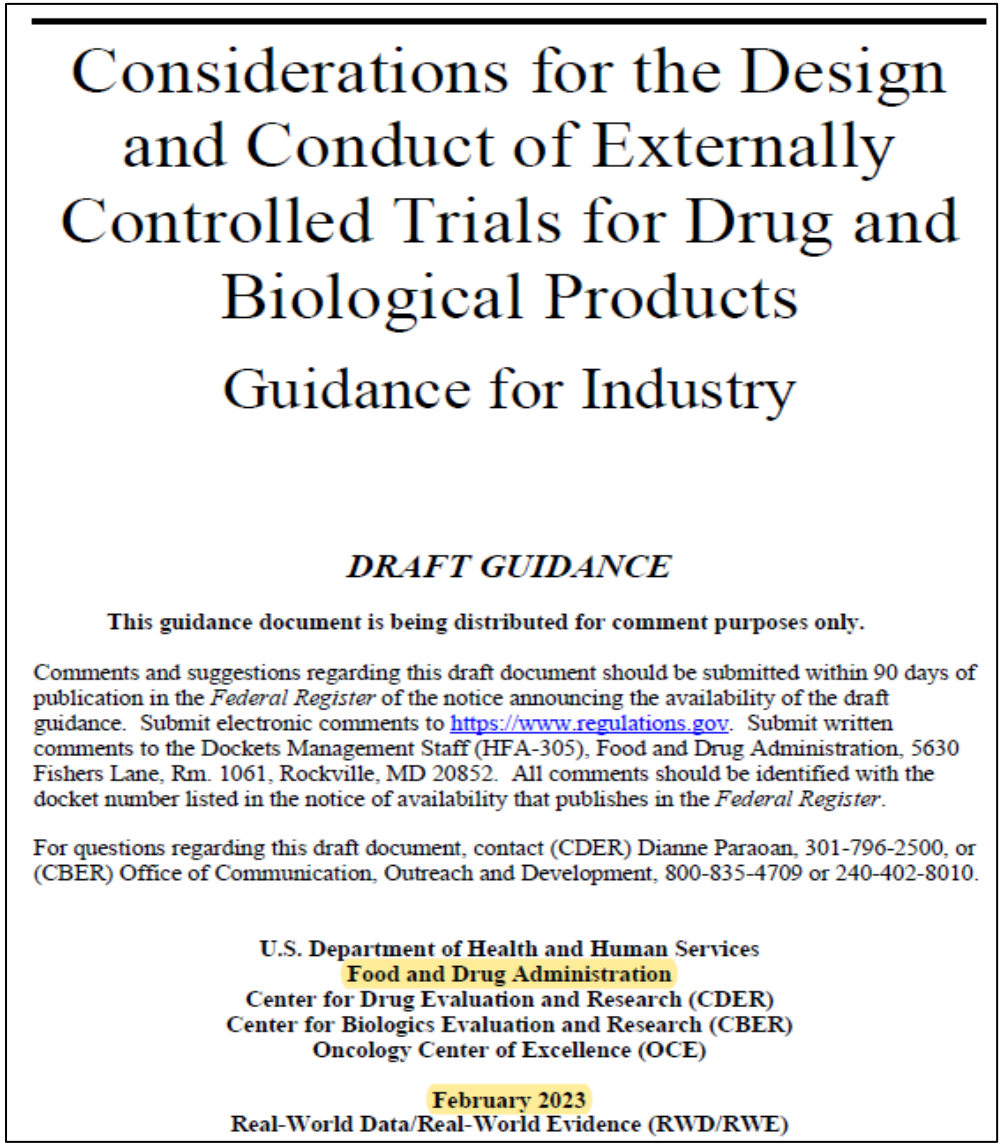
EXTERNAL CONTROL ARM (ECA)

External Control Arm

- RWD/E is used as an external benchmark for context or as a formal comparator to a single-arm trial.
- Potentially support new product approvals and new indications.

Situation: A randomized clinical trial may not be feasible or may be very challenging to conduct

- **Accelerated approval**/Label inclusion or expansion
  - **Potential time & resource saving**
- ECA to provide **comparative effectiveness**:
  - Competitive advantage
  - Contextualize and interpret the outcomes of the **single-arm** clinical trials



Considerations for the Design and Conduct of Externally Controlled Trials for Drug and Biological Products

Guidance for Industry

DRAFT GUIDANCE

This guidance document is being distributed for comment purposes only.

Comments and suggestions regarding this draft document should be submitted within 90 days of publication in the *Federal Register* of the notice announcing the availability of the draft guidance. Submit electronic comments to https://www.regulations.gov. Submit written comments to the Dockets Management Staff (HFA-305), Food and Drug Administration, 5630 Fishers Lane, Rm. 1061, Rockville, MD 20852. All comments should be identified with the docket number listed in the notice of availability that publishes in the *Federal Register*.

For questions regarding this draft document, contact (CDER) Dianne Paraoan, 301-796-2500, or (CBER) Office of Communication, Outreach and Development, 800-835-4709 or 240-402-8010.

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)
Oncology Center of Excellence (OCE)

February 2023
Real-World Data/Real-World Evidence (RWD/RWE)

CONDITIONS FOR SUCCESSFUL ECA CASES

Daiichi-Sankyo

In General

- Rare disease/ biomarker/ high unmet needs
- Randomization unethical, infeasible
- Novel therapeutic approach/technology
- Large expected effect size for the new TRT (e.g., >20% higher response, HR=.54)
- Single-arm pivotal trial for registration
- No recognized or scarce information of SOC

For Regulatory Decision Making

- Sufficient time to plan an ECA study in the data package and inform regulatory agencies to get their buy-in for the protocol/analysis plan before the study conduct
- Regulatory agencies recognize the use of ECA to evaluate relative benefit
- Fit-for-purpose database (**sufficient sample size and power**, available exposures, outcomes and key confounders, compatible time period) is obtainable

Can support HTA, HEOR/Market Access and Commercial

Example of enough sample size and power demonstrated significant p<0.01 in an external control arm.

Endpoint	Estimate sACC ($n = 257$)	Estimate LTAC ($n = 257$)	RR (95% CI), *p*
ORR, %	38.8	73.8	1.9 (1.6–2.3), <0.0001
CR rate, %	24.1	50.1	2.1 (1.6–2.8), <0.0001
			HR (95% CI), *p*
Median OS, mo	6.8	23.5	0.52 (0.40–0.68), <0.0001
Median PFS, mo	2.2	3.5	0.60 (0.48–0.75), <0.0001

LEUKEMIA & LYMPHOMA
2023, VOL. 64, NO. 3, 573–585
https://doi.org/10.1080/10428194.2022.2160200

Taylor & Francis
Taylor & Francis Group

ORIGINAL ARTICLE

OPEN ACCESS Check for updates

Use of a real-world synthetic control arm for direct comparison of lisocabtagene maraleucel and conventional therapy in relapsed/refractory large B-cell lymphoma

Hoa Van Le[a*], Kim Van Naarden Braun[b], Grzegorz S. Nowakowski[c], David Sermer[d], John Radford[e],

BEST PRACTICES IN RWE USE – PROSPECTIVE RESEARCH FRAMEWORK

1. **Identify and define the problem**
2. **A priori**, determine and declare that a study purpose, is it a Hypothesis Evaluation Treatment Effectiveness (HETE) study or an Exploratory study.
3. Post a HETE study protocol and analysis plan on a public study registration site prior to conducting the study analysis.
4. Include key stakeholders (patients, caregivers, clinicians, clinical administrators, HTA/payers, regulators, manufacturers) in designing, conducting, and disseminating HETE studies.
5. Perform HETE studies on a different data source and population than the one used to generate the hypotheses to be tested unless it is not feasible (e.g., another data set is not available)
6. Publish HETE study results with attestation to conformance and/or deviation from the study protocol and original analysis plan.

Source: ISPOR-ISPE 2017

Passion for Innovation.
Compassion for Patients.™

Daiichi-Sankyo

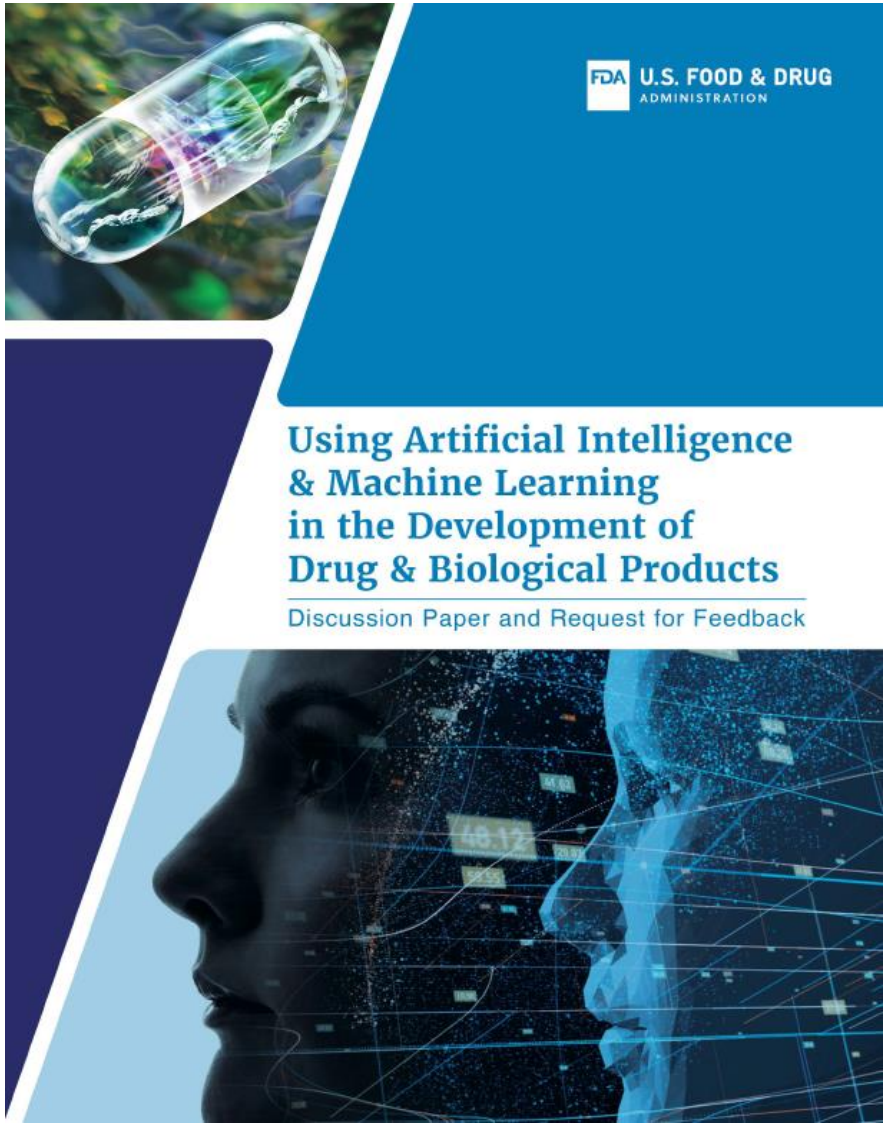
AI/ML in the Product Development

AI, ML & NLP*

Artificial Intelligence (**AI**): A branch of computer science, statistics, and engineering that uses algorithms or models to perform tasks and exhibit behaviors such as learning, making decisions, and making predictions.

