



Study Report

P4-C1-007 / P4-C2-011

DARWIN EU® - Assessing the feasibility of investigating acute myeloid leukaemia: incidence, patient characteristics, treatments, and survival in the period 2015–2024

23/07/2026

Version 6.0

Authors: Anton Barchuk, Adam Black, Talita Duarte-Salles

Public

CONTENTS

LIST OF ABBREVIATIONS.....	6
1. TITLE	8
2. DESCRIPTION OF THE STUDY TEAM	8
3. ABSTRACT	10
4. AMENDMENTS AND UPDATES	13
5. MILESTONES	13
6. RATIONALE AND BACKGROUND	13
7. RESEARCH QUESTION AND OBJECTIVES.....	14
8. RESEARCH METHODS.....	14
8.1. Study design	14
Figure 1. Graphical depiction of the characterisation study.	15
8.2. Follow-up	15
Figure 2. Included observation time for the denominator population.	16
8.3. Study population with inclusion and exclusion criteria.....	16
8.4. Study setting and data sources	16
Table 1. Data sources.....	17
8.5. Study period	19
8.6. Variables	19
8.6.1. Exposure	19
8.6.2. Outcome	19
8.6.3. Covariates, including confounders, effect modifiers, intercurrent events, and other variables .	19
Table 2. Treatment regimens in AML, which were considered when assessing results of treatment patterns.	21
8.7. Study size	21
8.8. Data transformation	21
8.9. Statistical methods	22
8.9.1. Main summary measures	22
8.9.2. Main statistical methods	22
Objective 1: Incidence of AML.....	22
Objective 2–4: Patient-level characterisation	22
8.9.3. Missing values.....	23
8.9.4. Sensitivity analysis	23
9. RESULTS	23
9.1. Participants.....	23
Table 3. Participant attrition in the study.....	23
9.2. Descriptive data	24
Table 4. Participants characteristics at the time of acute myeloid leukaemia diagnosis.....	25
9.3. Main results	26
9.3.1. Incidence of AML.....	26
Table 5. Incidence per 100,000 of acute myeloid leukaemia, overall and in subgroups by age and calendar years, with 95% confidence interval.....	27
Figure 3. Incidence per 100,000 of acute myeloid leukaemia by calendar year in Croatia (NAJS), Denmark (DK-DHR), Finland (FinOMOP-THL), Norway (CRN), and the Netherlands (NCR) (applying	

the Eurostat national denominator population) and in InGef RDB and SIDIAP (applying data source denominator population).	28
Figure 4. Incidence per 100,000 of acute myeloid leukaemia by calendar year and age group in Croatia (NAJS), Denmark (DK-DHR), Finland (FinOMOP-THL), Norway (CRN), and the Netherlands (NCR) (applying the Eurostat national denominator population) and in InGef RDB and SIDIAP (applying data source denominator population).	29
9.3.2. Characterisation	30
Table 6. History of myelodysplastic syndrome and malignancies prior to diagnosis of acute myeloid leukaemia.	30
Table 7. Presence of measurement of genetic mutations in individuals with acute myeloid leukaemia in the first 30 days after diagnosis.....	30
Table 8. Presence of concepts for measurement of blasts in bone marrow, blood, and lactate dehydrogenase in blood in the first 30 days after initial diagnosis.....	31
Figure 5. Timing of measurement of blasts in bone marrow in individuals with acute myeloid leukaemia.	31
Figure 6. Timing of measurement of blasts in blood in individuals with acute myeloid leukaemia. ..	32
Figure 7. Timing of measurement of lactate dehydrogenase in blood in individuals with acute myeloid leukaemia.	32
Table 9. Presence of information on WHO/ECOG performance status in individuals with acute myeloid leukaemia in the first 30 days after diagnosis.	33
9.3.3. Prevalence and timing of treatment	33
Figure 8. Timing of azacitidine, cytarabine, daunorubicin, idarubicin, and venetoclax use and allogeneic stem cell transplantation in individuals with AML in the age group 40 to 59 years old....	35
Figure 9. Timing of azacitidine, cytarabine, daunorubicin, idarubicin, and venetoclax use and allogeneic stem cell transplantation in individuals with acute myeloid leukaemia in the age group 60 to 79 years old.	36
9.3.4. Treatment patterns	36
Table 10. Most common treatment pathways recorded in different data sources in individuals with acute myeloid leukaemia (all age groups).	38
Figure 10. Most common treatment pathways (recorded in different data sources in individuals with acute myeloid leukaemia).	42
9.3.5. Overall survival of individuals with AML	42
Figure 11. Overall survival probabilities of individuals with acute myeloid leukaemia (all age groups).	43
Table 11. Overall survival estimates of individuals with acute myeloid leukaemia (all age groups). .	43
Figure 12. Overall survival probabilities of individuals with acute myeloid leukaemia by sex.	44
Figure 13. Overall survival probabilities of individuals with acute myeloid leukaemia by age (<18 and 18 and older).	45
Table 12. Overall survival estimates of individuals with acute myeloid leukaemia by age group.	45
Figure 14. Overall survival probabilities of individuals with acute myeloid leukaemia by age (0 to 17; ≥18; 18 to 39; 40 to 59; 60 to 79; 80 and over).	47
10. DISCUSSION	47
10.1. Key results	47
10.2. Strengths and limitations of the research methods.....	48
10.3. Interpretation	50
11. CONCLUSION.....	52
12. REFERENCES	53
13. ANNEXES.....	55

ANNEX I. Description of data sources.....	55
ANNEX II. Fitness for use assessment.....	58
ANNEX III. Operational and reporting considerations.....	60
ANNEX IV. List of stand-alone documents.....	61
Table S1. List of concepts to describe acute myeloid leukaemia (excluding acute promyelocytic leukaemia).	61
Table S2. List of concepts to describe acute promyelocytic leukaemia.	86
Table S3. List of concepts to describe medicines.	88
Table S4. List of concepts to describe procedures.	89
Table S5. List of concepts to describe measurement (blood).	90
Table S6. List of concepts to describe WHO/ECOG performance status.....	91
Table S7. List of concepts to describe genetic mutations.	92
ANNEX V. Glossary.....	97

Study title	DARWIN EU® - Assessing the feasibility of investigating acute myeloid leukaemia: incidence, patient characteristics, treatments, and survival in the period 2015–2024
Study report version	V6.0
Date	23/07/2026
EUPAS number	EUPAS1000000773
Active substance	N/A
Medicinal product	N/A
Research question and objectives	What Real-World Data and evidence is currently available to contextualise treatment choices and outcomes in AML patients? This study aimed to estimate the incidence of AML and characterise patients with AML in terms of comorbidities, diagnostic tests, treatments, and survival in DARWIN EU® network. Specific objectives were: 1. To estimate the annual incidence of AML in children (<18 years old) and adults (≥18). 2. To characterise AML patients in terms of demographics, performance status, conditions, disease characteristics, measurements, and procedures. 3. To describe treatment patterns in AML patients. 4. To estimate the overall survival at 1, 3, and 5 years for patients with newly diagnosed AML.
Countries of study	Croatia, Denmark, Finland, Germany, the Netherlands, Norway, Spain
Authors	Anton Barchuk (a.barchuk@darwin-eu.org) Adam Black (a.black@darwin-eu.org) Talita Duarte-Salles (t.duarte@darwin-eu.org)

LIST OF ABBREVIATIONS

Acronyms/term	Description
AML	Acute myeloid leukaemia
APL	Acute Promyelocytic Leukaemia
CC	Coordination centre
CDM	Common Data Model
CI	Confidence Interval
CRN	Cancer Registry Norway
DARWIN EU®	Data Analysis and Real-World Interrogation Network
DK-DHR	Danish Health Data Registries
DRE	Digital Research Environment
ECOG	Eastern Cooperative Oncology Group
EHR	Electronic Health Records
EMA	European Medicines Agency
EU	European Union
EUPAS	EU Post-Authorisation Studies Register
FinOMOP-HUS	Hospital District of Helsinki and Uusimaa
FinOMOP-THL	Finnish Care Register for Health Care (THL)
GDPR	General Data Protection Regulation
H12O	Hospital Universitario 12 de Octubre
HARMONY-AML	HARMONY - Acute Myeloid Leukaemia
HCUP-NIS	Healthcare Cost and Utilisation Project National Inpatient Sample
ICC	International Consensus Classification
ICD	International Classification of Diseases
ICDO	International Classification of Diseases for Oncology
InGef RDB	InGef – Institut für angewandte Gesundheitsforschung Berlin GmbH
IP	Inpatient
IRB	Institutional Review Board
LDH	Lactate dehydrogenase
MDS	Myelodysplastic Syndrome/Myelodysplastic neoplasia
NAJS	Croatian National Public Health Information System
NCR	Netherlands Cancer Registry
OHDSI	Observational Health Data Sciences and Informatics
OMOP	Observational Medical Outcomes Partnership
OP	Outpatient
RMST	Restricted mean survival time
RWE	Real-World Evidence

Acronyms/term	Description
SD	Standard deviation
SHI	Statutory health insurance
SIDIAP	The Information System for Research in Primary Care
SNOMED	Systematised Nomenclature of Medicine
WHO	World Health Organisation

1. TITLE

DARWIN EU® - Assessing the feasibility of investigating acute myeloid leukaemia: incidence, patient characteristics, treatments, and survival in the period 2015–2024

2. DESCRIPTION OF THE STUDY TEAM

Study team role	Names	Organisation
Principal Investigator	Anton Barchuk Talita Duarte-Salles	Erasmus MC
Data Scientist	Adam Black Cesar Barboza	Erasmus MC
Clinical Domain Expert	Anton Barchuk	Erasmus MC
Statistician	Adam Black	Erasmus MC
Study Manager	Natasha Yefimenko Nosova	Erasmus MC
Data source	Names	Data Partner Organisation*
Study P4-C1-007		
NAJS	Jakov Vukovic Ivan Pristas Anamaria Jurcevic Antea Jezdzic Marko Cavlina Karlo Pintaric Mario Sekerija	Croatian Institute of Public Health
InGef RDB	Raeleesha Norris Annika Vivirito Alexander Harms	InGef – Institut für angewandte Gesundheitsforschung Berlin GmbH
NCR	Jelle Evers Maaïke van Swieten Avinash Dinmohamed	Netherlands Comprehensive Cancer Organisation (IKNL)
HARMONY-AML	María Abáigar Ángela Villaverde Mireya Morote	HARMONY Alliance Foundation
SIDIAP	Laura Granés González Irene López Sanchez Agustina Guiliadori Picco Elena Roel Herranz	IDIAPJGol
Data source	Names	Data Partner Organisation*
Study P4-C2-011		
DK-DHR	Elvira Bräuner Susanne Bruun	Danish Medicines Agency
FinOMOP-HUS	Eric Fey Marianna Niemi Kimmo Porkka	Helsinki University Hospital
FinOMOP-THL	Petteri Hovi Toni Lehtonen Laura Salonen	The Finnish Institute for Health and Welfare (THL)

CRN	Espen Enerly Lise Enerstvedt Anna Skog	Norwegian Institute of Public Health
H12O	Paula Rubio Mayo Javier de la Cruz Juan Luis Cruz Bermúdez Noelia García Barrio	Hospital Universitario 12 de Octubre

*Data partners do not have an investigator role. Data partners execute code at their data source, review, and approve their results.

3. ABSTRACT

Title

DARWIN EU® - Assessing the feasibility of investigating acute myeloid leukaemia: incidence, patient characteristics, treatments, and survival in the period 2015–2024

Rationale and background

Acute myeloid leukaemia (AML) is an aggressive haematological malignancy characterised by the uncontrolled proliferation of myeloid precursors. It primarily affects older adults but is also diagnosed in children. Despite advances in diagnostics and treatment, survival remains poor, especially among elderly patients. Diagnosis relies on blast counts and genetic markers, with evolving classifications emphasising molecular features. Standard treatment includes induction chemotherapy and, in eligible patients, allogeneic stem cell transplant. Although novel targeted therapies show promise, further evidence is needed to support their safety.

Research question and objectives

What Real-World Data and evidence are currently available to contextualise treatment choices and outcomes in AML patients?

This study aimed to estimate the incidence of AML and characterise patients with AML in terms of comorbidities, diagnostic tests, treatments, and survival in the DARWIN EU® network.

Methods

Study design

Retrospective cohort studies were conducted using routinely collected health data from 10 data sources across 7 European countries.

Population

For estimating incidence of AML, we included all individuals present in the respective data sources during the study period (January 1, 2015, to December 31, 2024, or until the end of available data). For characterising AML patients, treatments, and survival, we included all patients diagnosed with AML within the same period.

Variables

Sex (female/male), age (continuous and age groups 0 to 17; ≥18; 18 to 39; 40 to 59; 60 to 79; ≥80 years), calendar year, performance status, cytogenetic risk type, and genetic mutations.

Outcomes

Death from any cause.

Data sources

A total of 10 data sources across 7 countries were used: two cancer registries (Netherlands Cancer Registry (NCR), Cancer Registry Norway (CRN)), three public health nationwide registries (Danish Data Health Registries (DK-DHR), Population Registers at the Finnish Institute for Health and Welfare (FinOMOP-THL), and Information System (NAJS), (Croatian National Public Health); two hospitals (Hospital District of Helsinki and Uusimaa (FinOMOP-HUS) and Hospital Universitario 12 de Octubre (H12O)); one data source that combines information from several hospital (HARMONY - Acute Myeloid Leukaemia (HARMONY-AML)), one primary care data source (The Information System for Research in Primary Care (SIDIAP)); and one claims data source (InGef Research Database (InGef RDB)).

Statistical analysis

Annual incidence rates per 100,000 person-years were estimated overall and stratified by sex, age, and calendar year. The statistical analyses were performed on the Observational Medical Outcomes Partnership Common Data Model (OMOP CDM)-mapped data using the *IncidencePrevalence* package. For data sources with national coverage, official national population estimates were used to calculate incidence; for others, we used the data source population denominator.

Demographics (age and sex) were described at the time of AML diagnosis. Large-scale characterisation was conducted using the *CohortCharacteristics* package at any time before, at the date of AML diagnosis, and two years after, and included demographics, selected comorbidities, medications, procedures, and results were reviewed for the presence of expected characteristics of AML patients (cytogenetic risk profile, genetic mutations, procedures).

All pre-specified medicines at the ingredient (active compound) level and their combinations initiated in patients with AML were described using the *TreatmentPatterns* package.

Overall survival probabilities and restricted mean survival time (RMST) at 1, 3, and 5 years, as well as median survival, were calculated using the *CohortSurvival* package. Results were presented with 95% confidence intervals.

A minimum cell count of 5 was applied to all reporting results, with any smaller count reported as “<5”.

Results

Overall, 30,542 individuals were included in the study (ranging from 631 in FinOMOP-HUS to 7,784 in NCR). The mean age at diagnosis ranged from 55 years in HARMONY-AML to 68 years in DK-DHR.

Acute Promyelocytic Leukaemia (APL) represented a small proportion (2.4–4.9%) of all identified AML patients in all data sources, except in H12O, where the majority of individuals with AML had APL (84.1%).

Estimates of the overall incidence of AML per 100,000 stayed in the range between 2.9 per 100,000 (95% confidence interval (CI) [2.8; 3.1]) in SIDIAP to 8.6 per 100,000 (95% CI [7.7; 9.6]) in NAJS. Incidence increased markedly with age across all data sources, and incidence was higher in males than in females. Calendar-year trends from 2015 to 2024 generally showed stable incidence rates across most data sources, with no clear temporal trend. Incidence in the paediatric population (less than 18 years old) ranged from 0.3 per 100,000 (95% CI [0.2; 0.5]) in SIDIAP to 1.8 per 100,000 (95% CI [0.9; 3.2]) in NAJS.

The proportion of individuals with a recorded history of myelodysplastic syndrome (MDS) any time until 90 days before AML diagnosis ranged from 7.8% in SIDIAP to 16.1% in FinOMOP-THL.

Recorded genetic mutation measurements within the first 30 days after diagnosis were observed in individuals with AML across three data sources: FinOMOP-HUS, HARMONY-AML, and H12O.

World Health Organisation/Eastern Cooperative Oncology Group (WHO/ECOG) performance status scores for AML patients were available in only two data sources: HARMONY-AML and NCR. In HARMONY-AML, performance status was recorded for 63.3% of individuals. Among them, ECOG 0 was the most common, in 36.5% of individuals (n=587), followed by ECOG 1 in 23.6% (n=380) and ECOG 2 in 3.1% (n=51); there were no records of individuals with ECOG 3 or 4 in this data source. In NCR, the majority of individuals had no recorded performance status score (57.9%). Of those with a recorded ECOG status, ECOG 0 was observed in 17.6% of individuals with AML (n=1,370), followed by ECOG 1 in 15.3% (n=1,191), ECOG 2 in 5.5% (n=429), ECOG 3 in 2.53% (n=197), and ECOG 4 in 1.12% of individuals with AML (n=87).

Across all data sources, azacitidine, cytarabine, daunorubicin, idarubicin, venetoclax, fludarabine, etoposide, gemtuzumab, ozogamicin, and allogeneic stem cell transplantation were the most frequently recorded treatments over different 30-day periods. In several data sources (CRN, DK-DHR, FinOMOP-HUS, NCR), cytarabine, daunorubicin, and idarubicin had the most records in the initial 30-day period post

diagnosis, followed by a sharp decline, whilst venetoclax and azacitidine maintained relatively high proportions throughout subsequent 30-day periods. Allogeneic stem cell transplantation was more common 90 days following AML diagnosis. The AML treatment pathways could be broadly grouped into azacitidine and venetoclax, cytarabine, and transplantation as part of the combined treatment and varied across data sources.

Stratified results in paediatric populations with AML were not available due to the small sample size.

Overall survival estimates varied considerably across different data sources and tended to be higher in hospital data sources than in cancer registry data sources. Median survival was not reached in InGef RDB, H12O, and HARMONY-AML. Among data sources with an estimable median, it ranged from 8.2 months (95% CI [7.5; 9.1]) in DK-DHR to 15.7 months (95% CI [13.4; 20.5]) in FinOMOP-HUS and 15.6 months (95% CI [14.0; 18.0]) in SIDIAP.

One-year survival probability in registries ranged from 42.4% in DK-DHR to 50.6% [48.8; 52.4] in NAJS, with 48.1% [46.2; 50.1] in FinOMOP-THL, 48.8% [46.5; 51.1] in CRN, and 44.6% [43.4; 45.7] in NCR. It was 55.9% [53.3; 58.6] in SIDIAP, a primary healthcare data source, and 55.0% [51.2; 59.1] in FinOMOP-HUS, a hospital data source.

Three data sources were outliers with one-year survival probabilities substantially higher (both overall and in the paediatric population) than the values reported in the registry data sources: 93.0% [92.1; 93.8] in H12O (predominantly APL patients), 72.6% [70.4; 74.9] in HARMONY-AML, and 78.8% [77.3; 80.2] in InGef RDB.

Discussion

In this study, we estimated the incidence of AML, characterised AML patients by demographics, performance status, comorbidities, disease characteristics, measurements, and procedures, described treatments, and estimated overall survival for patients newly diagnosed with AML.

Overall, AML epidemiological estimates are broadly consistent with established knowledge, taking into account the types of data available and the characteristics of individuals with AML across different data sources. Our findings also suggest that some data sources are not representative of the overall AML population and that the generalisability of findings from them is limited.

However, substantial missingness in diagnostic, molecular, and treatment-pathway data are observed, highlighting challenges for AML research using observational data. We also observed limitations in the treatment pathway characterisation tools due to the inherent complexity of the treatment guidelines. This study can inform the fit-for-purpose assessment of future studies on AML, the current needs for improvements in both data quality and analytical tools for treatment patterns in oncology, and the creation of more advanced disease definitions, such as relapse/refractory AML.

4. AMENDMENTS AND UPDATES

None.

5. MILESTONES

Study deliverable	Timelines (planned)	Timelines (actual)
Final Study Protocol	August 2025	August 2025
Creation of Analytical code	September 2025	September/October 2025
Execution of Analytical Code on the data	October/November 2025	October 2025–March 2026
Draft Study Report	November 2025	April 2026
Final Study Report	December 2025	May 2026

*Planned dates are dependent on obtaining approvals from the internal review boards of the data sources and deliverables review rounds.

6. RATIONALE AND BACKGROUND

Acute Myeloid Leukaemia (AML) is an aggressive haematological malignancy characterised by the clonal expansion of myeloid precursor cells in the bone marrow and peripheral blood.[1] AML affects both adults and children, though its incidence increases markedly with age. It is the most common form of acute leukaemia in adults, with a median age at diagnosis of approximately 68 years in high-income countries.[2] In children, AML is relatively rare and accounts for only about 15–20% of acute leukaemias.[1,3]

The pathogenesis of AML involves a series of genetic and epigenetic alterations that disrupt normal myeloid differentiation and proliferation. This leads to the accumulation of immature blast cells in the bone marrow, which suppresses normal haematopoiesis. AML has a rapid onset of symptoms, including fatigue and shortness of breath on exertion due to anaemia; recurrent infections due to neutropenia; and an increased tendency to bruise and bleed due to thrombocytopenia.[4]

The estimated five-year overall survival probability in AML was approximately 30% in studies based on data from the US and Japan. 5-year survival probabilities vary significantly by age, with younger patients achieving survival rates of up to 50%, while rates fall below 10% in older patients.[5,6] The relative overall 5-year survival rate in AML in Europe from around 2012 has been reported to be 17–18%, which is among the lowest rates compared to other myeloid malignancies.[7,8] In light of advances in diagnostics and risk stratification, the collection of detailed information (immunophenotype, genetic abnormalities, molecular data, and clinical data) in AML is becoming increasingly important for surveillance.[7] Relapse remains common, particularly in patients with high-risk subtypes or those who fail to achieve deep remissions after induction therapy.

AML diagnosis is based on a specific blast threshold (indicating the presence of blasts in the bone marrow or peripheral blood), when there is evidence of extramedullary myeloblasts, and/or presence of AML-defining genetic abnormalities. Various AML classification systems exist.[9,10] A recent comparative study of the 5th edition of the World Health Organisation (WHO) Classification (2022) and the International Consensus Classification (ICC) for AML highlighted a shift toward genetically defined subtypes. Morphologically defined AML categories decreased significantly from 13% to 5%, while myelodysplasia-related AML increased to 28% in the WHO classification and to 26% in the ICC classification, with other genetically defined AML remaining the largest group.[11] In parallel with changes in disease classification, adjustments have also been made to AML risk classification.

First-line treatment for AML in fit patients typically consists of induction chemotherapy (commonly cytarabine and anthracyclines), followed by consolidation with high-dose chemotherapy (typically

cytarabine) and, if indicated and a response to induction is obtained, then an allogeneic stem cell transplant in transplant-eligible patients.

Younger patients without significant comorbidities are often considered fit for chemotherapy, but defining clear eligibility remains complex. The goal of intensive treatment is to induce complete remission and follow with post-remission therapy (chemotherapy and/or allogeneic stem cell transplantation) guided by clinical factors, molecular risk profile, donor availability, and patient preferences. The introduction of new agents, such as FLT3, IDH1/2, BCL-2, DNMT, and hedgehog pathway inhibitors, has increased therapeutic options [1]; however, these therapies often lack long-term safety data in routine clinical settings. In addition, patients frequently experience substantial side effects from treatment, including infections, organ toxicity, mucositis, and prolonged cytopenia.

Given these factors, conducting post-marketing studies on AML treatments is crucial. They also provide valuable Real-World Evidence (RWE) that can inform clinical practice and improve patient outcomes.

This study aimed to generate RWE on the characteristics, treatments, and survival of patients with AML using routinely collected healthcare and cancer registry data.

By analysing treatment trajectories and survival across diverse populations, this study provides evidence to contextualise clinical decision-making and support regulatory evaluations of emerging therapies. The results of this study can also be the basis for further work on dedicated algorithms to improve data quality in AML.

7. RESEARCH QUESTION AND OBJECTIVES

What Real-World Data and evidence are currently available to contextualise treatment choices and outcomes in AML patients?

This study aimed to estimate the incidence of AML and characterise patients with AML in terms of comorbidities, diagnostic tests, treatments, and survival in the DARWIN EU® network.

Specific objectives were:

1. To estimate the annual incidence of AML in children (<18 years old) and adults (≥18).
2. To characterise AML patients in terms of demographics, performance status, conditions, disease characteristics, measurements, and procedures.
3. To describe treatment patterns in AML patients.
4. To estimate the overall survival at 1, 3, and 5 years for patients with newly diagnosed AML.

8. RESEARCH METHODS

8.1. Study design

Retrospective cohort studies were conducted using routinely collected health data from 10 data sources across 7 European countries:

- A descriptive epidemiology study to address objective 1, assessing the incidence of AML.
- A characterisation study to address objectives 2–4, assessing characteristics of patients newly diagnosed with a condition of interest (objective 2), treatment patterns (objective 3), and survival of patients with AML (objective 4).

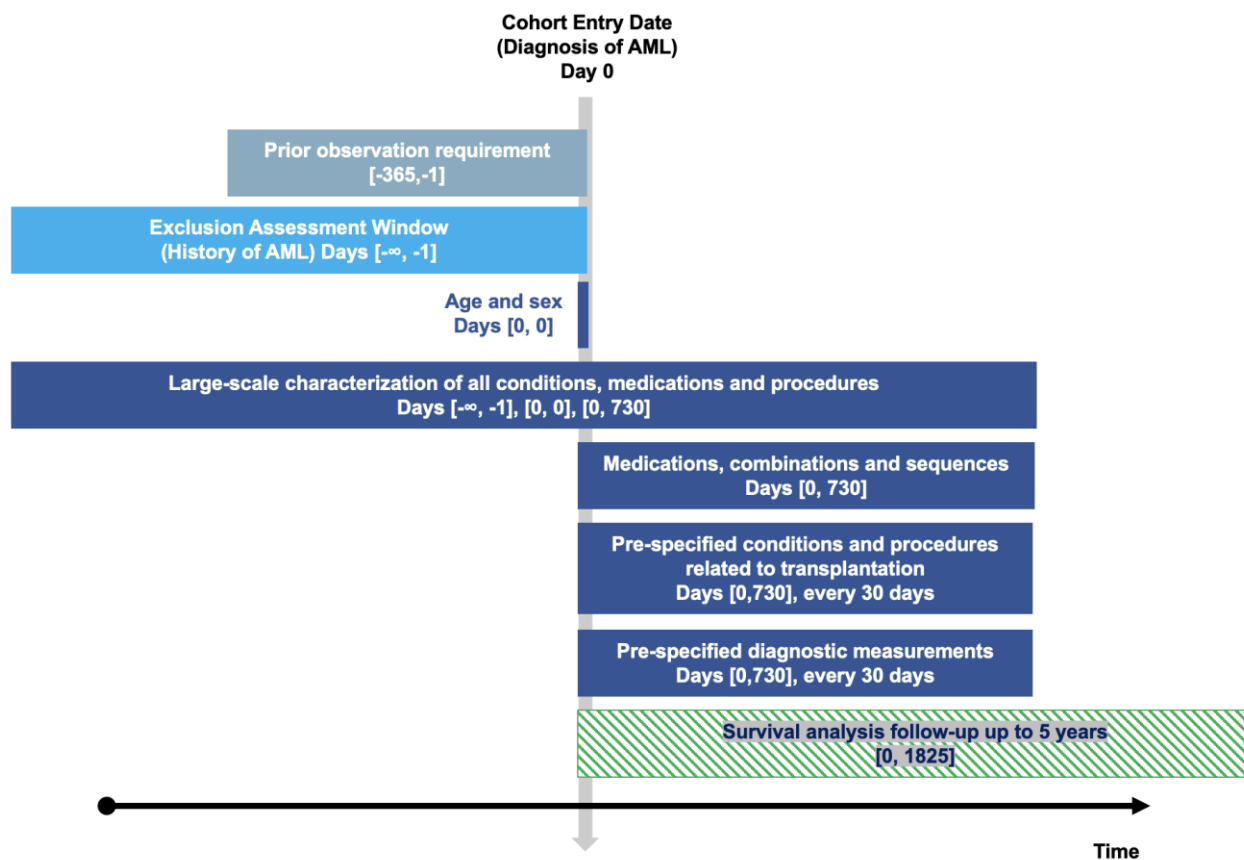


Figure 1. Graphical depiction of the characterisation study.

a) patients with 365 days of prior history and no prior history of AML were included; b) demographics was described at the date of AML diagnosis; c) assessment windows for large-scale characterisation were based on the AML diagnosis date: any time before one day before AML diagnosis $[-\infty, -1]$, at the date of AML diagnosis $[0, 0]$, up to two years or 730 days after AML diagnosis $[1, 730]$; d) treatment combinations and sequences were described within 2 years (730 days) after AML diagnosis; e) pre-specified procedures and measurement were described within sequential 30-day periods after the AML diagnosis date until two years (730 days); f) follow-up up to 5 years (1825 days) was used in survival analysis.
AML=Acute Myeloid Leukaemia

8.2. Follow-up

For the descriptive epidemiological study (objective 1 - incidence of AML using data source denominator population), follow-up started from the latest of the following:

- the study start date (01/01/2015),
- the date at which individuals in the data source had 365 days of prior history.

The end of follow-up was defined as the earliest of the following:

- the date of AML diagnosis,
- loss to follow-up,
- death,
- the end of the observation period (31/12/2024).

An example of entry and exit into the denominator population is shown in **Figure 2**. In this example, person ID 1 already had a sufficient prior history before the study start date, and the observation period ended after the study end date; therefore, this person contributed to the study throughout its entire duration. Person IDs 2 and 4 entered the study only when they had sufficient prior history. Person ID 3 left the data source upon exiting (the end of the observation period). Lastly, person ID 5 had two observation periods in the data source. The first period spanned from the study's start to the end of the observation period. The second period resumed once sufficient prior history was reached and ended at the study end date.

