Rigshospitalet
Centre for Cancer and Organ Diseases

Danish Cancer Institute

DALY-CARE Danish Lymphoid Cancer Research

UNIVERSITY OF COPENHAGEN

DAnish LYmphoid-lineage CAncer REsearch
DALY-CARE

– an example cohort of +65000 patients with +3000 routine data variables per patient for development of data-driven hematology

Carsten Utoft Niemann, MD, PhD, associate professor
Head of CLL Lab, Chair of Nordic CLL study group
Department of Hematology, Rigshospitalet, Copenhagen, Denmark

Carsten.utoft.niemann@regionh.dk; www.rigshospitalet.dk/CLL-lab

REGION H

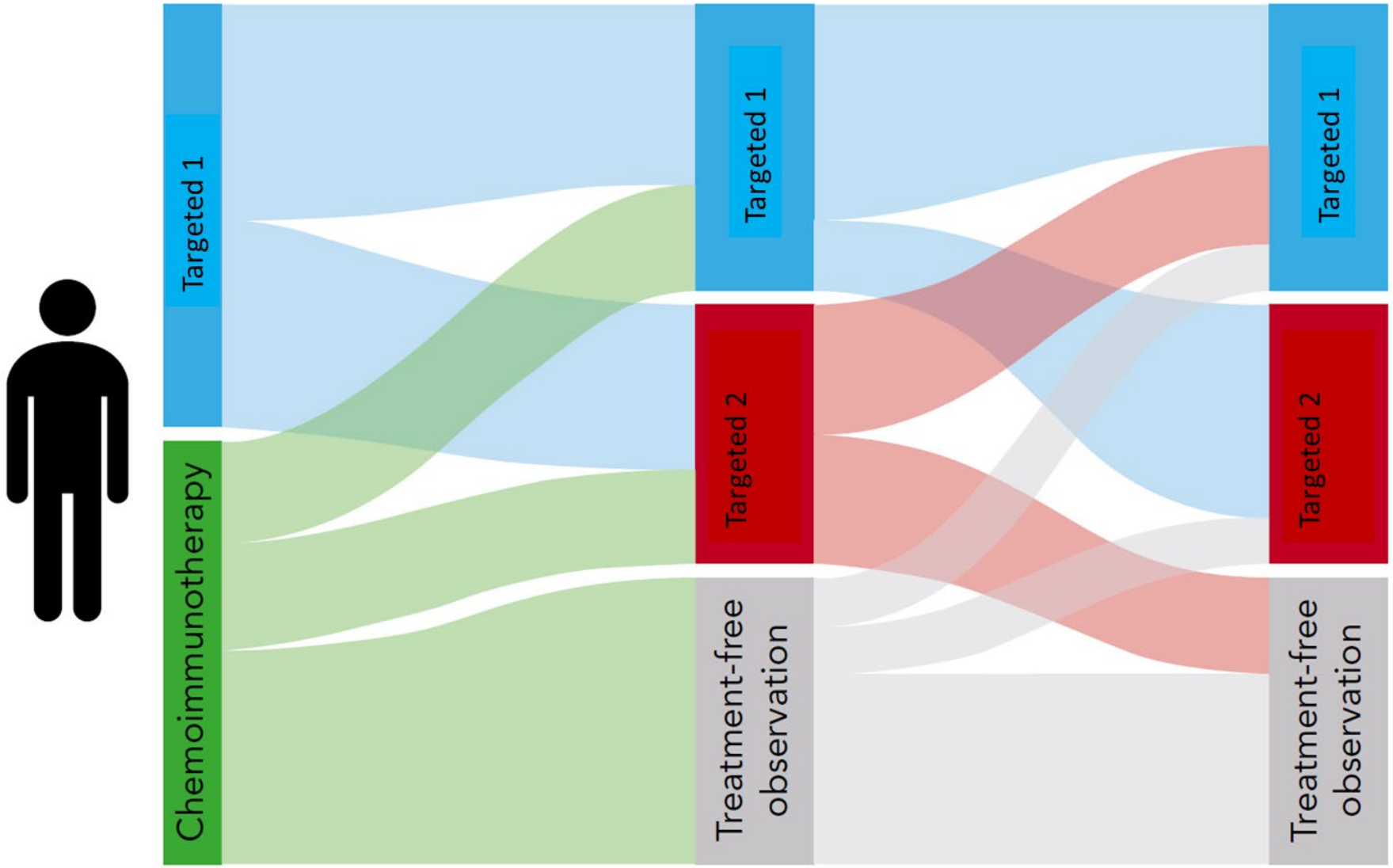
Disclosures:

Research Support	Danish Cancer Society, Novo Nordisk Foundation, Copenhagen University Hospital, Persimune, Arvid Nilsson's Foundation, Abbvie, Janssen, AstraZeneca, Alfred Benzon Foundation
P.I.	Roche, Novartis, Janssen, AstraZeneca, Genmab, Beigene
Consultancy, grants	Roche, Gilead, Janssen, Abbvie, Novartis, CSL Behring, AstraZeneca, Takeda, Octapharma, Genmab, Beigene, Lilly, MSD

Collaborators:

HOVON, Arnon Kater, Amsterdam
GCLLSG, Michael Hallek, Barbara Eichhorst, Köln
Nordic CLLSG, Anders Österborg, Richard Rosenquist, Karolinska, Stockholm
ERICLL.org, Paolo Ghia
EuroMRD, Christiane Pott
CLL-CLUE: Sigrid Skånland, Rikshospitalet, Oslo; Thorsten Zenz, Zürich
Mathew Davids, DFCI, Boston
Adrian Wiestner, NHLBI, NIH
PERSIMUNE, Jens Lundgren, Rigshospitalet
Genetics, COVID, Sisse Ostrowski, Rigshospitalet, Henrik Hjalgrim, Danish Cancer Society

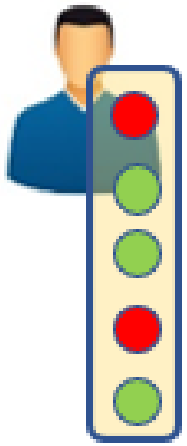
Treatment trajectories in hematology / oncology – Not enough to consider one line of treatment!