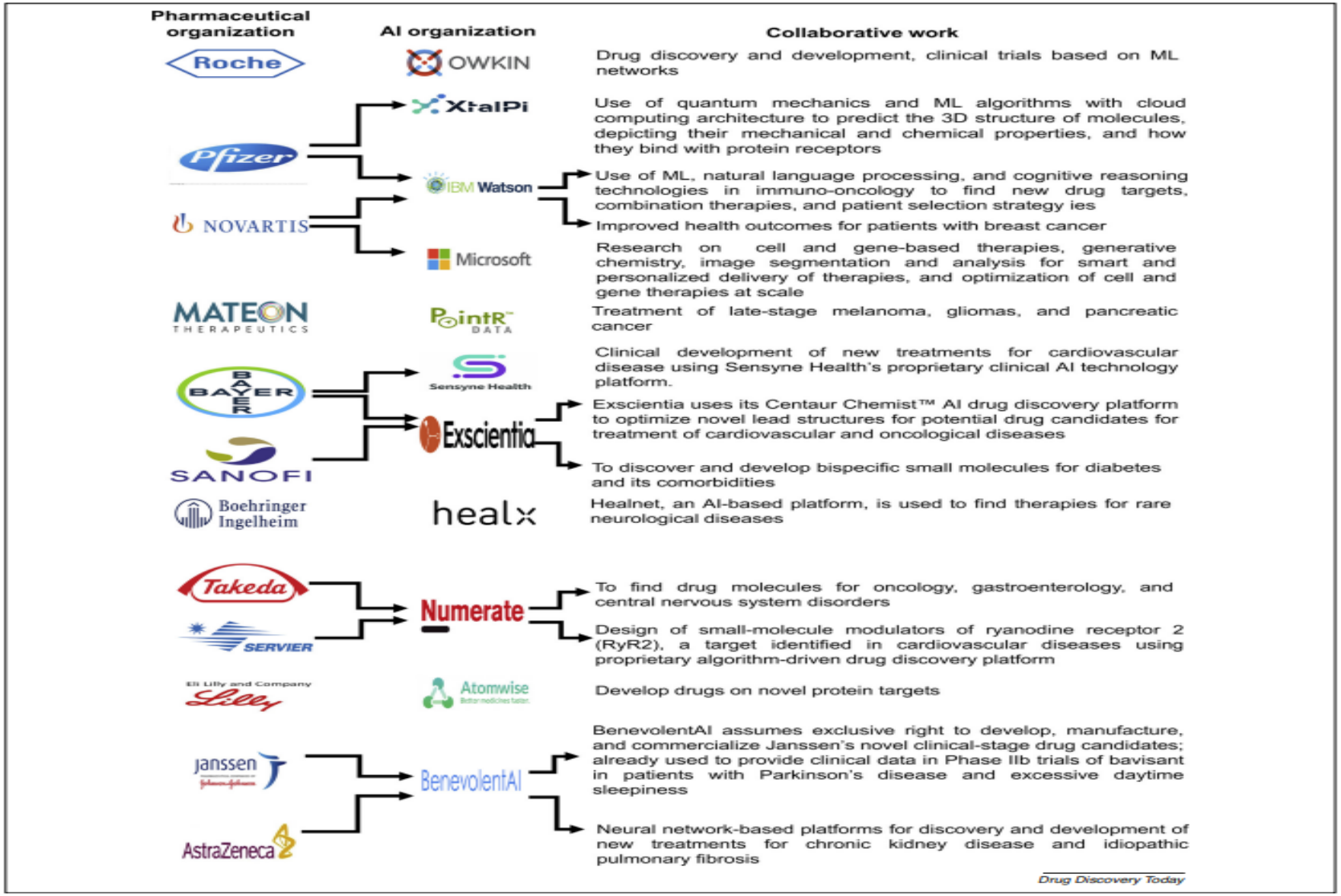
Machine Learning (**ML**): A subset of AI that allows ML models to be developed by ML training algorithms through analysis of data, without being explicitly programmed.

Natural Language Processing (**NLP**): The branch of computer science, specifically the branch of AI, concerned with giving computers the ability to understand text and spoken words in much the same way human beings can.



*Source: *FDA 19 May 2023, https://www.fda.gov/science-research/science-and-research-special-topics/artificial-intelligence-and-machine-learning-aiml-drug-development*

PHARMA COMPANIES & AI/ML ORGANIZATIONS



Source: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7577280/pdf/main.pdf

AI/ML IN ENTIRE VALUE CHAIN

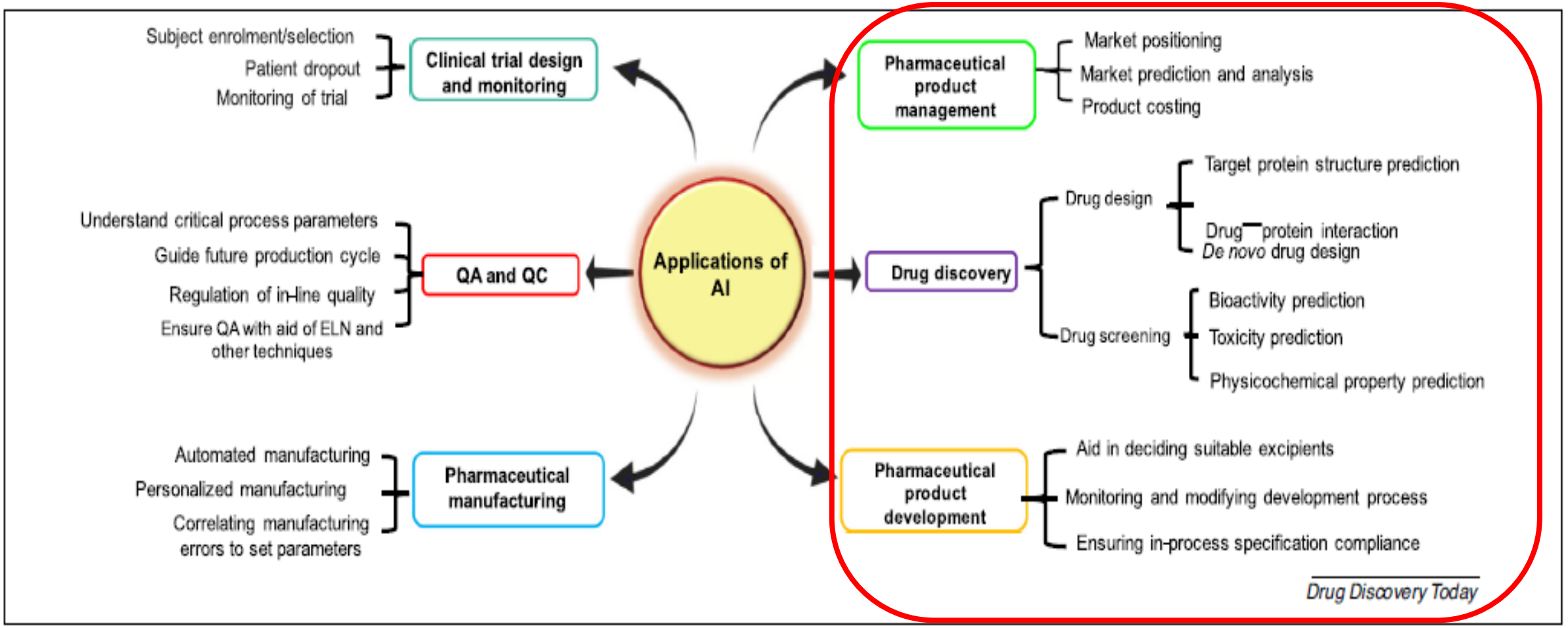


FIGURE 2

Applications of artificial intelligence (AI) in different subfields of the pharmaceutical industry, from drug discovery to pharmaceutical product management.