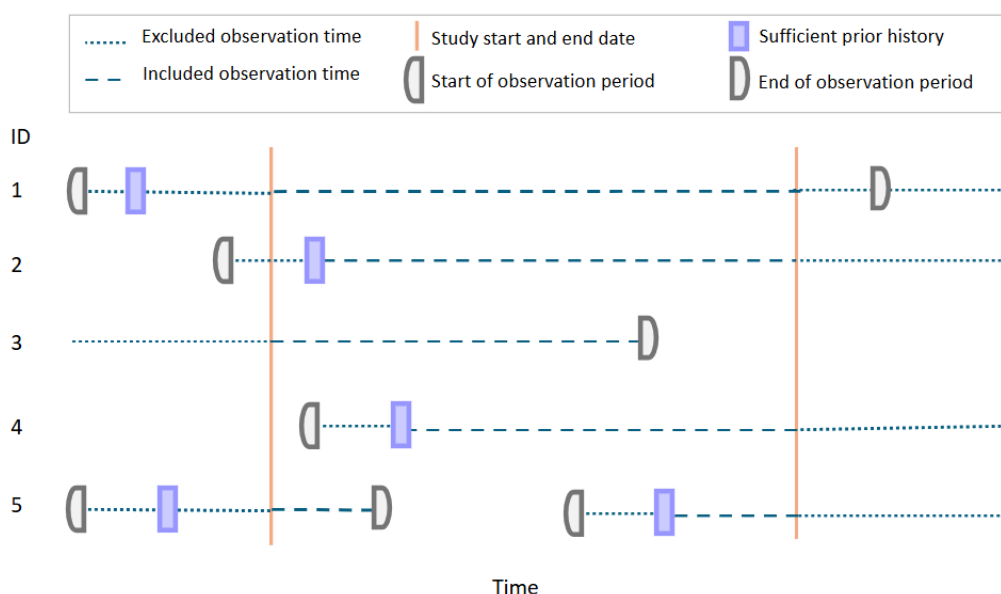


Figure 2. Included observation time for the denominator population.

8.3. Study population with inclusion and exclusion criteria

For the estimation of incidence, all individuals present in the data source were included in the denominator, in accordance with the requirements described in **Section 8.2**. Only individuals with newly diagnosed AML were included in the study for objectives 2–4.

Objective 1 inclusion criteria

- Observation in the period 01/01/2015 to 31/12/2024.
- Minimum 365 days of available history before contributing to person-time-at-risk (except for children <1 year).

Objective 2–4 inclusion criteria

- First diagnosis of AML in the period of 01/01/2015 to 31/12/2024.
- Minimum 365 days of available history before diagnosis date (except for children <1 year).
- No prior history of AML before the study period.

8.4. Study setting and data sources

This study was conducted using routinely collected data from 10 data sources in the DARWIN EU® network of data partners across 7 European countries (**Table 1**). All data were a priori mapped to the Observational Medical Outcomes Partnership Common Data Model (OMOP CDM).

Table 1. Data sources.

Country	Name of Data source	Health Care setting	Type of Data	Number of individuals ¹	Period covered ²	Contributing to	Denominator source ³
Public health registries							
Croatia	NAJS	Primary care, Hospital care (IP, OP)	Registries, EHRs	4.79 M	2014 to 2025	Objectives 1–4	External (Eurostat)
Denmark	DK-DHR	Primary care (prescriptions as a proxy), Hospital care (IP, OP)	Registries, EHRs	8.59 M	1995 to 2024	Objectives 1–4	External (Eurostat)
Finland	FinOMOP-THL	Primary care, Hospital care (IP, OP)	Registry	6.62 M	2011 to 2024	Objectives 1–4	External (Eurostat)
Cancer registries							
Netherlands (the)	NCR	Hospital care (IP, OP)	Registry	2.76 M	1993 to 2024	Objectives 1–4	External (Eurostat)
Norway	CRN	Hospital care (IP, OP)	Registry	1.25 M	1953 to 2024	Objectives 1–4	External (Eurostat)
Claims							
Germany	InGef RDB	Primary care, Hospital care (IP, OP)	Claims	10.7 M	2016 to 2025	Objectives 1–3	Internal
Hospitals or Hospital Data Registry							
Finland	FinOMOP-HUS	Primary care, Hospital care (IP, OP)	EHRs	2.92 M	1990 to 2025	Objectives 2–4	N/A
Spain	HARMONY-AML	Hospital care (IP, OP)	Disease registry based on EHR	7.9 K	1996 to 2022	Objectives 2–3	N/A
Spain	H12O	Hospital care (IP, OP)	EHRs	1.87 M	1991 to 2024	Objectives 2–4	N/A
Primary care							
Spain	SIDIAP	Primary care	EHRs	8.89 M	2006 to 2024	Objectives 1–4	Internal

CRN=Cancer Registry Norway, DK-DHR=Danish Health Data Registries, EHR=Electronic health record, FinOMOP-HUS=Hospital District of Helsinki and Uusimaa, FinOMOP-THL=Population Registers at the Finnish Institute for Health and Welfare, H12O=Hospital Universitario 12 de Octubre, HARMONY-AML=HARMONY - Acute Myeloid Leukaemia, InGef RDB=InGef Research

Database, NAJS=Croatian National Public Health Information System, NCR=Netherlands Cancer Registry, SIDIAP=The Information System for Research in Primary Care, N/A=Not applicable, IP=inpatient, OP=outpatient.

¹ Includes all individuals recorded in the data source

² Calendar period with a consistent number of over 1000 individuals who contribute at least one day in a specific month in this observation period (HARMONY-AML – any number of individuals)

³ External denominators were obtained from publicly available demographic data from Eurostat (https://ec.europa.eu/eurostat/databrowser/product/page/DEMO_PJAN)

Data sources selection

These data sources met the criteria for data quality, completeness, timeliness, and representativeness for a population-level descriptive epidemiological study and a patient-level characterisation study and covered different regions of Europe ([Annexes I and II](#)).

In terms of relevance, the selection of data sources was based on the availability of data for the selected outcomes (AML diagnosis and death from any cause) and the drugs of interest (specific AML medicines) required to perform the described analyses. Additionally, the data sources were selected for their ability to support timely Institutional Review Board (IRB) approvals, thereby ensuring alignment with the timelines established by stakeholders for this study.

Data source justification and key characteristics are presented in [Annexes I and II](#).

8.5. Study period

The study period was from 01/01/2015 to 31/12/2024 (estimation of incidence), or the latest available date in each data source (characterisation).

8.6. Variables

8.6.1. Exposure

Not applicable. This was a non-interventional descriptive study with no pre-specified exposure of interest.

8.6.2. Outcome

Objective 1

The outcome for this objective was the first diagnosis of AML.

Objective 4:

The outcome for this objective was death from any cause (overall survival).

The concept sets used for identifying outcomes are described in [Annex IV](#). These sets were created during the study execution following the DARWIN EU® phenotyping standard processes, which involved the review of code lists by clinical experts and the review of phenotypes after their execution in the participating data sources.

8.6.3. Covariates, including confounders, effect modifiers, intercurrent events, and other variables

For all objectives, sex, age, and calendar year were used to stratify results:

- Sex: female, male
- The following age groups were defined: 0 to 17; ≥18; 18 to 39; 40 to 59; 60 to 79; 80 and over.
- Calendar year: between 2015 and 2024
- All results were stratified by the type of AML (acute promyelocytic leukaemia (APL) and non-APL).

Objective 2: patient-level characterisation:

The following variables were described.

- Sex: female, male – at time of diagnosis
- Age groups: 0 to 17; ≥18; 18 to 39; 40 to 59; 60 to 79; 80 and over – at time of diagnosis
- World Health Organisation/Eastern Cooperative Oncology Group (WHO/ECOG) performance status
- History of myelodysplastic syndrome (MDS) [-inf,-90]

- History of any cancer [-inf,-1], excluding non-melanoma skin cancer

Large-scale characterisation

This objective included large-scale characterisation that identified all conditions, medications, measurements, and procedures within specific time windows, before and after diagnosis, to test the quality of the AML definitions using different nomenclatures and classification concepts in Real-World Data (phenotype), and data in general, and to inform more complex future studies.

In large-scale characterisation the presence of following characteristics was assessed (type of AML, genetic mutations: FLT3, IDH1, IDH2, NPM1, CEBPA, DDX41, TP53.)

In addition, measurements of blast in bone marrow, blood, and blood counts (e.g. leucocyte count, neutrophil count, platelet count), blast cell count, and lactate dehydrogenase blood measurement were assessed within sequential 30-day periods after the AML diagnosis date until two years (730 days) (see [Annex IV](#)).

Objective 3: Treatment Patterns

For this objective, medications used in AML therapy, including allogeneic and autologous stem cell transplant, were assessed in the period up to two years (730 days) after initial AML diagnosis and reported as cumulative frequency. The aforementioned treatments were combined into treatment combinations and compared to the ones in the current guidelines ([Table 2](#)).[3]

Treatment sequences and combinations were created and described using the *TreatmentPatterns* package and the following ingredients were assessed: arsenic trioxide, azacitidine, cyclophosphamide (not part of AML treatment, but may be used as part of transplantation), cytarabine, daunorubicin, doxorubicin, enasidenib, gemtuzumab ozogamicin, gilteritinib, glasdegib, idarubicin, ivosidenib, midostaurin, mitoxantrone, olutasidenib, pemigatinib, quizartinib, thioguanine, treosulfan, tretinoin, venetoclax (See [Annex II](#)). Supportive care treatments, such as blood product transfusions, antimicrobials, and corticosteroids, were not included. Investigational agents were also not included.

Table 2. Treatment regimens in AML, which were considered when assessing results of treatment patterns.

Type of the therapy	Therapies
Induction Therapy	"7+3" regimen: 7 days of cytarabine + 3 days of an anthracycline High-dose cytarabine (HiDAC) +/- targeted therapy FLT3 inhibitors (midostaurin) IDH1/2 inhibitors (ivosidenib/enasidenib) Gemtuzumab ozogamicin (anti-CD33 antibody-drug conjugate) Mitoxantrone
Consolidation Therapy	High-dose cytarabine (HiDAC) Allogeneic stem cell transplant Midostaurin Azacitidine
Other	FLAG-IDA CLAG-M Gilteritinib Ivosidenib Enasidenib Venetoclax

AML=Acute Myeloid Leukaemia

Treatment codes were refined during the study implementation following the DARWIN EU® phenotyping standard processes, which involved the review of code lists by clinical experts and the review of phenotypes after their execution in the participating data sources.[12]

8.7. Study size

No sample size has been calculated, as this was a descriptive disease epidemiology study which did not test a specific hypothesis. In addition, we used already collected available data to perform large-scale characterisation, described treatment patterns, and estimated incidence and survival. Thus, the sample size was driven by the availability of data for patients with AML. Based on a preliminary feasibility assessment conducted before the study execution, the number of person counts for AML, without accounting for the study period in the data sources included in this study, ranged from 3,600 in NAJS to 16,400 in NCR.

8.8. Data transformation

Analyses were conducted separately for each data source. Before study initiation, test runs of the analyses were performed on a subset of the data sources and quality control checks were performed. Once all the tests passed (see [Annex III](#)), the final study codes package was released in the version-controlled Study Repository for execution against all the participating data sources.

The data partners locally executed the analytics against the OMOP CDM in R Studio and reviewed and approved the aggregated results (by default).

The study results of all data sources were checked, after which they were made available to the team, and the dissemination phase started. All results were locked and timestamped for reproducibility and transparency.

8.9. Statistical methods

8.9.1. Main summary measures

In objective one, we report incidence rates per 100,000 with 95% confidence intervals (CIs); absolute counts and percentages are reported in characterisation study; survival probabilities (%) and restricted mean survival time (RMST) at 1, 3, and 5 years after AML diagnosis, and median survival is reported with 95% CIs.

8.9.2. Main statistical methods

Objective 1: Incidence of AML

Calculation of population-based incidence rates is part of the DARWIN EU® pipeline for population-level descriptive epidemiology. This analysis was performed using the *IncidencePrevalence* R package and provided incidence rates per 100,000 person-years. Annual incidence rates of AML were calculated as the number of newly diagnosed AML cases per 100,000 person-years of the population at risk during each calendar year. Individuals enter the denominator population at the start of each calendar year within the study period, or once the eligibility criteria were fulfilled. Those study participants who enter the denominator population contributed time-at-risk up until their first diagnosis during each calendar year. If they did not have the condition of interest, they contributed time-at-risk until the end of the calendar year. Time-at-risk of individuals who died or were diagnosed with AML were censored at the time of death or diagnosis date, respectively. Similarly, the time-at-risk for individuals who were lost to follow-up were censored at the time of loss to follow-up (i.e. the last contact). Individuals with data until the end of the calendar period without a record of the condition were administratively censored at the end of each calendar period. Incidence rates were given together with 95% Poisson confidence intervals.

For data sources that collect population-level data, e.g. cancer and public health registries, to ensure the reliable estimation of incidence rates using an appropriate denominator population, external denominators were used from Eurostat (https://ec.europa.eu/eurostat/databrowser/product/page/DEMO_PJAN) Mid-year population was used to derive annual incidence rates per 100,000, also stratified by age and sex.

Objective 2–4: Patient-level characterisation

Patient-level characterisation was conducted before the diagnosis, at the diagnosis date, and within 730 days after diagnosis, including patient demographics, comorbidities, medications, procedures, and measurements using the *CohortCharacteristics* package.

Large-scale characterisation was used to extract comorbidities, measurements, medications (at ingredient level), and procedures. Relevant conditions, medications, and procedures were reported any time before one day before AML diagnosis, at the date of AML diagnosis, and up to two years or 730 days after AML diagnosis.

Diagnostic measurements (genetic mutations, blast in bone marrow, blood, and blood counts) and procedures were reported within sequential 30-day periods after the AML diagnosis date until two years (730 days). All characteristics were reported as absolute counts and proportions.

All treatments and treatment combinations of pre-specified medicines of interest starting from the index date after AML diagnosis were described using the *TreatmentPatterns* package (*computePathways* function). All treatment cohorts were extended to a minimum of 30 days prior to analysis. A minimum treatment duration of 14 days was required for a treatment episode to be included. Treatment durations were collapsed when the gap between two treatment episodes was 30 days or less. A combination of treatments was defined as two or more treatments that overlapped by at least 14 days, with a minimum post-combination duration of 21 days. Only the first occurrence of each treatment sequence was retained. Pathways were split at 30 days, and the maximum pathway length was set to 5 treatment lines.

Treatment combinations were also compared to treatment guidelines for AML.[3]

Overall survival probabilities and restricted mean survival time (RMST) at 1, 3, and 5 years after AML diagnosis and median survival were calculated using the Kaplan-Meier method (*CohortSurvival* package) and presented with 95% confidence intervals.

A minimum cell count of 5 was used when reporting results, with any smaller count reported as “<5”.

8.9.3. Missing values

Records with missing values were excluded from the analysis; no imputation was performed.

8.9.4. Sensitivity analysis

Sensitivity analyses were not performed.

9. RESULTS

The full set of results for this study is available through an interactive web application, Shiny app at [EUPAS1000000773](https://eupas1000000773).

9.1. Participants

Overall, 30,542 individuals were included in the study across 10 data sources from 7 countries. The number of individuals with an AML diagnosis included in the study ranged from 631 in FinOMOP-HUS to 7,784 in NCR (**Table 3**).

Table 3. Participant attrition in the study.

	NAJS	DK-DHR	FinOM OP-HUS	FinOM OP-THL	InGef RDB	NCR	CRN	SIDIAP	HARM ONY- AML ¹	H12O
First diagnosis of AML in the period of 01/01/2015 to 31/12/2024;	3,641	4,113	1,003	3,324	5,315	7,819	2,002	1,498	1,607	4,575
Minimum 365 days of available history before diagnosis date (except for children <1 year)	3,401	4,097	934	3,318	4,808	7,788	1,978	1,458	1,607	4,151
No prior AML (final study cohort)	3,099	3,456	631	2,620	4,101	7,784	1,978	1,439	1,607	3,827

AML=Acute myeloid leukaemia, CRN=Cancer Registry Norway, DK-DHR=Danish Data Health Registries, FinOMOP-HUS=Hospital District of Helsinki and Uusimaa, FinOMOP-THL=Population Registers at the Finnish Institute for Health and Welfare, H12O=Hospital Universitario 12 de Octubre, HARMONY-AML=HARMONY - Acute Myeloid Leukaemia, InGef RDB=InGef Research Database, IP=inpatient, NAJS=Croatian National Public Health Information System, NCR=Netherlands Cancer Registry, SIDIAP=The Information System for Research in Primary Care.

¹ Prior observation of the 365 days criteria was not applied in HARMONY-AML

9.2. Descriptive data

Demographic characteristics of study participants at the time of diagnosis are presented in [Table 4](#). The mean age ranged from 55 years in HARMONY-AML to 68 years in DK-DHR. No individuals under 18 years were observed in HARMONY-AML, and the proportion of those under 18 years old in the rest of the data sources ranged from 1.8% in SIDIAP to 4.3% in FinOMOP-HUS.

APL represented a small proportion (2.4–4.9%) of all identified AML patients within the included data sources, except in H12O, where the majority of individuals with AML had APL (84.1%).

Table 4. Participants characteristics at the time of acute myeloid leukaemia diagnosis.

		NAJS	DK-DHR	FinOMOP-HUS	FinOMOP-THL	InGef RDB	NCR	CRN	SIDIAP	HARMONY-AML	H12O
Overall		3,099	3,456	631	2,620	4,101	7,784	1,978	1,439	1,607	3,827
Age	Mean (SD)	66 (17)	68 (17)	63 (19)	67 (18)	66 (17)	65 (18)	66 (20)	67 (18)	55 (13)	57 (17)
	Median (q25; q75)	69 (68; 78)	72 (62; 79)	68 (54; 76)	71 (60; 79)	68 (57; 79)	69 (57; 77)	71 (58; 79)	72 (58; 80)	57 (47; 64)	58 (48; 68)
Age groups (N; %)	<18	73 (2.4%)	72 (2.1%)	27 (4.3%)	60 (2.3%)	97 (2.4%)	218 (2.8%)	73 (3.7%)	26 (1.8%)	N/A	139 (3.6%)
	≥18	3026 (97.6%)	3384 (97.9%)	604 (95.7%)	2560 (97.7%)	4004 (97.6%)	7566 (97.2%)	1905 (96.3%)	1413 (98.2%)	1607 (100%)	3688 (96.4%)
	18–39	185 (6.0%)	185 (5.4%)	51 (8.1%)	164 (6.3%)	238 (5.8%)	496 (6.4%)	139 (7.0%)	90 (6.3%)	226 (14.1%)	408 (10.7%)
	40–59	568 (18.3%)	481 (13.9%)	136 (21.6%)	413 (15.8%)	879 (21.4%)	1,496 (19.2%)	330 (16.7%)	272 (18.9%)	728 (45.3%)	1,500 (39.2%)
	60–79	1,664 (53.7%)	1,878 (54.3%)	317 (50.2%)	1,393 (53.2%)	1,965 (47.9%)	4,106 (52.7%)	989 (50.0%)	666 (46.3%)	631 (39.2%)	1,499 (39.2%)
	80 and over	609 (19.7%)	840 (24.3%)	100 (15.8%)	590 (22.5%)	922 (22.5%)	1,468 (18.9%)	447 (22.6%)	385 (26.8%)	22 (1.4%)	281 (7.3%)
Sex (N; %)	Female	1,397 (45.1%)	1,524 (44.1%)	325 (51.5%)	1,213 (46.3%)	1,711 (41.7%)	3,297 (42.4%)	877 (44.3%)	696 (48.4%)	750 (46.6%)	1,693 (44.2%)
	Male	1,702 (54.9%)	1,932 (55.9%)	306 (48.5%)	1,407 (53.7%)	2,390 (58.3%)	4,487 (57.6%)	1,101 (55.7%)	743 (51.6%)	858 (53.4%)	2,133 (55.7%)
Type of AML (N; %)	APL	118 (3.8%)	125 (3.6%)	29 (4.6%)	86 (3.3%)	122 (3.0%)	382 (4.9%)	87 (4.4%)	34 (2.4%)	0 (0)	3,220 (84.1%)

AML=Acute Myeloid Leukaemia, APL=Acute Promyelocytic Leukaemia, CRN=Cancer Registry Norway, DK-DHR=Danish Data Health Registries, FinOMOP-HUS=Hospital District of Helsinki and Uusimaa, FinOMOP-THL=Population Registers at the Finnish Institute for Health and Welfare, H12O=Hospital Universitario 12 de Octubre, HARMONY-AML=HARMONY - Acute Myeloid Leukaemia, InGef RDB=InGef Research Database, N=absolute number, N/A=Not available, NAJS=Croatian National Public Health Information System, NCR=Netherlands Cancer Registry, q25=25th percentile, q75=75th percentile, SD=Standard Deviation, SIDIAP=The Information System for Research in Primary Care.

9.3. Main results

9.3.1. Incidence of AML

Table 5 presents the estimated incidence of acute myeloid leukaemia (AML) per 100,000 person-years, overall and by age group, sex, and calendar year. Data source population denominators were used to estimate annual and overall incidence in FinOMOP-HUS, InGef RDB, and SIDIAP, and population data from Eurostat were used to estimate annual incidence in NAJS, DK-DHR, FinOMOP-THL, CRN, and NCR (**Figure 3**).

Overall AML incidence estimates ranged from 2.9 per 100,000 (95% CI [2.8; 3.1]) in SIDIAP to 7.1 per 100,000 (95% CI [6.9; 7.3]) in InGef RDB. The most recent annual AML incidence (2024) recorded in five of the 8 included data sources ranged from 3.5 per 100,000 (95% CI [3.1; 4.1]) in CRN to 8.6 per 100,000 (95% CI [7.7; 9.6]) in NAJS (**Table 5**).

Incidence increased markedly with age across all data sources (**Table 5**).

When available, estimates for those under 18 years were low, ranging from 0.3 [0.2; 0.5] in SIDIAP to 1.8 [0.9; 3.2] in NAJS. In the 18–39 age group, estimates remained low across data sources, from 0.7 [0.6; 0.9] in SIDIAP to 1.8 [1.1; 2.9] in NAJS. Among those aged 40–59, incidence estimates increased, ranging from 1.8 [1.6; 2.0] in SIDIAP to 4.9 [3.7; 6.4] in NAJS. A notable increase in incidence was observed in the 60–79 age group, with estimates ranging from 7.2 [6.7; 7.8] in SIDIAP, 19 [16.3; 21.9] in NAJS, 17.5 [16.8; 18.3] in InGef RDB, and 14.8 [12.8; 17.1] in DK-DHR. For those aged 80 and above, incidence was highest across all data sources, ranging from 12.7 [11.4; 14.0] in SIDIAP to 35.0 [32.8; 37.3] in InGef RDB, with high estimates in NAJS at 33.1 [25.8; 41.9] and FinOMOP-HUS at 13.1 [10.6; 15.9] as well.

Incidence estimates by sex showed consistently higher incidence in males than in females across all data sources (**Table 5**). Estimates in females ranged from 2.8 [2.6; 3.0] in SIDIAP to 7.2 [6; 8.4] in NAJS, while in males, estimates ranged from 3.1 [2.9; 3.3] in SIDIAP to 10.2 [8.8; 11.7] in NAJS.

Calendar-year trends from 2015 to 2024 generally showed stable incidence rates across most data sources (**Figure 3**), with no obvious time trends. In NAJS, annual estimates were relatively low in 2015 (3.3 per 100,000), but from 2016 onwards, estimates were consistently high, typically ranging from 7.6 [6.8; 8.5] to 8.9 [8; 9.9]. In DK-DHR, estimates ranged from 5.6 [5; 6.3] in 2016 to 6.7 [6; 7.4] in 2022. In FinOMOP-HUS, incidence rates showed a moderate upward trend throughout the study period, from 3.0 [2.2; 4.1] in 2016 to 6.4 [4.9; 8.1] in 2024; however, confidence intervals largely overlapped for the annual incidence estimates. FinOMOP-THL incidence estimates remained between 4.5 and 5.5 per 100,000 during the study period. In InGef RDB, incidence estimates varied from 6.7 to 7.6 between 2015 and 2024. Incidence estimates in NCR were also stable, ranging from 4.0 [3.6; 4.5] in 2016 to 4.7 [4.4; 5.1] in 2021, with no clear trend. CRN incidence estimates ranged from 3.2 [2.7; 3.7] to 4.2 [3.6; 4.8]. Finally, in SIDIAP, the incidence of AML was lowest among all data sources, ranging from 2.4 [2.0; 2.8] in 2016 to 3.3 [2.9; 3.8] in 2022. Overall, no data source showed a clear trend in AML incidence rates over the study period. When stratified by age group, AML incidence rates were also stable over time across all data sources (**Figures 4**).

Table 5. Incidence per 100,000 of acute myeloid leukaemia, overall and in subgroups by age and calendar years, with 95% confidence interval.

		NAJS ¹	DK-DHR ¹	FinOMOP-THL ¹	InGef RDB ²	NCR ¹	CRN ¹	SIDIAP ²
Overall		8.6 (7.7; 9.6)	5.9 (5.3; 6.6)	5.0 (4.4; 5.6)	7.1 (6.9; 7.3)	4.6 (4.2; 4.9)	3.5 (3.1; 4.1)	2.9 (2.8; 3.1)
Age groups	<18	1.8 (0.9; 3.2)	0.6 (0.2; 1.3)	N/A	1.0 (0.8; 1.3)	0.5 (0.3; 0.8)	N/A	0.3 (0.2; 0.5)
	18–39	1.8 (1.1; 2.9)	1.6 (1; 2.3)	0.8 (0.4; 1.4)	1.5 (1.3; 1.7)	1.2 (0.9; 1.5)	1 (0.6; 1.6)	0.7 (0.6; 0.9)
	40–59	4.9 (3.7; 6.4)	3.0 (2.2; 4)	3.2 (2.3; 4.3)	4.6 (4.3; 4.9)	3 (2.5; 3.6)	1.9 (1.3; 2.8)	1.8 (1.6; 2.0)
	60–79	19.0 (16.3; 21.9)	14.8 (12.8; 17.1)	12.0 (10.2; 14.1)	17.5 (16.8; 18.3)	11.1 (10.1; 12.1)	9.3 (7.6; 11.3)	7.2 (6.7; 7.8)
	80 and over	33.1 (25.8; 41.9)	27.2 (21.8; 33.7)	17.3 (13.1; 22.4)	35.0 (32.8; 37.3)	17 (14.5; 19.9)	17.5 (12.8; 23.3)	12.7 (11.4; 14.0)
Sex	Female	7.2 (6; 8.4)	5.2 (4.5; 6.1)	4.9 (4.1; 5.7)	5.9 (5.6; 6.2)	4 (3.6; 4.4)	3.1 (2.5; 3.8)	2.8 (2.6; 3.0)
	Male	10.2 (8.8; 11.7)	6.6 (5.7; 7.6)	5.1 (4.3; 6)	8.3 (7.9; 8.6)	5.1 (4.7; 5.6)	4 (3.3; 4.8)	3.1 (2.9; 3.3)
Years	2015	3.3 (2.8; 3.9)	6.1 (5.4; 6.7)	5.1 (4.5; 5.8)	6.8 (6.2; 7.5)	4 (3.7; 4.3)	3.5 (3; 4.1)	3.0 (2.5; 3.4)
	2016	8.3 (7.5; 9.3)	5.6 (5; 6.3)	4.6 (4; 5.2)	7.3 (6.7; 7.9)	4.2 (3.9; 4.5)	4.1 (3.6; 4.7)	2.4 (2.0; 2.8)
	2017	8.7 (7.9; 9.7)	6.3 (5.7; 7)	4.5 (4; 5.1)	7.2 (6.6; 7.9)	4.7 (4.4; 5)	4.2 (3.6; 4.8)	2.9 (2.5; 3.4)
	2018	8.5 (7.6; 9.5)	5.8 (5.2; 6.4)	4.8 (4.2; 5.4)	6.7 (6.2; 7.4)	4.3 (4; 4.6)	3.5 (3; 4.1)	3.2 (2.8; 3.7)
	2019	7.7 (6.9; 8.6)	6.3 (5.6; 6.9)	4.8 (4.2; 5.4)	6.9 (6.3; 7.5)	4.6 (4.3; 5)	3.7 (3.2; 4.3)	3.0 (2.5; 3.4)
	2020	8.9 (8; 9.9)	6.1 (5.5; 6.8)	4.6 (4.1; 5.2)	6.7 (6.1; 7.3)	4.4 (4.1; 4.7)	3.9 (3.4; 4.5)	2.7 (2.3; 3.1)
	2021	8.3 (7.4; 9.2)	5.4 (4.8; 6)	5.5 (4.9; 6.2)	7.4 (6.7; 8.0)	4.7 (4.4; 5.1)	3.6 (3.1; 4.1)	3.1 (2.7; 3.6)
	2022	7.6 (6.8; 8.5)	6.7 (6; 7.4)	5 (4.4; 5.6)	7.6 (7.0; 8.3)	4.4 (4.1; 4.7)	3.6 (3.2; 4.2)	3.3 (2.9; 3.8)
	2023	8.4 (7.5; 9.3)	5.9 (5.3; 6.6)	5 (4.4; 5.6)	6.8 (6.2; 7.5)	4.7 (4.4; 5.1)	3.2 (2.7; 3.7)	N/A
	2024	8.6 (7.7; 9.6)	N/A	N/A	7.3 (6.7; 7.9)	4.6 (4.2; 4.9)	3.5 (3.1; 4.1)	N/A

CRN=Cancer Registry Norway, DK-DHR=Danish Data Health Registries, FinOMOP-THL=Population Registers at the Finnish Institute for Health and Welfare, H12O=Hospital Universitario 12 de Octubre, HARMONY-AML=HARMONY - Acute Myeloid Leukaemia, InGef RDB=InGef Research Database, N/A=Not available, NAJS=Croatian National Public Health Information System, NCR=Netherlands Cancer Registry, SIDIAP=The Information System for Research in Primary Care.