Adapted from Niemann, Blood, 2020

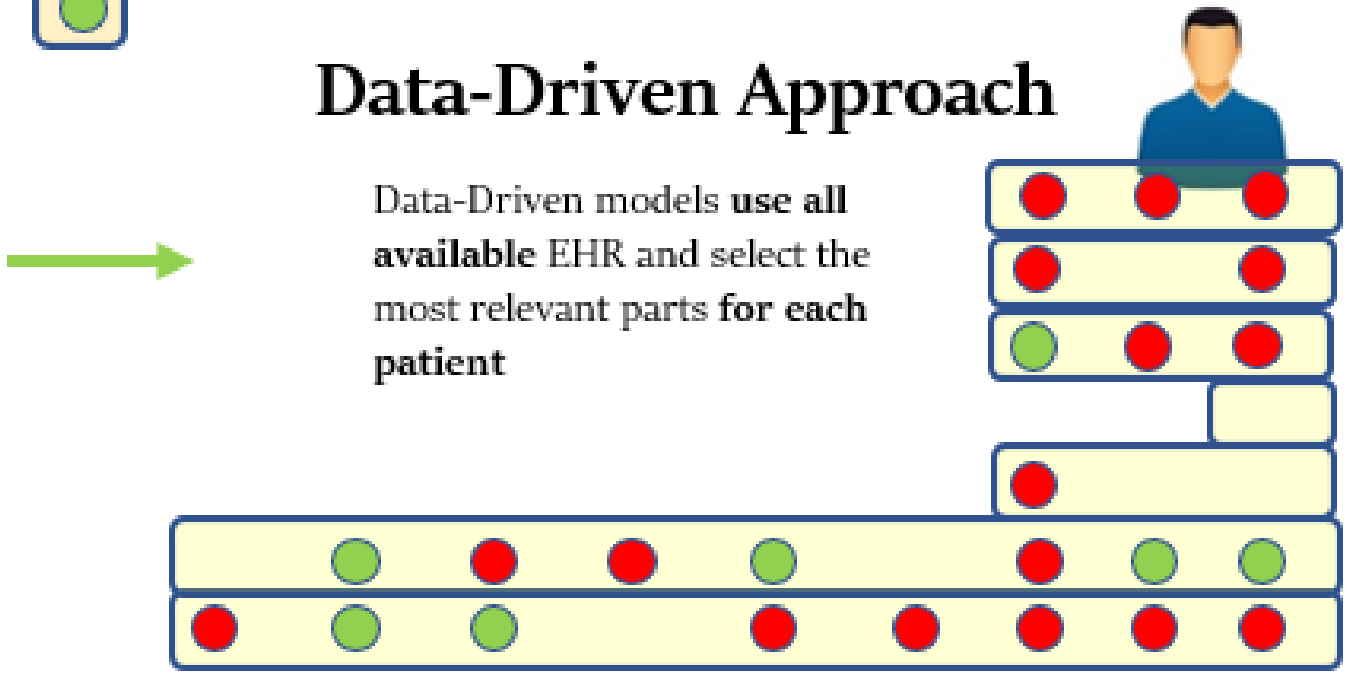
What we mean by having more data

Traditional Approach

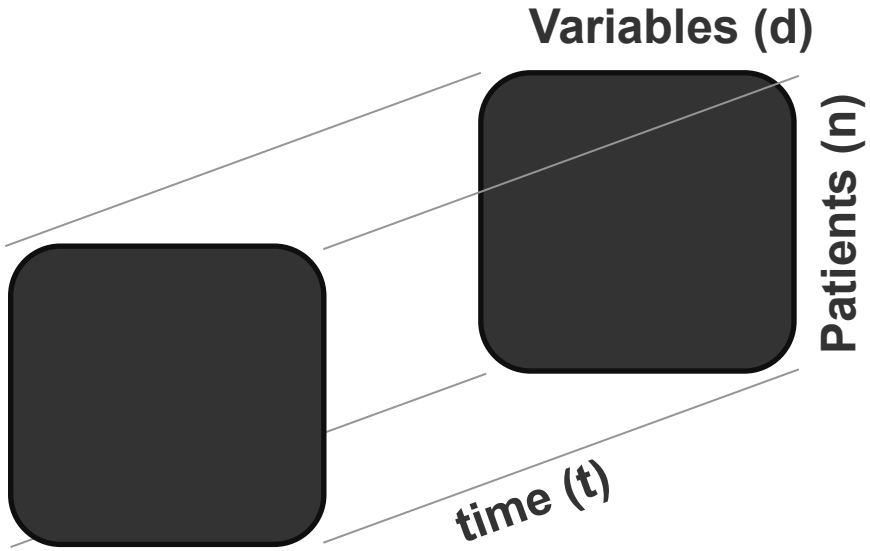


Risk-Scores only use a **snap-shot** of a Patient EHR with a few variables for **all patients**

Data-Driven Approach

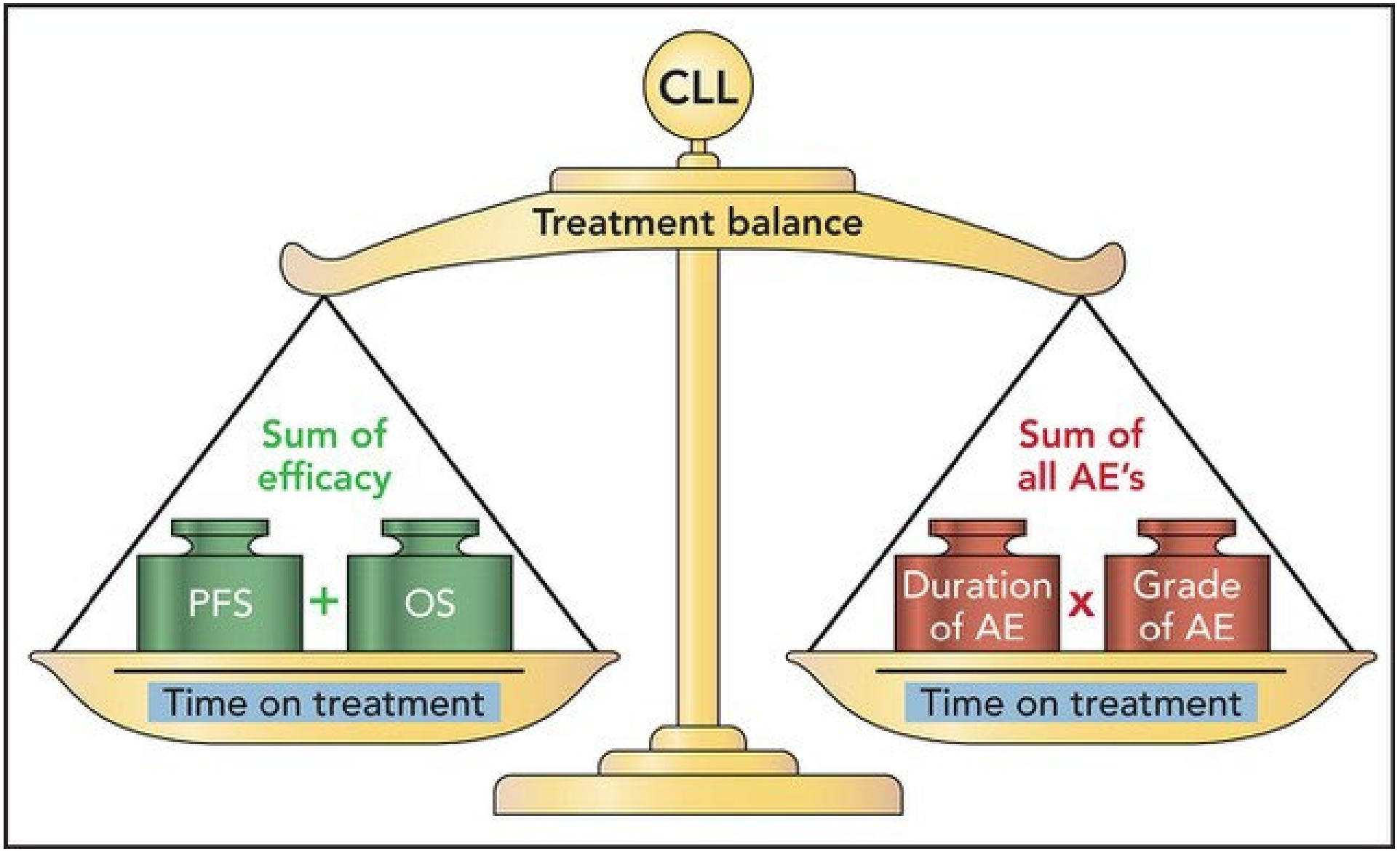


Data-Driven models **use all available** EHR and select the most relevant parts **for each patient**



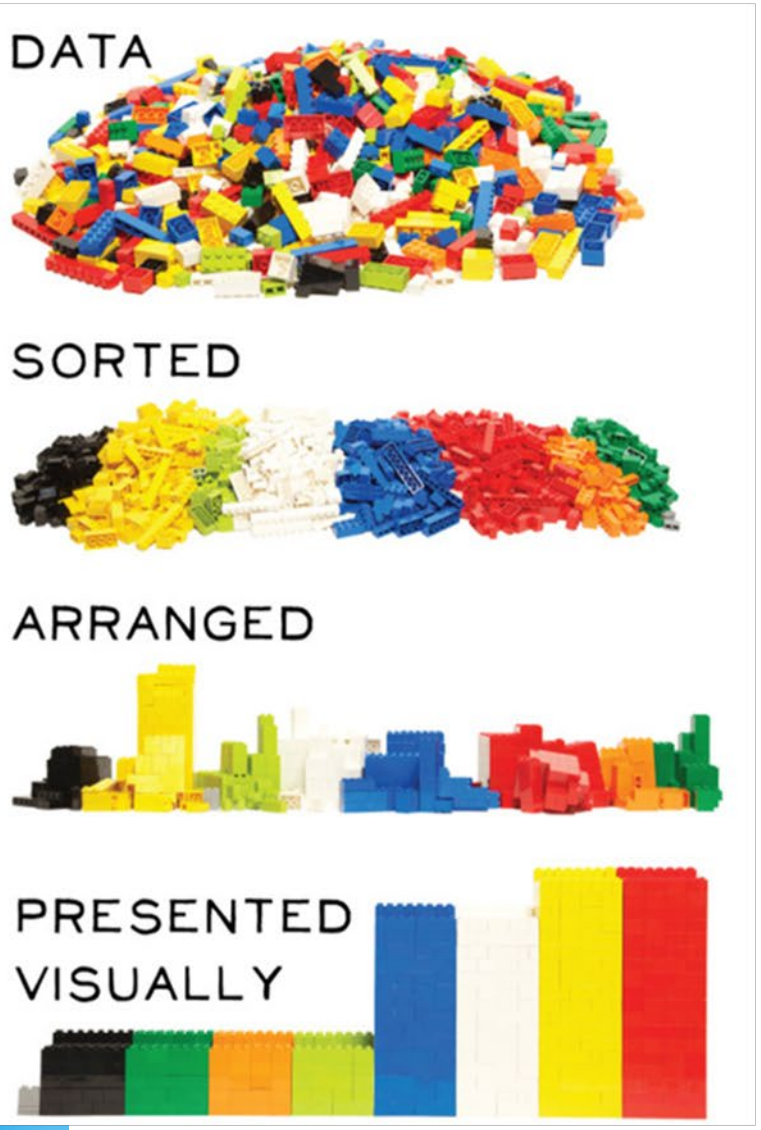
Courtesy of Rudi Agius

Treatment outcomes – Efficacy AND toxicity AND lenght matter!