Source: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7577280/pdf/main.pdf

AI/ML – GENOME-WIDE

RESEARCH ARTICLE

Machine learning on genome-wide association studies to predict the risk of radiation-associated contralateral breast cancer in the WECARE Study

Sangkyu Lee[1], Xiaolin Liang[2], Meghan Woods[2], Anne S. Reiner[2], Patrick Concannon[3], Leslie Bernstein[4], Charles F. Lynch[5], John D. Boice[6], Joseph O. Deasy[1], Jonine L. Bernstein[2‡], Jung Hun Oh[1‡*]

ARTICLE

Genome-wide Modeling of Polygenic Risk Score in Colorectal Cancer Risk

Minta Thomas,[1] Lori C. Sakoda,[1,2] Michael Hoffmeister,[3] Elisabeth A. Rosenthal,[4] Jeffrey K. Lee,[2]

AI/ML – RESPONSES

BMJ Open

Identification of responders to inhaled corticosteroids in a chronic obstructive pulmonary disease population using cluster analysis

David R Hinds,[1] Rachael L DiSantostefano,[1] Hoa V Le,[1,2] Steven Pascoe[3]

CLINICAL CANCER RESEARCH | PRECISION MEDICINE AND IMAGING

Using Machine Learning Algorithms to Predict Immunotherapy Response in Patients with Advanced Melanoma AC

Paul Johannet[1], Nicolas Coudray[2,3], Douglas M. Donnelly[4], George Jour[5], Irineu Illa-Bochaca[4], Yuhe Xia[6],

AI/ML – CANCER RISKS

Machine learning identifies interacting genetic variants contributing to breast cancer risk: A case study in Finnish cases and controls

Hamid Behravan[1], Jaana M. Hartikainen[1], Maria Tengström[2,3], Katri Pylkäs[4], Robert Winqvist[4], Veli-Matti Kosma[1,5] & Arto Mannermaa[1,5]

We propose an effective machine learning approach to identify group of interacting single nucleotide polymorphisms (SNPs), which contribute most to the breast cancer (BC) risk by assuming dependencies among BCAC iCOGS SNPs. We adopt a gradient tree boosting method followed by an adaptive iterative SNP search to capture complex non-linear SNP-SNP interactions and consequently, obtain group of interacting SNPs with high BC risk-predictive potential. We also propose a support vector machine formed by the identified SNPs to classify BC cases and controls. Our approach achieves mean average precision (mAP) of 72.66, 67.24 and 69.25 in discriminating BC cases and controls in KBCP, OBCS and merged KBCP-OBCS sample sets, respectively. These results are better than the mAP of 70.08, 63.61 and 66.41 obtained by using a polygenic risk score model derived from 51 known BC-associated SNPs, respectively, in KBCP, OBCS and merged KBCP-OBCS sample sets. BC subtype analysis further reveals that the 200 identified KBCP SNPs from the proposed method performs favorably in classifying estrogen receptor positive (ER+) and negative (ER−) BC cases both in KBCP and OBCS data. Further, a biological analysis of the identified SNPs reveals genes related to important BC-related mechanisms, estrogen metabolism and apoptosis.

Machine Learning Algorithms for Predicting the Recurrence of Stage IV Colorectal Cancer After Tumor Resection

Yucan Xu, Lingsha Ju, Jianhua Tong, Cheng-Mao Zhou* & Jian-Jun Yang*

The aim of this study is to explore the feasibility of using machine learning (ML) technology to predict postoperative recurrence risk among stage IV colorectal cancer patients. Four basic ML algorithms were used for prediction—logistic regression, decision tree, GradientBoosting and lightGBM. The research samples were randomly divided into a training group and a testing group at a ratio of 8:2. 999 patients with stage 4 colorectal cancer were included in this study. In the training group, the GradientBoosting model's AUC value was the highest, at 0.881. The Logistic model's AUC value was the lowest, at 0.734. The GradientBoosting model had the highest F1_score (0.912). In the test group, the AUC Logistic model had the lowest AUC value (0.692). The GradientBoosting model's AUC value was 0.734, which can still predict cancer progress. However, the gbm model had the highest AUC value (0.761), and the gbm model had the highest F1_score (0.974). The GradientBoosting model and the gbm model performed better than the other two algorithms. The weight matrix diagram of the GradientBoosting algorithm shows that chemotherapy, age, LogCEA, CEA and anesthesia time were the five most influential risk factors for tumor recurrence. The four machine learning algorithms can each predict the risk of tumor recurrence in patients with stage IV colorectal cancer after surgery. Among them, GradientBoosting and gbm performed best. Moreover, the GradientBoosting weight matrix shows that the five most influential variables accounting for postoperative tumor recurrence are chemotherapy, age, LogCEA, CEA and anesthesia time.

AI/ML – NLP

VALUE IN HEALTH 22 (2019) 77–84

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/jval

Methodology

Identifying Patients With Relapsing-Remitting Multiple Sclerosis Using Algorithms Applied to US Integrated Delivery Network Healthcare Data

Hoa Van Le, MD, PhD [1], Chi Thi Le Truong, MD, PhD [2], Aaron W.C. Kamauu, MD, MS, MPH [1,3], John Holmén, PhD, MS [4], Christopher Fillmore, MD, MS [4], Monica G. Kobayashi, PhD, MBMA [1], Canter Martin, MBA [1], Meritxell Sabidó, MD, MS [5], Schiffon L. Wong, MPH [6,]*

Drug Saf (2017) 40:317–331
DOI 10.1007/s40264-016-0491-0

ORIGINAL RESEARCH ARTICLE

Evaluation of Facebook and Twitter Monitoring to Detect Safety Signals for Medical Products: An Analysis of Recent FDA Safety Alerts

Carrie E. Pierce[1] · Khaled Bouri[2] · Carol Pamer[2] · Scott Proestel[2] · Harold W. Rodriguez[1] · Hoa Van Le[1] · Clark C. Freifeld[1,3] · John S. Brownstein[1] · Mark Walderhaug[2] · I. Ralph Edwards[4] · Nabarun Dasgupta[1]

KEY STEPS FOR ML PROJECT

1. **Identify and define the problem**
2. **Identify the relevant data to solve the problem**
3. Prepare your data and make sure it is clean, secured, and governed
4. Select several machine learning algorithms applicable to your data and business challenge
5. Train the algorithm to create the models
6. Evaluate your models to find the best performing algorithm
7. Make predictions based on new, incoming data
8. Assess the validity of your predictions
9. Refine and repeat; the information you gather from analyzing the validity of predictions is then fed back into the machine learning cycle to help improve accuracy in the next iteration

Building a machine learning application is an iterative process. One can't simply train a model once and leave it alone

Source: IBM Machine Learning for Dummies (IBM Limited Edition 2018)

External Control Arm – Case Example

A COMPARISON OF IDE-CEL (ABECMA) WITH REAL-WORLD OUTCOMES IN RELAPSED AND REFRACTORY MULTIPLE MYELOMA

MOTIVATION

- **External Control Arm (ECA) is one of the effective approaches** for evaluating the **relative benefits and risks** of a new therapy that direct evidence only comes from **single-arm** or uncontrolled studies.
- **Regulatory agencies recognize the option of using ECA** to demonstrate new treatment efficacy **for accelerated approval** when
  - a disease is **rare**
  - has **no satisfactory treatment**
  - the new treatment **appears very promising** based on preliminary clinical data.
- **Multiple myeloma** is a **rare hematologic malignancy** with ~160,000 newly diagnosed cases and 106,000 deaths worldwide in 2018.
- There is **no clear consensus on the optimal therapy or standard of care**.

MOTIVATION & GOAL

- Phase 1 ide-cel (**Abecma**) clinical trial (NCT02658929) reported a confirmed **ORR** of **76%** and a median **PFS** of **8.8 months** in triple-class exposed patients with relapsed and refractory multiple myeloma (RRMM)**.**
- Previous retrospective data for patients with RRMM from 14 different US academic institutions reported a median **PFS of only 3−4 months**.
- **Patient-level data** in triple-class exposed RRMM patients **are not well characterized** and the **limited data** vary across geographies.
- **Use of ECA was advised by EMA** for demonstration of significant benefit as lack of control arm poses a challenge with establishment of significant clinical benefit based on the single arm trial.

OBJECTIVES OF THE ECA STUDY

- **To describe** the **demographics, disease characteristics, treatment patterns,** and **clinical outcomes** in real-world RRMM patients with characteristics similar to the study population of phase 2 **KarMMa** single arm trial (NCT03361748).

- **To compare clinical outcomes** from the real-world RRMM patients treated with currently available therapies and the patients treated with ide-cel in the KarMMa study.

EXPERIMENTAL & EXTERNAL CONTROL ARMS

EXPERIMENTAL ARM (n=128)

- The clinical trial population in the phase 2 **KarMMa** (NCT03361748).
- Of **140** patients enrolled, **128** received ide-cel with a median follow-up of 13.3 months.

EXTERNAL CONTROL ARM (n=190)

- Real-world patient-level data were collected from 6 data sources and merged into a **single common data model** (**ADaM** format).
- Initiated subsequent (index) therapy and had disease assessment.

+

Inclusion and exclusion criteria included:

- ≥18 years old
- received ≥3 prior regimens
- was refractory to their last regimen
- had measurable disease
- had adequate organ function.

RWD SOURCES

- **Internal Connect® MM registry** which currently follows 3011 patients from ~250 sites in the US.
- **Clinical trial site data was acquired and managed by IQVIA** and obtained through manual chart abstraction into the study electronic case report form.
- **The external electronic health records (EHRs)** included**:**
  - **Flatiron Health**
  - **COTA Health**
  - **Guardian Research Network** - a nationwide clinical and molecular oncology network*
  - **M2Gen -** an alliance of 19 of the nation's leading cancer centers**

*https://www.guardianresearch.org/; **https://www.m2gen.com/our-approach?hsLang=en

PROPENSITY SCORE BALANCING

To ensure a balance of baseline characteristics, **PS trimmed IPTW** was used via following steps:

Step 1: Selection of covariates for PS model

- **Variables** were selected based on clinical importance.

- Covariates included age, sex, bone lesions, time from diagnosis, prior regimens, cytogenetic risk, refractoriness, proteasome inhibitors, anti-CD38 antibodies, and tests.

- Allowed up to **30% missing data** for highly prognostic covariates with **multiple imputation**.

Step 2: PS modelling

The final PS obtained for each imputed data set and the IPTW were then used to perform the balancing.

Step 3: Assessment of balance between cohorts

Pooled standardized mean differences (SMD) were computed using Rubin's rules before and after balancing.