¹ The population from Eurostat was used to estimate incidence in NAJS, DK-DHR, FinOMOP-THL, CRN, and NCR. Incidence rates per 100,000 by age group and sex for data sources where estimates were based on the Eurostat population taken from the most recent available year: 2024 for NAJS, NCR, and CRN; 2023 for FinOMOP-THL and DK-DHR.

² Data source denominator was used to estimate incidence in InGef RDB and SIDIAP.

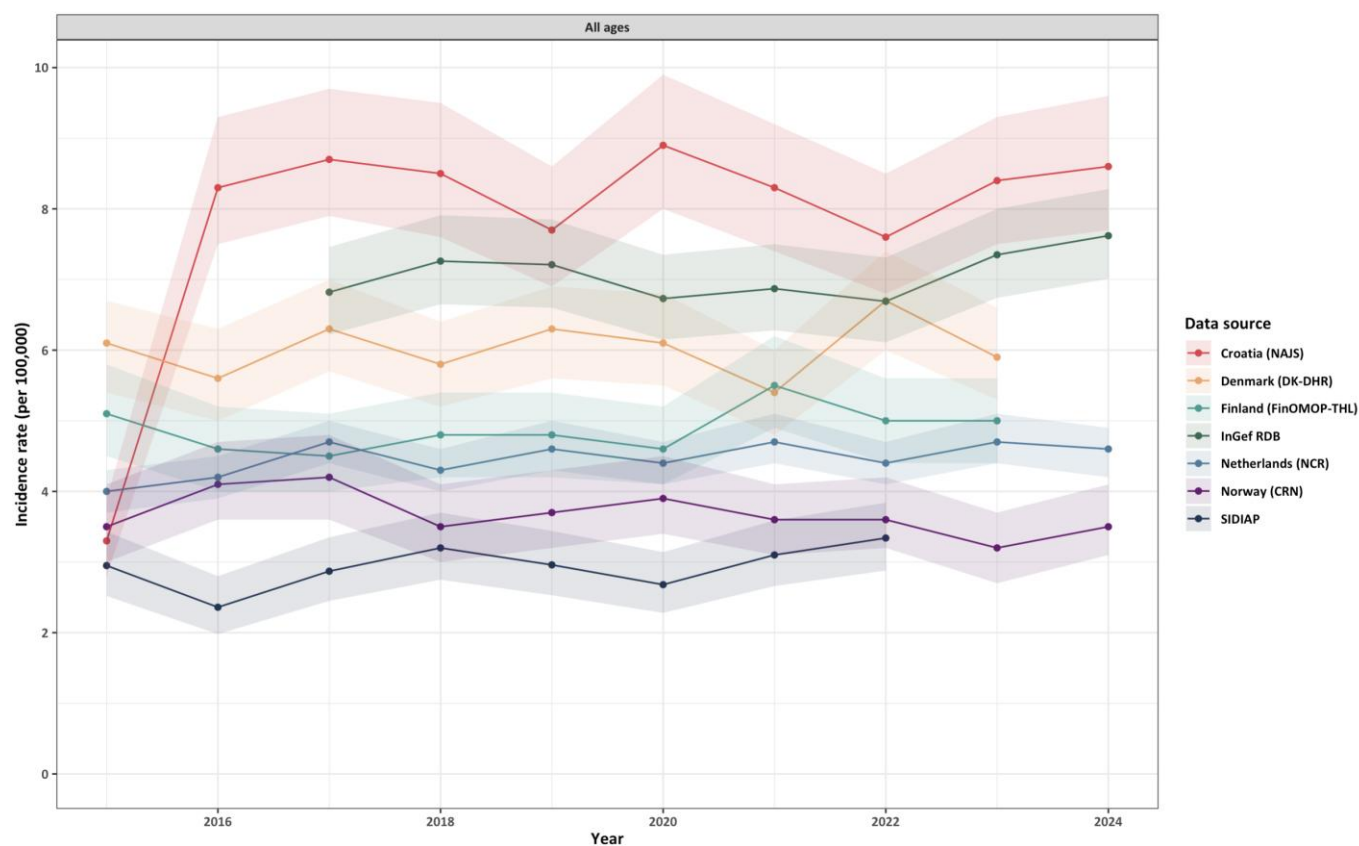


Figure 3. Incidence per 100,000 of acute myeloid leukaemia by calendar year in Croatia (NAJS), Denmark (DK-DHR), Finland (FinOMOP-THL), Norway (CRN), and the Netherlands (NCR) (applying the Eurostat national denominator population) and in InGef RDB and SIDIAP (applying data source denominator population).

NAJS=Croatian National Public Health Information System, DK-DHR=Danish Data Health Registries, FinOMOP-THL=Population Registers at the Finnish Institute for Health and Welfare, InGef RDB=InGef Research Database, NCR=Netherlands Cancer Registry, CRN=Cancer Registry Norway, SIDIAP=The Information System for Research in Primary Care

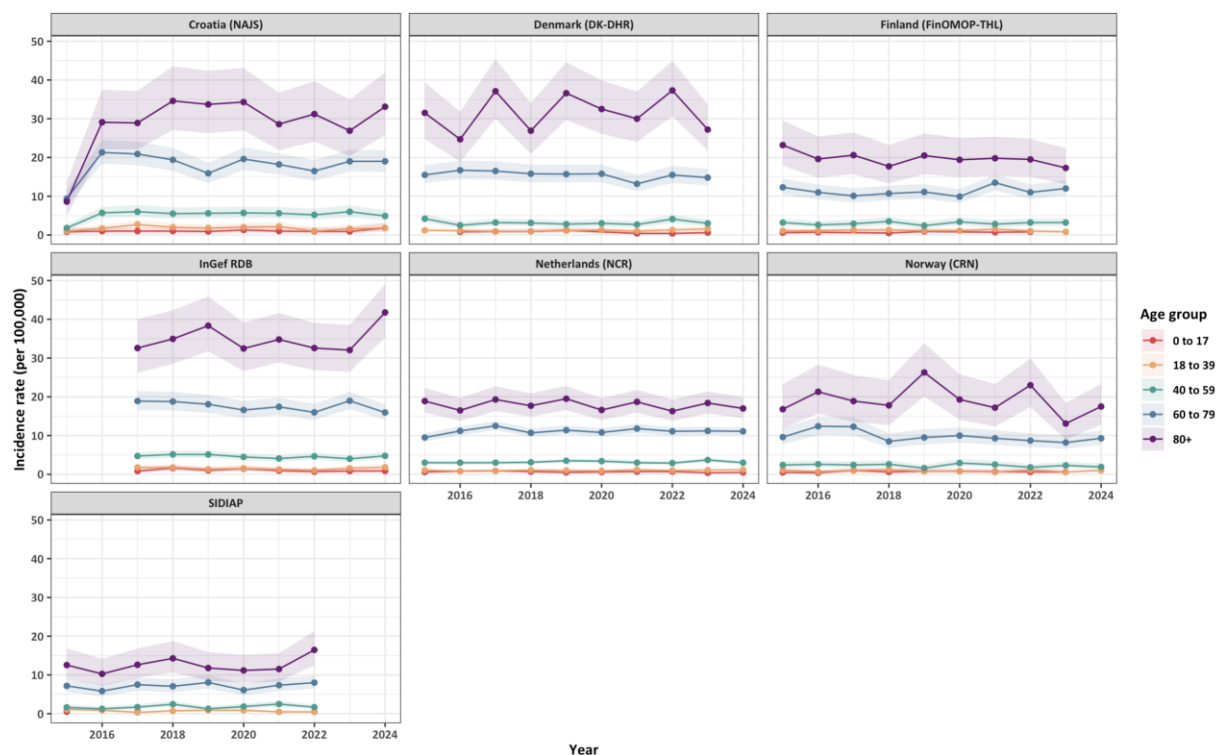


Figure 4. Incidence per 100,000 of acute myeloid leukaemia by calendar year and age group in Croatia (NAJS), Denmark (DK-DHR), Finland (FinOMOP-THL), Norway (CRN), and the Netherlands (NCR) (applying the Eurostat national denominator population) and in InGef RDB and SIDIAP (applying data source denominator population).

NAJS=Croatian National Public Health Information System, DK-DHR=Danish Data Health Registries, FinOMOP-THL=Population Registers at the Finnish Institute for Health and Welfare, InGef RDB=InGef Research Database, NCR=Netherlands Cancer Registry, CRN=Cancer Registry Norway, SIDIAP=The Information System for Research in Primary Care

9.3.2. Characterisation

The proportion of individuals with a recorded history of MDS in the 90 days prior to AML diagnosis ranged from 2.4% (n=90) in H12O to 16.1% (n=421) in FinOMOP-THL (**Table 6**). The proportion of individuals with any general concepts describing cancer prior to specific AML diagnosis concepts was relatively high across all data sources and showed greater variation, ranging from 17.1% (n=1,330) in NCR to 55.0% (n=1,705) in NAJS.

Table 6. History of myelodysplastic syndrome and malignancies prior to diagnosis of acute myeloid leukaemia.

	NAJS	DK-DHR	FinOMOP-HUS	FinOMOP-THL	InGef RDB	NCR	CRN	SIDIAP	H12O
AML patients	3,099	3,456	631	2,620	4,101	7,784	1,978	1,439	3,827
History of Myelodysplastic syndrome anytime up to 90 days before AML diagnosis	364 (11.8%)	463 (13.4%)	64 (10.1%)	421 (16.1%)	396 (9.7%)	N/A ¹	N/A ¹	112 (7.8%)	90 (2.4%)
History of cancer anytime up to one day before AML diagnosis	1,705 (55.0%)	1,528 (44.2%)	242 (38.4%)	1,410 (53.8%)	1,666 (40.6%)	1,330 (17.1%)	445 (22.5%)	443 (30.8%)	754 (19.7%)

AML=Acute Myeloid Leukaemia, CRN=Cancer Registry Norway, DK-DHR=Danish Data Health Registries, FinOMOP-HUS=Hospital District of Helsinki and Uusimaa, FinOMOP-THL=Population Registers at the Finnish Institute for Health and Welfare, H12O=Hospital Universitario 12 de Octubre, InGef RDB=InGef Research Database, N/A=Not available, NAJS=Croatian National Public Health Information System, NCR=Netherlands Cancer Registry, SIDIAP=The Information System for Research in Primary Care.

1- excluded from this analysis because they do not contain prior data history

Recorded genetic mutation measurements within the first 30 days after diagnosis had a high degree of missingness and were observed in individuals with AML across three data sources: FinOMOP-HUS, HARMONY-AML, and H12O (**Table 7**). Across all three genetic markers available in records of individuals with AML, HARMONY-AML consistently showed the highest proportions of recorded measurements, with up to 90% of patients having a record of NPM1 mutation.

Table 7. Presence of measurement of genetic mutations in individuals with acute myeloid leukaemia in the first 30 days after diagnosis.

	FinOMOP-HUS	HARMONY-AML	H12O
AML patients	631	1,608	3,827
CEBPA	15 (2.4%)	1388 (86.3%)	N/A
DDX41	N/A	254 (15.8%)	N/A
NPM1	69 (10.9%)	1448 (90.1%)	25 (0.7%)

FinOMOP-HUS=Hospital District of Helsinki and Uusimaa, H12O=Hospital Universitario 12 de Octubre, HARMONY-AML=HARMONY - Acute Myeloid Leukaemia; N/A=Not available

Table 8 shows the availability of recorded measurements of blasts in blood, blasts in bone marrow, and lactate dehydrogenase (LDH) levels in blood within the first 30 days after initial AML diagnosis across six data sources: NAJS, DK-DHR, FinOMOP-HUS, InGef RDB, HARMONY-AML, and H12O.

For blasts in bone marrow measurements, recorded proportions ranged from 1.5% (n=56) in H12O to 69.6% (n=439) in FinOMOP-HUS. In all data sources, except InGef RDB, the highest proportions were

observed in the first 30-day period after diagnosis (**Figure 5**). For blasts in blood in the first 30-day period after diagnosis, 71.0% (n=2,455) of individuals with AML had this measurement in DK-DHR, 74.3% (n=469) in FinOMOP-HUS, and 6.0% (n=228) in H12O. In the subsequent 30-day periods, this proportion decreased across all three data sources (**Figure 6**). Finally, LDH measurements in blood during the same time period were available in NAJS, DK-DHR, FinOMOP-HUS, InGef RDB, HARMONY-AML, and H12O (**Figure 7**).

Table 8. Presence of concepts for measurement of blasts in bone marrow, blood, and lactate dehydrogenase in blood in the first 30 days after initial diagnosis.

	NAJS	DK-DHR	FinOMOP-HUS	InGef RDB	HARMONY-AML	H12O
Blasts in bone marrow	275 (8.9%)	NA	439 (69.6%)	92 (2.2%)	1,030 (64.1%)	56 (1.5%)
Blasts in blood	N/A	2,455 (71%)	469 (74.3%)	N/A	N/A	228 (6%)
Lactate dehydrogenase blood test	1,899 (61.3%)	3,237 (93.7%)	542 (85.9%)	1,070 (26.1%)	19 (1.2%)	758 (19.8%)

DK-DHR=Danish Data Health Registries, FinOMOP-HUS=Hospital District of Helsinki and Uusimaa, H12O=Hospital Universitario 12 de Octubre, HARMONY-AML=HARMONY - Acute Myeloid Leukaemia, InGef RDB=InGef Research Database, N/A=Not available, NAJS=Croatian National Public Health Information System.

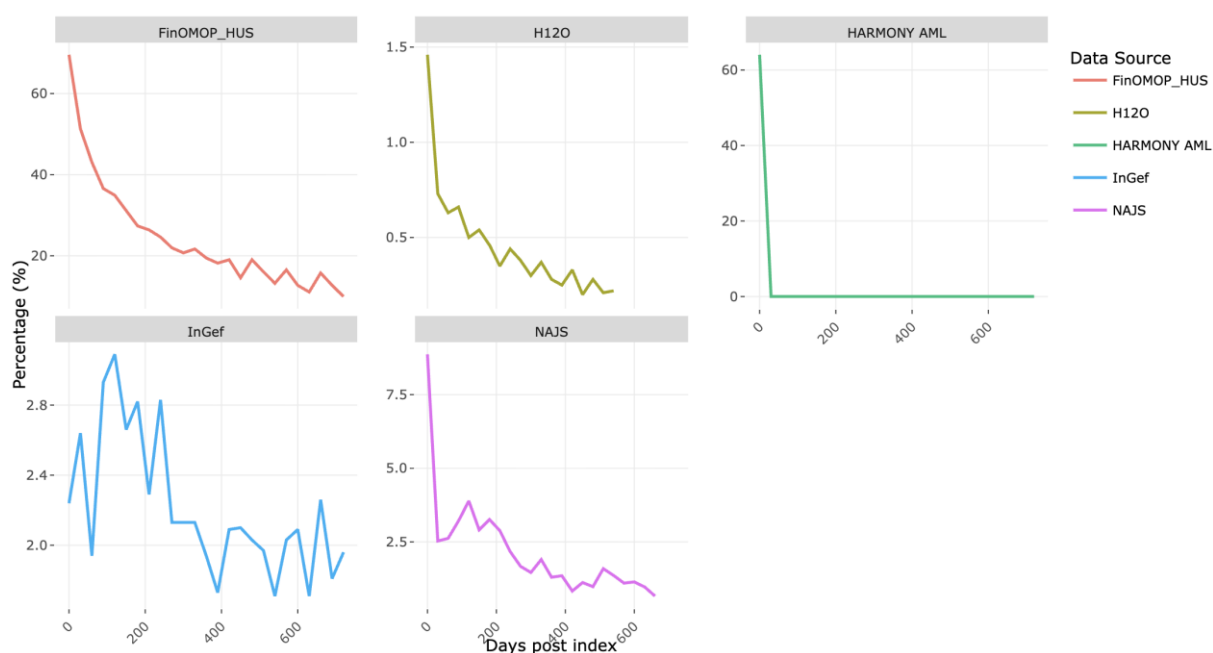


Figure 5. Timing of measurement of blasts in bone marrow in individuals with acute myeloid leukaemia.

FinOMOP-HUS=Hospital District of Helsinki and Uusimaa, H12O=Hospital Universitario 12 de Octubre, HARMONY-AML=HARMONY - Acute Myeloid Leukaemia, InGef RDB=InGef Research Database, NAJS=Croatian National Public Health Information System.

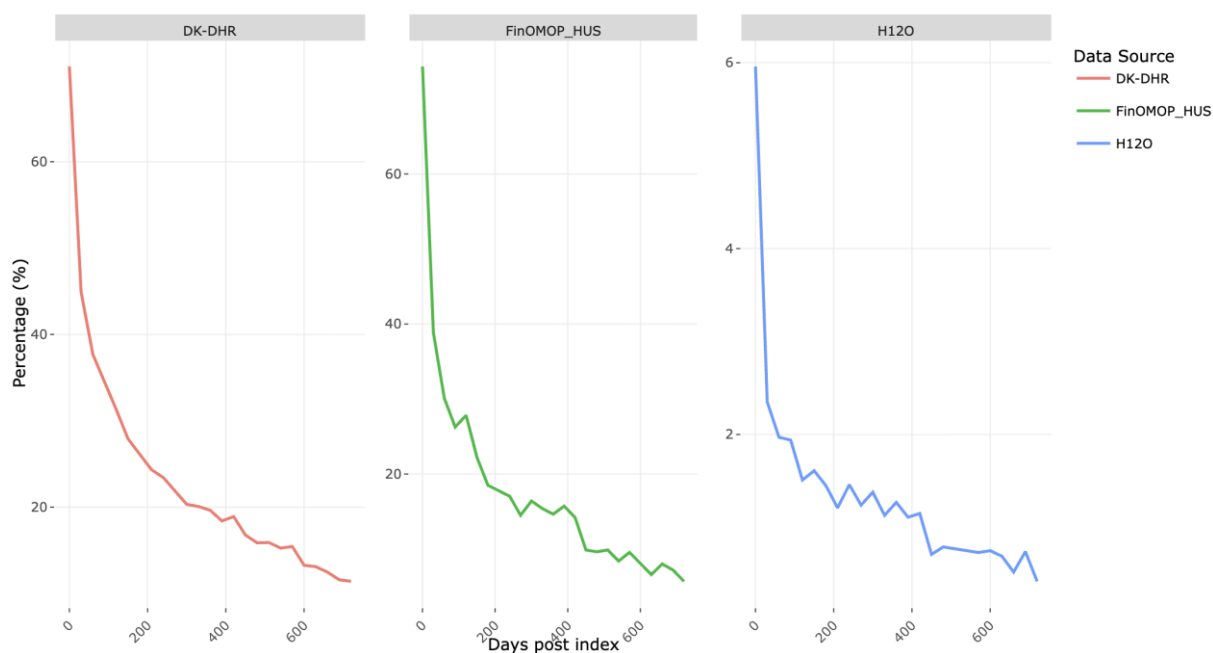


Figure 6. Timing of measurement of blasts in blood in individuals with acute myeloid leukaemia.

DK-DHR=Danish Data Health Registries, FinOMOP-HUS=Hospital District of Helsinki and Uusimaa, H12O=Hospital Universitario 12 de Octubre.

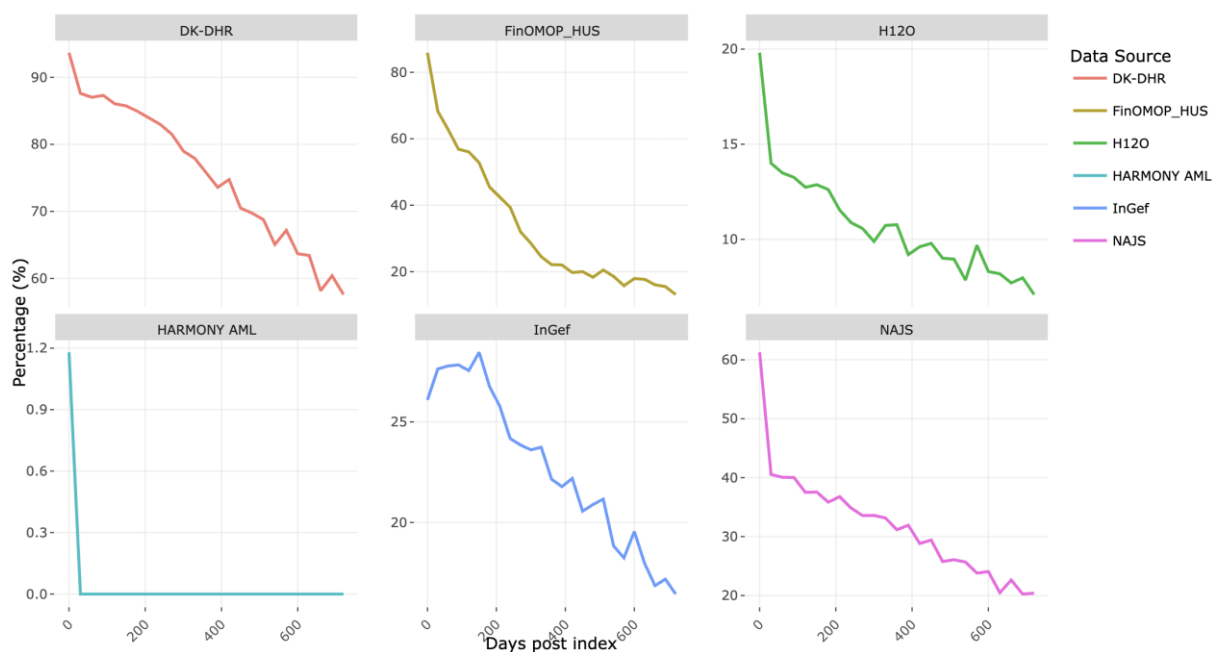


Figure 7. Timing of measurement of lactate dehydrogenase in blood in individuals with acute myeloid leukaemia.

DK-DHR=Danish Data Health Registries, FinOMOP-HUS=Hospital District of Helsinki and Uusimaa, H12O=Hospital Universitario 12 de Octubre, HARMONY-AML=HARMONY - Acute Myeloid Leukaemia, InGef RDB=InGef Research Database, NAJS=Croatian National Public Health Information System.

WHO/ECOG performance status scores for AML patients were available in two data sources: HARMONY-AML and NCR (**Table 9**). In HARMONY-AML, performance status was recorded for 63.3% of individuals. Among them, ECOG 0 was the most common at 36.5% (n=587), followed by ECOG 1 at 23.6% (n=380) and ECOG 2 at 3.1% (n=51); there were no records of individuals with ECOG 3 or 4 in this data source. In NCR, performance status was recorded for 42.1% of individuals. Of those with a recorded ECOG status, ECOG 0 was observed in 17.6% (n=1,370 of individuals with AML), followed by ECOG 1 at 15.3% (n=1,191), ECOG 2 at 5.5% (n=429), ECOG 3 at 2.53% (n=197), and ECOG 4 at 1.12% (n=87).

Table 9. Presence of information on WHO/ECOG performance status in individuals with acute myeloid leukaemia in the first 30 days after diagnosis.

WHO/ECOG performance status	HARMONY-AML	NCR
Number of individuals with AML	1,608	7,784
0	587 (36.5%)	1,370 (17.6%)
1	380 (23.6%)	1,191 (15.3%)
2	51 (3.1%)	429 (5.5%)
3	0	197 (2.5%)
4	0	87 (1.1%)
Unknown	590 (36.7%)	4,510 (57.9%)

AML=Acute myeloid leukaemia, HARMONY-AML=HARMONY - Acute Myeloid Leukaemia, NCR=Netherlands Cancer Registry, WHO/ECOG=World Health Organisation/Eastern Cooperative Oncology Group

9.3.3. Prevalence and timing of treatment

The proportion of individuals with AML who had records of specific treatments (including both specific ingredients and stem cell transplants) within the first 120 days after AML diagnosis is described below.

Treatments were analysed in only 9 of 10 data sources, as SIDIAP was not fit for purpose for this analysis due to its healthcare setting (primary care). Furthermore, some results are limited to the most common treatments, documented in more than 5% of cases; therefore, some data sources (FinOMOP-THL and H12O) are not listed.

Treatments in paediatric patients were very infrequent and only observed in FinOMOP-HUS, please see the Shiny app at [EUPAS1000000773](https://eupas1000000773).

Overall, across data sources with any recorded treatment, we observed a substantial proportion of patients with no records of the pre-specified therapies. In addition, unlike drugs, stem cell transplantation has not been mapped across multiple data sources.

In HARMONY-AML, cytarabine peaked at 64.7% of individuals with AML during days 0–29 and 49.0% during days 30–59, with this proportion decreasing to 11.2% later on. Idarubicin was recorded in 56.4% of individuals during days 0–29, declining to 13.4% during days 30–59. The proportion of individuals with daunorubicin records increased from 12.6% in days 0–29 to 32.6% in days 30–59, with no records thereafter. Etoposide and gemtuzumab ozogamicin were recorded above the 5% threshold only in HARMONY-AML, with 13.5% and 5.4% of individuals in the first 30-day period having records, respectively. Allogeneic stem cell transplantation was only documented within the 90–119-day window in 12.8% of individuals with AML.

In NCR, cytarabine records during the first 120 days were identified in 35.2% of individuals with AML on days 0–29 and 34.0% on days 30–59, then sharply declined to 12.0% on days 60–89. Similarly, daunorubicin was recorded in 21.7% and 24.8% of patients in the first two 30-day periods, decreasing to 6.9% thereafter. Idarubicin was only recorded in days 0–29 in 14.5% of individuals with AML. The proportion of individuals

with records of azacitidine remained stable throughout all periods, ranging from 13.4% to 14.8%. The proportion of individuals with records of venetoclax in the first three periods ranged from 5.1% to 5.7%, while the proportion of individuals with records of midostaurin in the first two periods was 5.6% and 6.0%. Allogeneic stem cell transplantation was recorded during the 90–119-day period in 11.8% of individuals with AML.

In DK-DHR, azacitidine and cytarabine were the only ingredients that exceeded the reporting threshold of at least 5% per reporting stratum in all four of the first 30-day periods. The proportion of individuals with azacitidine records remained stable, ranging from 14.1% to 16.6%, while that of individuals with cytarabine records declined modestly from 26.2% in days 0–29 to 19.5% in days 90–119. The proportions of individuals with daunorubicin records were 6.5% and 6.7% in the first two 30-day time periods. Allogeneic stem cell transplantation was either not recorded or recorded in less than 5% of patients.

In FinOMOP-HUS, cytarabine was recorded in 35.3% of individuals with AML during days 0–29, then this proportion decreased to 17.4% by days 90–119. Idarubicin was recorded in 28.2% of individuals with AML in only the earliest period, dropping to 10.2% between days 30–59. The proportion of individuals with records of azacitidine remained relatively stable across all periods, ranging from 22.9% to 26.1%, while the proportion of individuals with venetoclax records ranged from 10.5% to 12.0%. The proportion of individuals with fludarabine records only exceeded the threshold of 5% during days 30–59 and 90–119, at 5.2% and 9.2%, respectively. Allogeneic stem cell transplantation was either not recorded or recorded in less than 5% of patients.

In CRN, cytarabine was recorded in 20.3% and 18.4% of individuals with AML in the first 30-day periods, then decreased to 6.5%. The proportion of individuals with azacitidine records remained stable at 11.0%–14.2% across all periods. Daunorubicin was recorded in 10.2% and 9.9% of individuals with AML in the initial two windows, while idarubicin was recorded in 7.7% of individuals with AML in days 0–29. The proportion of individuals with records of venetoclax was consistently around 5% threshold across all periods. The proportion of individuals with fludarabine and gemtuzumab ozogamicin records only exceeded the threshold in the 90–119 days after diagnosis, with 6.9% and 5.4% of individuals having records of these drugs, respectively. Allogeneic stem cell transplantation was either not recorded or recorded in less than 5% of patients.

In InGef RDB, azacitidine and venetoclax were the most commonly recorded treatments, with the proportion of individuals with azacitidine ranging from 9.7% to 14.7% and the proportion of records of venetoclax from 9.5% to 13.2% over time. Allogeneic stem cell transplantation was documented from the 30–59-day window onwards in up to 7.5% of individuals with AML.

Proportions of individuals treated with the most common ingredients are shown stratified by younger (40–59 years old, [Figure 8](#)) and older (60–79 years old, [Figure 9](#)) individuals. The presence of records for cytarabine, daunorubicin, and idarubicin, followed by allogeneic stem cell transplantation, was more common across all data sources in the 40–59 age group, whereas azacitidine and venetoclax were more prevalent in individuals with AML aged 60–79, with a much lower proportion of individuals who had allogeneic stem cell transplantation.