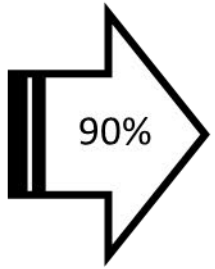
Niemann, Blood, 2023

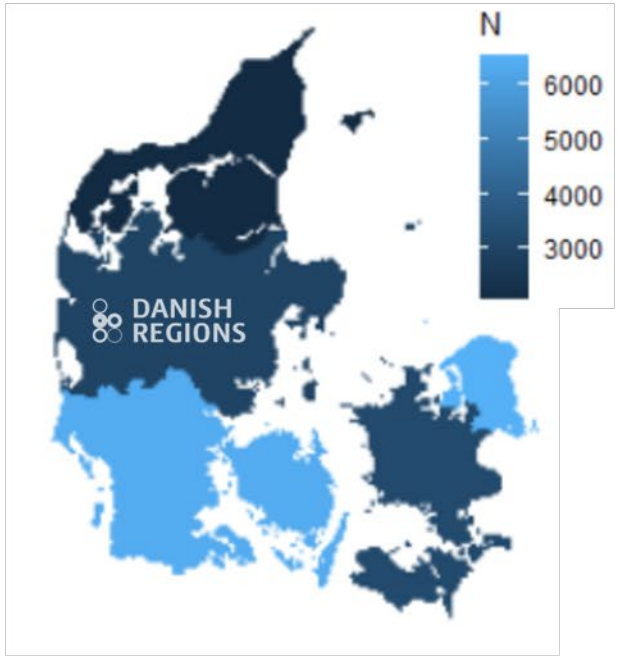
Building a standardized health data resource

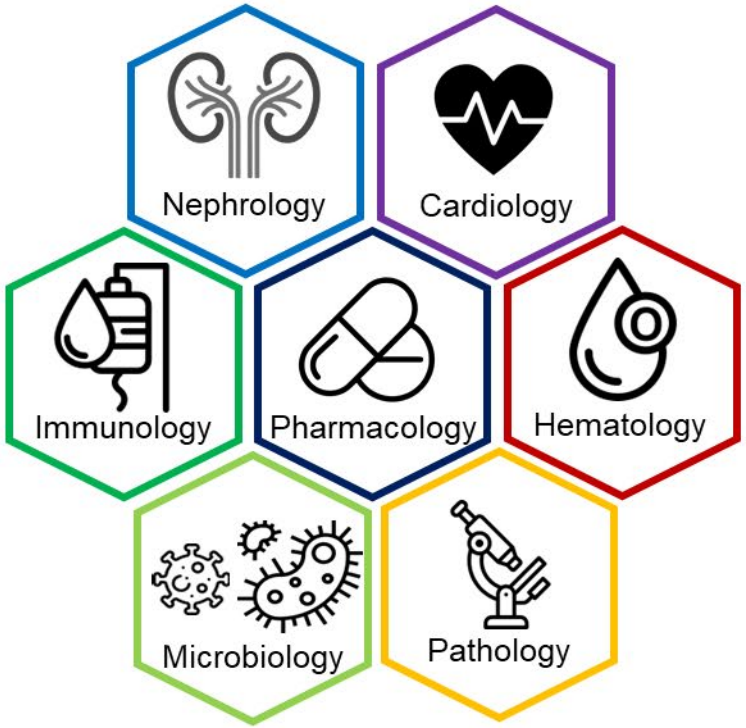


Encoding/Mapping

- Medication/Rx using ATC codes
- Dx using ICD10 codes
- Pathology using SNOMED/ICD-O-3 codes
- Biochemistry using NPU codes
- Microbiology using MORG codes
- Location using SHAK codes



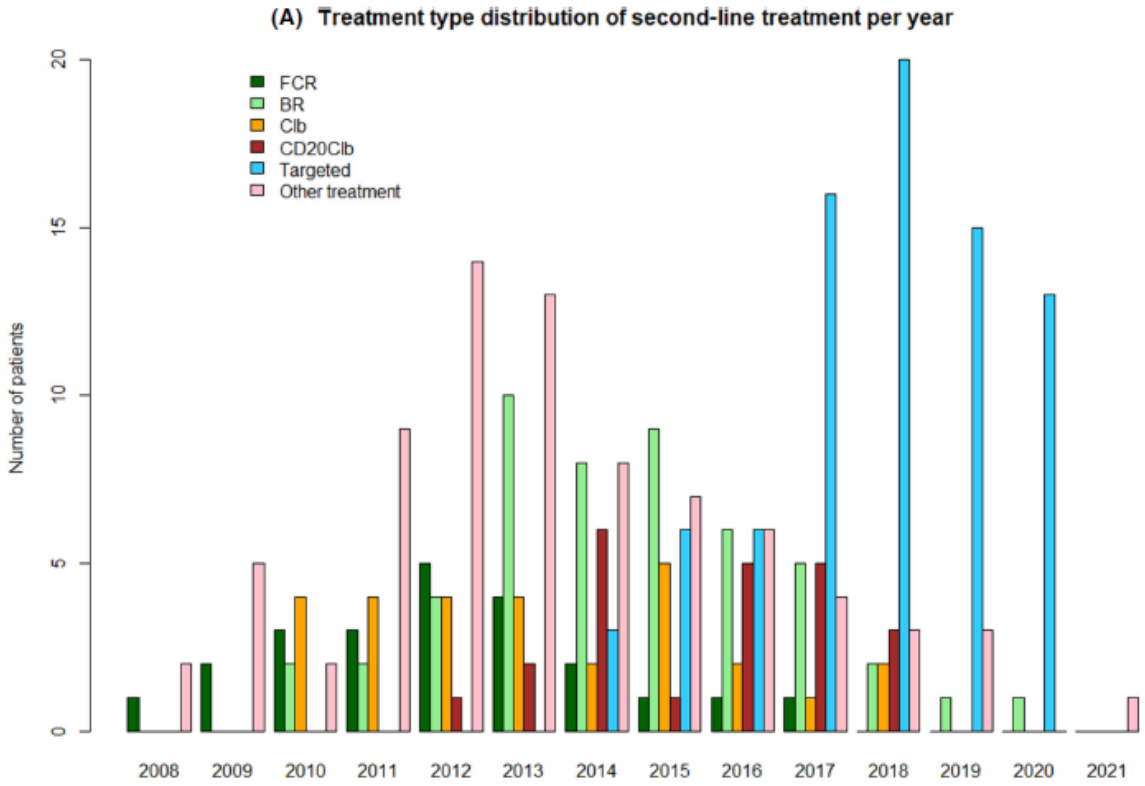


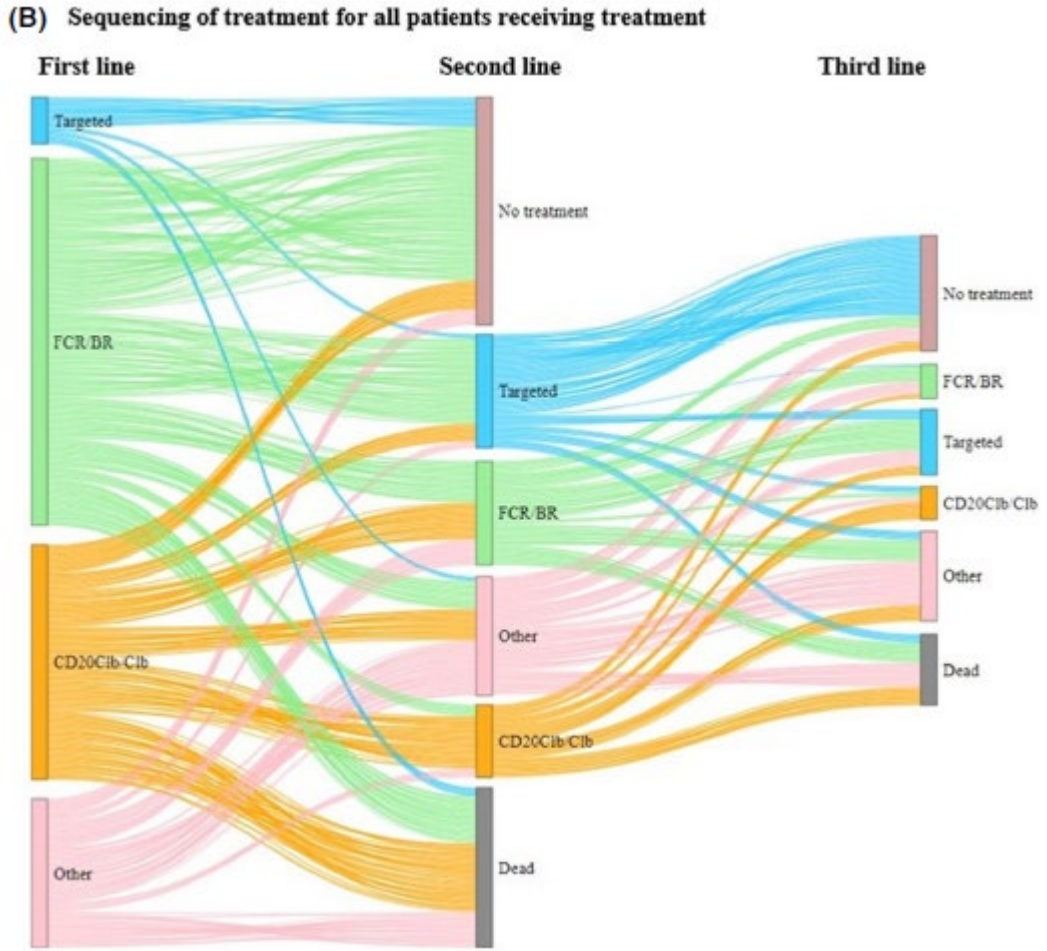


Briehgel et al., medRxiv, 2024

Your patient's disease trajectory?