Step 4: Balancing methods and criteria

- It was originally planned to provide **2:1 match**; however, it actually included the same number of RW patients as in the KarMMa cohort.

- IPTW was chosen as primary methodology with a **sensitivity analysis** using nearest neighbor **matching**.

OUTCOME ANALYSIS

Risk ratios and hazards ratios

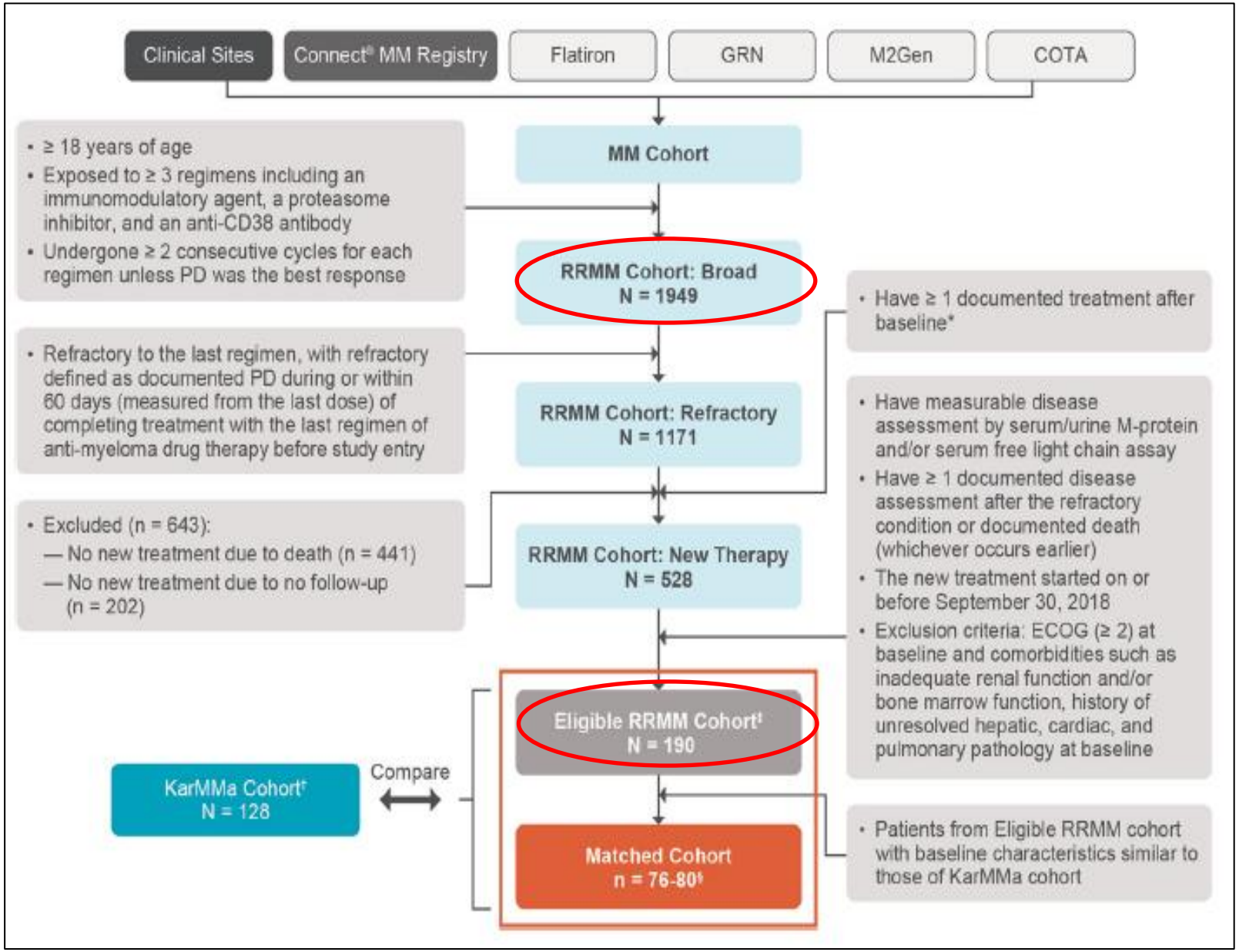
- RR and HR were then estimated for each of the 30 datasets.
- **Overall estimates** were then obtained using Rubin's rules. This approach provided less biased estimates with appropriate confidence intervals.

Analysis firewall

- For the initial wave of analysis, there was a **firewall** between the PS balancing process and the analysis of the post-baseline outcome data.
- The balancing process was conducted by statisticians and programmers **without knowledge of outcome data**.
- Once the **IPTW** and **the matching flags** were generated and frozen, the statisticians and programmers were given access to the outcome data.

RRMM COHORT ATTRITION

- The ECA was pooled from 6 data sources in US and Europe.
- **1,949** real-world patients with RRMM who received ≥3 prior regimens.
- **1,171** were refractory to their last regimen at baseline.
- **528** had received a subsequent line of therapy.
- **190** patients in the **Eligible RRMM** cohort.
- A total of **128 ide-cel** treated patients from the KarMMa study were compared with the Eligible RRMM cohort.



BASELINE CHARATERISTICS BEFORE & AFTER MATCHING

- Differences in disease characteristics between the Eligible RRMM and KarMMa cohorts were statistically adjusted.
- **Trimmed stabilized IPTW** or **greedy nearest neighbor matching** based on PS improved the balance of demographic features and patient characteristics across cohorts.

	Before matching			After matching		
Covariate*	KarMMa Cohort† (N=128)	Eligible RRMM Cohort (N=190)	SMD (KarMMa-Eligible RRMM)	KarMMa Cohort† (N=76–80)	Eligible RRMM Cohort (N=76–80)	SMD (KarMMa – Eligible RRMM)
Age, years	59.8	64.5	**–0.5068**	61.7	62.2	**–0.0520**
Male, %	60.0	60.0	0.0194	60.0	60.0	0.0007
Corrected calcium, mmol/L	3.0	2.4	**0.4302**	2.5	2.5	**0.0003**
Time since initial diagnosis	6.9	4.9	**0.5814**	5.8	5.6	**0.0856**
Number of prior regimens	5.6	4.8	**0.5288**	5.3	5.1	**0.0851**
Number of prior regimens per year since diagnosis	1.2	1.3	−0.1423	1.2	1.2	0.0148
Triple-class refractory status,‡ %	80.0	40.0	**0.9491**	80.0	80.0	**0.0165**

SMD, standardized mean difference. Multiple imputation procedures created 30 datasets and overall estimates were obtained using Rubin's rules to combine the individual estimates.

*A covariate was not included in the balancing if it had > 30% missing for the Eligible RRMM cohort. Means are presented for continuous variables and proportions are presented for categorical variables. Standardized mean difference was obtained from the KarMMa cohort minus the Eligible RRMM cohort and used trimmed stabilized weights when combining the mean and standard deviation.

†Across all target doses.

‡Triple-class refractory was defined as refractory to an immunomodulatory agent, a proteasome inhibitor, and an anti-CD38 antibody.

OVERALL RESPONSE RATE – PRIMARY ANALYSIS

- Efficacy parameters were significantly improved in the KarMMa cohort across **all target doses**, compared with the Eligible RRMM cohort:
  - **ORR** (76.4% vs 32.2%) and **≥VGPR** (57.9% vs 13.7%).
- Improvements with ide-cel were greater in patients who received the **highest target dose**:
  - **ORR** (82.0% vs 31.4) and **≥VGPR** (67.4% vs 13.5%).

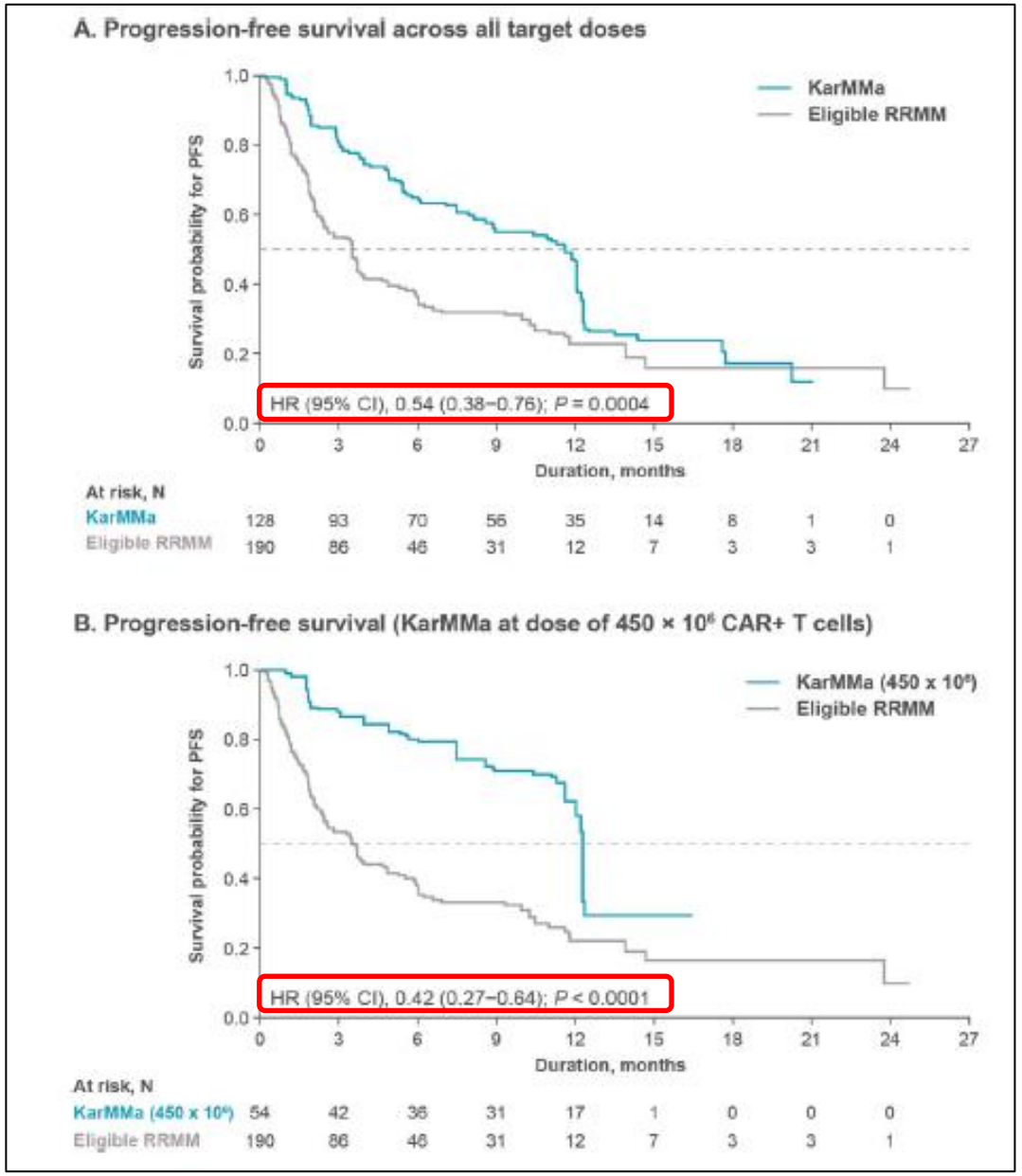
Table 2. Response rates adjusted for stabilized trimmed inverse probability treatment weighting.