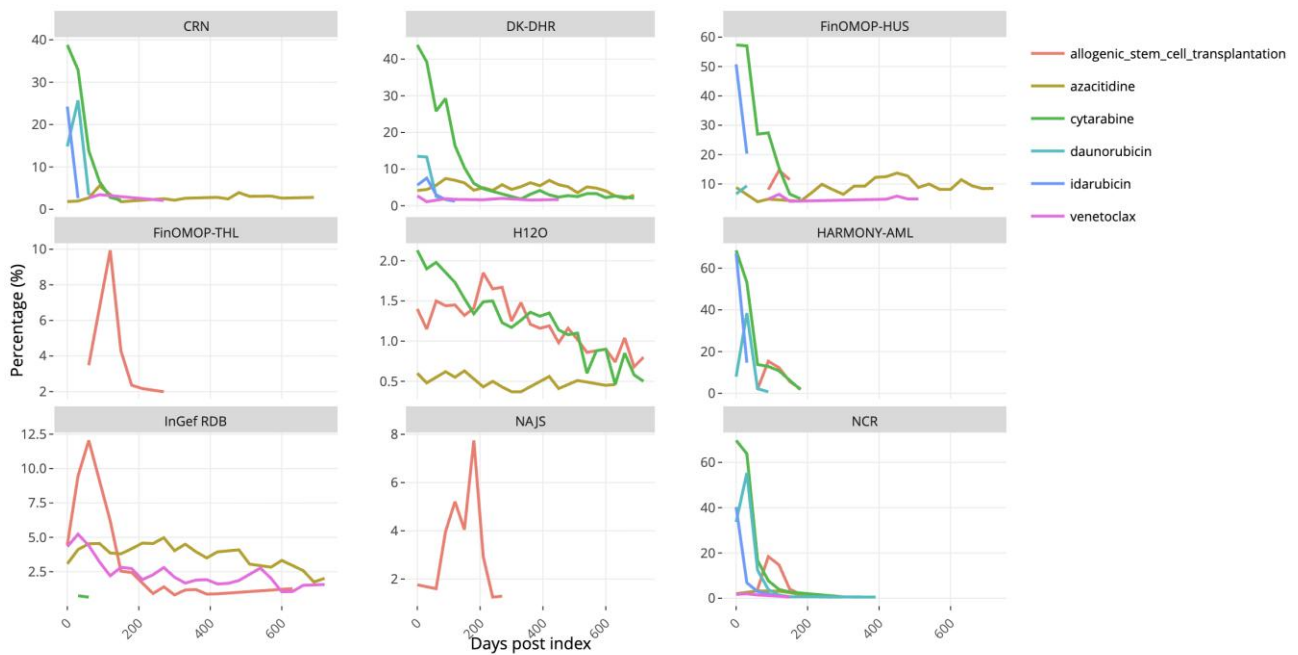


Figure 8. Timing of azacitidine, cytarabine, daunorubicin, idarubicin, and venetoclax use and allogeneic stem cell transplantation in individuals with AML in the age group 40 to 59 years old.

CRN=Cancer Registry Norway, DK-DHR=Danish Data Health Registries, FinOMOP-HUS=Hospital District of Helsinki and Uusimaa, FinOMOP-THL=Population Registers at the Finnish Institute for Health and Welfare, H12O=Hospital Universitario 12 de Octubre, HARMONY-AML=HARMONY - Acute Myeloid Leukaemia, InGef RDB=InGef Research Database, NAJS=Croatian National Public Health Information System, NCR=Netherlands Cancer Registry

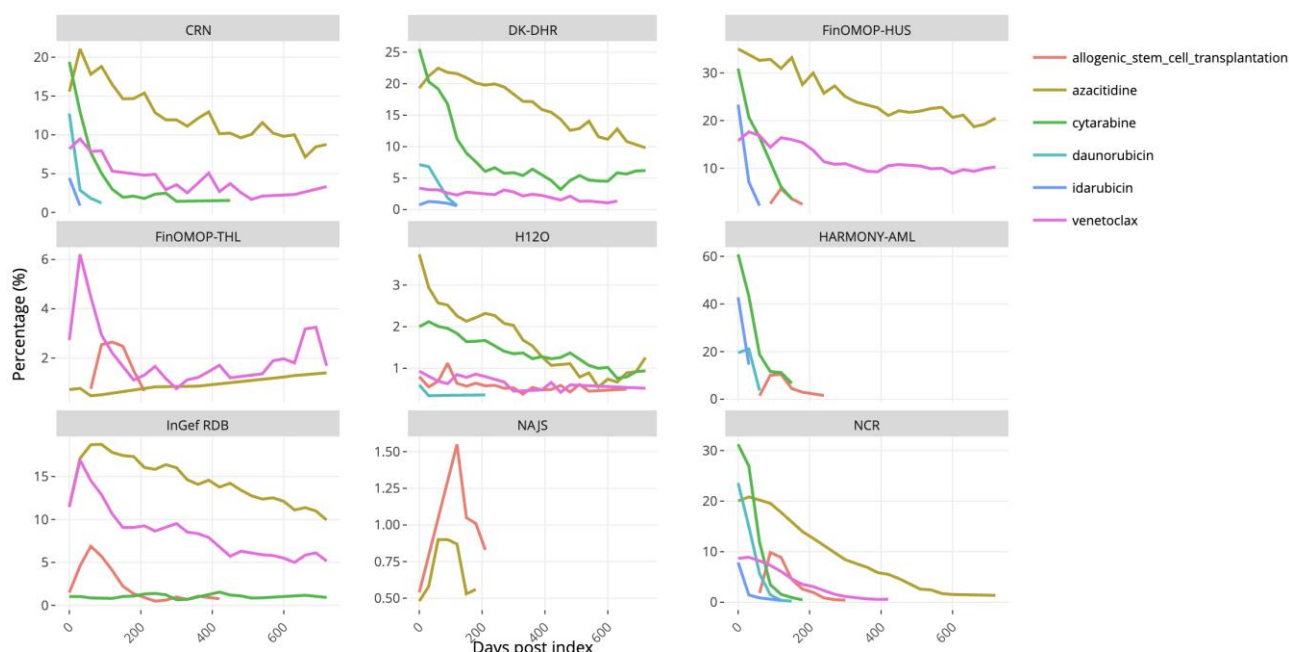


Figure 9. Timing of azacitidine, cytarabine, daunorubicin, idarubicin, and venetoclax use and allogeneic stem cell transplantation in individuals with acute myeloid leukaemia in the age group 60 to 79 years old.

CRN=Cancer Registry Norway, DK-DHR=Danish Data Health Registries, FinOMOP-HUS=Hospital District of Helsinki and Uusimaa, FinOMOP-THL=Population Registers at the Finnish Institute for Health and Welfare, H12O=Hospital Universitario 12 de Octubre, HARMONY-AML=HARMONY - Acute Myeloid Leukaemia, InGef RDB=InGef Research Database, NAJS=Croatian National Public Health Information System, NCR=Netherlands Cancer Registry

9.3.4. Treatment patterns

The recording of AML treatment pathways varied across data sources. Pathways could be broadly grouped into azacitidine and venetoclax, cytarabine, and transplantation-containing regimens, with some overlap observed in sequential combinations (e.g. when one combination is followed by another). The most common pathways (presented in at least 3% of cases) are presented in Table 10 and in Figure 10. Separate results for children with AML were not available due to small sample sizes.

Treatment pathways, at least partially aligned with known recommendations and guidelines, could only be identified in HARMONY-AML, CRN, NCR, and FinOMOP-HUS. For the remaining data sources, the data were neither granular enough nor adequately coded to allow extraction of the pathways.

Azacitidine-containing pathways were common across most data sources.

- In CRN, azacitidine accounted for 44.4% of cases, spanning azacitidine monotherapy (21.9%), azacitidine and venetoclax (7.1%), azacitidine and venetoclax followed by azacitidine (6.6%), and azacitidine, followed by azacitidine and venetoclax (5.3%).
- In FinOMOP-HUS, azacitidine monotherapy was recorded in 38.9% of cases, with combinations with azacitidine adding a further 19.3%.
- In DK-DHR, azacitidine monotherapy alone accounted for 27.9%, with azacitidine then cytarabine adding a further 3.2%.
- In H12O, azacitidine monotherapy was recorded in 31.1% and azacitidine and venetoclax in a further 4.4%.
- In NAJS, azacitidine monotherapy was recorded in 26.1%.

- In NCR, azacitidine monotherapy was recorded in 20.5% and azacitidine and venetoclax in a further 7.9%.
- In InGef RDB, azacitidine monotherapy accounted for 12.2% and azacitidine and venetoclax, then azacitidine, for a further 3.5%.
- In FinOMOP-THL, azacitidine monotherapy was recorded in 8.7% of cases.
- In HARMONY-AML, this pathway was below 3%

In addition, venetoclax-containing pathways were particularly common in the FinOMOP-THL and CRN.

- In FinOMOP-THL, venetoclax monotherapy was the second most common pathway overall, accounting for 27.0% of cases.
- In CRN, venetoclax was observed across three pathways, totalling 19.0%: azacitidine and venetoclax (7.1%), azacitidine and venetoclax then azacitidine (6.6%), and azacitidine then azacitidine and venetoclax (5.3%), all of which also contained azacitidine.
- In NCR, azacitidine and venetoclax accounted for 7.9%.
- In InGef RDB, venetoclax monotherapy was recorded in 9.5%.
- In H12O, the azacitidine and venetoclax combination was recorded in 4.4% of cases.
- In HARMONY-AML, DK-DHR, and FinOMOP-HUS this pathway was below 3%.

Cytarabine-containing pathways were among the most common in DK-DHR, H12O, NCR, HARMONY-AML, and CRN.

- In DK-DHR, cytarabine monotherapy was the second most common pathway at 24.4%, with cytarabine and daunorubicin adding 3.9%, cytarabine and daunorubicin then cytarabine adding 3.3%, and cytarabine followed by azacitidine adding 3.3%.
- In CRN, cytarabine-containing pathways accounted for 27.2%, including cytarabine and daunorubicin (7.9%), cytarabine monotherapy (5.7%), and several sequential cytarabine-based combinations, each accounting for between 3.1% and 3.7% of cases.
- In NAJS, cytarabine monotherapy was recorded in 14.0%, and cytarabine with daunorubicin in 6.2%.
- In H12O, the proportion of individuals receiving cytarabine monotherapy was 14.2%, and overall, the proportion of individuals with cytarabine concomitant pathways was approximately 21.3%.
- In NCR, cytarabine and daunorubicin were observed in 8.1% of cases, with several sequential pathways containing cytarabine contributing additional cases, all below 5% of cases.
- In HARMONY-AML, cytarabine appeared in most recorded pathways, most frequently with idarubicin, followed by combinations of cytarabine and daunorubicin, collectively accounting for approximately 71.4% of cases in that data source.

Transplantation-containing pathways were common in HARMONY-AML, FinOMOP-THL, InGef RDB, H12O, and NCR.

- In HARMONY-AML, allogeneic stem cell transplantation was observed in the three most frequent pathways: cytarabine and idarubicin then cytarabine and daunorubicin then allogeneic stem cell transplantation (19.2%), allogeneic stem cell transplantation alone (18.0%), and cytarabine and idarubicin then cytarabine and daunorubicin then autologous stem cell transplantation (12.2%), collectively accounting for approximately half of all recorded cases.

- In FinOMOP-THL, allogeneic stem cell transplantation alone was the single most frequent pathway at 49.6%.
- In InGef RDB, allogeneic stem cell transplantation alone accounted for 35.6% of cases.
- In H12O, allogeneic stem cell transplantation was recorded in 27.1% of cases, with sequential cytarabine followed by allogeneic stem cell transplantation pathways accounting for an additional 3.6%.
- In NCR, cytarabine and daunorubicin, then allogeneic stem cell transplantation was recorded in 6.6% and cytarabine and idarubicin, then cytarabine and daunorubicin, then allogeneic stem cell transplantation in 5.8%, with allogeneic stem cell transplantation alone in a further 3.1%.
- In NAJS, allogeneic stem cell transplantation alone was recorded in 7.6%.
- No transplantation records were identified in three data sources: FinOMOP-HUS, DK-DHR, and CRN.

Table 10. Most common treatment pathways recorded in different data sources in individuals with acute myeloid leukaemia (all age groups).

Data sources	Most common pathways ¹ (recorded in 3% or more cases)	N (%)
NAJS	<i>azacitidine</i>	186 (26.09)
	<i>thioguanine</i>	104 (14.6)
	cytarabine	100 (14.0)
	allogeneic stem cell transplantation	54 (7.6)
	cytarabine and daunorubicin	44 (6.2)
	other	225 (31.6)
DK-DHR	<i>azacitidine</i>	408 (27.9)
	cytarabine	357 (24.4)
	cytarabine and daunorubicin	57 (3.9)
	cytarabine and daunorubicin followed by cytarabine	48 (3.3)
	cytarabine followed by azacitidine	48 (3.3)
	<i>azacitidine followed by cytarabine</i>	47 (3.2)
	other	496 (34.0)
FinOMOP-HUS	<i>azacitidine</i>	93 (38.9)
	other	37 (15.5)
	<i>azacitidine and venetoclax</i>	27 (11.3)
	<i>azacitidine and venetoclax followed by azacitidine</i>	21 (8.8)
	<i>azacitidine followed by azacitidine and venetoclax</i>	15 (6.3)
	cytarabine and idarubicin	13 (5.4)
	allogeneic stem cell transplantation and fludarabine and treosulfan	12 (5.0)
	cytarabine and idarubicin followed by cytarabine followed by allogeneic stem cell transplantation and fludarabine and treosulfan	11 (4.6)
	<i>azacitidine and venetoclax followed by venetoclax</i>	10 (4.2)
FinOMOP-THL	allogeneic stem cell transplantation	257 (49.6)
	<i>venetoclax</i>	140 (27.0)

Data sources	Most common pathways ¹ (recorded in 3% or more cases)	N (%)
	<i>azacitidine</i>	45 (8.7)
	midostaurin	19 (3.7)
	tretinoin	18 (3.5)
	other	39 (7.5)
InGef RDB	allogeneic stem cell transplantation	595 (35.6)
	azacitidine	204 (12.2)
	venetoclax	159 (9.5)
	azacitidine and venetoclax followed by azacitidine	58 (3.5)
	other	655 (39.2)
CRN	<i>azacitidine</i>	120 (21.9)
	cytarabine and daunorubicin	43 (7.9)
	<i>azacitidine and venetoclax</i>	39 (7.1)
	arsenic trioxide	38 (7.0)
	<i>azacitidine and venetoclax followed by azacitidine</i>	36 (6.6)
	cytarabine	31 (5.7)
	<i>azacitidine followed by azacitidine and venetoclax</i>	29 (5.3)
	cytarabine and daunorubicin followed by cytarabine followed by fludarabine and treosulfan	20 (3.7)
	cytarabine and daunorubicin followed by azacitidine	19 (3.5)
	cytarabine and daunorubicin followed by cytarabine	19 (3.5)
	cytarabine and idarubicin followed by cytarabine and daunorubicin followed by fludarabine and treosulfan	17 (3.1)
	other	136 (24.9)
NCR	<i>azacitidine</i>	913 (20.5)
	cytarabine and daunorubicin	361 (8.1)
	<i>azacitidine and venetoclax</i>	353 (7.9)
	cytarabine and daunorubicin followed by allogeneic stem cell transplantation	294 (6.6)
	cytarabine and idarubicin followed by cytarabine and daunorubicin followed by allogeneic stem cell transplantation	260 (5.8)
	cytarabine and daunorubicin followed by cytarabine	155 (3.5)
	cytarabine and daunorubicin followed by cytarabine followed by allogeneic stem cell transplantation	148 (3.3)
	cytarabine and idarubicin followed by cytarabine and daunorubicin followed by autologous stem cell transplantation	145 (3.3)
	allogeneic stem cell transplantation	138 (3.1)
	other	1,691 (37.9)
HARMONY-AML	cytarabine and idarubicin followed by cytarabine and daunorubicin followed by allogeneic stem cell transplantation	239 (19.2)
	allogeneic stem cell transplantation	224 (18.0)

Data sources	Most common pathways ¹ (recorded in 3% or more cases)	N (%)
	cytarabine and idarubicin followed by cytarabine and daunorubicin followed by autologous stem cell transplantation	152 (12.2)
	cytarabine and idarubicin followed by cytarabine and daunorubicin	138 (11.1)
	cytarabine and etoposide and gemtuzumab ozogamicin and idarubicin and tretinoin	101 (8.1)
	cytarabine and etoposide and idarubicin and tretinoin	98 (7.9)
	cytarabine and daunorubicin	63 (5.1)
	cytarabine and idarubicin	55 (4.4)
	cytarabine and daunorubicin followed by cytarabine	42 (3.4)
	other	132 (10.6)
H12O	<i>azacitidine</i>	<i>70 (31.1)</i>
	allogeneic stem cell transplantation	61 (27.1)
	cytarabine	32 (14.2)
	doxorubicin	19 (8.4)
	tretinoin	11 (4.9)
	<i>azacitidine and venetoclax</i>	<i>10 (4.4)</i>
	cytarabine followed by allogeneic stem cell transplantation	8 (3.6)
	cytarabine followed by azacitidine and cytarabine	8 (3.6)
	other	6 (2.7)

CRN=Cancer Registry Norway, DK-DHR=Danish Data Health Registries, FinOMOP-HUS=Hospital District of Helsinki and Uusimaa, FinOMOP-THL=Population Registers at the Finnish Institute for Health and Welfare, H12O=Hospital Universitario 12 de Octubre, HARMONY-AML=HARMONY - Acute Myeloid Leukaemia, InGef RDB=InGef Research Database, NAJS=Croatian National Public Health Information System, NCR=Netherlands Cancer Registry.

¹ Combination of treatments are described with “and”, and sequences are described with “followed by”.

Bolded are combinations and drugs likely to be used in curative intent, *Italic are combinations and drugs likely to be used in palliative settings.*



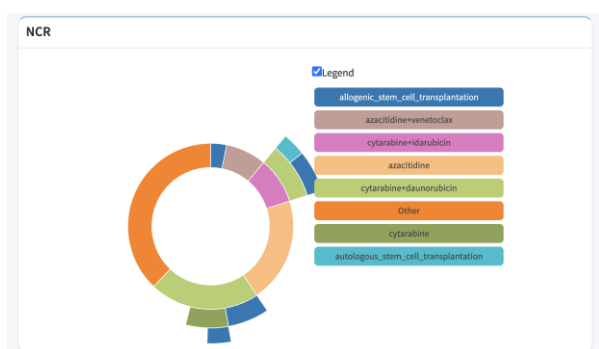


Figure 10. Most common treatment pathways (recorded in different data sources in individuals with acute myeloid leukaemia).

CRN=Cancer Registry Norway, DK-DHR=Danish Data Health Registries, FinOMOP-HUS=Hospital District of Helsinki and Uusimaa, FinOMOP-THL=Population Registers at the Finnish Institute for Health and Welfare, H12O=Hospital Universitario 12 de Octubre, HARMONY-AML=HARMONY - Acute Myeloid Leukaemia, InGef RDB=InGef Research Database, NAJS=Croatian National Public Health Information System, NCR=Netherlands Cancer Registry

9.3.5. Overall survival of individuals with AML

Overall survival estimates varied considerably across different data source types. The number of individuals included in the analysis ranged from 630 in FinOMOP-HUS to 7,725 in NCR, with the number of event (death from any cause) counts ranging from 391 in FinOMOP-HUS to 5,299 in NCR. Overall survival estimates tended to be higher in hospital data sources than in cancer registry data sources ([Table 11](#), [Figure 11](#)).

Overall, median survival was not reached in HARMONY-AML, InGef RDB, and H12O, indicating that fewer than half of the patients had died within the observation period in these sources. Among data sources with an estimable median, it ranged from 8.2 months (95% CI [7.5; 9.1]) in DK-DHR to 15.7 months (95% CI [13.4; 20.5]) in FinOMOP-HUS and 15.6 months (95% CI [14.0; 18.0]) in SIDIAP. In NAJS, FinOMOP-THL, CRN, and NCR, individuals with AML had median survival estimates of 12.7 (11.3; 14.0), 11.3 (10.2; 12.2), 11.3 (9.8; 12.9), and 8.9 (8.3; 9.6) months, respectively.

The restricted mean survival over five years ranged from 20.5 months [19.6; 21.3] in DK-DHR to 54.5 [53.9; 55.1] in H12O. It was 44.4 [43.6; 45.3] months in InGef RDB, 38.2 [36.9; 39.5] months in HARMONY-AML, 28.4 [26.3; 30.4] months in FinOMOP-HUS, 27.2 [25.8; 28.6] months in SIDIAP, 25.4 [24.5; 26.4] months in NAJS, 25.3 [24.0; 26.5] months in CRN, 24.6 [23.6; 25.6] months in FinOMOP-THL, and 22.6 [22.0; 23.1] in NCR.

Among registries, one-year survival ranged from 42.4% in DK-DHR to 50.6% [48.8; 52.4] in NAJS, with 48.1% [46.2; 50.1] in FinOMOP-THL, 48.8% [46.5; 51.1] in CRN, and 44.6% [43.4; 45.7] in NCR. It was 55.9% [53.3; 58.6] in SIDIAP, a primary healthcare data source, and 55.0% [51.2; 59.1] in FinOMOP-HUS, a hospital data source.

One-year survival probability was 93.0% [92.1; 93.8] in H12O, 72.6% [70.4; 74.9] in HARMONY-AML, and 78.8% [77.3; 80.2] in InGef RDB, substantially higher than the registry data sources.

The distribution of 3-year and 5-year survival probabilities was similar across the data sources. The lowest 5-year survival probability was observed in DK-DHR: 21.7% [20.2; 23.3]. Comparisons of survival probabilities by sex consistently showed higher survival estimates among females than males across all data sources ([Figure 12](#)). Throughout all time points, the highest survival estimates were observed in H12O and InGef RDB.

When overall survival was examined by age group for individuals under 18 years, it was high across all data sources (Figure 13) and decreased gradually with age (Figure 14 and Table 12).

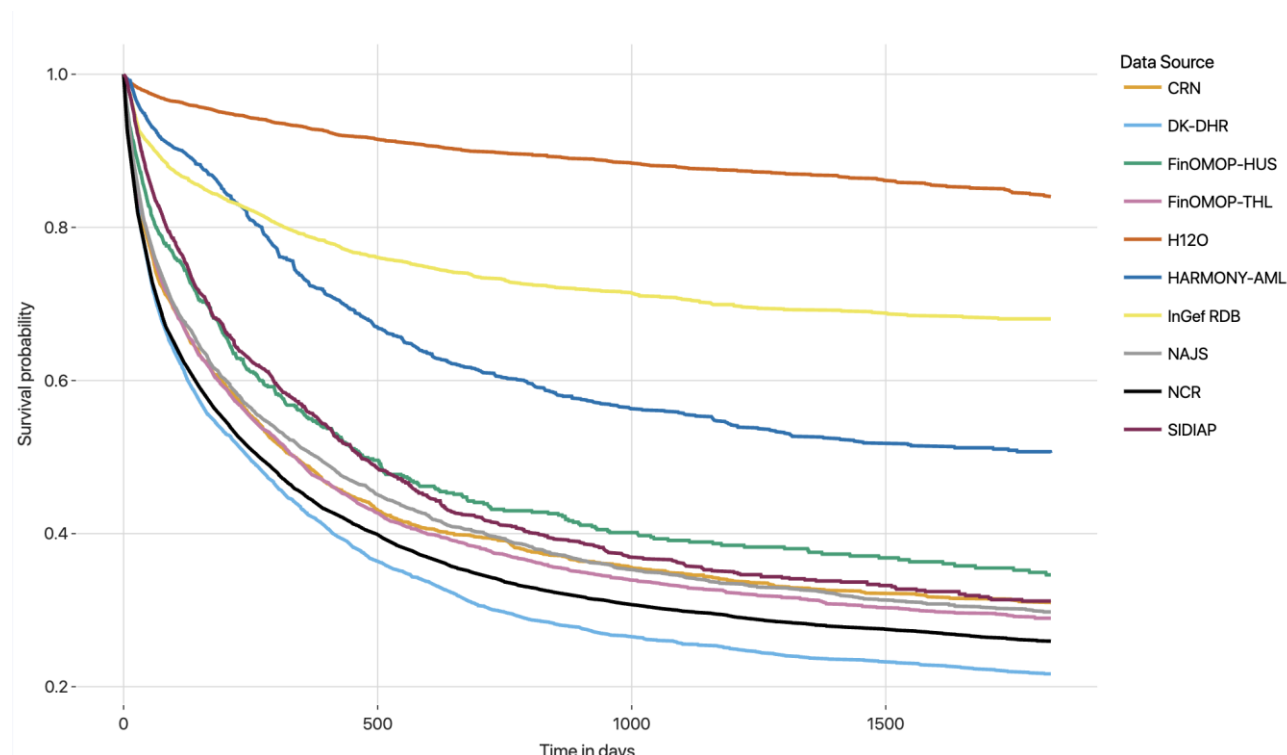


Figure 11. Overall survival probabilities of individuals with acute myeloid leukaemia (all age groups).

CRN=Cancer Registry Norway, DK-DHR=Danish Data Health Registries, FinOMOP-HUS=Hospital District of Helsinki and Uusimaa, FinOMOP-THL=Population Registers at the Finnish Institute for Health and Welfare, H12O=Hospital Universitario 12 de Octubre, HARMONY-AML=HARMONY - Acute Myeloid Leukaemia, InGef RDB=InGef Research Database, NAJS=Croatian National Public Health Information System, NCR=Netherlands Cancer Registry, SIDIAP=The Information System for Research in Primary Care.

Table 11. Overall survival estimates of individuals with acute myeloid leukaemia (all age groups).

	Records in the analysis	Number of events	Median survival (95% CI)	Restricted mean survival (95% CI) over 5 years	1-year survival probability (95% CI)	3-year survival probability (95% CI)	5-year survival probability (95% CI)
Hospitals							
FinOMOP-HUS	630	391	15.7 (13.4; 20.5)	28.4 (26.3; 30.4)	55.0% (51.2; 59.1)	39.1% (35.4; 43.2)	34.6% (30.9; 38.9)
HARMONY-AML	1,586	702	N/R	38.2 (36.9; 39.5)	72.6% (70.4; 74.9)	55.8% (53.3; 58.4)	50.5% (47.8; 53.3)
H12O ¹	3,774	444	N/R	54.5 (53.9; 55.1)	93.0% (92.1; 93.8)	87.9% (86.7; 89.0)	84.0% (82.5; 85.6)
Registries							
NAJS	3,030	1,889	12.7 (11.3; 14.0)	25.4 (24.5; 26.4)	50.6% (48.8; 52.4)	34.5% (32.7; 36.4)	29.8% (27.9; 31.7)
DK-DHR	3,433	2,467	8.2 (7.5; 9.1)	20.5 (19.6; 21.3)	42.4% (40.7; 44.1)	25.7% (24.2; 27.3)	21.7% (20.2; 23.3)

	Records in the analysis	Number of events	Median survival (95% CI)	Restricted mean survival (95% CI) over 5 years	1-year survival probability (95% CI)	3-year survival probability (95% CI)	5-year survival probability (95% CI)
FinOMOP-THL	2,614	1,704	11.3 (10.2; 12.2)	24.6 (23.6; 25.6)	48.1% (46.2; 50.1)	33.1% (31.2; 35.1)	28.9% (27.1; 30.9)
NCR	7,725	5,299	8.9 (8.3; 9.6)	22.6 (22.0; 23.1)	44.6% (43.4; 45.7)	29.9% (28.8; 31.0)	25.9% (24.9; 27.1)
CRN	1,870	1,185	11.3 (9.8; 12.9)	25.3 (24.0; 26.5)	48.8% (46.5; 51.1)	34.8% (32.6; 37.2)	31.0% (28.8; 33.4)
Primary care							
SIDIAP	1,435	849	15.6 (14.0; 18.0)	27.2 (25.8; 28.6)	55.9% (53.3; 58.6)	36.1% (33.5; 39.0)	31.2% (28.5; 34.2)
Claims							
InGef RDB	4,070	865	N/R	44.4 (43.6; 45.3)	78.8% (77.3; 80.2)	70.7% (68.9; 72.5)	68.0% (66.1; 70.0)

APL=Acute Promyelocytic Leukaemia, CI=Confidence interval, CRN=Cancer Registry Norway, DK-DHR=Danish Data Health Registries, FinOMOP-HUS=Hospital District of Helsinki and Uusimaa, FinOMOP-THL=Population Registers at the Finnish Institute for Health and Welfare, H12O=Hospital Universitario 12 de Octubre, HARMONY-AML=HARMONY - Acute Myeloid Leukaemia, InGef RDB=InGef Research Database; N/R=Not reached, NAJS=Croatian National Public Health Information System, NCR=Netherlands Cancer Registry, SIDIAP=The Information System for Research in Primary Care

¹ The majority of patients in H12O have Acute Promyelocytic Leukaemia (APL).