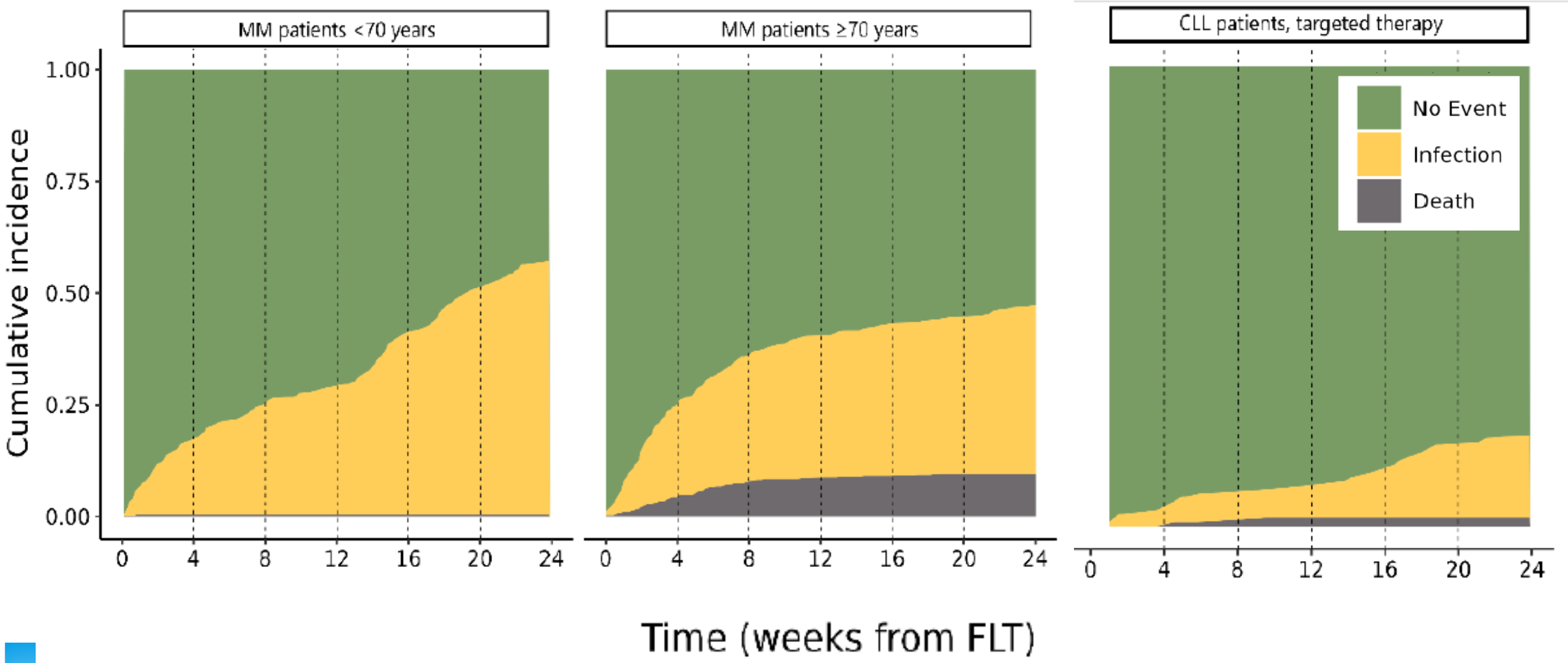
Treatment Landscape in CLL





Vainer et al., BJH, 2024

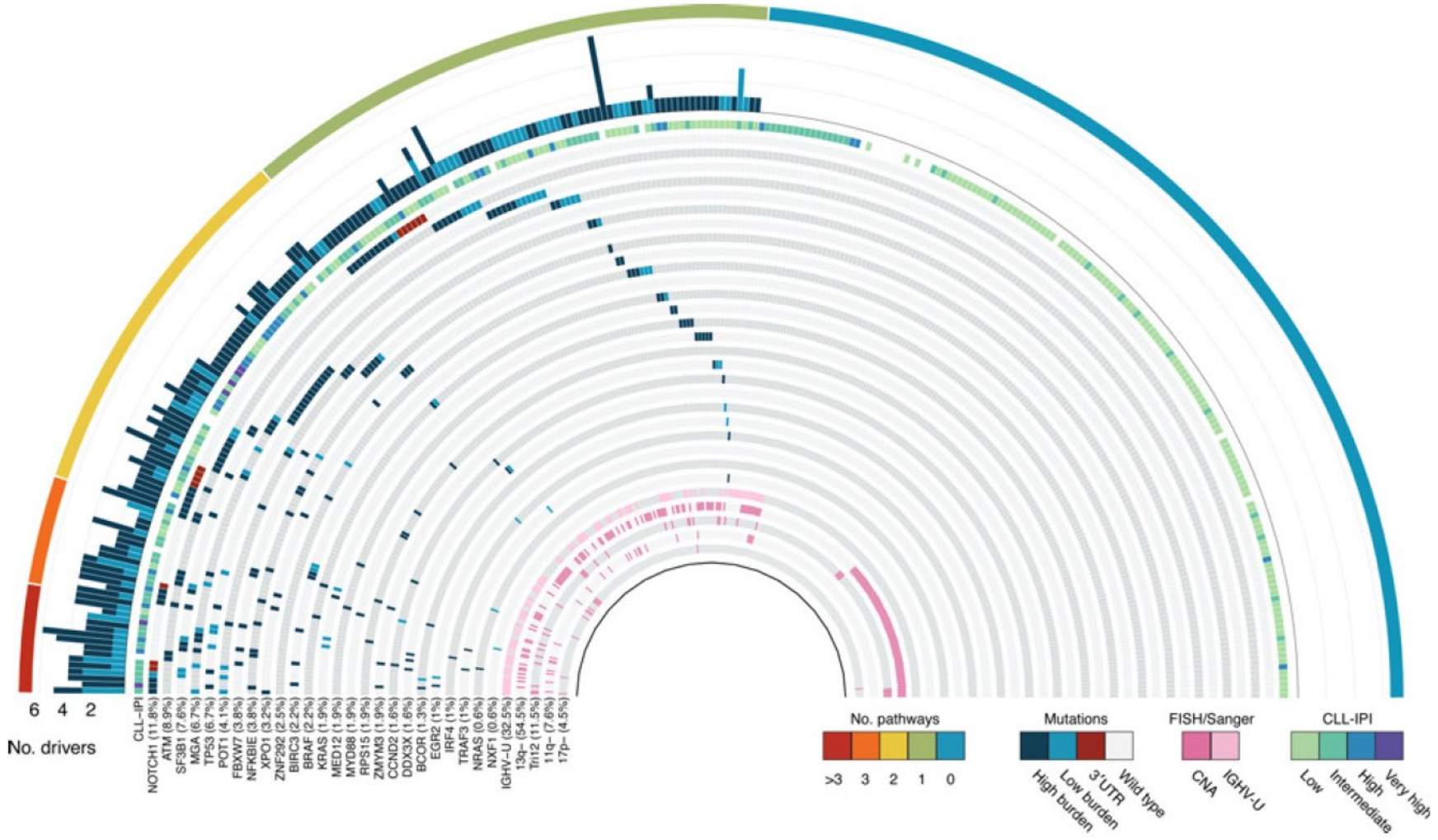
Infections upon treatment of lymphoid cancers