Response[a]	KarMMa cohort[b] (*N* = 128)	Eligible RRMM cohort (*N* = 190)	KarMMa cohort 450 × 10^6 CAR + T Cells (*N* = 54)	Eligible RRMM cohort (*N* = 190)
ORR (95% CI), %	76.4 (67.8–86.1)	32.2 (24.4–42.3)	82.0 (70.3–95.7)	31.4 (25.0–39.4)
RR (95% CI)	2.4 (1.7–3.3)		2.6 (2.0–3.5)	
P	<0.0001		<0.0001	
≥VGPR[c] rate (95% CI), %	57.9 (47.8–70.1)	13.7 (8.6–21.9)	67.4 (52.6–86.4)	13.5 (9.1–20.1)
RR (95% CI)	4.2 (2.5–7.2)		5.0 (3.1–8.0)	
P	<0.0001		<0.0001	

ORR was defined as percentage of patients who achieved a best response of partial response or better.
≥VGPR rate was defined as percentage of patients who achieved a best response of VGPR or better.
CI confidence interval, *IPTW* inverse probability treatment weighting, *ORR* overall response rate, *RR* risk ratio, *VGPR* very good partial response.
[a]Derived for the KarMMa and Eligible RRMM cohorts using trimmed stabilized inverse probability treatment weighting propensity score.
[b]Across all target doses.
[c]Complete response not reported due to missing biopsy data in the Eligible RRMM cohort to confirm response.

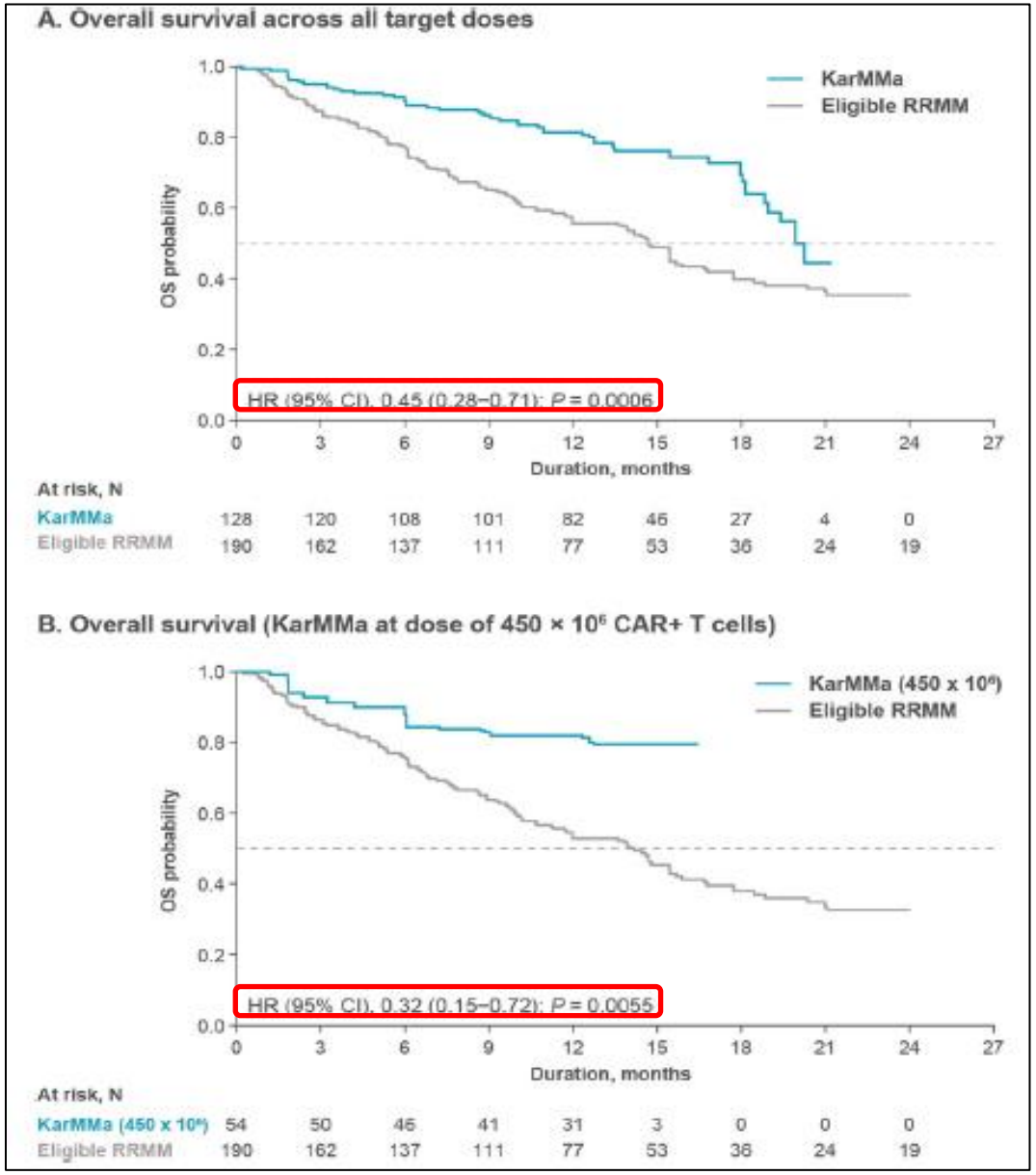
PROGRESSION-FREE SURVIVAL

- **Median PFS** was significantly prolonged in KarMMa patients across **all target doses**, compared with Eligible RRMM patients (**11.6 vs 3.5 months**; HR=0.54, 95% CI: 0.38–0.76)
- **Median PFS** at the **highest target dose** in the KarMMa cohort was **12.3 vs 3.5 months** in the Eligible RRMM cohort (HR=0.42, 95% CI: 0.27–0.64)
- **Median follow-up** was 12.9 months (range, 0.2–21.2) and 11.1 months (range, 0.2–24.0), respectively.



OVERALL SURVIVAL

- **Median OS** was significantly improved with ide-cel in KarMMa across all target doses, vs the Eligible RRMM cohort (20.2 vs 14.7 months; HR=0.45, 95% CI: 0.28–0.71)
- **Median OS** was not reached at the highest target dose versus 14.2 months, HR=0.32, 95% CI:0.15–0.72)
- The **estimated 12-month** probability of surviving was:
  - 80% vs 56% for all doses
  - 82% vs 53% at the high dose.
- **Median follow-up** among surviving patients was 14.4 vs15.0 months.



DISCUSSION (1)

- The ECA study is **a large-scale**, patient-level examination of outcomes with currently available treatments in real-world RRMM patients and the **first study to directly compare results** in the KarMMa and ECA.

- **Responses** and **survival** outcomes were **significantly improved** in ide-cel−treated patients in the KarMMa study vs ECA (ORR and ≥VGPR, $p < 0.0001$; PFS, $p = 0.0004$; OS, $p = 0.0006$).

- OS data were considered **immature** at the time of analysis. In addition, patients in the KarMMa study were **more heavily pre-treated** and more had double- or triple-class refractory disease than ECA.

DISCUSSION (2)

- Stringent inclusion/exclusion criteria and PS methods ensured **robust** and **reliable** comparisons.
- **Subgroup** and **sensitivity analyses** confirmed the overall results.
- ORR, ≥VGPR, median OS and PFS in patients treated with ide-cel were **approximately double** those in the ECA.
- The ECA study demonstrated a **clear benefit** with ide-cel treatment over currently available therapies, with significant increases in efficacy.

KEY EMA FEEDBACK

- **"The EMA was satisfied** that **the comparison showed clinically relevant and statistically significant benefit for Abecma across all pre-defined efficacy endpoints**."

- **"The EMA was also very clear about some of the key limitations of this comparison**. These limiting factors included: a long time period (up to 60 days from the index date) allowed for collection of baseline data; overlapping recruitment periods of the real-world study and the MM-001 study at the same study centers; and a large proportion of missing data (up to 30%) for some included co-variates and several co-variates excluded from the propensity score model due to >30% missing data."

- **"Despite the limitations, the EMA considered that the results indicated that Abecma was associated with responses well above the current standard of care. The treatment was approved** by the European Commission, becoming the first cell therapy authorized in the EU for treating fifth-line multiple myeloma."

Source: EMA Assessment Report

KEY MESSAGES OF THE ECA CASE

- **Early communications** and sharing the study protocol and SAP with regulatory agencies are critical.
- **Pre-specify robust** methodology to address potential biases, endpoints, comparators, subgroups and covariates in the **priori** study protocol and SAP.
- Appropriate flow for **feasibility, baseline and outcome data phases is important.**
- To ensure **blinding to outcomes**, a **firewall** should be set up.
- **Diversified perceptions** from regulatory agencies on the missingness of RWD elements.
- Early published ECA results in the **Nature Blood Cancer** journal (Impact Factor >10) demonstrate scientific integrity and minimize queries during agency assessment.