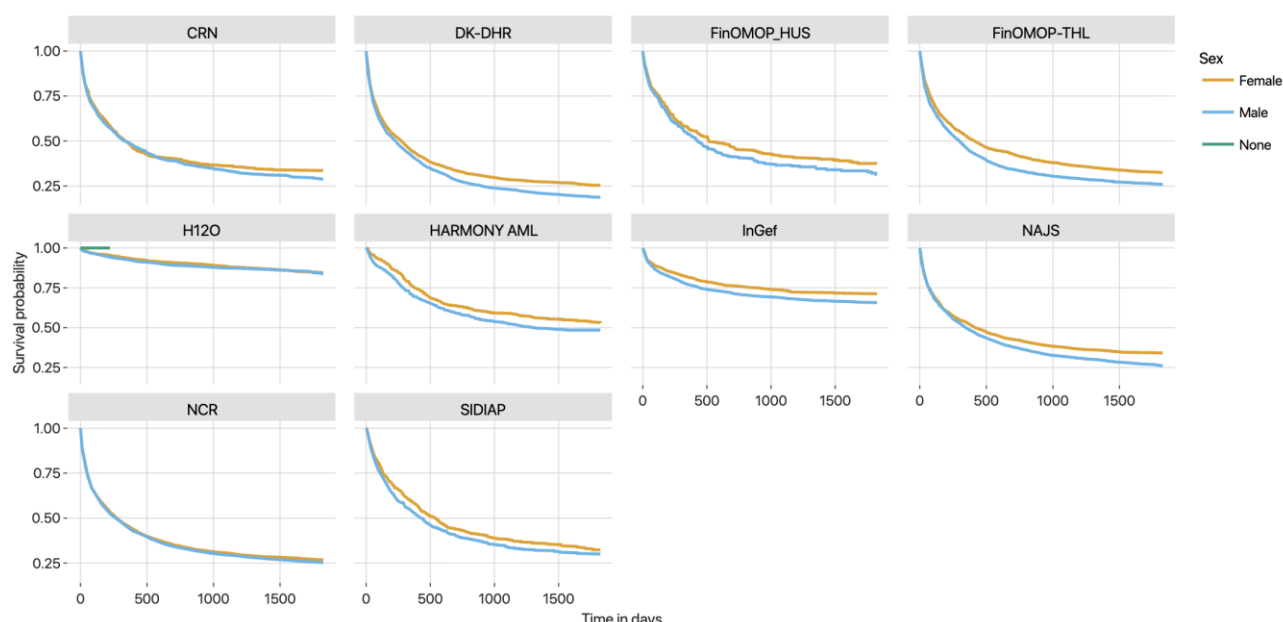


Figure 12. Overall survival probabilities of individuals with acute myeloid leukaemia by sex.

CRN=Cancer Registry Norway, DK-DHR=Danish Data Health Registries, FinOMOP-HUS=Hospital District of Helsinki and Uusimaa, FinOMOP-THL= Population Registers at the Finnish Institute for Health and Welfare , H12O=Hospital Universitario 12 de Octubre,

HARMONY-AML=HARMONY - Acute Myeloid Leukaemia, InGef RDB=InGef Research Database, NAJS=Croatian National Public Health Information System, NCR=Netherlands Cancer Registry, SIDIAP=The Information System for Research in Primary Care.

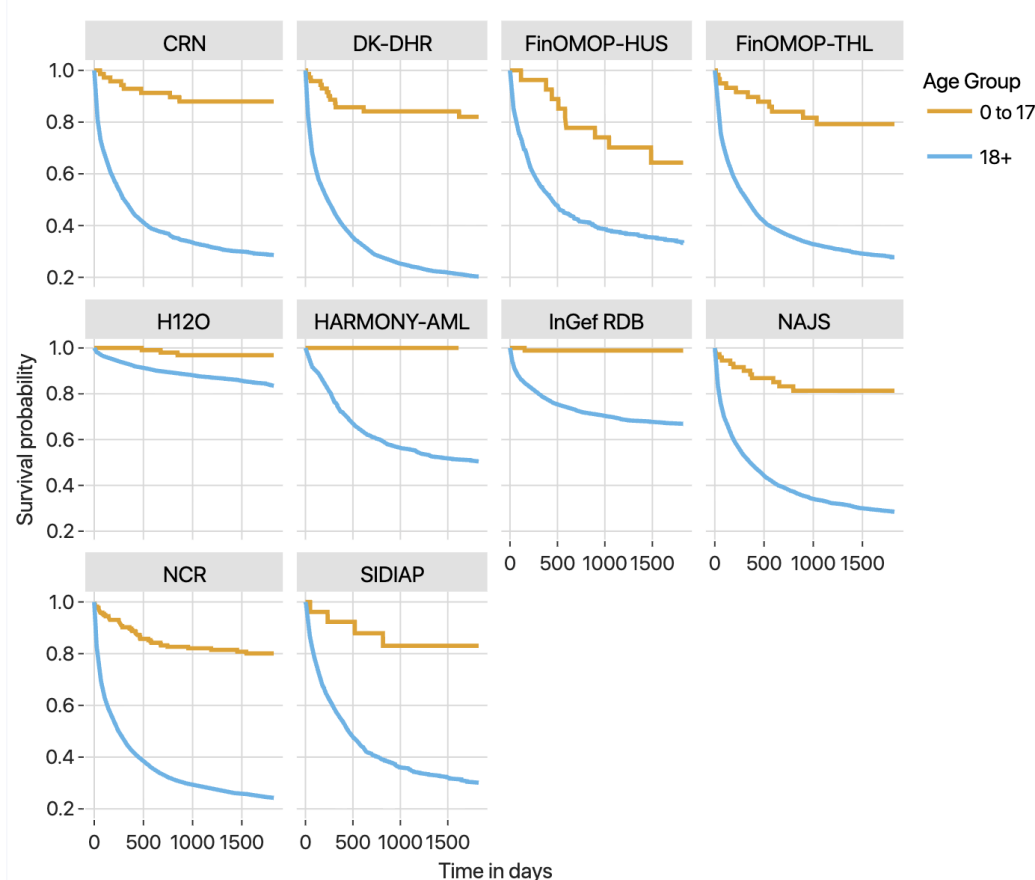


Figure 13. Overall survival probabilities of individuals with acute myeloid leukaemia by age (<18 and 18 and older).

CRN=Cancer Registry Norway, DK-DHR=Danish Data Health Registries, FinOMOP-HUS=Hospital District of Helsinki and Uusimaa, FinOMOP-THL= Population Registers at the Finnish Institute for Health and Welfare, H12O=Hospital Universitario 12 de Octubre, HARMONY-AML=HARMONY - Acute Myeloid Leukaemia, InGef RDB=InGef Research Database, NAJS=Croatian National Public Health Information System, NCR=Netherlands Cancer Registry, SIDIAP=The Information System for Research in Primary Care.

Table 12. Overall survival estimates of individuals with acute myeloid leukaemia by age group.

Data Source	Age group	1-year survival probability (95% CI)	3-year survival probability (95% CI)	5-year survival probability (95% CI)
NAJS	0 to 17	88.4% (81.2; 96.3)	81.3% (72.3; 91.6)	81.3% (72.3; 91.6)
	18 to 39	89.2% (84.8; 93.9)	77.3% (71.1; 84.1)	75.6% (69.1; 82.7)
	40 to 59	69.8% (66.0; 73.8)	56.0% (51.8; 60.6)	52.5% (48.1; 57.2)
	60 to 79	46.1% (43.6; 48.6)	27.5% (25.1; 30.0)	21.3% (19.0; 23.9)
	80 and older	27.4% (23.9; 31.4)	10.7% (8.0; 14.3)	5.1% (3.1; 8.4)
DK-DHR	0 to 17	85.8% (77.9; 94.3)	84.2% (76.0; 93.2)	82.1% (73.2; 91.9)
	18 to 39	86.3% (81.4; 91.6)	71.0% (64.2; 78.4)	66.5% (59.1; 74.9)
	40 to 59	73.5% (69.6; 77.7)	57.2% (52.7; 62.2)	53.9% (49.2; 59.1)
	60 to 79	41.7% (39.5; 44.1)	21.3% (19.3; 23.4)	15.6% (13.8; 17.7)
	80 and older	12.3% (10.2; 14.9)	2.2% (1.3; 3.7)	1.5% (0.8; 3.0)

Data Source	Age group	1-year survival probability (95% CI)	3-year survival probability (95% CI)	5-year survival probability (95% CI)
FinOMOP-HUS	0 to 17	96.3% (89.4; 100.0)	70.2% (54.8; 89.8)	64.3% (47.6; 86.9)
	18 to 39	96.0% (90.7; 100.0)	82.7% (72.4; 94.4)	73.4% (60.6; 88.8)
	40 to 59	71.7% (64.5; 79.8)	65.3% (57.6; 74.0)	60.7% (52.6; 70.1)
	60 to 79	48.3% (43.1; 54.1)	28.9% (24.1; 34.6)	25.1% (20.4; 31.0)
	80 to 150	22.0% (15.2; 31.8)	6.0% (2.7; 13.5)	2.4% (0.6; 9.2)
FinOMOP-THL	0 to 17	89.7% (82.2; 97.9)	79.2% (68.8; 91.2)	79.2% (68.8; 91.2)
	18 to 39	91.7% (87.4; 96.1)	83.6% (77.8; 89.9)	79.8% (73.2; 86.9)
	40 to 59	75.7% (71.6; 80.1)	65.8% (61.1; 70.8)	60.5% (55.4; 66.0)
	60 to 79	46.1% (43.5; 48.9)	27.3% (24.8; 29.9)	22.3% (19.9; 25.0)
	80 and older	17.2% (14.3; 20.6)	5.1% (3.5; 7.5)	3.0% (1.7; 5.1)
InGef RDB	0 to 17	98.9% (96.8; 100.0)	98.9% (96.8; 100.0)	98.9% (96.8; 100.0)
	18 to 39	97.2% (95.1; 99.5)	91.4% (87.4; 95.6)	89.7% (85.2; 94.5)
	40 to 59	89.6% (87.4; 91.8)	84.4% (81.7; 87.3)	82.2% (79.1; 85.4)
	60 to 79	77.0% (74.9; 79.2)	68.0% (65.4; 70.8)	63.9% (60.8; 67.1)
	80 and older	56.1% (51.7; 61.0)	33.4% (27.3; 40.8)	32.1% (26.0; 39.8)
NCR	0 to 17	89.7% (85.7; 93.9)	82.1% (77.0; 87.5)	80.1% (74.7; 85.9)
	18 to 39	81.9% (78.5; 85.5)	68.7% (64.4; 73.3)	65.3% (60.7; 70.2)
	40 to 59	66.8% (64.4; 69.3)	54.7% (52.1; 57.4)	50.1% (47.4; 53.0)
	60 to 79	40.6% (39.1; 42.2)	22.7% (21.3; 24.1)	17.9% (16.6; 19.3)
	80 and older	13.2% (11.6; 15.1)	3.0% (2.1; 4.2)	1.1% (0.6; 2.1)
CRN	0 to 17	92.9% (87.0; 99.1)	87.9% (80.4; 96.2)	87.9% (80.4; 96.2)
	18 to 39	91.6% (87.0; 96.5)	80.4% (73.6; 87.8)	80.4% (73.6; 87.8)
	40 to 59	75.3% (70.7; 80.2)	63.7% (58.5; 69.5)	57.9% (52.3; 64.0)
	60 to 79	43.9% (40.8; 47.2)	26.5% (23.6; 29.7)	21.4% (18.6; 24.7)
	80 and older	13.6% (10.4; 17.6)	1.8% (0.7; 4.5)	0.4% (0.1; 3.1)
SIDIAP	0 to 17	92.3% (82.6; 100.0)	83.0% (69.0; 99.8)	83.0% (69.0; 99.8)
	18 to 39	91.9% (86.3; 97.9)	76.2% (67.3; 86.2)	76.2% (67.3; 86.2)
	40 to 59	78.0% (73.1; 83.3)	63.3% (57.3; 70.0)	57.6% (50.9; 65.1)
	60 to 79	55.9% (52.1; 59.9)	31.9% (28.1; 36.2)	25.9% (22.1; 30.4)
	80 and older	28.5% (24.1; 33.7)	9.7% (6.7; 13.9)	5.1% (2.9; 9.2)
HARMONY-AML	NA	NA	NA	NA
	18 to 39	83.7% (78.9; 88.8)	64.1% (57.7; 71.1)	61.1% (54.5; 68.5)
	40 to 59	79.3% (76.3; 82.3)	64.0% (60.5; 67.8)	59.4% (55.6; 63.4)
	60 to 79	61.7% (57.9; 65.7)	44.3% (40.3; 48.6)	37.2% (33.1; 41.9)
	80 and older	30.5% (13.7; 67.6)	NA	NA
H12O	0 to 17	100.0% (100.0; 100.0)	96.8% (93.3; 100.0)	96.8% (93.3; 100.0)
	18 to 39	98.4% (97.1; 99.7)	96.2% (94.0; 98.5)	93.1% (89.5; 96.9)
	40 to 59	97.2% (96.3; 98.0)	94.6% (93.3; 95.8)	92.6% (90.9; 94.4)
	60 to 79	91.7% (90.3; 93.2)	85.5% (83.5; 87.5)	80.3% (77.5; 83.2)

Data Source	Age group	1-year survival probability (95% CI)	3-year survival probability (95% CI)	5-year survival probability (95% CI)
	80 and older	66.4% (61.0; 72.2)	47.7% (41.1; 55.3)	33.7% (25.8; 43.8)

CI=Confidence interval, CRN=Cancer Registry Norway, DK-DHR=Danish Data Health Registries, FinOMOP-HUS=Hospital District of Helsinki and Uusimaa, FinOMOP-THL=Population Registers at the Finnish Institute for Health and Welfare, H12O=Hospital Universitario 12 de Octubre, HARMONY-AML=HARMONY - Acute Myeloid Leukaemia, InGef RDB=InGef Research Database, NA=Not available, NAJS=Croatian National Public Health Information System, NCR=Netherlands Cancer Registry, SIDIAP=The Information System for Research in Primary Care.

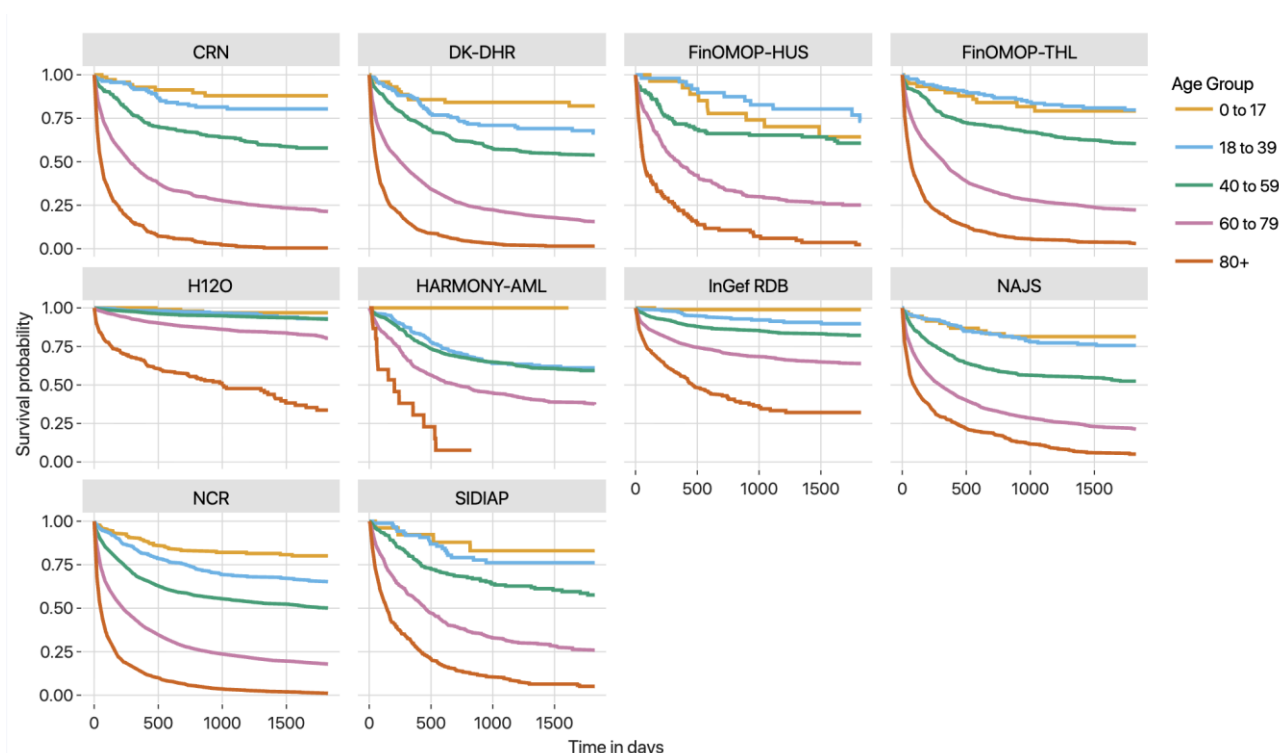


Figure 14. Overall survival probabilities of individuals with acute myeloid leukaemia by age (0 to 17; ≥18; 18 to 39; 40 to 59; 60 to 79; 80 and over).

CRN=Cancer Registry Norway, DK-DHR=Danish Data Health Registries, FinOMOP-HUS=Hospital District of Helsinki and Uusimaa, FinOMOP-THL= Population Registers at the Finnish Institute for Health and Welfare, H12O=Hospital Universitario 12 de Octubre, HARMONY-AML=HARMONY - Acute Myeloid Leukaemia, InGef RDB=InGef Research Database, NAJS=Croatian National Public Health Information System, NCR=Netherlands Cancer Registry, SIDIAP=The Information System for Research in Primary Care.

10. DISCUSSION

10.1. Key results

In this study, we estimated the incidence of AML, characterised AML patients by demographics, performance status, comorbidities, disease characteristics, measurements, and procedures, described treatments, and estimated overall survival for patients with newly diagnosed AML. The results should not be used to compare the quality of treatment, competence or survival between the different data sources, and the institutions represented in this collaboration. However, results of the study will further help to understand and improve data quality, completeness, and assist work on more advanced AML research, such as treatment patterns and the identification of relapse/refractory AML.

Overall, 30,542 individuals with AML diagnosis were included in the study (631 in FinOMOP-HUS to 7,784 in NCR). The mean age varied from 55 years in HARMONY-AML to 68 years in DK-DHR.

Estimates of the overall incidence of AML per 100,000 are within the range of 2.9 (95% CI [2.8; 3.1]) in SIDIAP to 8.6 (95% CI [7.7; 9.6]) in NAJS. Incidence increases with age across all data sources. Incidence estimates by sex consistently show higher rates in males than in females across all data sources. Calendar-year trends from 2015 to 2024 generally demonstrate stable incidence rates across most data sources, with no clear time trends.

High missingness in recording of genetic mutation (CEBPA, DDX41, NPM1) was observed, with only three data sources recording this data: FinOMOP-HUS, HARMONY-AML, and H12O; and only one of them, namely HARMONY-AML, with high completeness.

Records of blasts in bone marrow measurement, blasts in blood measurements, and LDH measurements in blood were recorded very variably across data sources.

High missingness in recording of ECOG status was observed across most data sources, with some WHO/ECOG performance status scores available in two data sources, although not for all patients: HARMONY-AML and NCR (majority with ECOG=0 and 1). The absence of ECOG=3 and 4 in HARMONY-AML suggests that in this data source individuals with AML with poor performance status were not included.

Across all data sources, azacitidine, cytarabine, daunorubicin, idarubicin, venetoclax, fludarabine, etoposide, gemtuzumab ozogamicin, and allogeneic stem cell transplantation were the most frequently recorded ingredients over different 30-day periods. The AML treatment pathways varied across data sources; however, they could not be correctly estimated in all data sources due to inadequate mapping in regimens and/or missing data. Pathways can be broadly grouped into azacitidine and venetoclax, cytarabine, and transplantation-containing regimens, with some overlap observed in sequential combinations (e.g. when one combination is followed by another).

Allogeneic stem cell transplantation is more common after 90 days following AML diagnosis, with the proportion of patients recorded for allogeneic stem cell transplantation peaking after the decline in cytarabine use across several data sources (FinOMOP-THL, FinOMOP-HUS, HARMONY-AML, NCR), reflecting current recommendations on the treatment of AML, when induction therapy is followed by consolidation with allogeneic stem cell transplantation.

Overall survival estimates vary widely across different data source types.

In registries, one-year survival probability ranges from 42.4% in DK-DHR to 50.6% [48.8; 52.4] in NAJS. It was 55.9% [53.3; 58.6] in SIDIAP, the primary healthcare data source, and 55.0% [51.2; 59.1] in FinOMOP-HUS.

A few data sources were outliers with regards to one-year survival probability: 93.0% [92.1; 93.8] in H12O (where the majority of individuals had APL), 72.6% [70.4; 74.9] in HARMONY-AML, and 78.8% [77.3; 80.2] in InGef RDB, markedly higher than in registry data sources. There could be several explanations for this, including the type of patients in the specific data source, possible presence of false-positive diagnoses, and selective loss to follow-up.

Overall survival is substantially high for individuals under 18 years across all data sources and decreases gradually with age.

10.2. Strengths and limitations of the research methods

The study relied on routinely collected healthcare data, so it is important to consider potential data quality issues. In our study, incidence rates per 100,000 calculated from DK-DHR, InGef RDB, and NAJS differed from official national statistics and tended to be moderately overestimated, possibly due to the inclusion of provisional diagnoses, especially in the claims data sources. This will be further explored with the data partners to understand how we can overcome the limitation.

Secondly, hospital data sources contain selected patients depending on the hospital profile (e.g. H12O contains mostly patients with APL, a distinct subtype of AML characterised by a favourable prognosis); therefore, the data on patient characteristics and treatment patterns from hospital data might differ from more general registries.

Third, information on the use of different treatments depends on the data source's ability to capture it and then map it to OMOP CDM. For example, information on allogeneic stem cell transplantation was unavailable or very limited in several data sources (e.g. CRN, DK-DHR), likely reflecting incomplete mapping of procedures to the OMOP CDM. Similarly, a substantial proportion of individuals with AML had no records of treatment, indicating missingness. For example, in FinOMOP-THL, we could not identify therapy prior to allogeneic stem cell transplantation, which is inconsistent with AML guidelines and indicates missingness. Some of this data exists in the raw, unmapped data, and therefore the missingness can be decreased with improved mapping. The Danish AML registry, which is not governed by the Danish Health Data Authority and contains detailed information on treatment, is not part of the DK-DHR data source, as the DK-DHR has access only to registries governed by the Danish Health Data Authority.

The completeness of treatments recorded was highest in HARMONY-AML (a specialised AML registry designed to capture treatment), followed by CRN and NCR (cancer registries); in other data sources a substantial amount of information on AML treatment was missing.

Fourth, because the *TreatmentPatterns* analytical package depends on both the settings in the algorithm and the variation in how treatments are registered across data sources (e.g. treatment duration, start and end dates), it was sometimes impossible to meaningfully reconstruct sequences and regimens from available data. The package also failed to identify correct sequences in NAJS, InGef RDB, and FinOMOP-THL. The results of this analysis did not yield logical sequences and combinations of treatments that could be compared with known regimens, as initially planned (for e.g. "7+3" regimen). This limitation can also be addressed through better mapping and improved algorithms suitable for oncology regimens.

Fifth, three data sources (H12O, HARMONY-AML, and InGef RDB) were clear outliers with regards to overall survival, overestimating it by a large degree. This can be explained by a selective (younger and with performance status not more than 2) population in HARMONY-AML and H12O, which is a highly specialised tertiary centre, where the majority of individuals had APL, and selective loss to follow-up and incomplete recording of death in InGef RDB. For InGef RDB, the issue will be investigated further to determine whether data quality can be improved.

Sixth, it should also be noted that there is some overlap between data sources. Individuals with AML in FinOMOP-HUS are also likely to be present HARMONY-AML and FinOMOP-THL. In the federated analysis, it is not possible to account for this overlap; however, the results of the survival analysis suggest that information on most individuals with AML in FinOMOP-HUS is also available in FinOMOP-THL, where less detailed information is recorded. At the same time, information from a rather limited sample of individuals with AML from FinOMOP-HUS was shared with HARMONY-AML.

Seventh, although the study was age-agnostic, there remains very limited information on children with AML, largely due to the small sample sizes in the different data sources. Treatment patterns could not be identified in children with AML due to small sample size. In addition, similarly to adults, incidence and survival estimates were overestimated NAJS, H12O, InGef RDB and HARMONY AML. AML in the paediatric population requires treatment with different regimens, with different intentions, having different outcomes, and the lack of information on children with AML in this study is a clear limitation for future research.

While data recorded in cancer registries are generally of better quality in terms of cancer-specific information (for e.g. here they were the only sources with some regimens mapped, biomarkers, and ECOG scores), even these data sources have limitations.

Finally, information on genetic mutation and ECOG/PS WHO remains limited across data sources in this study, continuing to pose a challenge in categorising AML patients into different risk groups. In future studies, the period to describe genetic mutation can be extended to 30 days prior diagnosis date.

10.3. Interpretation

In this study, hospital data sources (FinOMOP-HUS and HARMONY-AML) were analysed alongside combined nationwide registries (FinOMOP-THL, DK-DHR, NAJS), cancer registries (NCR, CRN), claims data (InGef RDB), and primary healthcare data (SIDIAP).

Estimates of AML incidence rates per 100,000 from various sources are available for the countries included in the study: for Finland, Denmark, and Norway from Nordcan [13], for the Netherlands from the national cancer registry portal [14], and for Croatia from the national report.[15]

The AML incidence estimates in our study were slightly higher than those in public sources for most countries, except the Netherlands. These differences may arise from the definition of the disease (or the list of ICD-O-3 codes included in the definition) used in public sources that report cancer registry data and mapping, as well as from differences in the structures of the data sources. In addition, information collected even in high-quality cancer registries may be incomplete and may include data from other available sources, for example, in Finland or Denmark, which may add cases not captured in the cancer registry, resulting in an increased risk of false positives or duplicate cases.

Two data sources largely overestimated incidence (NAJS and InGef RDB) and to a lesser extent, DK-DHR. It seems that public health registries and claims data pick preliminary, unconfirmed diagnoses, leading to an increase in incidence rates. This is a known limitation of those data sources.

Observed incidence rates of AML in the paediatric population in this study were similar to known rates; however, as with overall incidence, AML incidence in children was also high in NAJS and InGef RDB, likely reflecting specifics of data capture in these data sources rather than a true difference in incidence in the underlying populations. [13,14,15]

In conclusion, it is clear that for population-based incidence, cancer registries are the best source and should be used for this type of analysis.

Globally, AML was the most common subtype of leukaemia in 2022, with estimates of age-standardised incidence rates (Segi-Doll world standard [16,17]) of 3.1 per 100,000 in males and 2.4 per 100,000 in females. There was also a clear increase in AML incidence with age.[18] A similar pattern of incidence by sex and age was observed in our study across all data sources.

In H120 the majority of patients had APL, subtypes of AML with a favourable prognosis.[19] In addition, the patients in HARMONY-AML and H120 were, on average, younger, and, based on the comparison of performance status with NCR, had fewer contraindications to aggressive curative treatment. This suggests a more selected sample in these two data sources compared with population-based sources, such as cancer registries. Caution should be exercised when using these data sources, and, in general, hospital data sources, for AML studies.

The proportion of individuals with a recorded history of MDS in the 90 days prior to AML diagnosis ranged from 7% to 16% in population-based data sources in our study. Previous studies reported varying probabilities of progression from MDS to AML. For example, a large study examining the incidence and complications of MDS among Medicare beneficiaries in the United States showed that approximately 10% of patients with MDS progressed to AML.[20] In a large Healthcare Cost and Utilisation Project National Inpatient Sample (HCUP-NIS), out of 79,579 patients with AML who were hospitalised, 4,650 (6%) had a secondary diagnosis of MDS.[21] The history of cancer prior to diagnosis (excluding non-melanoma skin cancer) varied across data sources and likely captures both information of prior primary malignancies (e.g. in NCR), but also more general concepts describing malignancies that can be used to capture AML diagnosis

in EHR in the initial steps of diagnostic work-up (e.g. in FinOMOP-THL). While up to 15% of AML cases had a relation to previous therapy for another disease, this proportion is likely to rise due to increasing numbers of cancer survivors, especially in older age groups.[22] For example, one study showed that there was an increased risk of AML following radiotherapy for prostate cancer in men.[23]

Recorded genetic mutation measurements within the first 30 days after diagnosis were observed in individuals with AML across only three data sources: FinOMOP-HUS, HARMONY-AML, and H12O, with only HARMONY-AML having a high completeness rate. The mutation status measurements were available for CEBPA, DDX41, and NPM1. As AML represents a heterogeneous group of genetically distinct disorders, the information on the status of genetic abnormalities is required for proper classification and risk stratification.[9]

It is worth mentioning that in some countries, genetic mutation measurements and other characteristics are available in other data sources, e.g. the Danish AML registry in Denmark, which is not part of DK-DHR, as it is not governed by the Danish Health Data Authority.

Information on the blasts in bone marrow or blood is essential to the diagnosis of AML and follow-up after treatment, and it was captured in at least six data sources in our study, mainly in the first 30-day period after AML diagnosis. In at least two data sources in our study (DK-DHR, FinOMOP-HUS), this information was captured for the majority of AML patients, suggesting that it can be used in future studies.

Across all data sources, azacitidine, cytarabine, daunorubicin, idarubicin, venetoclax, fludarabine, etoposide, gemtuzumab ozogamicin, and allogeneic stem cell transplantation were the most frequently recorded treatments. It should be noted that stem cell transplantation is mapped to the Procedures domain of the OMOP CDM, and it was not mapped in multiple data sources, despite the drugs being mapped.

Overall, the information on treatment capture in our study mainly reflected the two main scenarios of treatment for AML patients: treatment with curative intent, including induction therapy with a combination of anthracyclines and cytarabine followed by stem cell transplantation; and, for patients not suitable for intensive therapy: venetoclax, azacitidine, low-dose cytarabine, or other agents.[22,24]

When the analyses on treatment were stratified by younger and older age groups, the pattern of cytarabine and anthracyclines followed by allogeneic stem cell transplantation was consistent across all data sources in the younger age group, whereas azacitidine and venetoclax were more common in older individuals with AML. Our results probably reflect the current population-based landscape of AML therapy, where most patients are not treated with curative intent.

In the randomised trial, the median overall survival for patients treated with the venetoclax and azacitidine combination was 14.7 months, while with azacitidine alone it was 9.6 months.[24] For data sources with an identifiable median in our study, it ranged from 8.2 months in DK-DHR to 15.7 months in FinOMOP-HUS and 15.6 months in SIDIAP. In a recent analysis of the Surveillance, Epidemiology, and End Results data source in the United States, the overall survival rate was 20.8% in the 2010s compared to roughly 5% in 1970.[25] In our study, five-year survival probabilities from population-based data sources ranged from 22% in DK-DHR to 31% in CRN, indicating the average prognosis of individuals with AML in Europe in the more recent period.