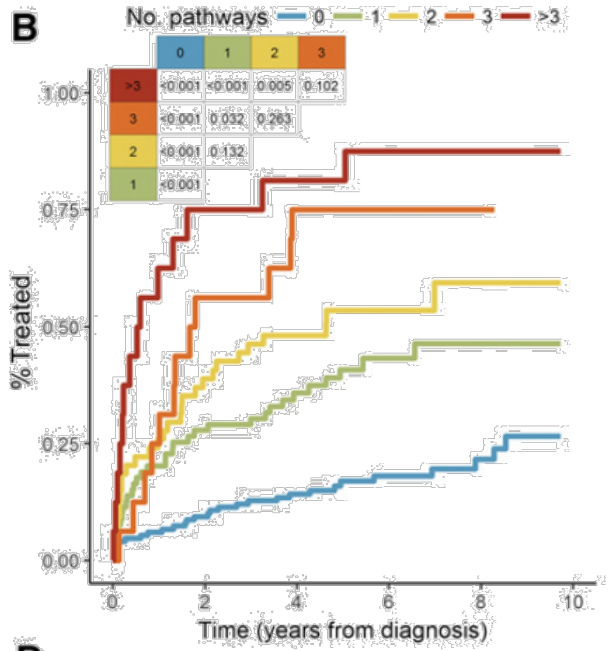


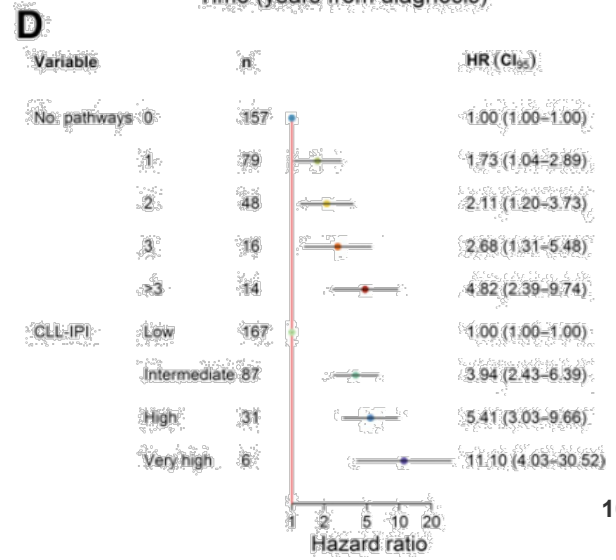
Lacoppidan et al., EHA 2024

Driver mutations in CLL

Allocation by signaling pathway







Brieghel et al, CCR, 2020

Ending Watch & Wait in CLL – smart prioritization?

CLL-IPI

Age >65 years	(1 point)
Binet stage B/C	(1 point)
B2M >3.5 mg/L	(2 points)
U-CLL	(2 points)
TP53 aberration	(4 points)

≤3

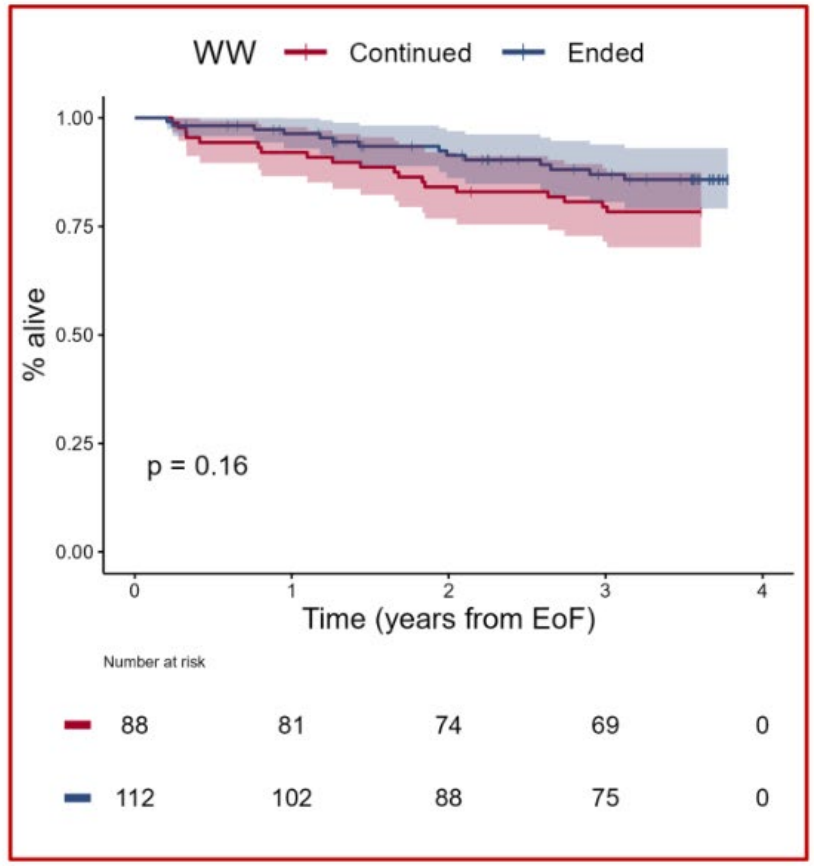
CLL-WONT

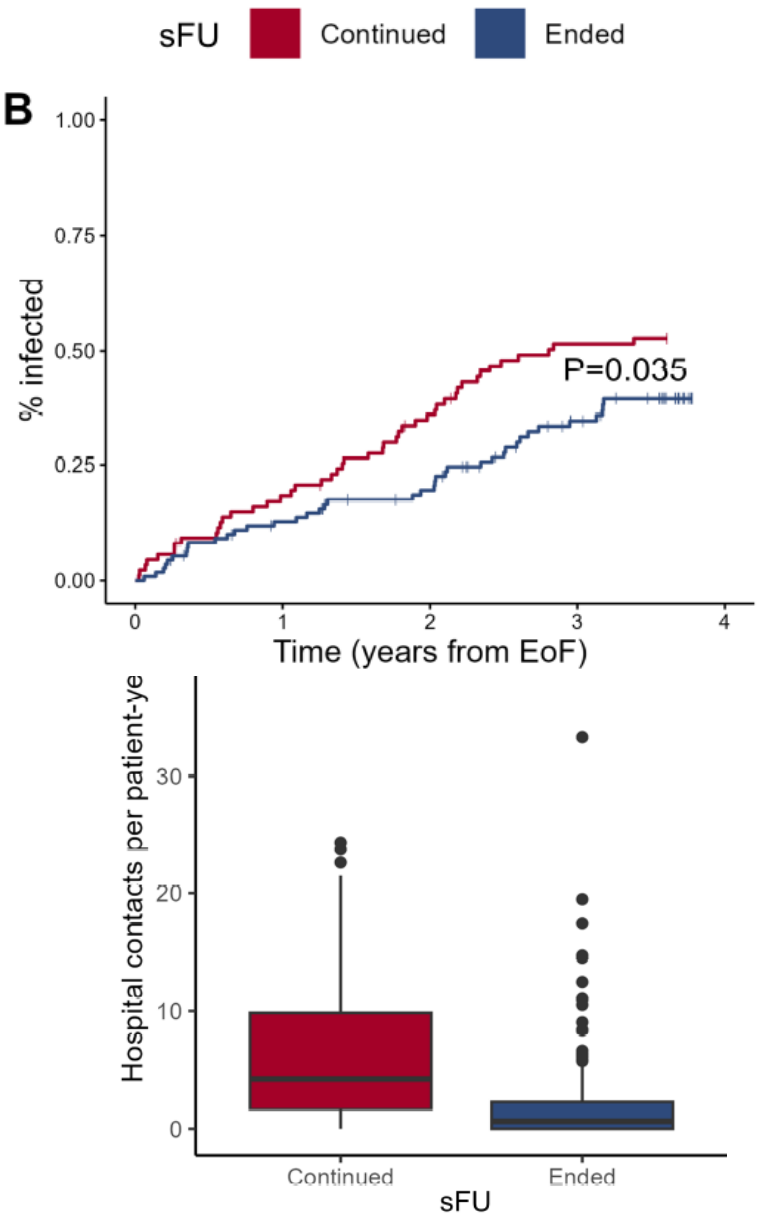
Age >65 years	(1 point)
LDH elevated	(1 point)
B2M >4.0 mg/L	(1 point)
del(11q)	(1 point)
ALC 15-30 or	(1 point)
ALC >30	(2 points)
U-CLL	(2 points)
Binet stage B	(2 points)