Blood Cancer Journal www.nature.com/bcj

ARTICLE OPEN Check for updates

KarMMa-RW: comparison of idecabtagene vicleucel with real-world outcomes in relapsed and refractory multiple myeloma

Sundar Jagannath[1✉], Yi Lin[2], Hartmut Goldschmidt[3], Donna Reece[4], Ajay Nooka[5], Alicia Senin[6], Paula Rodriguez-Otero[7], Ray Powles[8], Kosei Matsue[9], Nina Shah[10], Larry D. Anderson Jr[11], Matthew Streetly[12], Kimberly Wilson[13], Hoa Van Le[13], Arlene S. Swern[13], Amit Agarwal[13] and David S. Siegel[14]

Source: https://www.nature.com/articles/s41408-021-00507-2.pdf

Machine Learning – Case Example

Using Machine Learning Algorithm to Predict Medication Initiations and Discontinuations*

**Hoa V Le et al. Presentation # 474, 35th ICPE, 2019*

BACKGROUND, OBJECTIVE & STUDY DESIGN

Daiichi-Sankyo

Background

- Current treatments for chronic kidney disease (**CKD**) focus on slowing disease progression but do not address the underlying inflammation and cause of kidney disease.
- Oral mineralocorticoid receptor antagonists (**MRA**) include spironolactone and eplerenone. A new MRA may have some beneficial effect on CKD patients.
- The use of **machine learning** (**ML**) methods to predict factors associated with **MRA initiation/start and discontinuation** in CKD patients has not been well known.

Objective

To predict and determine factors associated with **initiation** and **discontinuation** of MRA in CKD patients in the UK Clinical Practice Research Datalink (CPRD).

Study population

- A retrospective cohort of CKD patients was created from the **CPRD GOLD** version dataset linked with Hospital Episode Statistics (**HES**) and the Office of National Statistics (**ONS**) for mortality. CPRD includes EMR and medications recorded by GPs.
- Age ≥18 years
- CKD diagnosis by READ/ICD-10 codes without heart failure in Jan 2008-March 2012
- ≥ 1-year baseline history and ≥ 2-year follow-up
- **Cohort 1:** Diabetic CKD for predicting **MRA initiation** by using MRA start date
- **Cohort 2:** CKD MRA users for predicting **MRA discontinuation** by using MRA start and end dates

STATISTICAL METHODS

- For all models, potential predictors for MRA initiation and discontinuation included demographics, lifestyle factors, laboratory tests, CKD stages, time from the first CKD diagnosis, comorbidities and medications.

- To predict and determine **factors** (**variables of importance**) associated with the **initiation** and **discontinuation** of MRA during the 2-year follow-up period, the following models were pre-specified before the internal/derivation stage:
  - Gradient Boosting Classifier (**GBC**)
  - Random Forest (**RF**)
  - Least absolute shrinkage and selection operator penalized logistic regression (**LASSO)**
  - Ridge regression **(RIDGE)**
  - Logistic regression (**LR**).

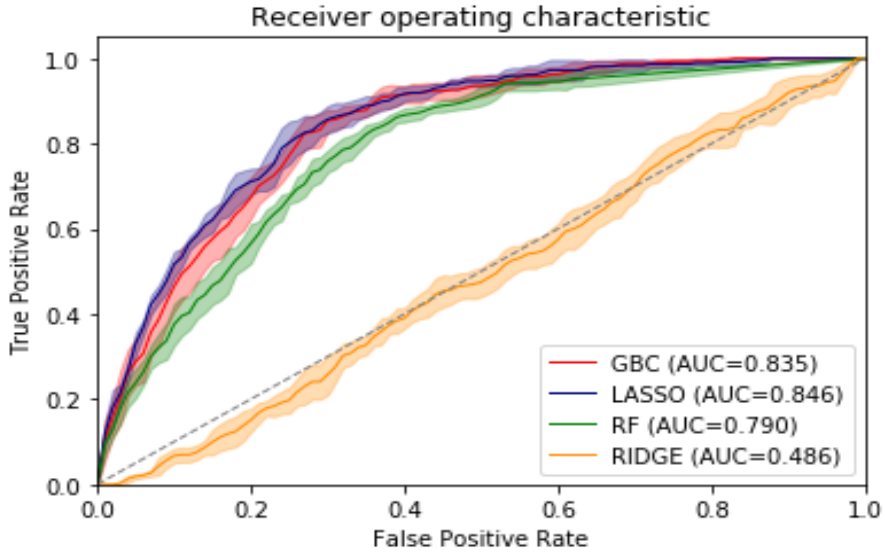
- Each cohort was randomly divided into a **derivation** cohort (70%) and an **external validation** cohort (30%).
- For internal validation among derivation cohort, **5-fold cross-validation** was used.
- **Average** (**95% CI**) Area under the Receiver Operating Characteristic curve (**AUC**) was used to evaluate the performance of the methods.
- **Discrimination** and **calibration** metrics as well as **precision (PPV) and recall (sensitivity)** for all models were also estimated.

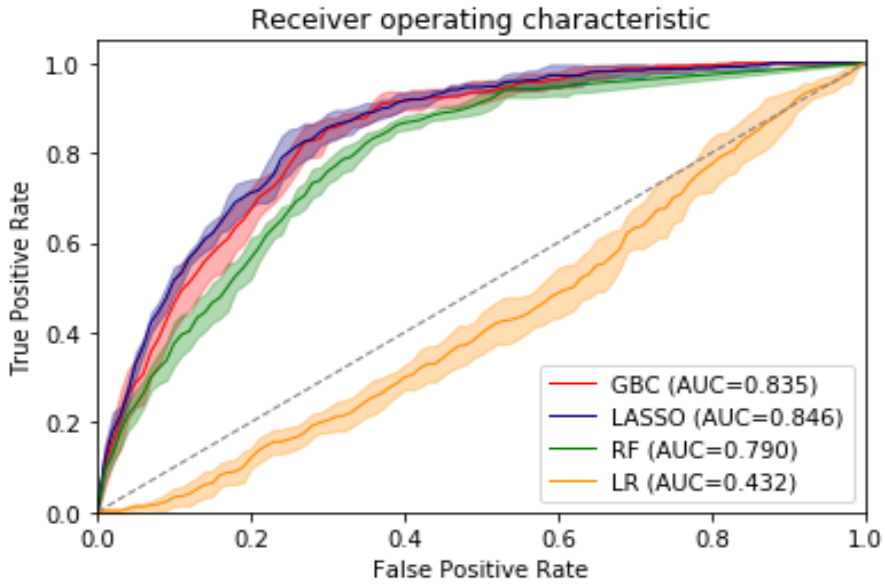
RESULTS: MRA INITIATIONS

- There were **11,053 patients** in the study cohort
- Median (mean) age was 72 (70.4) years old
- 44.8% were women
- **4.9% initiated MRA**
- In internal validation, **GBC, LASSO** and **RF** outperformed **RIDGE** and **LR** in predicting MRA initiation with an average AUC (95% CI) of:

GBC	: 0.839 (0.837-0.841)
LASSO	: 0.842 (0.840-0.844)
RF	: 0.812 (0.809-0.814)
RIDGE	: 0.475 (0.472-0.478)
LR	: 0.460 (0.457-0.462)

Internal 5-fold cross-validation





MRA INITIATIONS: CALIBRATION

- **GBC, LASSO** and **RF** exhibited better calibrations than **RIDGE** and **LR**

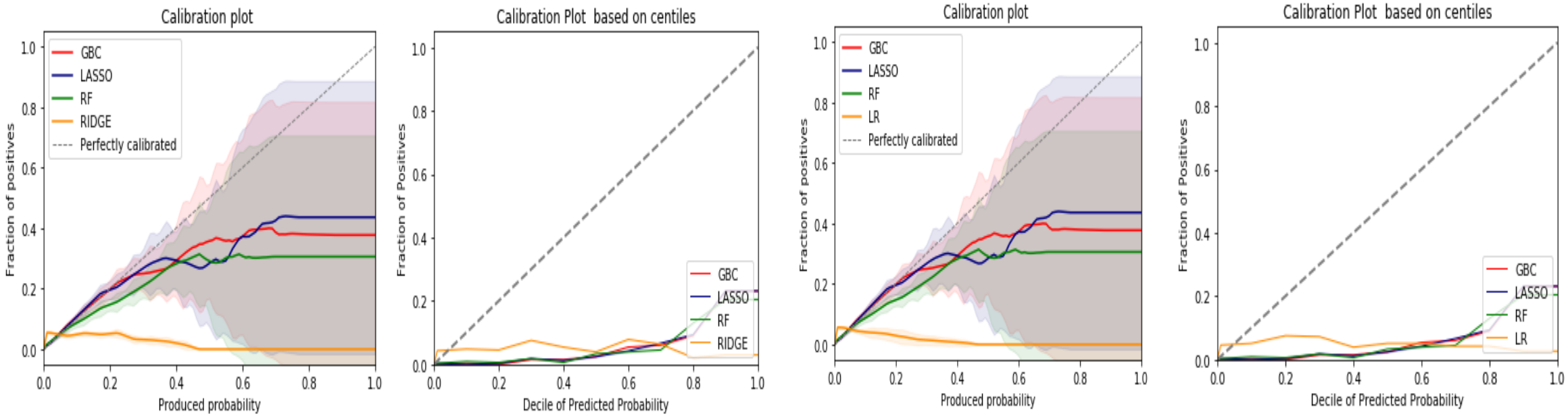
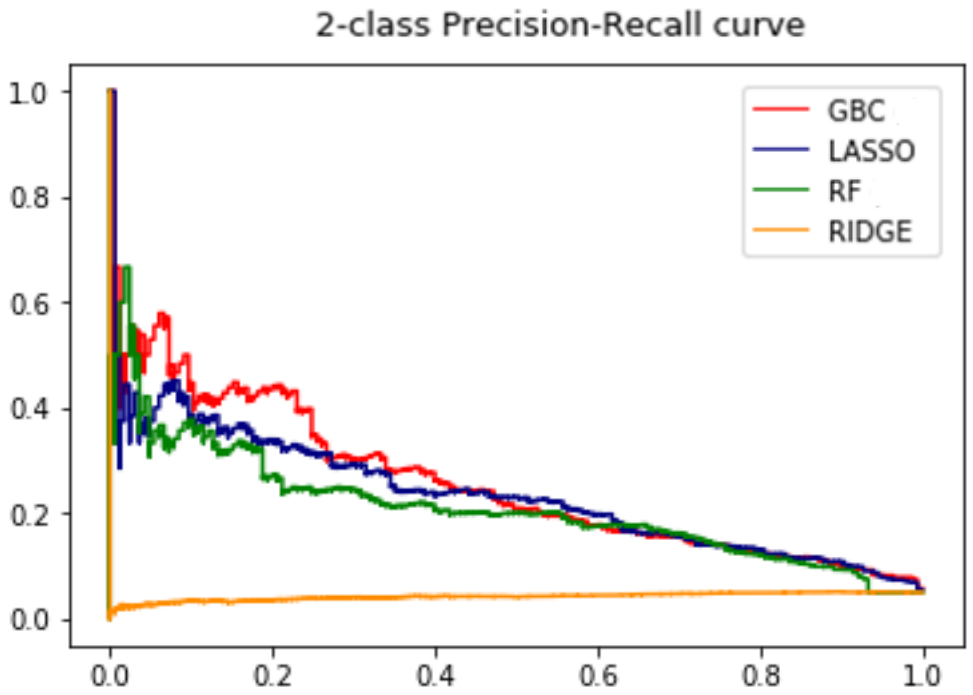