Three data sources (H12O, HARMONY-AML, and InGef RDB) were clear outliers with regards to overall survival, overestimating it to a large degree and making them less suitable for this analysis. For the paediatric AML population, overall survival in H12O and InGef RDB was also exceptionally and unrealistically high compared with other data sources, suggesting clear patient selection (H12O) or potentially other factors that require further investigation, e.g. possible problems with loss to follow-up or the presence of unconfirmed diagnoses.

Finally, similar to other studies, we observed a clear difference in prognosis by age, with older age groups having lower survival.[25,26]

11. CONCLUSION

Haemato-oncology is a complex topic to study using observational data due to the complexity of treatment and the requirements for highly specialised data, such as genetic data, procedures, and laboratory information. Most data sources are not equipped for this purpose.

In the 10 data sources, we observed substantial variability in the completeness of diagnostic, molecular, and treatment-pathway data, highlighting challenges for AML research using observational data. More importantly, a substantial proportion of individuals with AML lacked records of any treatment, or it was impossible to understand the sequence of treatment due to data and analytical limitations. This issue was even more pronounced in paediatric populations due to small sample sizes across different data sources.

Survival estimates were in line with the literature across most data sources but were relatively high in data sources with selected populations (APL, fit patients, eligible for clinical trials) and in claims data sources.

The results of this study should not be used to assess the quality of treatment, professional competence, or prognosis in AML in general, because we were describing results from different data sources with varying completeness that represent different AML populations. This study can inform the assessment of the fitness-for-use of available data for future studies on AML and the identification of data sources likely to contribute to such studies, depending on the question of interest. This study also highlights the current need for improvements in data quality, completeness, and analytical tools for treatment patterns in AML, as well as the steps needed to address more advanced research questions, such as those related to relapse/refractory AML.

12. REFERENCES

1. Shimony S, Stahl M, Stone RM. Acute myeloid leukemia: 2023 update on diagnosis, risk-stratification, and management. *Am J Hematol*. 2023;98(3):502-526. doi:10.1002/ajh.26822
2. Juliusson G, Antunovic P, Derolf Å, et al. Age and acute myeloid leukemia: real world data on decision to treat and outcomes from the Swedish Acute Leukemia Registry. *Blood*. 2009;113(18):4179-4187. doi:10.1182/blood-2008-07-172007
3. Döhner H, Estey E, Grimwade D, et al. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. *Blood*. 2017;129(4):424-447. doi:10.1182/blood-2016-08-733196
4. Khwaja A, Bjorkholm M, Gale RE, et al. Acute myeloid leukaemia. *Nat Rev Dis Primer*. 2016;2(1):16010. doi:10.1038/nrdp.2016.10
5. Linet MS, Curtis RE, Schonfeld SJ, Vo JB, Morton LM, Dores GM. Survival of adult AML patients treated with chemotherapy in the U.S. population by age, race and ethnicity, sex, calendar-year period, and AML subgroup, 2001–2019. *eClinicalMedicine*. 2024;71:102549. doi:10.1016/j.eclinm.2024.102549
6. Usuki K, Ohtake S, Honda S, et al. Real-world data of AML in Japan: results of JALSG clinical observational study-11 (JALSG-CS-11). *Int J Hematol*. 2024;119(1):24-38. doi:10.1007/s12185-023-03677-w
7. Visser O, Trama A, Maynadié M, et al. Incidence, survival and prevalence of myeloid malignancies in Europe. *Eur J Cancer*. 2012;48(17):3257-3266. doi:10.1016/j.ejca.2012.05.024
8. Maynadie M, De Angelis R, Marcos-Gragera R, et al. Survival of European patients diagnosed with myeloid malignancies: a HAEMACARE study. *Haematologica*. 2013;98(2):230-238. doi:10.3324/haematol.2012.064014
9. Arber DA, Orazi A, Hasserjian RP, et al. International Consensus Classification of Myeloid Neoplasms and Acute Leukemias: integrating morphologic, clinical, and genomic data. *Blood*. 2022;140(11):1200-1228. doi:10.1182/blood.2022015850
10. Khoury JD, Solary E, Abal O, et al. The 5th edition of the World Health Organization Classification of Haematolymphoid Tumours: Myeloid and Histiocytic/Dendritic Neoplasms. *Leukemia*. 2022;36(7):1703-1719. doi:10.1038/s41375-022-01613-1
11. Huber S, Baer C, Hutter S, et al. AML classification in the year 2023: How to avoid a Babylonian confusion of languages. *Leukemia*. 2023;37(7):1413-1420. doi:10.1038/s41375-023-01909-w
12. Dernie F, Corby G, Robinson A, et al. Standardised and Reproducible Phenotyping Using Distributed Analytics and Tools in the Data Analysis and Real World Interrogation Network (DARWIN EU). *Pharmacoepidemiol Drug Saf*. 2024;33(11):e70042. doi:10.1002/pds.70042
13. Engholm G, Ferlay J, Christensen N, et al. NORDCAN – a Nordic tool for cancer information, planning, quality control and research. *Acta Oncol*. 2010;49(5):725-736. doi:10.3109/02841861003782017
14. Netherlands Comprehensive Cancer Organisation (IKNL). Netherlands Cancer Registry (NCR) Data. Accessed February 4, 2026. http://nkr-cijfers.iknl.nl/viewer/incidentie-per-jaar?language=en_GB
15. Mario Šekerija. *Hrvatski Zavod Za Javno Zdravstvo. Rak u Hrvatskoj 2023./2024*. 2026. Accessed July 4, 2026. <https://www.hzjz.hr/wp-content/uploads/2026/02/Bilten-2023.pdf>
16. Doll R. The geographical distribution of cancer. *Br J Cancer*. 1969;23(1):1-8.

17. Segi M, Fujisaku S, Kurihara M, Narai Y, Sasajima K. The age-adjusted death rates for malignant neoplasms in some selected sites in 23 countries in 1954-1955 and their geographical correlation. *Tohoku J Exp Med*. 1960;72(1):91-103.
18. Daltveit DS, Morgan E, Colombet M, et al. Global patterns of leukemia by subtype, age, and sex in 185 countries in 2022. *Leukemia*. 2025;39(2):412-419. doi:10.1038/s41375-024-02452-y
19. Voso MT, Guarnera L, Lehmann S, et al. Acute promyelocytic leukemia: long-term outcomes from the HARMONY project. *Blood*. 2025;145(2):234-243. doi:10.1182/blood.2024026186
20. Goldberg SL, Chen E, Corral M, et al. Incidence and Clinical Complications of Myelodysplastic Syndromes Among United States Medicare Beneficiaries. *J Clin Oncol*. 2010;28(17):2847-2852. doi:10.1200/JCO.2009.25.2395
21. Obeidat K, Turk E, Deenadayalan V, Litvin R, Zia M, Rifkin SD. Demographic Characteristics and Outcomes in Patients with AML and Previous Diagnosis of CML or MDS. *Blood*. 2022;140(Supplement 1):8942-8943. doi:10.1182/blood-2022-170618
22. Döhner H, Wei AH, Appelbaum FR, et al. Diagnosis and management of AML in adults: 2022 recommendations from an international expert panel on behalf of the ELN. *Blood*. 2022;140(12):1345-1377. doi:10.1182/blood.2022016867
23. Wang R, Zeidan AM, Yu JB, et al. Myelodysplastic Syndromes and Acute Myeloid Leukemia After Radiotherapy for Prostate Cancer: A Population-Based Study. *The Prostate*. 2017;77(5):437-445. doi:10.1002/pros.23281
24. DiNardo CD, Jonas BA, Pullarkat V, et al. Azacitidine and Venetoclax in Previously Untreated Acute Myeloid Leukemia. *N Engl J Med*. 2020;383(7):617-629. doi:10.1056/NEJMoa2012971
25. Görlich D, Lanwer L, Sauerland C, Krug U, Faldum A. Characterizing Overall Survival of Patients with Acute Myeloid Leukemia: A Competing Risk Analysis of SEER Data Covering 46 Years. *Cancers*. 2025;17(23):3735. doi:10.3390/cancers17233735
26. Hemminki K, Zitricky F, Försti A, et al. Age-specific survival in acute myeloid leukemia in the Nordic countries through a half century. *Blood Cancer J*. 2024;14(1):44. doi:10.1038/s41408-024-01033-7

13. ANNEXES

ANNEX I. Description of data sources

Croatian National Public Health Information System (NAJS)

The National Public Health Information System (Nacionalni javnozdravstveni informacijski sustav - NAJS) is an organised system of information services by the Croatian Institute of Public Health (CIPH). NAJS enables data collecting, processing, recording, managing, and storing of health-related data from health care providers as well as production and management of health information. NAJS contains medical and public health data collected and stored in health registries and other health data collections, including cancer registry, mortality, work injuries, occupational diseases, communicable and non-communicable diseases, health events, disabilities, psychosis and suicide, diabetes, drug abuse, and others.

Danish Data Health Registries (DK-DHR)

Danish health data is collected, stored, and managed in national health registers at the Danish Health Data Authority and covers the entire population. There are no gaps in terms of gender, age, and geography in Danish health data due to mandatory reporting on all patients. Personal identification numbers enable linking of data across registers, regardless of whether citizens have moved around the country. The Danish Health Data Authority is responsible for the national health registers and for maintaining and developing standards and classifications in the Danish healthcare system. Legislation ensures balance between personal data protection and use. In the present data base, we have access to the following registries for the entire Danish population of 5.9 million persons from 1995: the central Person Registry, the National Patient Registry, the Register of Pharmaceutical Sales, the National Cancer Register, the Cause of Death registry, the Clinical Laboratory Information Register, COVID-19 test and vaccination registries, the complete vaccination registry.

Hospital District of Helsinki and Uusimaa (FinOMOP-HUS)

The HUS data lake is a comprehensive, integrated data source derived from all patients who visit the HUS hospitals and receive treatment. HUS is responsible for specialised healthcare in the Uusimaa region of Finland and the treatment of many rare and severe diseases that are nationally centralised to HUS. HUS' catchment area covers a population of over two million people. All visits, examinations, laboratory tests, procedures, and treatments are recorded in the HUS IT systems and integrated into the data lake. The data lake stores decades of clinical information in digital format, data from both past and current source systems are available.

Population Registers at the Finnish Institute for Health and Welfare (FinOMOP-THL)

The THL HILMO data source encompasses all healthcare encounters in Finland. The main content of the THL CDM is the Finnish Care Register for Health Care (HILMO). It is a continuation of the former Hospital Discharge Register, which initially gathered data on patients discharged from hospitals. The Care Register has comprehensive data on the use of services and service users from Finnish public inpatient and outpatient primary and specialised care nationwide. In addition, the CDM also contains the vaccination data from the Finnish National Vaccination Register and positive COVID-19 test results from the Finnish National Infectious Diseases Register, which THL maintains. The CDM is currently produced from the sources above and is limited to observation periods commencing after 2011. The National Population registry data forms the basis for creating the patient population in the data source. This ensures the up-to-date location (municipality of residence) of patients, as well as complete records of deaths (although not the cause of death). Using the full population as a basis for the person table also facilitates calculations on a population level, such as incidence rates. The current CDM population comprises all persons who have been alive and residing in Finland since 2011.

InGef Research Database (InGef RDB)

The InGef database comprises anonymised longitudinal claims data of about 10 million individuals across more than 50 statutory health insurance providers (SHIs) throughout Germany. Data have been longitudinally linked over the past decade. Patients can be traced across the healthcare sectors. All patient-level and provider-level data in the InGef research database are anonymised to comply with German data protection regulations and German federal law. German SHI claims data available in the InGef database includes information on demographics (year and quarter of birth, gender, death date if applicable, region of residence on administrative district level), hospitalisations, outpatient services (diagnoses, treatments, specialities of physicians), dispensing of drugs, dispensing of remedies and aids, and sick leave and sickness allowance times. Additionally, costs or cost estimates from the SHI perspective are available for all key cost elements. All diagnoses in Germany are coded using the International Classification of Diseases, version 10 in the German Modification (ICD-10-GM). The persistence (membership over time) is relatively high in the InGef database. Between 2009 and 2013, 70.6% of insurance members survived and remained insured with the same SHI, without any gaps in their observational time. Persons leaving one of the participating SHIs and entering another participating SHI can be linked during yearly database consistency updates and are thus not lost over time. The InGef database is dynamic, i.e. claims data are updated in an ongoing process, and new SHIs may join or leave the database. By law, only the last 10 years of data are allowed to be used. At every new release, this window shifts, dropping older data and adding new data.

Netherlands Cancer Registry (NCR)

The NCR compiles clinical data of all individuals newly diagnosed with cancer in the Netherlands. Trained cancer registration clerks have been registering newly diagnosed cancer patients on a national basis since 1989, with a total of 3 million patients included. Over the past 30 years, this registry has provided clinicians and researchers with a wealth of clinical data (e.g. patient and tumour characteristics, primary treatment, survival) on cancer patients of all ages.

Cancer Registry Norway (CRN)

The Cancer Registry of Norway (CRN) collects national data and produces statistics on cancer incidence, prevalence, mortality, and survival in Norway. Since 2024, CRN is part of the Norwegian Institute of Public Health. CRN was established in 1951 and is one of the oldest national cancer registries in the world. Every citizen has a unique personal identification number, allowing for linkage between multiple sources of data. All medical doctors in the country are required by law to notify the Cancer Registry of new cancer cases, and registration does not require patient consent. The Statutory Regulations for the Cancer Registry of Norway also include the registration of treatment and follow-up of Norwegian cancer patients. Clinical registries—comprehensive registration schemes dedicated to specific cancers— have been established to include detailed information on diagnostic measures, therapy, and follow-up.

The Information System for Research in Primary Care (SIDIAP)

The Information System for Research in Primary Care (SIDIAP) is a clinical database of anonymised patient records in Catalonia, Spain. The Spanish public healthcare system covers more than 98% of the population, and more than two-thirds of the Catalan population see their general practitioner at least once a year. The computerisation of the primary care patient records of the Catalan Health Institute (CHI) was complete in 2005. SIDIAP was designed to provide a valid and reliable database of information from clinical records of patients registered in primary care centres for use in biomedical research. SIDIAP contains data from anonymised patients' healthcare records for nearly six million people (approximately 80% of the Catalan population) registered in 287 primary care practices throughout Catalonia since 2005. It includes data collected by health professionals during routine visits in primary care, including anthropometric measurements, clinical diagnoses (International Classification of Diseases 10th revision ICD-10), laboratory tests, prescribed and dispensed drugs, hospital referrals, demographic, and lifestyle information. It was

previously shown that SIDIAP population is highly representative of the entire Catalan region in terms of geographic, age, and sex distributions. The high quality of these data has been previously documented, and SIDIAP has been successfully applied to epidemiological studies of key exposures and outcomes. Quality checks to identify duplicate patient ids are performed centrally at each SIDIAP database update. Checks for logical values and data harmonisation are performed. For biochemistry data, consistency for measurements taken in different laboratories is assessed, and unit conversion is undertaken when needed.

Hospital Universitario 12 de Octubre (H12O)

The primary data source is the Hospital Universitario 12 de Octubre's electronic health records (EHR). The Hospital Universitario 12 de Octubre is a public district general hospital in the neighbourhood of Orcasur, in Madrid, Spain. It is part of the hospital network of the Servicio Madrileño de Salud (SERMAS). It contains information from the different health domains (laboratory, prescriptions, treatments, administrative, diagnoses, etc.). In addition, information is also obtained from other data sources, such as the pathological anatomy system, which provides information about sample analysis, and the cost system, containing information on the cost associated with a hospital contract.

HARMONY-Acute Myeloid Leukaemia (HARMONY-AML)

The HARMONY Big Data Platform is a secure, innovative research data ecosystem providing a large data lake and state-of-the-art analytic services. The anonymous patient data sets at the Big Data Platform are collected securely following all legal and ethical requirements, harmonised to the OMOP CDM, and then analysed. The data sets are provided by HARMONY Alliance Partners and Associated Members. As of May 2024, around 124,000 records of patients diagnosed with Haematological Malignancies have been transferred to the HARMONY Platform, and 70,000 records are available for analysis to execute the research questions that have been approved by the members of the HARMONY Alliance. The HARMONY Platform includes tools enabling research teams to analyse data at scale. The data scientists curating the Harmony BigData Platform developed ready-to-use models to answer research questions promptly. Interactive dashboards offer easy data analysis by applying statistical templates and pre-defined visualisation tools. On the dashboard, the user can customise the variables to start modelling data to get search results or understand trends, patterns, and insights within the data. On the HARMONY BigData Platform, researchers and data scientists can test their scientific hypotheses or explore the patterns in data. They gain new capabilities to characterise the molecular landscape of blood cancers, understand their pathophysiology, unlock the understanding of disease progression, and identify novel drug targets.

ANNEX II. Fitness for use assessment

Croatian National Public Health Information System (NAJS)

NAJS was included in this study because it contains national cancer registry data from Croatia which collects relevant information on patients with AML and allows for characterisation and treatment patterns assessment in the target population (AML patients) at the same time, as it collects information from other registries. This data source also contributed to survival estimation, as information about death is available. Based on a preliminary feasibility assessment, the number of person counts for AML patients (based on all available data) in NAJS exceeded 3,600. Data availability and follow-up in NAJS are a limitation for this study. Data was available starting from 2017. The last observation period month was January 2025.

Denmark: Danish Data Health Registries (DK-DHR)

DK-DHR was included in this study because it contains national cancer registry data that collects relevant information on patients with AML and allows for characterisation and treatment patterns in the target population (AML patients). This data source further contributed to survival estimation, as information about death is available. Based on a preliminary feasibility assessment, the number of person counts for AML patients (based on all available data) in DK-DHR exceeded 10,200. Data availability and follow-up in DK-DHR are sufficient. Data was available starting from 1995. The last observation period month was November 2024, which slightly deviated from the study period and represented a minor limitation.

Finland: Hospital District of Helsinki and Uusimaa (FinOMOP-HUS)

FinOMOP-HUS was included in the study because it contains information on all patients with AML treated at Helsinki University Hospital, which has a specialised centre for haematological malignancies. It contributed to all the objectives of the study, as it captures information on patients' characteristics, treatment, and procedures. This data source also contributed to survival estimation, as information about death is available. Based on a preliminary feasibility assessment, the number of person counts for AML patients (based on all available data) in FinOMOP-HUS exceeded 1,800. The last observation period month was December 2025, which slightly deviated from the study period and represented a minor limitation.

Finland: Population Registers at the Finnish Institute for Health and Welfare (FinOMOP-THL)

FinOMOP-THL was included in the study because it contains information on all hospitalised patients in Finland, including haematological patients. It contributed to incidence estimation, but information on characterisation and treatment patterns is limited to hospital discharge records. This data source also contributed to survival estimation, as information about death is available. Based on a preliminary feasibility assessment, the number of person counts for AML patients (based on all available data) in FinOMOP-THL exceeded 4,400. The last observation period month was October 2024, which slightly deviated from the study period and represented a minor limitation.

Germany: InGef Research Database (InGef RDB)

InGef RDB was included in the study because it contains claim data on AML patients. There are several limitations in relation to the study. Oncological diagnoses are over-reported in the InGef data source. This assessment is the result of comparing against available German cancer registry reports which show that prevalence and incidence of various cancer types are over-reported. There are also limitations for the characterisation and treatment patterns: InGef is limited in the ability to identify all individual treatments administered in the inpatient setting, as newer and very high-cost active ingredients may be better documented as they are billed separately from other treatments. The procedures, tests, and medications administered during a hospital stay may be bundled into one value for billing purposes and not listed individually, which is a clear limitation. Based on a preliminary feasibility assessment, the number of person counts for AML patients (based on all available data) in InGef exceeded 5,800. The last observation period month was September 2025, which aligned with the study period.

Netherlands: Netherlands Cancer Registry (NCR)

NCR was included in this study because it represents the national cancer registry in the Netherlands. It collects relevant information on patients with AML but allows for the assessment of primary therapy only (i.e. usually only within the first 9 months after diagnosis), which represents a limitation for the study. This data source also contributed to survival estimation, as information about death is available through yearly linkage with the national Personal Records Database. Based on a preliminary feasibility assessment, the number of person counts for AML patients (based on all available data) in NCR exceeded 16,400. The last observation period month for assessing survival was January 2025, which aligned with the study period.

Norway: Cancer Registry Norway (CRN)

CRN was included in this study because it represents the national cancer registry in Norway. It collects relevant information on patients with AML and allows for the assessment of therapy, but not characterisation, which is a limitation for CRN in this study. This data source also contributed to survival estimation, as information about death is available. Based on a preliminary feasibility assessment, the number of person counts for AML patients (based on all available data) in CRN exceeded 9,500. The last observation period month was December 2024, which deviated from the study period and represented a limitation for this data source.

Spain: The Information System for Research in Primary Care (SIDIAP)

SIDIAP was included in the study as it provides detailed information from primary care providers and contributes to incidence estimation. This data source also contributed to survival estimation, as information about death is available. However, there is limited information on treatment patterns for AML patients. Based on a preliminary feasibility assessment, the number of person counts for AML patients (based on all available data) in SIDIAP exceeded 5,300. It also captures information on patients' characteristics, treatment, and procedures. The last observation period month was June 2023, which deviated from the study period and represented a limitation for this data source.

Spain: HARMONY - Acute Myeloid Leukaemia (HARMONY-AML)

HARMONY-AML was included in the study because it encompasses datasets provided by HARMONY Alliance Partners and Associated Members and contains records of patients diagnosed with AML. Despite initial plan, this data source did not contribute to survival estimation, as information about death is not captured consistently in EHR, that represent minor deviation from the protocol. HARMONY is a platform for conducting studies on haematological malignancies and captures information on treatment, diagnostic tests, and procedures. Based on a preliminary feasibility assessment, the number of person counts for AML patients (based on all available data) in HARMONY-AML exceeded 11,800. The last observation period month was June 2022, which deviated from the study period and represented a limitation for this data source.

Spain: Hospital Universitario 12 de Octubre (H12O)

H12O was included in the study because it contains information on all patients with AML treated at Hospital Universitario 12 de Octubre, which has a specialised centre for haematological malignancies. It contributed to all the study objectives, except incidence calculation, as it captures information on patients' characteristics, treatment, and procedures. Based on a preliminary feasibility assessment, the number of person counts for AML patients (based on all available data) in H12O exceeded 7,200. The last observation period month was August 2024, which slightly deviated from the study period and represented a minor limitation.

ANNEX III. Operational and reporting considerations

DATA MANAGEMENT

Data management

All data sources have previously mapped their data to the OMOP common data model. This enabled the use of standardised analytics and using DARWIN EU® tools across the network since the structure of the data and the terminology system was harmonised. The OMOP CDM was developed and maintained by the Observational Health Data Sciences and Informatics (OHDSI) initiative and is described in detail on the wiki page of the CDM: <https://ohdsi.github.io/CommonDataModel> and in The Book of OHDSI: <http://book.ohdsi.org>

The analytic code for this study was developed in R and utilised standardised analytics wherever feasible. Each data partner ran the study code on their data source containing individual data and then returned the results (csv files) which only included aggregated data. The results from each contributing data site were then compiled into tables and figures for the study report.

Data storage and protection

For this study, personal data from individuals in various EU member states were processed, using information collected from national/regional electronic health record data sources. Due to the sensitive nature of this personal medical data, it is important to be fully aware of ethical and regulatory aspects and to strive to take all reasonable measures to ensure compliance with ethical and regulatory issues on privacy.

All data sources used in this study have been used for pharmaco-epidemiological research and have well-established mechanisms to ensure that European and local regulations governing the ethical use of data and adequate privacy controls are adhered to. In agreement with these regulations, rather than combining person-level data and performing only a central analysis, local analyses were run, which generate non-identifiable aggregate summary results.

The output files are stored in the DARWIN EU® Digital Research Environment (DRE). These output files do not contain any data that allows identification of individuals included in the study. The DRE implements additional security measures to ensure a high level of protection for stored data and to comply with the local implementation of the General Data Protection Regulation (GDPR) (EU) 679/2016 in the various member states.

QUALITY CONTROL

Data source quality control

When defining drug cohorts, non-systemic products were excluded from the list of included codes summarised at the ingredient level.

When defining cohorts for indications, the *CohortDiagnostics* (<https://github.com/OHDSI/CohortDiagnostics>) and *DrugExposureDiagnostics* (<https://cran.r-project.org/web/packages/DrugExposureDiagnostics/index.html>) R packages were used to assess the use of different codes across the data sources contributing to the study and to identify any codes possibly omitted in error. The *DrugExposureDiagnostics* package evaluates ingredient-specific attributes and patterns in drug exposure records.

The study code was based on DARWIN EU® R packages: *IncidencePrevalence* to estimate Incidence and Prevalence, *DrugUtilisation* to characterise drug use, and *CohortCharacteristics* to describe the cohort by indication. These packages include numerous automated unit tests to ensure code validity, along with software peer review and user testing. The R package will be made publicly available via GitHub.

ANNEX IV. List of stand-alone documents

Table S1. List of concepts to describe acute myeloid leukaemia (excluding acute promyelocytic leukaemia).