≤3

Clinical evaluation of low to intermediate risk

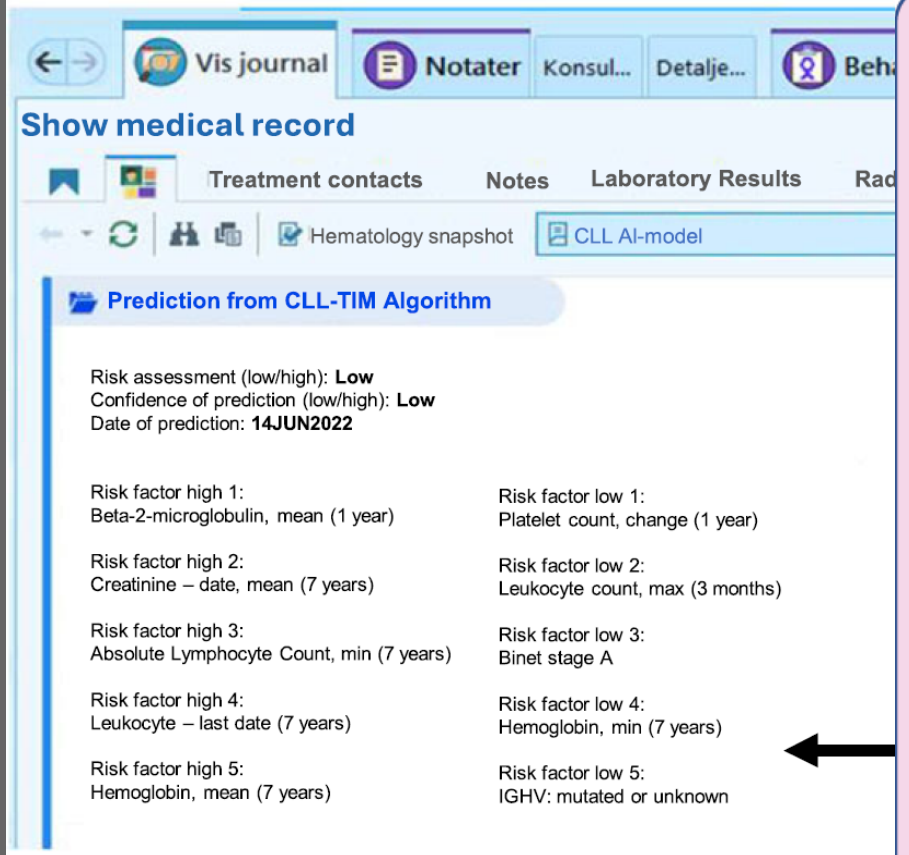
Continue WW | End WW

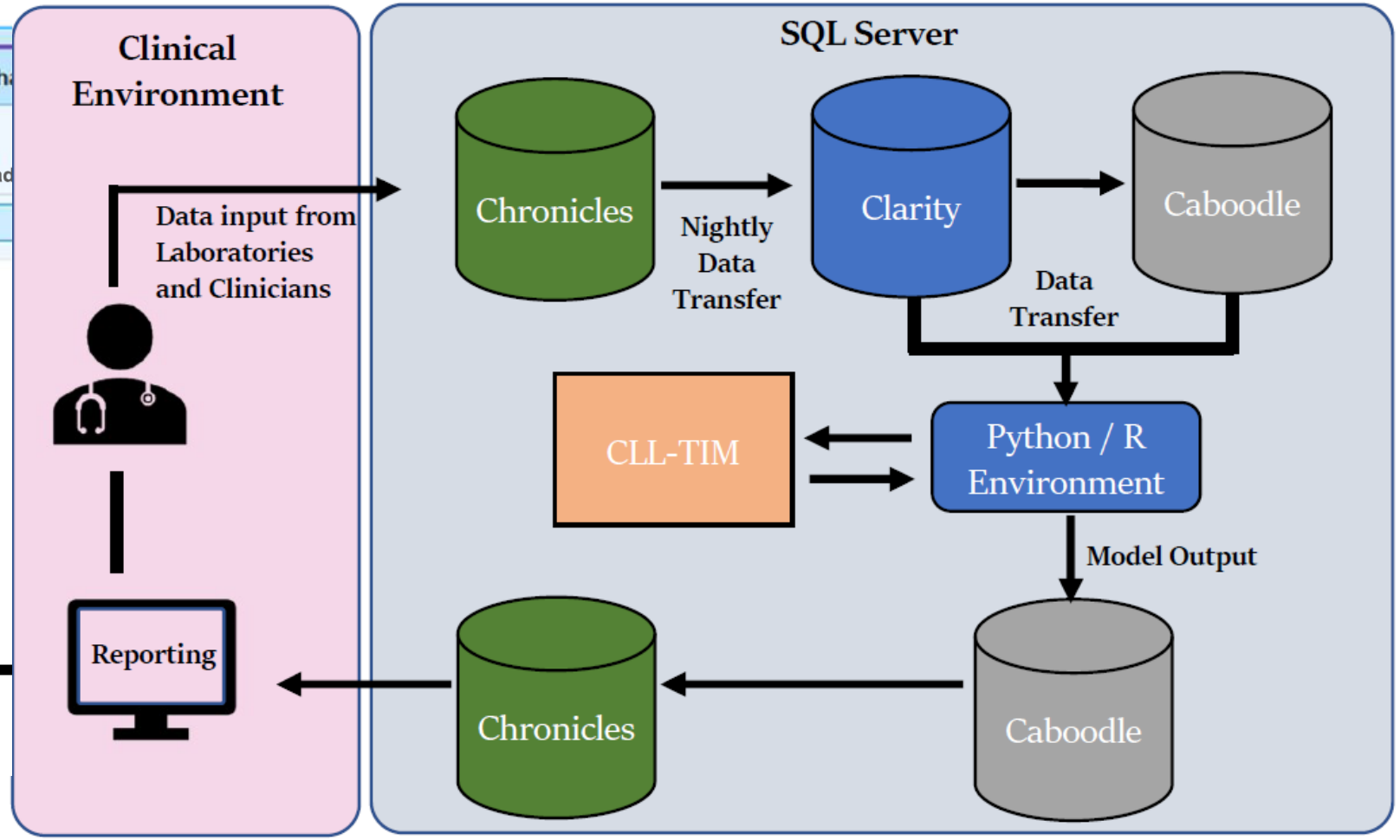




Brieghel et al., Blood Adv, 2024

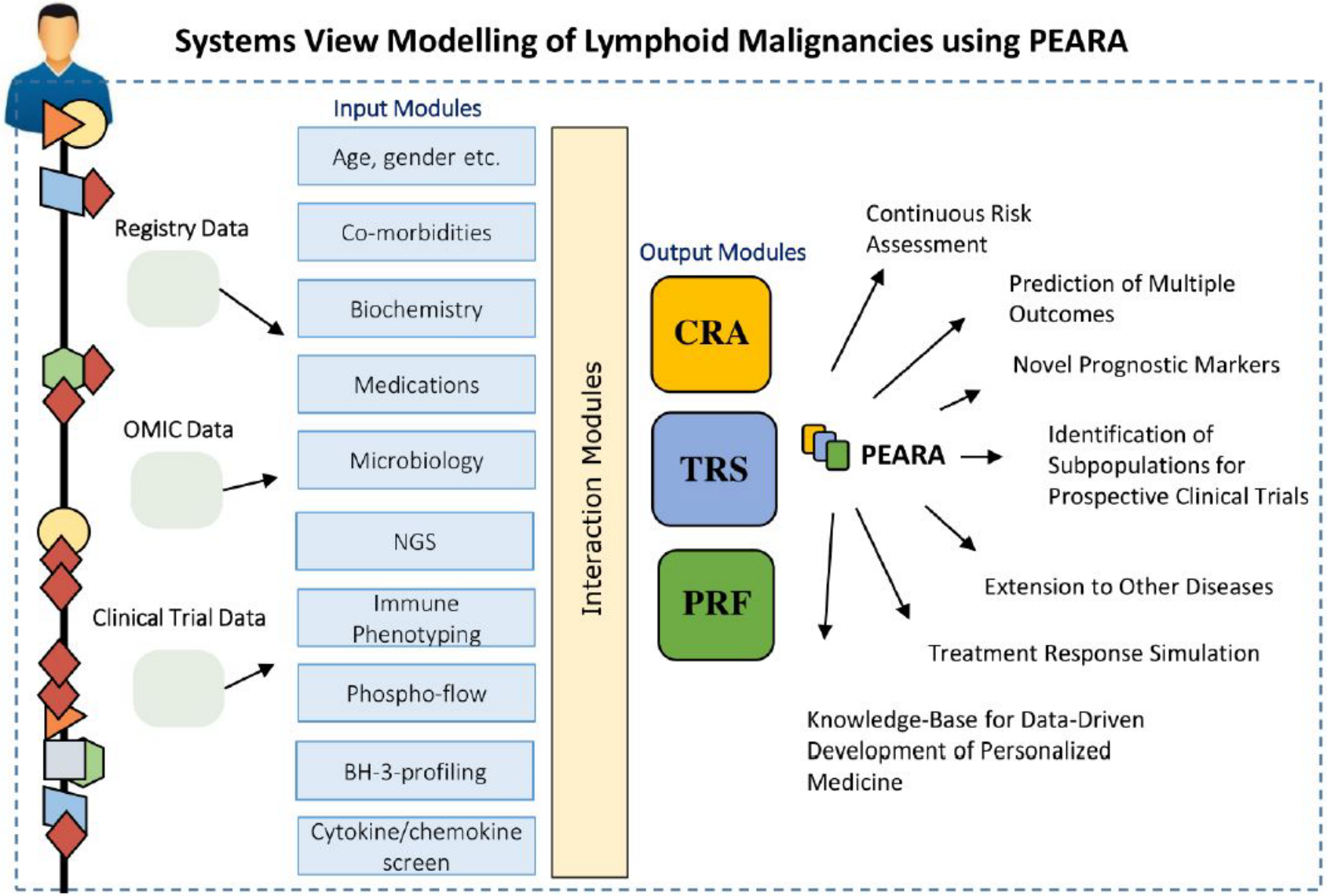
Deploying data-driven hematology – need to automatize and visualize in context!