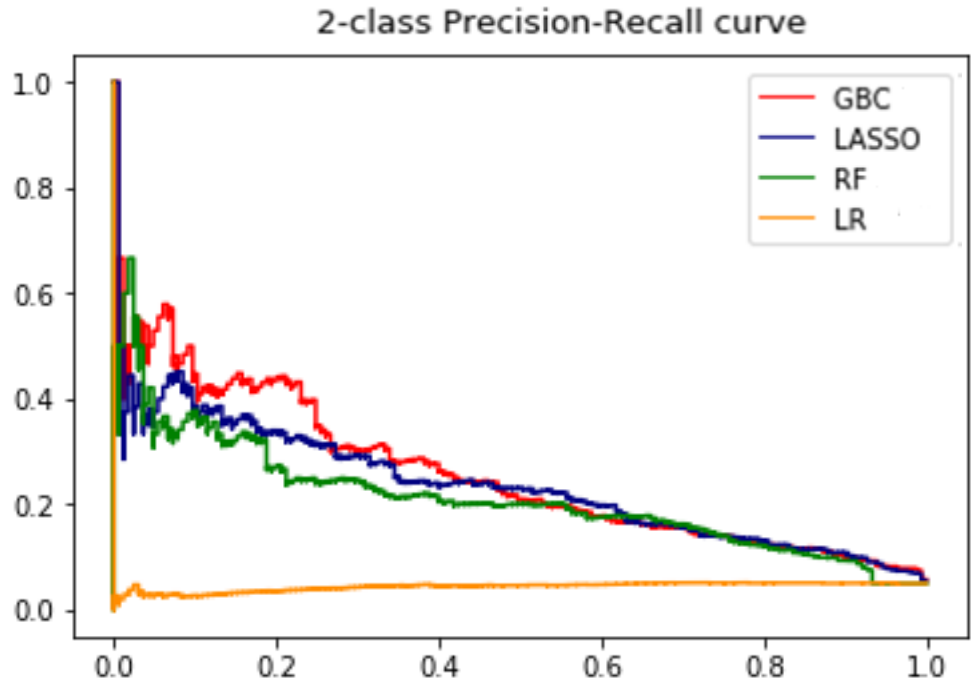
Figure. Cross-validated model calibration for different modelling techniques. The x-axis shows the predicted probability of MRA initiation, while the y-axis shows the fraction of actual MRA initiation for each predicted probability. The shaded areas depict the standard deviation across different folds in a 5-fold cross-validation.

GBC, gradient boosting classifier; **LASSO,** Least absolute shrinkage and selection operator penalized logistic regression; **RF**, random forest; **RIDGE,** Ridge regression; **LR**, logistic regression.

MRA INITIATIONS: PRECISION AND RECALL

- The trade-off between precision and recall were demonstrated. Within any one model, one can also decide to emphasize either precision or recall.





GBC, gradient boosting classifier; **LASSO,** Least absolute shrinkage and selection operator penalized logistic regression; **RF**, random forest; **RIDGE,** Ridge regression; **LR**, logistic regression.

MRA INITIATIONS: VARIABLE IMPORTANCE

Top 20 variables of importance for MRA initiation identified by GBC and RF included:

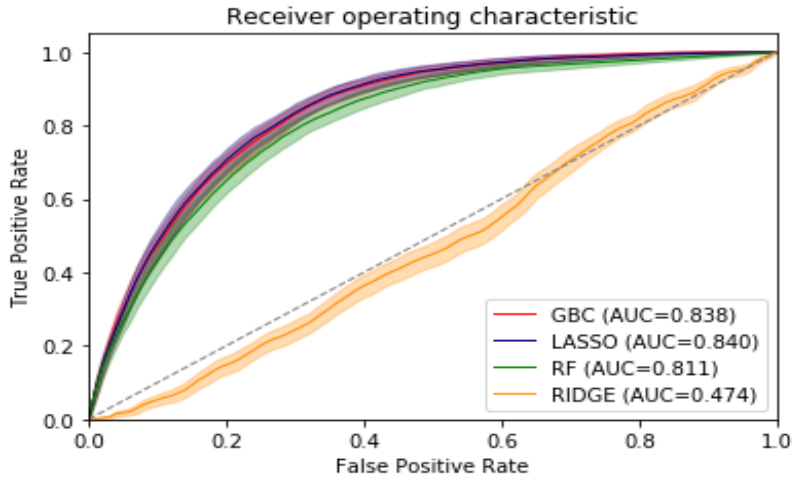
- Longer CKD history since its first diagnosis
- Older age
- A&E visits, practice nurse visits, general practitioner (GP) visits
- Hospitalizations
- Hypertension
- Prior use of
  - angiotensin receptor blocker (**ARB**)
  - calcium channel blocker (**CCB**)
  - beta blocker (**BB**)
  - angiotensin-converting-enzyme inhibitor (**ACEI**)
  - diuretics

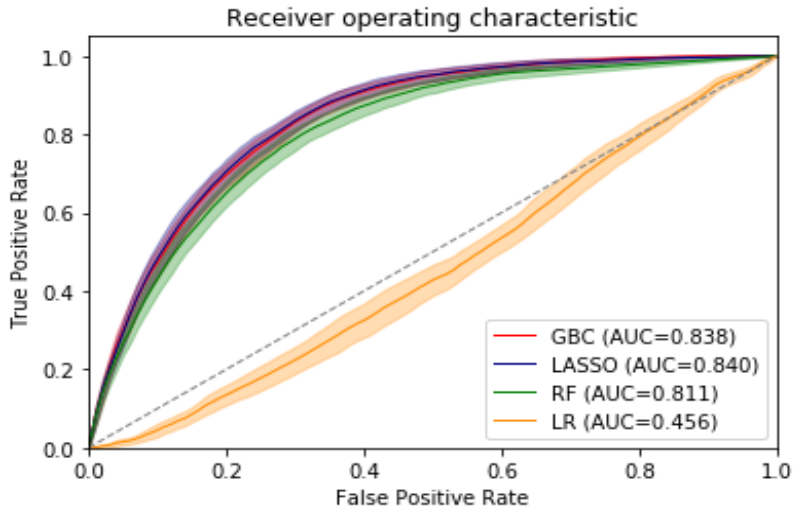
MRA INITIATIONS: EXTERNAL VALIDATION

- Similar findings were found for **external validation**

GBC	: 0.839 (0.837-0.841)
LASSO	: 0.842 (0.840-0.844)
RF	: 0.812 (0.809-0.814)
RIDGE	: 0.475 (0.472-0.478)
LR	: 0.460 (0.457-0.462)

External validation





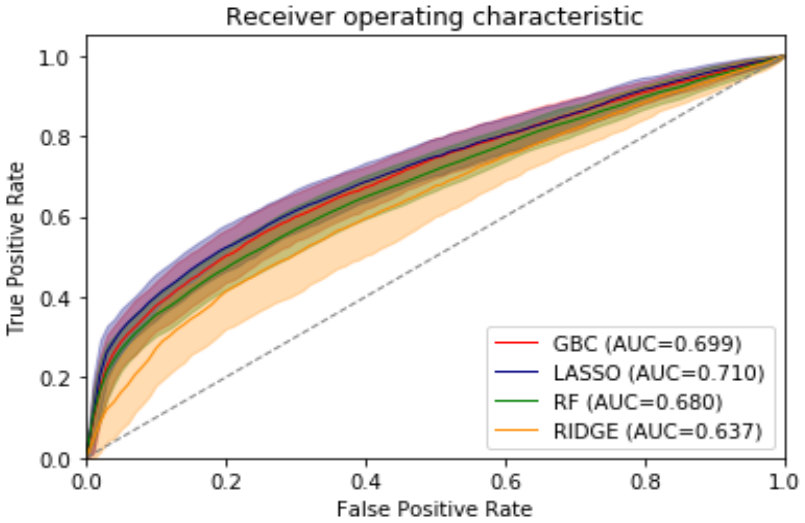
GBC, gradient boosting classifier; **LASSO,** Least absolute shrinkage and selection operator penalized logistic regression; **RF**, random forest; **RIDGE,** Ridge regression; **LR**, logistic regression.