ID	Code	Name	Class	Standard Concept Caption
4289318	69077002	Acute basophilic leukemia	Morph Abnormality	Standard
44499369	9870/3	Acute basophilic leukemia	ICDO Histology	Non-Standard
36402831	9870/3-C42.0	Acute basophilic leukemia of blood	ICDO Condition	Non-Standard
44504703	9870/3-C42.1	Acute basophilic leukemia of bone marrow	ICDO Condition	Non-Standard
36542716	9870/3-C42.4	Acute basophilic leukemia of hematopoietic system, NOS	ICDO Condition	Non-Standard
36558299	9870/3-C77.2	Acute basophilic leukemia of intra-abdominal lymph nodes	ICDO Condition	Standard
36562067	9870/3-C77.1	Acute basophilic leukemia of intrathoracic lymph nodes	ICDO Condition	Standard
36545239	9870/3-C77.9	Acute basophilic leukemia of lymph node, NOS	ICDO Condition	Standard
36520735	9870/3-C77.3	Acute basophilic leukemia of lymph nodes of axilla or arm	ICDO Condition	Standard
36550971	9870/3-C77.0	Acute basophilic leukemia of lymph nodes of head, face and neck	ICDO Condition	Standard
36538220	9870/3-C77.4	Acute basophilic leukemia of lymph nodes of inguinal region or leg	ICDO Condition	Standard
36548148	9870/3-C77.8	Acute basophilic leukemia of lymph nodes of multiple regions	ICDO Condition	Standard
36560511	9870/3-C77.5	Acute basophilic leukemia of pelvic lymph nodes	ICDO Condition	Standard
36538535	9870/3-C42.3	Acute basophilic leukemia of reticuloendothelial system, NOS	ICDO Condition	Non-Standard
36526880	9870/3-C42.2	Acute basophilic leukemia of spleen	ICDO Condition	Standard
44504687	9805/3-C42.1	Acute biphenotypic leukemia of bone marrow	ICDO Condition	Standard

ID	Code	Name	Class	Standard Concept Caption
37151888	1162993004	Acute eosinophilic leukemia	Morph Abnormality	Standard
44499363	9840/3	Acute erythroid leukemia	ICDO Histology	Non-Standard
4031360	14317002	Acute erythroid leukemia	Morph Abnormality	Standard
36402749	9840/3-C42.0	Acute erythroid leukemia of blood	ICDO Condition	Non-Standard
44504697	9840/3-C42.1	Acute erythroid leukemia of bone marrow	ICDO Condition	Non-Standard
36551008	9840/3-C42.4	Acute erythroid leukemia of hematopoietic system, NOS	ICDO Condition	Non-Standard
36519613	9840/3-C77.2	Acute erythroid leukemia of intra-abdominal lymph nodes	ICDO Condition	Standard
36526328	9840/3-C77.1	Acute erythroid leukemia of intrathoracic lymph nodes	ICDO Condition	Standard
36535672	9840/3-C77.9	Acute erythroid leukemia of lymph node, NOS	ICDO Condition	Standard
36545542	9840/3-C77.3	Acute erythroid leukemia of lymph nodes of axilla or arm	ICDO Condition	Standard
36548394	9840/3-C77.0	Acute erythroid leukemia of lymph nodes of head, face and neck	ICDO Condition	Standard
36555620	9840/3-C77.4	Acute erythroid leukemia of lymph nodes of inguinal region or leg	ICDO Condition	Standard
36527919	9840/3-C77.8	Acute erythroid leukemia of lymph nodes of multiple regions	ICDO Condition	Standard
36526460	9840/3-C77.5	Acute erythroid leukemia of pelvic lymph nodes	ICDO Condition	Standard
36553926	9840/3-C42.3	Acute erythroid leukemia of reticuloendothelial system, NOS	ICDO Condition	Non-Standard
36544596	9840/3-C42.2	Acute erythroid leukemia of spleen	ICDO Condition	Standard
36525432	9801/3-C77.9	Acute leukemia, NOS, of lymph node, NOS	ICDO Condition	Standard
44505175	9801/3-C44.9	Acute leukemia, NOS, of skin, NOS	ICDO Condition	Standard

ID	Code	Name	Class	Standard Concept Caption
44499380	9910/3	Acute megakaryoblastic leukemia	ICDO Histology	Non-Standard
4079686	277602003	Acute megakaryoblastic leukemia	Disorder	Standard
4180418	52220008	Acute megakaryoblastic leukemia	Morph Abnormality	Standard
36402838	9910/3-C42.0	Acute megakaryoblastic leukemia of blood	ICDO Condition	Non-Standard
44500837	9910/3-C42.1	Acute megakaryoblastic leukemia of bone marrow	ICDO Condition	Non-Standard
36529557	9910/3-C42.4	Acute megakaryoblastic leukemia of hematopoietic system, NOS	ICDO Condition	Non-Standard
36537884	9910/3-C77.2	Acute megakaryoblastic leukemia of intra-abdominal lymph nodes	ICDO Condition	Standard
36561974	9910/3-C77.1	Acute megakaryoblastic leukemia of intrathoracic lymph nodes	ICDO Condition	Standard
36518868	9910/3-C77.9	Acute megakaryoblastic leukemia of lymph node, NOS	ICDO Condition	Standard
36554471	9910/3-C77.3	Acute megakaryoblastic leukemia of lymph nodes of axilla or arm	ICDO Condition	Standard
36566351	9910/3-C77.0	Acute megakaryoblastic leukemia of lymph nodes of head, face and neck	ICDO Condition	Standard
36543552	9910/3-C77.4	Acute megakaryoblastic leukemia of lymph nodes of inguinal region or leg	ICDO Condition	Standard
36551928	9910/3-C77.8	Acute megakaryoblastic leukemia of lymph nodes of multiple regions	ICDO Condition	Standard
36522090	9910/3-C77.5	Acute megakaryoblastic leukemia of pelvic lymph nodes	ICDO Condition	Standard
36531965	9910/3-C42.3	Acute megakaryoblastic leukemia of reticuloendothelial system, NOS	ICDO Condition	Non-Standard
36560284	9910/3-C42.2	Acute megakaryoblastic leukemia of spleen	ICDO Condition	Standard
135768	413441006	Acute monocytic leukemia	Disorder	Standard
4319285	22331004	Acute monocytic leukemia	Morph Abnormality	Standard

ID	Code	Name	Class	Standard Concept Caption
44499376	9891/3	Acute monocytic leukemia	ICDO Histology	Non-Standard
36402597	9891/3-C42.0	Acute monocytic leukemia of blood	ICDO Condition	Non-Standard
44504709	9891/3-C42.1	Acute monocytic leukemia of bone marrow	ICDO Condition	Non-Standard
36568038	9891/3-C42.4	Acute monocytic leukemia of hematopoietic system, NOS	ICDO Condition	Non-Standard
36558475	9891/3-C77.2	Acute monocytic leukemia of intra-abdominal lymph nodes	ICDO Condition	Standard
36523266	9891/3-C77.1	Acute monocytic leukemia of intrathoracic lymph nodes	ICDO Condition	Standard
36536359	9891/3-C77.9	Acute monocytic leukemia of lymph node, NOS	ICDO Condition	Standard
36540175	9891/3-C77.3	Acute monocytic leukemia of lymph nodes of axilla or arm	ICDO Condition	Standard
36555831	9891/3-C77.0	Acute monocytic leukemia of lymph nodes of head, face and neck	ICDO Condition	Standard
36526209	9891/3-C77.4	Acute monocytic leukemia of lymph nodes of inguinal region or leg	ICDO Condition	Standard
36543033	9891/3-C77.8	Acute monocytic leukemia of lymph nodes of multiple regions	ICDO Condition	Standard
36567966	9891/3-C77.5	Acute monocytic leukemia of pelvic lymph nodes	ICDO Condition	Standard
36563481	9891/3-C42.3	Acute monocytic leukemia of reticuloendothelial system, NOS	ICDO Condition	Non-Standard
36542504	9891/3-C42.2	Acute monocytic leukemia of spleen	ICDO Condition	Standard
42511606	9912/3	Acute myeloid leukemia with BCR-ABL1	ICDO Histology	Standard
37151872	1162928000	Acute myeloid leukemia	Morph Abnormality	Standard
44498890	9911/3	Acute myeloid leukemia (megakaryoblastic) with t(1;22)(p13;q13); RBM15-MKL1	ICDO Histology	Non-Standard
36402815	9911/3-C42.0	Acute myeloid leukemia (megakaryoblastic) with t(1;22)(p13;q13); RBM15-MKL1 of blood	ICDO Condition	Non-Standard

ID	Code	Name	Class	Standard Concept Caption
44501830	9911/3-C42.1	Acute myeloid leukemia (megakaryoblastic) with t(1;22)(p13;q13); RBM15-MKL1 of bone marrow	ICDO Condition	Non-Standard
36517377	9911/3-C42.4	Acute myeloid leukemia (megakaryoblastic) with t(1;22)(p13;q13); RBM15-MKL1 of hematopoietic system, NOS	ICDO Condition	Non-Standard
4029177	128827005	Acute myeloid leukemia myelodysplasia related	Morph Abnormality	Standard
608277	1157157006	Acute myeloid leukemia with 11q23 abnormality	Disorder	Standard
44499370	9871/3	Acute myeloid leukemia with abnormal marrow eosinophils	ICDO Histology	Non-Standard
36402765	9871/3-C42.0	Acute myeloid leukemia with abnormal marrow eosinophils of blood	ICDO Condition	Non-Standard
44504704	9871/3-C42.1	Acute myeloid leukemia with abnormal marrow eosinophils of bone marrow	ICDO Condition	Non-Standard
36550859	9871/3-C42.4	Acute myeloid leukemia with abnormal marrow eosinophils of hematopoietic system, NOS	ICDO Condition	Non-Standard
36535883	9871/3-C77.2	Acute myeloid leukemia with abnormal marrow eosinophils of intra-abdominal lymph nodes	ICDO Condition	Standard
36550889	9871/3-C77.1	Acute myeloid leukemia with abnormal marrow eosinophils of intrathoracic lymph nodes	ICDO Condition	Standard
36532401	9871/3-C77.9	Acute myeloid leukemia with abnormal marrow eosinophils of lymph node, NOS	ICDO Condition	Standard
36532889	9871/3-C77.3	Acute myeloid leukemia with abnormal marrow eosinophils of lymph nodes of axilla or arm	ICDO Condition	Standard
36521596	9871/3-C77.0	Acute myeloid leukemia with abnormal marrow eosinophils of lymph nodes of head, face and neck	ICDO Condition	Standard
36549181	9871/3-C77.4	Acute myeloid leukemia with abnormal marrow eosinophils of lymph nodes of inguinal region or leg	ICDO Condition	Standard
36560043	9871/3-C77.8	Acute myeloid leukemia with abnormal marrow eosinophils of lymph nodes of multiple regions	ICDO Condition	Standard
36545420	9871/3-C77.5	Acute myeloid leukemia with abnormal marrow eosinophils of pelvic lymph nodes	ICDO Condition	Standard
36537061	9871/3-C42.3	Acute myeloid leukemia with abnormal marrow eosinophils of reticuloendothelial system, NOS	ICDO Condition	Non-Standard
36527103	9871/3-C42.2	Acute myeloid leukemia with abnormal marrow eosinophils of spleen	ICDO Condition	Standard

ID	Code	Name	Class	Standard Concept Caption
37165057	1237368006	Acute myeloid leukemia with BCR-ABL1	Disorder	Standard
37204375	783017000	Acute myeloid leukemia with BCR::ABL1 fusion	Morph Abnormality	Standard
37312067	788740009	Acute myeloid leukemia with biallelic CEBPA mutation	Morph Abnormality	Standard
42511604	9878/3	Acute myeloid leukemia with biallelic mutation of CEBPA	ICDO Histology	Standard
4304355	103688009	Acute myeloid leukemia with CBFB::MYH11 fusion	Morph Abnormality	Standard
3654662	838355002	Acute myeloid leukemia with CBFB::MYH11 fusion	Disorder	Standard
45771384	703819004	Acute myeloid leukemia with CEBPA mutation	Morph Abnormality	Standard
35622696	764855007	Acute myeloid leukemia with CEBPA somatic mutations	Disorder	Standard
42872933	450928003	Acute myeloid leukemia with DEK::NUP214 fusion	Morph Abnormality	Standard
44498898	9869/3	Acute myeloid leukemia with inv(3)(q21;q26.2) or t(3;3)(q21;q26.2); RPN1-EVI1	ICDO Histology	Non-Standard
36402820	9869/3-C42.0	Acute myeloid leukemia with inv(3)(q21;q26.2) or t(3;3)(q21;q26.2); RPN1-EVI1 of blood	ICDO Condition	Non-Standard
36551184	9869/3-C42.1	Acute myeloid leukemia with inv(3)(q21;q26.2) or t(3;3)(q21;q26.2); RPN1-EVI1 of bone marrow	ICDO Condition	Non-Standard
36552461	9869/3-C42.4	Acute myeloid leukemia with inv(3)(q21;q26.2) or t(3;3)(q21;q26.2); RPN1-EVI1 of hematopoietic system, NOS	ICDO Condition	Non-Standard
36683269	780844005	Acute myeloid leukemia with inv(3)(q21;q26.2) or t(3;3)(q21;q26.2); RPN1-EVI1	Disorder	Standard
4029663	128829008	Acute myeloid leukemia with KMT2A rearrangement	Morph Abnormality	Standard
4304051	103691009	Acute myeloid leukemia with maturation	Morph Abnormality	Standard
44499373	9874/3	Acute myeloid leukemia with maturation	ICDO Histology	Non-Standard
36402806	9874/3-C42.0	Acute myeloid leukemia with maturation of blood	ICDO Condition	Non-Standard
44504707	9874/3-C42.1	Acute myeloid leukemia with maturation of bone marrow	ICDO Condition	Non-Standard

ID	Code	Name	Class	Standard Concept Caption
36566422	9874/3-C42.4	Acute myeloid leukemia with maturation of hematopoietic system, NOS	ICDO Condition	Non-Standard
36558245	9874/3-C77.2	Acute myeloid leukemia with maturation of intra-abdominal lymph nodes	ICDO Condition	Standard
36529483	9874/3-C77.1	Acute myeloid leukemia with maturation of intrathoracic lymph nodes	ICDO Condition	Standard
36553526	9874/3-C77.9	Acute myeloid leukemia with maturation of lymph node, NOS	ICDO Condition	Standard
36563915	9874/3-C77.3	Acute myeloid leukemia with maturation of lymph nodes of axilla or arm	ICDO Condition	Standard
36535120	9874/3-C77.0	Acute myeloid leukemia with maturation of lymph nodes of head, face and neck	ICDO Condition	Standard
36519450	9874/3-C77.4	Acute myeloid leukemia with maturation of lymph nodes of inguinal region or leg	ICDO Condition	Standard
36552622	9874/3-C77.8	Acute myeloid leukemia with maturation of lymph nodes of multiple regions	ICDO Condition	Standard
36551575	9874/3-C77.5	Acute myeloid leukemia with maturation of pelvic lymph nodes	ICDO Condition	Standard
36528546	9874/3-C42.3	Acute myeloid leukemia with maturation of reticuloendothelial system, NOS	ICDO Condition	Non-Standard
36550331	9874/3-C42.2	Acute myeloid leukemia with maturation of spleen	ICDO Condition	Standard
4234749	359648001	Acute myeloid leukemia with maturation, FAB M2	Disorder	Standard
42872934	450929006	Acute myeloid leukemia with MECOM rearrangement	Morph Abnormality	Standard
4304199	103689001	Acute myeloid leukemia with minimal differentiation	Morph Abnormality	Standard
42511180	9877/3	Acute myeloid leukemia with mutated NPM1	ICDO Histology	Non-Standard
45766616	703820005	Acute myeloid leukemia with mutated NPM1	Morph Abnormality	Standard
42511605	9879/3	Acute myeloid leukemia with mutated RUNX1	ICDO Histology	Standard
40483761	445448008	Acute myeloid leukemia with myelodysplasia-related changes	Disorder	Standard

ID	Code	Name	Class	Standard Concept Caption
44499377	9895/3	Acute myeloid leukemia with myelodysplasia-related changes	ICDO Histology	Non-Standard
36402773	9895/3-C42.0	Acute myeloid leukemia with myelodysplasia-related changes of blood	ICDO Condition	Non-Standard
44504710	9895/3-C42.1	Acute myeloid leukemia with myelodysplasia-related changes of bone marrow	ICDO Condition	Non-Standard
36540730	9895/3-C42.4	Acute myeloid leukemia with myelodysplasia-related changes of hematopoietic system, NOS	ICDO Condition	Non-Standard
36520645	9895/3-C77.2	Acute myeloid leukemia with myelodysplasia-related changes of intra-abdominal lymph nodes	ICDO Condition	Standard
36528799	9895/3-C77.1	Acute myeloid leukemia with myelodysplasia-related changes of intrathoracic lymph nodes	ICDO Condition	Standard
36548112	9895/3-C77.9	Acute myeloid leukemia with myelodysplasia-related changes of lymph node, NOS	ICDO Condition	Standard
36522484	9895/3-C77.3	Acute myeloid leukemia with myelodysplasia-related changes of lymph nodes of axilla or arm	ICDO Condition	Standard
36550752	9895/3-C77.0	Acute myeloid leukemia with myelodysplasia-related changes of lymph nodes of head, face and neck	ICDO Condition	Standard
36523177	9895/3-C77.4	Acute myeloid leukemia with myelodysplasia-related changes of lymph nodes of inguinal region or leg	ICDO Condition	Standard
36555611	9895/3-C77.8	Acute myeloid leukemia with myelodysplasia-related changes of lymph nodes of multiple regions	ICDO Condition	Standard
36556432	9895/3-C77.5	Acute myeloid leukemia with myelodysplasia-related changes of pelvic lymph nodes	ICDO Condition	Standard
36528868	9895/3-C42.3	Acute myeloid leukemia with myelodysplasia-related changes of reticuloendothelial system, NOS	ICDO Condition	Non-Standard
36521194	9895/3-C42.2	Acute myeloid leukemia with myelodysplasia-related changes of spleen	ICDO Condition	Standard
35622003	763309005	Acute myeloid leukemia with NPM1 somatic mutation	Disorder	Standard
37158043	1251642007	Acute myeloid leukemia with NUP98 rearrangement	Morph Abnormality	Standard
42872942	450937003	Acute myeloid leukemia with RBM15::MKL1 fusion	Morph Abnormality	Standard
4028713	128828000	Acute myeloid leukemia with RUNX1::RUNX1T1 fusion	Morph Abnormality	Standard

ID	Code	Name	Class	Standard Concept Caption
37116722	733598001	Acute myeloid leukemia with t(6;9)(p23;q34) translocation	Disorder	Standard
44498920	9865/3	Acute myeloid leukemia with t(6;9)(p23;q34); DEK-NUP214	ICDO Histology	Non-Standard
36402789	9865/3-C42.0	Acute myeloid leukemia with t(6;9)(p23;q34); DEK-NUP214 of blood	ICDO Condition	Non-Standard
44499416	9865/3-C42.1	Acute myeloid leukemia with t(6;9)(p23;q34); DEK-NUP214 of bone marrow	ICDO Condition	Non-Standard
36547990	9865/3-C42.4	Acute myeloid leukemia with t(6;9)(p23;q34); DEK-NUP214 of hematopoietic system, NOS	ICDO Condition	Non-Standard
37110870	725390002	Acute myeloid leukemia with t(8;16)(p11;p13) translocation	Disorder	Standard
37110871	725391003	Acute myeloid leukemia with t(8;16)(p11;p13) translocation	Morph Abnormality	Standard
606958	1148906001	Acute myeloid leukemia with t(8;21)(q22;q22) RUNX1-RUNX1T1	Disorder	Standard
40481524	444911000	Acute myeloid leukemia with t(9;11)(p22;q23); MLLT3-MLL	Disorder	Standard
44499372	9873/3	Acute myeloid leukemia without maturation	ICDO Histology	Non-Standard
4304356	103690005	Acute myeloid leukemia without maturation	Morph Abnormality	Standard
36402808	9873/3-C42.0	Acute myeloid leukemia without maturation of blood	ICDO Condition	Non-Standard
44504706	9873/3-C42.1	Acute myeloid leukemia without maturation of bone marrow	ICDO Condition	Non-Standard
36548182	9873/3-C42.4	Acute myeloid leukemia without maturation of hematopoietic system, NOS	ICDO Condition	Non-Standard
36555336	9873/3-C77.2	Acute myeloid leukemia without maturation of intra-abdominal lymph nodes	ICDO Condition	Standard
36553288	9873/3-C77.1	Acute myeloid leukemia without maturation of intrathoracic lymph nodes	ICDO Condition	Standard
36524587	9873/3-C77.9	Acute myeloid leukemia without maturation of lymph node, NOS	ICDO Condition	Standard
36523246	9873/3-C77.3	Acute myeloid leukemia without maturation of lymph nodes of axilla or arm	ICDO Condition	Standard
36523523	9873/3-C77.0	Acute myeloid leukemia without maturation of lymph nodes of head, face and neck	ICDO Condition	Standard

ID	Code	Name	Class	Standard Concept Caption
36529641	9873/3-C77.4	Acute myeloid leukemia without maturation of lymph nodes of inguinal region or leg	ICDO Condition	Standard
36562529	9873/3-C77.8	Acute myeloid leukemia without maturation of lymph nodes of multiple regions	ICDO Condition	Standard
36525886	9873/3-C77.5	Acute myeloid leukemia without maturation of pelvic lymph nodes	ICDO Condition	Standard
36560973	9873/3-C42.3	Acute myeloid leukemia without maturation of reticuloendothelial system, NOS	ICDO Condition	Non-Standard
36538708	9873/3-C42.2	Acute myeloid leukemia without maturation of spleen	ICDO Condition	Standard
4230253	359640008	Acute myeloid leukemia without maturation, FAB M1	Disorder	Standard
764782	61321000119106	Acute myeloid leukemia without maturation, FAB M1 in remission	Disorder	Standard
44499379	9897/3	Acute myeloid leukemia, 11q23 abnormalities	ICDO Histology	Non-Standard
36402810	9897/3-C42.0	Acute myeloid leukemia, 11q23 abnormalities of blood	ICDO Condition	Non-Standard
44504711	9897/3-C42.1	Acute myeloid leukemia, 11q23 abnormalities of bone marrow	ICDO Condition	Non-Standard
36561343	9897/3-C42.4	Acute myeloid leukemia, 11q23 abnormalities of hematopoietic system, NOS	ICDO Condition	Non-Standard
36522761	9897/3-C77.2	Acute myeloid leukemia, 11q23 abnormalities of intra-abdominal lymph nodes	ICDO Condition	Standard
36551581	9897/3-C77.1	Acute myeloid leukemia, 11q23 abnormalities of intrathoracic lymph nodes	ICDO Condition	Standard
36519835	9897/3-C77.9	Acute myeloid leukemia, 11q23 abnormalities of lymph node, NOS	ICDO Condition	Standard
36536188	9897/3-C77.3	Acute myeloid leukemia, 11q23 abnormalities of lymph nodes of axilla or arm	ICDO Condition	Standard
36532533	9897/3-C77.0	Acute myeloid leukemia, 11q23 abnormalities of lymph nodes of head, face and neck	ICDO Condition	Standard
36525753	9897/3-C77.4	Acute myeloid leukemia, 11q23 abnormalities of lymph nodes of inguinal region or leg	ICDO Condition	Standard
36562801	9897/3-C77.8	Acute myeloid leukemia, 11q23 abnormalities of lymph nodes of multiple regions	ICDO Condition	Standard

ID	Code	Name	Class	Standard Concept Caption
36557716	9897/3-C77.5	Acute myeloid leukemia, 11q23 abnormalities of pelvic lymph nodes	ICDO Condition	Standard
36530056	9897/3-C42.3	Acute myeloid leukemia, 11q23 abnormalities of reticuloendothelial system, NOS	ICDO Condition	Non-Standard
36533252	9897/3-C42.2	Acute myeloid leukemia, 11q23 abnormalities of spleen	ICDO Condition	Standard
140352	91861009	Acute myeloid leukemia, disease	Disorder	Standard
44499371	9872/3	Acute myeloid leukemia, minimal differentiation	ICDO Histology	Non-Standard
36402819	9872/3-C42.0	Acute myeloid leukemia, minimal differentiation of blood	ICDO Condition	Non-Standard
44504705	9872/3-C42.1	Acute myeloid leukemia, minimal differentiation of bone marrow	ICDO Condition	Non-Standard
36539918	9872/3-C42.4	Acute myeloid leukemia, minimal differentiation of hematopoietic system, NOS	ICDO Condition	Non-Standard
36535976	9872/3-C77.2	Acute myeloid leukemia, minimal differentiation of intra-abdominal lymph nodes	ICDO Condition	Standard
36557966	9872/3-C77.1	Acute myeloid leukemia, minimal differentiation of intrathoracic lymph nodes	ICDO Condition	Standard
36546092	9872/3-C77.9	Acute myeloid leukemia, minimal differentiation of lymph node, NOS	ICDO Condition	Standard
36535450	9872/3-C77.3	Acute myeloid leukemia, minimal differentiation of lymph nodes of axilla or arm	ICDO Condition	Standard
36567759	9872/3-C77.0	Acute myeloid leukemia, minimal differentiation of lymph nodes of head, face and neck	ICDO Condition	Standard
36532530	9872/3-C77.4	Acute myeloid leukemia, minimal differentiation of lymph nodes of inguinal region or leg	ICDO Condition	Standard
36564205	9872/3-C77.8	Acute myeloid leukemia, minimal differentiation of lymph nodes of multiple regions	ICDO Condition	Standard
36566508	9872/3-C77.5	Acute myeloid leukemia, minimal differentiation of pelvic lymph nodes	ICDO Condition	Standard
36531772	9872/3-C42.3	Acute myeloid leukemia, minimal differentiation of reticuloendothelial system, NOS	ICDO Condition	Non-Standard
36533846	9872/3-C42.2	Acute myeloid leukemia, minimal differentiation of spleen	ICDO Condition	Standard

ID	Code	Name	Class	Standard Concept Caption
4233531	359631009	Acute myeloid leukemia, minimal differentiation, FAB M0	Disorder	Standard
764781	61311000119104	Acute myeloid leukemia, minimal differentiation, FAB M0 in remission	Disorder	Standard
44499365	9861/3	Acute myeloid leukemia, NOS	ICDO Histology	Non-Standard
36402776	9861/3-C42.0	Acute myeloid leukemia, NOS, of blood	ICDO Condition	Non-Standard
44504699	9861/3-C42.1	Acute myeloid leukemia, NOS, of bone marrow	ICDO Condition	Non-Standard
36549276	9861/3-C42.4	Acute myeloid leukemia, NOS, of hematopoietic system, NOS	ICDO Condition	Non-Standard
36556422	9861/3-C72.9	Acute myeloid leukemia, NOS, of nervous system, NOS	ICDO Condition	Standard
36565437	9861/3-C72.8	Acute myeloid leukemia, NOS, of overlapping lesion of brain and central nervous system	ICDO Condition	Standard
44505178	9861/3-C44.9	Acute myeloid leukemia, NOS, of skin, NOS	ICDO Condition	Standard
44499378	9896/3	Acute myeloid leukemia, t(8;21)(q22;q22)	ICDO Histology	Non-Standard
36402780	9896/3-C42.0	Acute myeloid leukemia, t(8;21)(q22;q22) of blood	ICDO Condition	Standard
44502507	9896/3-C42.1	Acute myeloid leukemia, t(8;21)(q22;q22) of bone marrow	ICDO Condition	Standard
36542780	9896/3-C42.4	Acute myeloid leukemia, t(8;21)(q22;q22) of hematopoietic system, NOS	ICDO Condition	Standard
36522759	9896/3-C77.2	Acute myeloid leukemia, t(8;21)(q22;q22) of intra-abdominal lymph nodes	ICDO Condition	Standard
36521545	9896/3-C77.1	Acute myeloid leukemia, t(8;21)(q22;q22) of intrathoracic lymph nodes	ICDO Condition	Standard
36530037	9896/3-C77.9	Acute myeloid leukemia, t(8;21)(q22;q22) of lymph node, NOS	ICDO Condition	Standard
36521303	9896/3-C77.3	Acute myeloid leukemia, t(8;21)(q22;q22) of lymph nodes of axilla or arm	ICDO Condition	Standard
36564542	9896/3-C77.0	Acute myeloid leukemia, t(8;21)(q22;q22) of lymph nodes of head, face and neck	ICDO Condition	Standard

ID	Code	Name	Class	Standard Concept Caption
36532145	9896/3-C77.4	Acute myeloid leukemia, t(8;21)(q22;q22) of lymph nodes of inguinal region or leg	ICDO Condition	Standard
36522560	9896/3-C77.8	Acute myeloid leukemia, t(8;21)(q22;q22) of lymph nodes of multiple regions	ICDO Condition	Standard
36554900	9896/3-C77.5	Acute myeloid leukemia, t(8;21)(q22;q22) of pelvic lymph nodes	ICDO Condition	Standard
36540255	9896/3-C42.3	Acute myeloid leukemia, t(8;21)(q22;q22) of reticuloendothelial system, NOS	ICDO Condition	Standard
36551864	9896/3-C42.2	Acute myeloid leukemia, t(8;21)(q22;q22) of spleen	ICDO Condition	Standard
4149110	30962008	Acute myelomonocytic leukemia	Morph Abnormality	Standard
4003187	110005000	Acute myelomonocytic leukemia, FAB M4	Disorder	Standard
44499368	9867/3	Acute myelomonocytic leukemia, NOS	ICDO Histology	Non-Standard
36402777	9867/3-C42.0	Acute myelomonocytic leukemia, NOS, of blood	ICDO Condition	Non-Standard
44504702	9867/3-C42.1	Acute myelomonocytic leukemia, NOS, of bone marrow	ICDO Condition	Non-Standard
36565586	9867/3-C42.4	Acute myelomonocytic leukemia, NOS, of hematopoietic system, NOS	ICDO Condition	Non-Standard
36533733	9867/3-C77.2	Acute myelomonocytic leukemia, NOS, of intra-abdominal lymph nodes	ICDO Condition	Standard
36521486	9867/3-C77.1	Acute myelomonocytic leukemia, NOS, of intrathoracic lymph nodes	ICDO Condition	Standard
36561090	9867/3-C77.9	Acute myelomonocytic leukemia, NOS, of lymph node, NOS	ICDO Condition	Standard
36566997	9867/3-C77.3	Acute myelomonocytic leukemia, NOS, of lymph nodes of axilla or arm	ICDO Condition	Standard
36536525	9867/3-C77.0	Acute myelomonocytic leukemia, NOS, of lymph nodes of head, face and neck	ICDO Condition	Standard
36545253	9867/3-C77.4	Acute myelomonocytic leukemia, NOS, of lymph nodes of inguinal region or leg	ICDO Condition	Standard
36561959	9867/3-C77.8	Acute myelomonocytic leukemia, NOS, of lymph nodes of multiple regions	ICDO Condition	Standard

ID	Code	Name	Class	Standard Concept Caption
36560202	9867/3-C77.5	Acute myelomonocytic leukemia, NOS, of pelvic lymph nodes	ICDO Condition	Standard
36558666	9867/3-C42.3	Acute myelomonocytic leukemia, NOS, of reticuloendothelial system, NOS	ICDO Condition	Non-Standard
36531508	9867/3-C42.2	Acute myelomonocytic leukemia, NOS, of spleen	ICDO Condition	Standard
4003184	109991003	Acute panmyelosis with myelofibrosis	Disorder	Standard
44499383	9931/3	Acute panmyelosis with myelofibrosis	ICDO Histology	Non-Standard
36402859	9931/3-C42.0	Acute panmyelosis with myelofibrosis of blood	ICDO Condition	Non-Standard
44504713	9931/3-C42.1	Acute panmyelosis with myelofibrosis of bone marrow	ICDO Condition	Non-Standard
36533001	9931/3-C42.4	Acute panmyelosis with myelofibrosis of hematopoietic system, NOS	ICDO Condition	Non-Standard
36560600	9931/3-C77.2	Acute panmyelosis with myelofibrosis of intra-abdominal lymph nodes	ICDO Condition	Standard
36522665	9931/3-C77.1	Acute panmyelosis with myelofibrosis of intrathoracic lymph nodes	ICDO Condition	Standard
36543239	9931/3-C77.9	Acute panmyelosis with myelofibrosis of lymph node, NOS	ICDO Condition	Standard
36561689	9931/3-C77.3	Acute panmyelosis with myelofibrosis of lymph nodes of axilla or arm	ICDO Condition	Standard
36526333	9931/3-C77.0	Acute panmyelosis with myelofibrosis of lymph nodes of head, face and neck	ICDO Condition	Standard
36552138	9931/3-C77.4	Acute panmyelosis with myelofibrosis of lymph nodes of inguinal region or leg	ICDO Condition	Standard
36536016	9931/3-C77.8	Acute panmyelosis with myelofibrosis of lymph nodes of multiple regions	ICDO Condition	Standard
36552403	9931/3-C77.5	Acute panmyelosis with myelofibrosis of pelvic lymph nodes	ICDO Condition	Standard
36548865	9931/3-C42.3	Acute panmyelosis with myelofibrosis of reticuloendothelial system, NOS	ICDO Condition	Non-Standard
36559852	9931/3-C42.2	Acute panmyelosis with myelofibrosis of spleen	ICDO Condition	Standard