Agius et al., NPJ dig med, 2024

Perspectives:
Personalized TrEAtment RecommendAtion (PEARA)



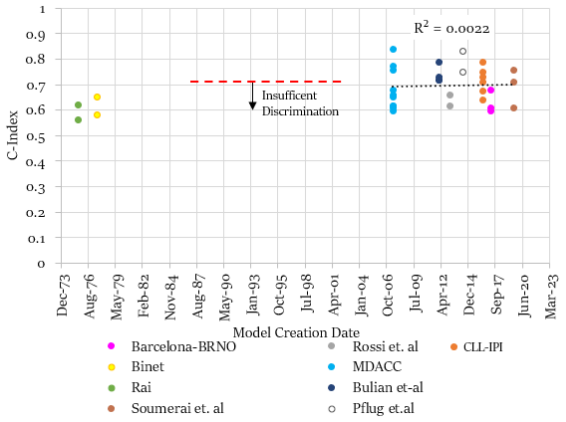
Standardized, open source data formats warranted

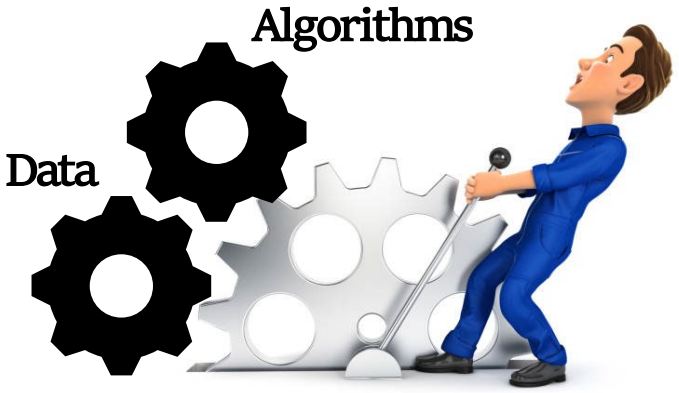
Data source	Official name (abbreviation)	Dataset name	Primary variable	Key mapping
RKKP	Danish National Lymphoma Registry (LYFO)	LYFO	Date_treatment_1st_start	patientid
	Danish National Multiple Myeloma Database (DaMyDa)	MM	Date_treatment_1_start_MPB	patientid
	Danish National Chronic Lymphocytic Leukemia Register (CLL)	CLL	Date_treatment	patientid
SDS	Register of Pharmaceutical Sales (LSR)	EPIKUR	atc	patientid
		EKOKUR	atc	patientid

Data source	Official name	Data set name	Description	Format	Raw data variable names
EPR/SP	AdministreretMedicin	Administreret_Medicin	Contains administered in-hospital pharmacological (w/ ATC codes) and non-pharmacolocial medicine including drug name (e.g. PARACETAMOL FILMOVERTRUKNE TABL. 500 MG) and date-time.	Long; UTF-8	PAT_KEY; PAT_ENC_CSN_KEY; TAKEN_TIME; ATC; MEDICATION_ID; SIG; ZC_MED_UNIT_name; ZC_ADMIN_ROUTE_name; MEDICATION_NAME; MAR_ACTION_C; NAME
	OrdineretMedicin	OrdineretMedicin	Contains prescribed in-hospital pharmacological (w/ ATC codes) and non-pharmacolocial medicine including drug name (e.g. PARACETAMOL FILMOVERTRUKNE TABL. 500 MG) and date-time.	Long; UTF-8	PAT_KEY; PAT_ENC_CSN_KEY; ORDER_MED_ID; PEND_ACTION_C; ORDER_STATUS_C; IS_PENDING_ORD_YN; IS_HELD_ORDER_C; HV_DISCRETE_DOSE; ORDER_START_TIME; ORDER_END_TIME; MEDICATION_ID; NAME; ATC
	ADTHændelser	ADT_Haendelser	Contains pre-hospital and in-hospital admisisons, discharges and transfers (ADT) with name of department/hospital, date-time and corresponding active ICD10 diagnosis codes. Further contains vital status. Caveat! Contains major registration bias before and after implementation of LPR3 on 2 Feb 2019	Long; UTF-8	PAT_KEY; DEATH_DATE; EVENT_TYPE_NAME; EFFECTIVE_TIME; HOSPITAL_AREA_NAME_IN; AFDELING_NAME; ID_NAME; HOSPITAL_AREA_ID; KONTAKT_START_LOCAL_DTTM; KONTAKT_END_LOCAL_DTTM; HOSPITAL_AREA_NAME_OUT; ADMISSION_TYPE; PATIENT_CLASS; CURRENT_ICD10_LIST_A; CURRENT_ICD10_LIST_B
	Flytningshistorik	Flytningshistorik	Contains cleaned information on in-hospital admisisons, discharges and transfers with name of department/hospital, date-time but without any corresponding ICD10 diagnosis codes. Caveat! Contains major registration bias before and after implementation of LPR3 on 2 Feb 2019	Long; UTF-8	PAT_KEY; Tidspunkt; Hospital; Overafdeling; Afsnit; Hændelse; PAT_ENC_CSN_KEY
	ITAOphold	ITAOphold	Contains cleaned information on ICU admission, discharges and transfers with name of department/hospital, date-time and corresponding ICD10 diagnosis codes. Caveat! Contains major registration bias before and after implementation of LPR3 on 2 Feb 2019	Long; UTF-8	PAT_KEY; Aktionsdiagnose; DX_NAME; RESPIRATORSTART; RESPIRATOREND; ICU_STAY_START; ICU_STAY_END; REGION_NAME; OVERAFDELING_NAME; OVERAFDELING_ID; ICU_DEP; AFSNIT_NAME; AFSNIT_KORTNAVN; AFSNIT_ID; ICU_STAY_START; ICU_STAY_END; RESPIRATORSTART; RESPIRATOREND; resp_YN
	AktiveProblemlisteDiagnoser	Aktive_Problemliste_Diagnoser	Contains active in-hospital diagnoses and corresponding ICD10 codes with date-time.	Long; UTF-8	PAT_KEY; CURRENT_ICD10_LIST; DX_NAME; NOTED_DATE
	BehandlingskontakterOgDiagnoser	Behandlingskontakter_diagnoser	Contains any active in-hospital diagnoses and corresponding A or B ICD10 codes with date-time.	Long; UTF-8	PAT_KEY; PAT_ENC_CSN_KEY; SKSKode; DX_NAME; Start date of diagnosis; Kontaktdato; DiagnoseType
	AllePrøvesvar	AlleProevesvar	Contains information on time-date for any perform test such as central lab biochemistry, POC test, ECG and radiology. The corresponding NPU code, value and unit is available for central lab biochemistry, whereas most non-biochemistry results simply refer to other modules.	Long; UTF-8	PAT_KEY; PAT_ENC_CSN_KEY; PROC_NAME; EXTERNAL_NAME; COMPONENT; SPECIMN_TAKEN_TIME; RESULT_TIME; ORD_VALUE
SP	Behandlingsplaner_del1	Behandlingsplaner_del1	Contains information on hem/onc treatment plans such as Rituximab, R-CHOP, or Velcade with corresponding line of therapy, start and end dates. The plan ID links information to Tx_plan 2	Long; UTF-8	PAT_KEY; Plan ID; Plan navn; Protokol navn; Plan type; Status; Behandlingslinje; Behandlingsmål; Planlagte serier; Plan start dato; Plan afbrudt dato; Afbrudt årsag; Behandlingsstartdato; Behandlingsslutdat
	Behandlingsplaner_del2	Behandlingsplaner_del2	Contains information on describing each cycle of treatment for each hem/onc treatment plan. The plan ID links information to Tx_plan 1	Long; UTF-8	PAT_KEY; Plan ID; Serie navn; Serie nummer; Serie status; Serie start
	Bloddyrkning_del1	Bloddyrkning_del1	Contains infomation on any type of microbiology result with date-time including PCR tests for respitatory virus, HBV, EBV, CMV, atypical pneumonia, VZV, and faecal virus as well as results on cultures performed on sputum, BAL, blood, urine, and faeces. Further, contains information on plasma titers for agens such as HBV, EBV, CMV and Asp. gallactomannan and antimicrobial P-concentrations. The Best/Ord ID links information to Microbiology results 2, 3, and 4.	Long; UTF-8	PAT_KEY; Best./Ord. ID; Type; Komponentnavn; Prøveresultat; Prøvetagningstidspunkt; Prøvesvarstidspunkt