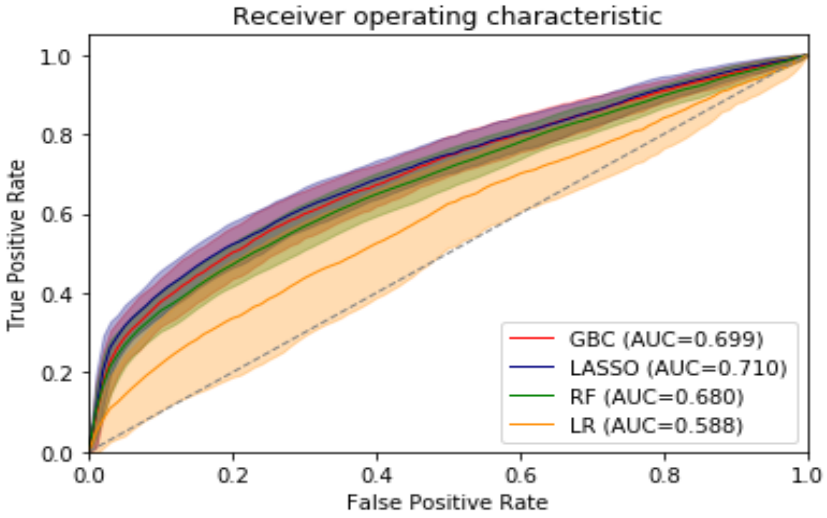
MRA DISCONTINUATIONS

- There were **729 naïve MRA** users:
  - 97% spironolactone
  - 3% eplerenone
- Median (mean) age was 80 (78.3) years old
- 58.2% were women
- Median (mean) duration of MRA use was 307 (390) days
- **65.2%** patients **discontinued MRA** use during the 2-year follow up period
- Similar to internal validation, in external validation **GBC, LASSO** and **RF** outperformed **RIDGE** and **LR** in predicting MRA discontinuation with an average AUC (95% CI) of:

GBC	: 0.699 (0.692-0.706)
LASSO	: 0.710 (0.704-0.716)
RF	: 0.680 (0.673-0.686)
RIDGE	: 0.637 (0.626-0.648)
LR	: 0.588 (0.569-0.607)

External validation





GBC, gradient boosting classifier; **LASSO,** Least absolute shrinkage and selection operator penalized logistic regression; **RF**, random forest; **RIDGE,** Ridge regression; **LR**, logistic regression.

MRA DISCONTINUATIONS: VARIABLE IMPORTANCE

Variables of importance for predicting MRA discontinuation included:

- Older age
- Dialysis
- Thrombosis
- Cramp
- CKD related A&E visits
- Hospitalizations
- Hyperlipidemia
- Hyperkalemia
- Prior use of renin

CONCLUSIONS OF THE ML CASE

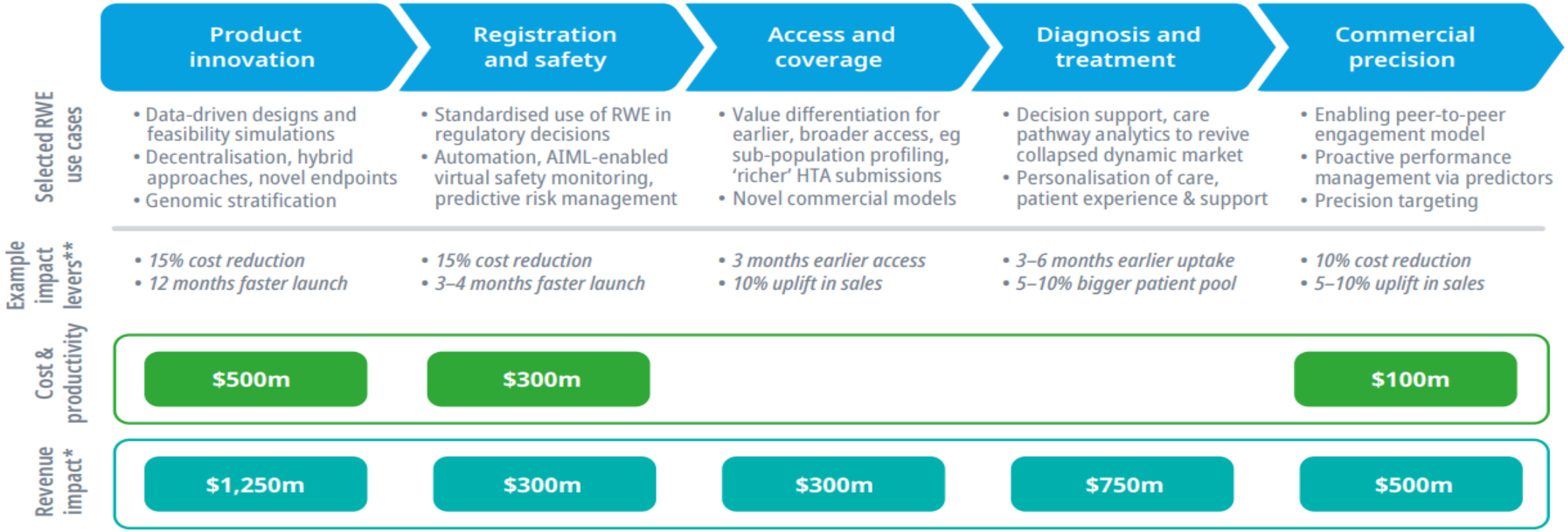
- **GBC, LASSO** and **RF** models performed better than the **RIDGE** and **LR** models in term of AUC in predicting the **MRA initiation** and **discontinuation.**
- **Variables of importance** for MRA initiation and discontinuation were determined.
- These findings support the **potential of incorporating machine learning** models into a study cohort to inform treatment pattern, medication initiation and discontinuation.

RWE & AI/ML OPPORTUNITY & CHALLENGES

Daiichi-Sankyo

Figure 2: The post-pandemic $4 billion RWE opportunity

Estimated value of RWE opportunities along the pharma value chain over three years



Source: IQVIA Thought Leadership
* Includes both accelerated and incremental revenue
** Selected levers, applied to typical addressable big pharma cost / revenue base; derived from IQVIA proof points of observed impact

Challenges

- Secondary data – Fit-for-purpose; Data access; Prospective infrastructure; Expertise

TAKE-HOME MESSAGES

- **Optimize the efficient use RWE & AI/ML** in the product development by maximizing opportunities and overcoming challenges.
- **Prospective research framework** is one of best practices for RWE and AI/ML.
- **Closely follow the FDA Guidance** on ECA to increase the chance of success.
- Building a ML application is an **iterative process**. One can’t simply train a model once and leave it alone.

Passion for Innovation.
Compassion for Patients.™

Daiichi-Sankyo

Thank you!

Email: hle@dsi.com

KEY TERMS & CONCEPTS (back-up)

Term	Definition
EHR studies	Studies that evaluate information from individual's electronic health records during routine delivery of health care stored in a digital format
Administrative and claims database studies	Studies that evaluate information from medical claims that health care providers submit to insurers to assess the impact of clinical treatments, procedures, and outcomes[1]
Open & Closed claims	Open claims data is derived from broad-based healthcare sources and can highlight a patient's activities over a longer timeframe, regardless of a patient's insurance provider. Closed payer claims data is derived from health insurance providers (payers), adjudicated patient's healthcare information during a specific enrollment period.
Registry studies	Studies of results collected from an organized system that uses observational methods to collect uniform data on specified outcomes in a population defined by a particular disease, condition, or exposure[2]
Patient survey studies	A study based on the responses of patients to specific questions contained in surveys administered to patients

1. US Food and Drug Administration. Accessed November 24, 2021. https://www.fda.gov/media/120060/download.
2. Zaletel M, et al. Accessed March 1, 2021. https://ec.europa.eu/health/sites/health/files/ehealth/docs/patient_registries_guidelines_en.pdf.

FDA 21ST CENTURY CURES ACT 2016 – RWE (back-up)

- FDA established a program to evaluate the potential use of real-world evidence to:
  - Support a new indication for a drug approved under section 505©
  - Satisfy post-approval study requirements
- Draft framework issued in December 2018:
  - Describe sources of RWE, challenges, pilot opportunities, etc.
- **Draft guidance for industry issued in Sep, Oct, Nov, and Dec 2021**

Real-World Data: Assessing Electronic Health Records and Medical Claims To Support Regulatory Decision-Making for Drug and Biological Products – SEP 2021

Data Standards for Drugs and Biological Product Submissions Containing Real-World Data – OCT 2021

Real-World Data: Assessing Registries to Support Regulatory Decision-Making for Drug and Biological Products – NOV 2021

Considerations for the Use of Real-World Data and Real-World Evidence to Support Regulatory Decision-Making for Drug and Biological Products – DEC 2021

KEY RWE APPLICATIONS (back-up)

External control arm – RWD/E is used as an external benchmark for context or as a formal comparator to a single-arm trial or indirect comparison.

Potentially support new product approvals and new indications.

Situation: A randomized clinical trial may not be feasible or may be very challenging to conduct

Population bridging – RWE is used to help bridge the applicability of clinical trial results to a particular sub-population that may not have been well represented in that original trial.

Potentially support product approval or label updates.

Situation: Different populations in different countries, Different standard of care, Adult to Pediatric populations, Infeasible to get enough patients in the clinical trial.

Comparative effectiveness – RWE is used to compare the benefit between treatments where the comparison of the investigational agent to the comparator or control is conducted entirely within the real-world data.

Situation: The investigational agent is being prescribed off-label, a pragmatic trial

Long-term effectiveness – RWE to measure longer-term outcomes than that studied in a clinical trial.

Endpoint establishment – RWE to support the establishment of a new surrogate endpoint being used in a clinical trial.

Safety monitoring – RWE for pharmacovigilance