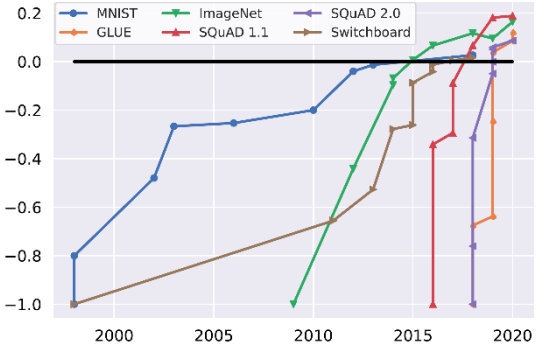
Brieghel et al., medRxiv, 2024

50 years of Prognostic Models for CLL





20 years of Modelling in NLP and Vision

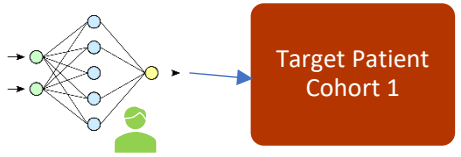


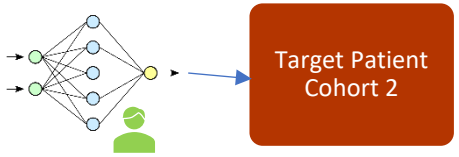
Responsibility for Patient Data = Privacy Protection + **Utilization/Exploitation**

High Privacy Protection
Low Utilization/Exploitation

We need to be somewhere here

Low Privacy Protection
High Utilization/Exploitation



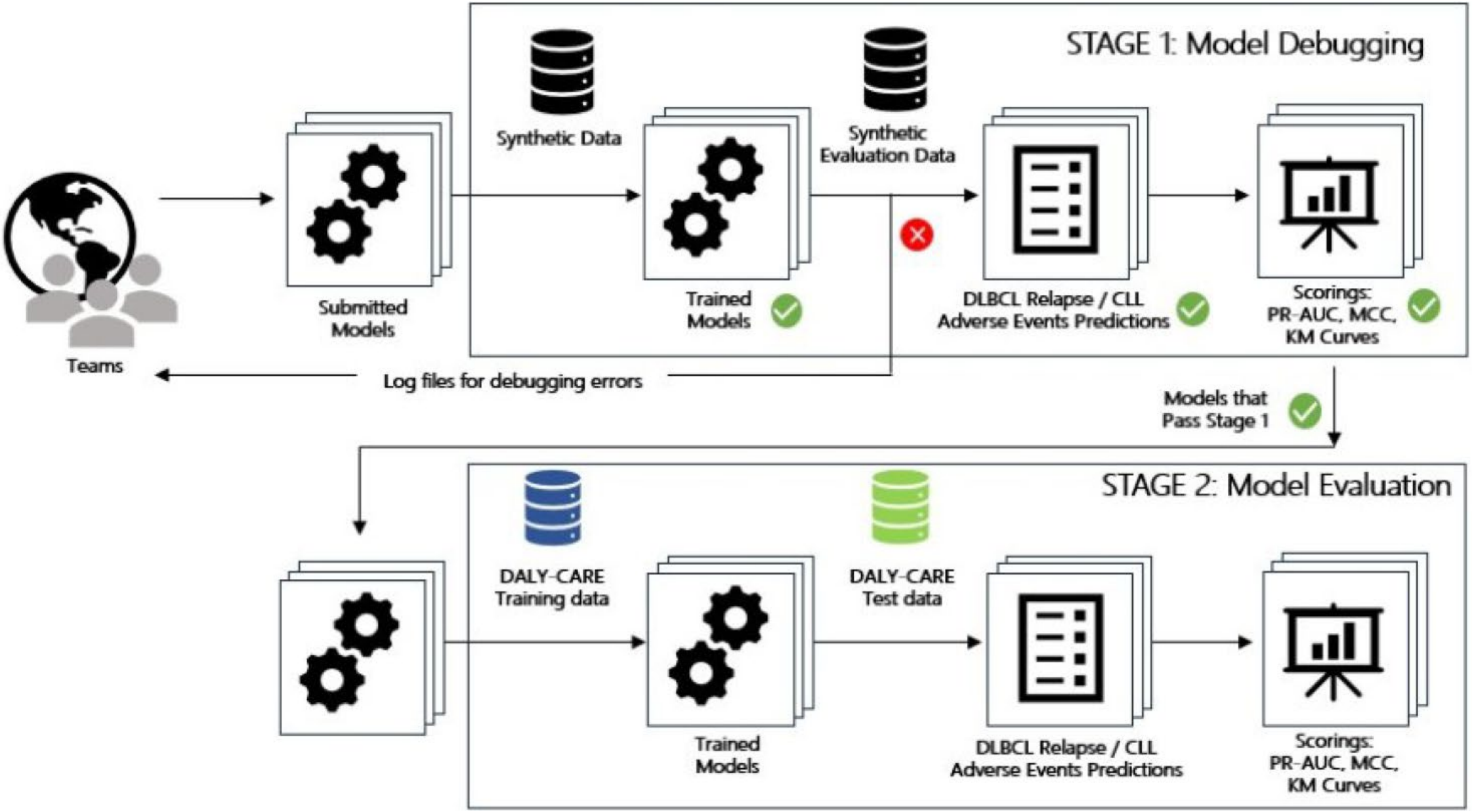


We need to crowd-source patient data

i) Introduce Competition: if you have data

– invite researchers to work on it

NEXT STEPS?


STAGE 1: Model Debugging
Synthetic Data
Synthetic Evaluation Data
Teams
Submitted Models
Trained Models
DLBCL Relapse / CLL Adverse Events Predictions
Scorings: PR-AUC, MCC, KM Curves
Log files for debugging errors
Models that Pass Stage 1
STAGE 2: Model Evaluation
DALY-CARE Training data
DALY-CARE Test data
Trained Models
DLBCL Relapse / CLL Adverse Events Predictions
Scorings: PR-AUC, MCC, KM Curves

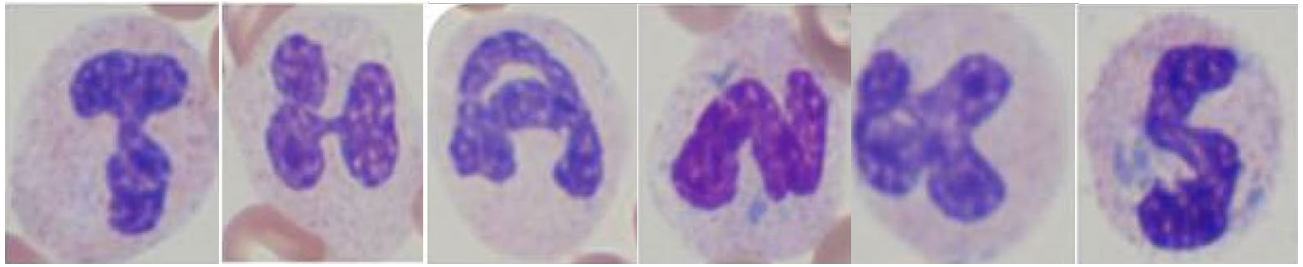
Implementing Data-driven medicine

- Provide open, competitive access to health data!
- Define high impact clinical outcomes to model!
- Merging secondary and primary use of health data!
- Communicate – data scientists AND physician researchers!
- Implement into the clinic

DALY-CARE Danish Lymphoid Cancer Research

Rigshospitalet

Centre for Cancer and Organ Diseases



UNIVERSITY OF COPENHAGEN

CLL Laboratory

Combining translational, epidemiological and clinical research to develop individually tailored supportive care and CLL specific treatment

Andreas Katsimigas
Caspar da Cunha-Bang
Casper M Frederiksen
Christian Brieghel
Emelie Rotbain
Hashim Elhussein
Lone Bredo Pedersen
Lotte Jacobsen
Mehdi Parviz
Mikkel Werling
Noomi Vainer
Rudi Agius
Rebecka Svanberg
Tereza Faitova
Thomas Lacoppidan

Nordic CLL Study Group

Carsten.utoft.niemann@regionh.dk
www.rigshospitalet.dk/CLL-lab

REGION

novo nordisk fonden
Danish Cancer Society
ERA PerMed
Benzon Foundation
Alfred Benzon Foundation