ID	Code	Name	Class	Standard Concept Caption
4002497	110004001	Acute promyelocytic leukemia, FAB M3	Disorder	Standard
4144191	307592006	Basophilic leukemia	Disorder	Standard
45765495	702446006	Core binding factor acute myeloid leukemia	Disorder	Standard
45766268	703387000	Cytogenetically normal acute myeloid leukemia	Disorder	Standard
44784140	700057001	Emberger syndrome	Disorder	Standard
138099	93451002	Erythroleukemia, FAB M6	Disorder	Standard
4301779	404155008	Granulocytic sarcoma affecting skin	Disorder	Standard
35622760	764940002	Inherited acute myeloid leukemia	Disorder	Standard
4173970	277604002	Malignant myeloid/lymphoid neoplasm	Disorder	Standard
35607963	763796007	Megakaryoblastic acute myeloid leukemia with t(1;22)(p13;q13)	Disorder	Standard
313159	188754005	Megakaryocytic leukemia	Disorder	Standard
36541314	9806/3-C42.1	Mixed phenotype acute leukemia with t(9;22)(q34;q11.2); BCR-ABL1 of bone marrow	ICDO Condition	Standard
36521921	9808/3-C42.1	Mixed phenotype acute leukemia, B/myeloid, NOS, of bone marrow	ICDO Condition	Standard
44502151	9809/3-C42.1	Mixed phenotype acute leukemia, T/myeloid, NOS, of bone marrow	ICDO Condition	Standard
42511842	9890/3-C80.9	Monocytic leukemia, NOS, (ICD-O-2) of unknown primary site	ICDO Condition	Standard
44498934	9898/3	Myeloid leukemia associated with Down syndrome	ICDO Histology	Non-Standard
42872940	450935006	Myeloid leukemia associated with Down Syndrome	Morph Abnormality	Standard
37119145	724644005	Myeloid leukemia associated with Down syndrome	Disorder	Standard
36402816	9898/3-C42.0	Myeloid leukemia associated with Down syndrome of blood	ICDO Condition	Standard
36539416	9898/3-C42.1	Myeloid leukemia associated with Down syndrome of bone marrow	ICDO Condition	Standard
36520517	9898/3-C42.4	Myeloid leukemia associated with Down syndrome of hematopoietic system, NOS	ICDO Condition	Standard
140967	94719007	Myeloid sarcoma	Disorder	Standard

ID	Code	Name	Class	Standard Concept Caption
4148718	35287006	Myeloid sarcoma	Morph Abnormality	Standard
44499382	9930/3	Myeloid sarcoma	ICDO Histology	Non-Standard
36529747	9930/3-C31.9	Myeloid sarcoma of accessory sinus, NOS	ICDO Condition	Standard
36539835	9930/3-C72.4	Myeloid sarcoma of acoustic nerve	ICDO Condition	Standard
36529829	9930/3-C02.3	Myeloid sarcoma of anterior 2/3 of tongue, NOS	ICDO Condition	Standard
36560834	9930/3-C04.0	Myeloid sarcoma of anterior floor of mouth	ICDO Condition	Standard
42511746	9930/3-C38.1	Myeloid sarcoma of anterior mediastinum	ICDO Condition	Standard
36549339	9930/3-C10.1	Myeloid sarcoma of anterior surface of epiglottis	ICDO Condition	Standard
36555072	9930/3-C47.9	Myeloid sarcoma of autonomic nervous system, NOS	ICDO Condition	Standard
36538012	9930/3-C01.9	Myeloid sarcoma of base of tongue, NOS	ICDO Condition	Standard
42512828	9930/3-C67.9	Myeloid sarcoma of bladder, NOS	ICDO Condition	Standard
36402876	9930/3-C42.0	Myeloid sarcoma of blood	ICDO Condition	Non-Standard
44504712	9930/3-C42.1	Myeloid sarcoma of bone marrow	ICDO Condition	Standard
36533739	9930/3-C02.1	Myeloid sarcoma of border of tongue	ICDO Condition	Standard
36542498	9930/3-C71.7	Myeloid sarcoma of brain stem	ICDO Condition	Standard
36549513	9930/3-C71.9	Myeloid sarcoma of brain, NOS	ICDO Condition	Standard
36567446	9930/3-C10.4	Myeloid sarcoma of branchial cleft	ICDO Condition	Standard
44502997	9930/3-C50.9	Myeloid sarcoma of breast, NOS	ICDO Condition	Standard

ID	Code	Name	Class	Standard Concept Caption
36538589	9930/3-C72.1	Myeloid sarcoma of cauda equina	ICDO Condition	Standard
36567529	9930/3-C71.6	Myeloid sarcoma of cerebellum, NOS	ICDO Condition	Standard
36525935	9930/3-C70.0	Myeloid sarcoma of cerebral meninges	ICDO Condition	Standard
36534433	9930/3-C71.0	Myeloid sarcoma of cerebrum	ICDO Condition	Standard
36555612	9930/3-C06.0	Myeloid sarcoma of cheek mucosa	ICDO Condition	Standard
36545086	9930/3-C69.3	Myeloid sarcoma of choroid	ICDO Condition	Standard
36534871	9930/3-C69.4	Myeloid sarcoma of ciliary body	ICDO Condition	Standard
36545799	9930/3-C00.6	Myeloid sarcoma of commissure of lip	ICDO Condition	Standard
36522733	9930/3-C69.0	Myeloid sarcoma of conjunctiva	ICDO Condition	Standard
36557752	9930/3-C49.4	Myeloid sarcoma of connective, Subcutaneous and other soft tissues of abdomen	ICDO Condition	Standard
36565598	9930/3-C49.0	Myeloid sarcoma of connective, Subcutaneous and other soft tissues of head, face, and neck	ICDO Condition	Standard
36547281	9930/3-C49.2	Myeloid sarcoma of connective, Subcutaneous and other soft tissues of lower limb and hip	ICDO Condition	Standard
36522587	9930/3-C49.5	Myeloid sarcoma of connective, Subcutaneous and other soft tissues of pelvis	ICDO Condition	Standard
36520660	9930/3-C49.3	Myeloid sarcoma of connective, Subcutaneous and other soft tissues of thorax	ICDO Condition	Standard
36526041	9930/3-C49.6	Myeloid sarcoma of connective, Subcutaneous and other soft tissues of trunk, NOS	ICDO Condition	Standard
36524111	9930/3-C49.1	Myeloid sarcoma of connective, Subcutaneous and other soft tissues of upper limb and shoulder	ICDO Condition	Standard
36565505	9930/3-C49.9	Myeloid sarcoma of connective, Subcutaneous and other soft tissues, NOS	ICDO Condition	Non-Standard
36564092	9930/3-C69.1	Myeloid sarcoma of cornea, NOS	ICDO Condition	Standard

ID	Code	Name	Class	Standard Concept Caption
36518257	9930/3-C72.5	Myeloid sarcoma of cranial nerve, NOS	ICDO Condition	Standard
36525418	9930/3-C02.0	Myeloid sarcoma of dorsal surface of tongue, NOS	ICDO Condition	Standard
36518384	9930/3-C31.1	Myeloid sarcoma of ethmoid sinus	ICDO Condition	Standard
36546647	9930/3-C00.2	Myeloid sarcoma of external lip, NOS	ICDO Condition	Standard
36552752	9930/3-C00.1	Myeloid sarcoma of external lower lip	ICDO Condition	Standard
36537893	9930/3-C00.0	Myeloid sarcoma of external upper lip	ICDO Condition	Standard
36522220	9930/3-C69.9	Myeloid sarcoma of eye, NOS	ICDO Condition	Standard
36542572	9930/3-C44.1	Myeloid sarcoma of eyelid	ICDO Condition	Standard
36518915	9930/3-C04.9	Myeloid sarcoma of floor of mouth, NOS	ICDO Condition	Standard
36532077	9930/3-C71.1	Myeloid sarcoma of frontal lobe	ICDO Condition	Standard
36520327	9930/3-C31.2	Myeloid sarcoma of frontal sinus	ICDO Condition	Standard
36562208	9930/3-C03.9	Myeloid sarcoma of gum, NOS	ICDO Condition	Standard
36543106	9930/3-C05.0	Myeloid sarcoma of hard palate	ICDO Condition	Standard
36520846	9930/3-C42.4	Myeloid sarcoma of hematopoietic system, NOS	ICDO Condition	Non-Standard
36552394	9930/3-C77.2	Myeloid sarcoma of intra-abdominal lymph nodes	ICDO Condition	Standard
36558994	9930/3-C77.1	Myeloid sarcoma of intrathoracic lymph nodes	ICDO Condition	Standard
36544072	9930/3-C69.5	Myeloid sarcoma of lacrimal gland	ICDO Condition	Standard
36545460	9930/3-C04.1	Myeloid sarcoma of lateral floor of mouth	ICDO Condition	Standard

ID	Code	Name	Class	Standard Concept Caption
36563002	9930/3-C10.2	Myeloid sarcoma of lateral wall of oropharynx	ICDO Condition	Standard
36537470	9930/3-C02.4	Myeloid sarcoma of lingual tonsil	ICDO Condition	Standard
36522441	9930/3-C00.9	Myeloid sarcoma of lip, NOS	ICDO Condition	Standard
44499619	9930/3-C22.0	Myeloid sarcoma of liver	ICDO Condition	Standard
36555175	9930/3-C03.1	Myeloid sarcoma of lower gum	ICDO Condition	Standard
36549345	9930/3-C77.9	Myeloid sarcoma of lymph node, NOS	ICDO Condition	Standard
36545021	9930/3-C77.3	Myeloid sarcoma of lymph nodes of axilla or arm	ICDO Condition	Standard
36530999	9930/3-C77.0	Myeloid sarcoma of lymph nodes of head, face and neck	ICDO Condition	Standard
36543985	9930/3-C77.4	Myeloid sarcoma of lymph nodes of inguinal region or leg	ICDO Condition	Standard
36531614	9930/3-C77.8	Myeloid sarcoma of lymph nodes of multiple regions	ICDO Condition	Standard
36555173	9930/3-C31.0	Myeloid sarcoma of maxillary sinus	ICDO Condition	Standard
44502508	9930/3-C38.3	Myeloid sarcoma of mediastinum, NOS	ICDO Condition	Standard
36528575	9930/3-C70.9	Myeloid sarcoma of meninges, NOS	ICDO Condition	Standard
36557627	9930/3-C06.9	Myeloid sarcoma of mouth, NOS	ICDO Condition	Standard
36526088	9930/3-C00.5	Myeloid sarcoma of mucosa of lip, NOS	ICDO Condition	Standard
36555753	9930/3-C00.4	Myeloid sarcoma of mucosa of lower lip	ICDO Condition	Standard
36525511	9930/3-C00.3	Myeloid sarcoma of mucosa of upper lip	ICDO Condition	Standard
36537899	9930/3-C30.0	Myeloid sarcoma of nasal cavity	ICDO Condition	Standard

ID	Code	Name	Class	Standard Concept Caption
36537589	9930/3-C72.9	Myeloid sarcoma of nervous system, NOS	ICDO Condition	Standard
36547836	9930/3-C71.4	Myeloid sarcoma of occipital lobe	ICDO Condition	Standard
36555084	9930/3-C72.2	Myeloid sarcoma of olfactory nerve	ICDO Condition	Standard
36518764	9930/3-C72.3	Myeloid sarcoma of optic nerve	ICDO Condition	Standard
36559030	9930/3-C69.6	Myeloid sarcoma of orbit, NOS	ICDO Condition	Standard
36563611	9930/3-C10.9	Myeloid sarcoma of oropharynx, NOS	ICDO Condition	Standard
42511905	9930/3-C56.9	Myeloid sarcoma of ovary	ICDO Condition	Standard
36556910	9930/3-C31.8	Myeloid sarcoma of overlapping lesion of accessory sinuses	ICDO Condition	Standard
36528834	9930/3-C71.8	Myeloid sarcoma of overlapping lesion of brain	ICDO Condition	Standard
36547062	9930/3-C72.8	Myeloid sarcoma of overlapping lesion of brain and central nervous system	ICDO Condition	Standard
42512321	9930/3-C50.8	Myeloid sarcoma of overlapping lesion of breast	ICDO Condition	Standard
36520696	9930/3-C49.8	Myeloid sarcoma of overlapping lesion of connective, subcutaneous and other soft tissues	ICDO Condition	Standard
36522695	9930/3-C69.8	Myeloid sarcoma of overlapping lesion of eye and adnexa	ICDO Condition	Standard
36533428	9930/3-C04.8	Myeloid sarcoma of overlapping lesion of floor of mouth	ICDO Condition	Standard
36539514	9930/3-C00.8	Myeloid sarcoma of overlapping lesion of lip	ICDO Condition	Standard
36566649	9930/3-C14.8	Myeloid sarcoma of overlapping lesion of lip, oral cavity and pharynx	ICDO Condition	Standard
36546905	9930/3-C06.8	Myeloid sarcoma of overlapping lesion of other and unspecified parts of mouth	ICDO Condition	Standard
36517400	9930/3-C05.8	Myeloid sarcoma of overlapping lesion of palate	ICDO Condition	Standard

ID	Code	Name	Class	Standard Concept Caption
36550579	9930/3-C47.8	Myeloid sarcoma of overlapping lesion of peripheral nerves and autonomic nervous system	ICDO Condition	Standard
36538941	9930/3-C48.8	Myeloid sarcoma of overlapping lesion of retroperitoneum and peritoneum	ICDO Condition	Standard
36563790	9930/3-C02.8	Myeloid sarcoma of overlapping lesion of tongue	ICDO Condition	Standard
36523395	9930/3-C09.8	Myeloid sarcoma of overlapping lesion of tonsil	ICDO Condition	Standard
36530524	9930/3-C10.8	Myeloid sarcoma of overlapping lesions of oropharynx	ICDO Condition	Standard
36526573	9930/3-C05.9	Myeloid sarcoma of palate, NOS	ICDO Condition	Standard
42512990	9930/3-C25.9	Myeloid sarcoma of pancreas, NOS	ICDO Condition	Standard
36543025	9930/3-C71.3	Myeloid sarcoma of parietal lobe	ICDO Condition	Standard
42512673	9930/3-C41.4	Myeloid sarcoma of pelvic bones, sacrum, coccyx and associated joints	ICDO Condition	Standard
36561875	9930/3-C77.5	Myeloid sarcoma of pelvic lymph nodes	ICDO Condition	Standard
42511854	9930/3-C76.3	Myeloid sarcoma of pelvis, NOS	ICDO Condition	Standard
36535017	9930/3-C47.4	Myeloid sarcoma of peripheral nerves and autonomic nervous system of abdomen	ICDO Condition	Standard
36535690	9930/3-C47.0	Myeloid sarcoma of peripheral nerves and autonomic nervous system of head, face, and neck	ICDO Condition	Standard
36549195	9930/3-C47.2	Myeloid sarcoma of peripheral nerves and autonomic nervous system of lower limb and hip	ICDO Condition	Standard
36554495	9930/3-C47.5	Myeloid sarcoma of peripheral nerves and autonomic nervous system of pelvis	ICDO Condition	Standard
36539706	9930/3-C47.3	Myeloid sarcoma of peripheral nerves and autonomic nervous system of thorax	ICDO Condition	Standard
36532657	9930/3-C47.6	Myeloid sarcoma of peripheral nerves and autonomic nervous system of trunk, NOS	ICDO Condition	Standard
36567812	9930/3-C47.1	Myeloid sarcoma of peripheral nerves and autonomic nervous system of upper limb and shoulder	ICDO Condition	Standard

ID	Code	Name	Class	Standard Concept Caption
36551688	9930/3-C48.2	Myeloid sarcoma of peritoneum, NOS	ICDO Condition	Standard
36561727	9930/3-C14.0	Myeloid sarcoma of pharynx, NOS	ICDO Condition	Standard
42512372	9930/3-C75.1	Myeloid sarcoma of pituitary gland	ICDO Condition	Standard
36527512	9930/3-C10.3	Myeloid sarcoma of posterior wall of oropharynx	ICDO Condition	Standard
36524798	9930/3-C42.3	Myeloid sarcoma of reticuloendothelial system, NOS	ICDO Condition	Non-Standard
36553431	9930/3-C69.2	Myeloid sarcoma of retina	ICDO Condition	Standard
36528195	9930/3-C06.2	Myeloid sarcoma of retromolar area	ICDO Condition	Standard
36530213	9930/3-C48.0	Myeloid sarcoma of retroperitoneum	ICDO Condition	Standard
44500838	9930/3-C44.3	Myeloid sarcoma of skin of other and unspecified parts of face	ICDO Condition	Standard
44501874	9930/3-C44.4	Myeloid sarcoma of skin of scalp and neck	ICDO Condition	Standard
44505241	9930/3-C44.9	Myeloid sarcoma of skin, NOS	ICDO Condition	Non-Standard
42511686	9930/3-C17.9	Myeloid sarcoma of small intestine, NOS	ICDO Condition	Standard
36528022	9930/3-C05.1	Myeloid sarcoma of soft palate, NOS	ICDO Condition	Standard
36544680	9930/3-C48.1	Myeloid sarcoma of specified parts of peritoneum	ICDO Condition	Standard
36545102	9930/3-C31.3	Myeloid sarcoma of sphenoid sinus	ICDO Condition	Standard
36567164	9930/3-C72.0	Myeloid sarcoma of spinal cord	ICDO Condition	Standard
36530975	9930/3-C70.1	Myeloid sarcoma of spinal meninges	ICDO Condition	Standard
36557182	9930/3-C42.2	Myeloid sarcoma of spleen	ICDO Condition	Standard

ID	Code	Name	Class	Standard Concept Caption
36547951	9930/3-C71.2	Myeloid sarcoma of temporal lobe	ICDO Condition	Standard
36556956	9930/3-C37.9	Myeloid sarcoma of thymus	ICDO Condition	Standard
36552226	9930/3-C02.9	Myeloid sarcoma of tongue, NOS	ICDO Condition	Standard
36561775	9930/3-C09.9	Myeloid sarcoma of tonsil, NOS	ICDO Condition	Standard
36557942	9930/3-C09.0	Myeloid sarcoma of tonsillar fossa	ICDO Condition	Standard
36536563	9930/3-C09.1	Myeloid sarcoma of tonsillar pillar	ICDO Condition	Standard
36560565	9930/3-C03.0	Myeloid sarcoma of upper Gum	ICDO Condition	Standard
36541843	9930/3-C05.2	Myeloid sarcoma of uvula	ICDO Condition	Standard
42512685	9930/3-C52.9	Myeloid sarcoma of vagina, NOS	ICDO Condition	Standard
36556658	9930/3-C10.0	Myeloid sarcoma of vallecule	ICDO Condition	Standard
36534606	9930/3-C02.2	Myeloid sarcoma of ventral surface of tongue, NOS	ICDO Condition	Standard
36535453	9930/3-C71.5	Myeloid sarcoma of ventricle, NOS	ICDO Condition	Standard
36519211	9930/3-C06.1	Myeloid sarcoma of vestibule of mouth	ICDO Condition	Standard
36520924	9930/3-C14.2	Myeloid sarcoma of waldeyer ring	ICDO Condition	Standard
42514010	9870/3-NUL	Neoplasm defined only by histology: Acute basophilic leukemia	ICDO Condition	Standard
42514002	9840/3-NUL	Neoplasm defined only by histology: Acute erythroid leukemia	ICDO Condition	Standard
42513977	9801/3-NUL	Neoplasm defined only by histology: Acute leukemia, NOS	ICDO Condition	Standard
42514026	9910/3-NUL	Neoplasm defined only by histology: Acute megakaryoblastic leukemia	ICDO Condition	Standard

ID	Code	Name	Class	Standard Concept Caption
42514020	9891/3-NULL	Neoplasm defined only by histology: Acute monocytic leukemia	ICDO Condition	Standard
42514028	9912/3-NULL	Neoplasm defined only by histology: Acute myeloid leukemia with BCR-ABL1	ICDO Condition	Standard
42514027	9911/3-NULL	Neoplasm defined only by histology: Acute myeloid leukemia (megakaryoblastic) with t(1;22)(p13;q13); RBM15-MKL1	ICDO Condition	Standard
42514011	9871/3-NULL	Neoplasm defined only by histology: Acute myeloid leukemia with abnormal marrow eosinophils	ICDO Condition	Standard
42514018	9878/3-NULL	Neoplasm defined only by histology: Acute myeloid leukemia with biallelic mutation of CEBPA	ICDO Condition	Standard
42514009	9869/3-NULL	Neoplasm defined only by histology: Acute myeloid leukemia with inv(3)(q21;q26.2) or t(3;3)(q21;q26.2); RPN1-EVI1	ICDO Condition	Standard
42514014	9874/3-NULL	Neoplasm defined only by histology: Acute myeloid leukemia with maturation	ICDO Condition	Standard
42514017	9877/3-NULL	Neoplasm defined only by histology: Acute myeloid leukemia with mutated NPM1	ICDO Condition	Standard
42514019	9879/3-NULL	Neoplasm defined only by histology: Acute myeloid leukemia with mutated RUNX1	ICDO Condition	Standard
42514021	9895/3-NULL	Neoplasm defined only by histology: Acute myeloid leukemia with myelodysplasia-related changes	ICDO Condition	Standard
42514006	9865/3-NULL	Neoplasm defined only by histology: Acute myeloid leukemia with t(6;9)(p23;q34); DEK-NUP214	ICDO Condition	Standard
42514013	9873/3-NULL	Neoplasm defined only by histology: Acute myeloid leukemia without maturation	ICDO Condition	Standard
42514023	9897/3-NULL	Neoplasm defined only by histology: Acute myeloid leukemia, 11q23 abnormalities	ICDO Condition	Standard
42514012	9872/3-NULL	Neoplasm defined only by histology: Acute myeloid leukemia, minimal differentiation	ICDO Condition	Standard
42514004	9861/3-NULL	Neoplasm defined only by histology: Acute myeloid leukemia, NOS	ICDO Condition	Standard
42514022	9896/3-NULL	Neoplasm defined only by histology: Acute myeloid leukemia, t(8;21)(q22;q22)	ICDO Condition	Standard
42514008	9867/3-NULL	Neoplasm defined only by histology: Acute myelomonocytic leukemia, NOS	ICDO Condition	Standard
42514031	9931/3-NULL	Neoplasm defined only by histology: Acute panmyelosis with myelofibrosis	ICDO Condition	Standard

ID	Code	Name	Class	Standard Concept Caption
42514340	9890/3-NUL	Neoplasm defined only by histology: Monocytic leukemia, NOS (ICD-O-2)	ICDO Condition	Standard
42514025	9898/3-NUL	Neoplasm defined only by histology: Myeloid leukemia associated with Down syndrome	ICDO Condition	Standard
42514030	9930/3-NUL	Neoplasm defined only by histology: Myeloid sarcoma	ICDO Condition	Non-Standard
42514056	9987/3-NUL	Neoplasm defined only by histology: Therapy-related myelodysplastic syndrome, NOS	ICDO Condition	Standard
42514029	9920/3-NUL	Neoplasm defined only by histology: Therapy-related myeloid neoplasm	ICDO Condition	Standard
36717461	721306009	Therapy related acute myeloid leukemia and myelodysplastic syndrome	Disorder	Standard
44499402	9987/3	Therapy-related myelodysplastic syndrome, NOS	ICDO Histology	Non-Standard
36402837	9987/3-C42.0	Therapy-related myelodysplastic syndrome, NOS, of blood	ICDO Condition	Non-Standard
44504727	9987/3-C42.1	Therapy-related myelodysplastic syndrome, NOS, of bone marrow	ICDO Condition	Non-Standard
36536073	9987/3-C42.4	Therapy-related myelodysplastic syndrome, NOS, of hematopoietic system, NOS	ICDO Condition	Non-Standard
44499381	9920/3	Therapy-related myeloid neoplasm	ICDO Histology	Non-Standard
36402846	9920/3-C42.0	Therapy-related myeloid neoplasm of blood	ICDO Condition	Non-Standard
44500395	9920/3-C42.1	Therapy-related myeloid neoplasm of bone marrow	ICDO Condition	Non-Standard
36567072	9920/3-C42.4	Therapy-related myeloid neoplasm of hematopoietic system, NOS	ICDO Condition	Non-Standard
4082485	277601005	Acute monoblastic leukemia	Disorder	Non-Standard
4189938	413442004	Acute monocytic/monoblastic leukemia	Disorder	Non-Standard
36715587	721305008	Acute myeloid leukemia due to recurrent genetic abnormality	Disorder	Non-Standard
4264447	397341000	Acute myeloid leukemia with multilineage dysplasia following a myelodysplastic syndrome or myelodysplastic syndrome/myeloproliferative disorder	Morph Abnormality	Non-Standard

ID	Code	Name	Class	Standard Concept Caption
4265011	397340004	Acute myeloid leukemia with recurrent genetic abnormality	Morph Abnormality	Non-Standard
4116880	285839005	Acute myelomonocytic leukemia - eosinophilic variant	Disorder	Non-Standard
4213196	80570006	Acute panmyelosis with myelofibrosis	Morph Abnormality	Non-Standard
44807009	815361000000107	Acute myeloid leukaemia with 11q23 abnormality	Disorder	Non-Standard
4184848	413443009	Acute myeloid leukemia	Morph Abnormality	Non-Standard
4073533	17788007	Acute myeloid leukemia, no ICD-O subtype	Morph Abnormality	Non-Standard
42511184	9890/3	Monocytic leukemia, NOS (ICD-O-2)	ICDO Histology	Non-Standard
44505521	9861/3-C44	Acute myeloid leukemia of skin	ICDO Condition	Non-Standard
44505522	9930/3-C44	Myeloid sarcoma of skin	ICDO Condition	Non-Standard

ICDO=International Classification of Diseases for Oncology

Table S2. List of concepts to describe acute promyelocytic leukaemia.

ID	Code	Name	Class	Standard Concept Caption
4112803	285769009	Acute promyelocytic leukemia - hypogranular variant	Disorder	Standard
4103718	28950004	Acute promyelocytic leukemia with PML::RARA fusion	Morph Abnormality	Standard
4002497	110004001	Acute promyelocytic leukemia, FAB M3	Disorder	Standard
4137687	425869007	Acute promyelocytic leukemia, FAB M3, in remission	Disorder	Standard
44499367	9866/3	Acute promyelocytic leukemia, t(15;17)(q22;q11-12)	ICDO Histology	Non-Standard
36402849	9866/3-C42.0	Acute promyelocytic leukemia, t(15;17)(q22;q11-12) of blood	ICDO Condition	Non-Standard
44504701	9866/3-C42.1	Acute promyelocytic leukemia, t(15;17)(q22;q11-12) of bone marrow	ICDO Condition	Non-Standard
36555269	9866/3-C42.4	Acute promyelocytic leukemia, t(15;17)(q22;q11-12) of hematopoietic system, NOS	ICDO Condition	Non-Standard

ID	Code	Name	Class	Standard Concept Caption
36567989	9866/3-C77.2	Acute promyelocytic leukemia, t(15;17)(q22;q11-12) of intra-abdominal lymph nodes	ICDO Condition	Standard
36529189	9866/3-C77.1	Acute promyelocytic leukemia, t(15;17)(q22;q11-12) of intrathoracic lymph nodes	ICDO Condition	Standard
36539568	9866/3-C77.9	Acute promyelocytic leukemia, t(15;17)(q22;q11-12) of lymph node, NOS	ICDO Condition	Standard
36523650	9866/3-C77.3	Acute promyelocytic leukemia, t(15;17)(q22;q11-12) of lymph nodes of axilla or arm	ICDO Condition	Standard
36530188	9866/3-C77.0	Acute promyelocytic leukemia, t(15;17)(q22;q11-12) of lymph nodes of head, face and neck	ICDO Condition	Standard
36521967	9866/3-C77.4	Acute promyelocytic leukemia, t(15;17)(q22;q11-12) of lymph nodes of inguinal region or leg	ICDO Condition	Standard
36536345	9866/3-C77.8	Acute promyelocytic leukemia, t(15;17)(q22;q11-12) of lymph nodes of multiple regions	ICDO Condition	Standard
36522823	9866/3-C77.5	Acute promyelocytic leukemia, t(15;17)(q22;q11-12) of pelvic lymph nodes	ICDO Condition	Standard
36529731	9866/3-C42.3	Acute promyelocytic leukemia, t(15;17)(q22;q11-12) of reticuloendothelial system, NOS	ICDO Condition	Non-Standard
36522662	9866/3-C42.2	Acute promyelocytic leukemia, t(15;17)(q22;q11-12) of spleen	ICDO Condition	Standard
36676614	773537001	Differentiation syndrome due to and following chemotherapy co-occurrent with acute promyelocytic leukemia	Disorder	Standard
4175688	278189009	Hypergranular promyelocytic leukemia	Disorder	Standard
42514007	9866/3-NULL	Neoplasm defined only by histology: Acute promyelocytic leukemia, t(15;17)(q22;q11-12)	ICDO Condition	Standard

ICDO=International Classification of Diseases for Oncology

Table S3. List of concepts to describe medicines.

ID	Code	Name	Class	Standard Concept Caption
1333379	18330	arsenic trioxide	Ingredient	Standard
1314865	1251	azacitidine	Ingredient	Standard
1333357	1828	busulfan	Ingredient	Standard
1310317	3002	cyclophosphamide	Ingredient	Standard
1311078	3041	cytarabine	Ingredient	Standard
1311799	3109	daunorubicin	Ingredient	Standard
1338512	3639	doxorubicin	Ingredient	Standard
793797	1940332	enasidenib	Ingredient	Standard
1350504	4179	etoposide	Ingredient	Standard
1395557	24698	fludarabine	Ingredient	Standard
42799711	1294580	gemtuzumab ozogamicin	Ingredient	Standard
1355513	2105806	gilteritinib	Ingredient	Standard
1355795	2105845	glasdegib	Ingredient	Standard
19078097	5650	idarubicin	Ingredient	Standard
1560123	2049873	ivosidenib	Ingredient	Standard
1301267	6718	melphalan	Ingredient	Standard
1594003	1919083	midostaurin	Ingredient	Standard
1309188	7005	mitoxantrone	Ingredient	Standard
742050	2623641	olutasidenib	Ingredient	Standard
37499576	2359268	pemigatinib	Ingredient	Standard
746225	2643048	quizartinib	Ingredient	Standard
1437379	10485	thioguanine	Ingredient	Standard
19042545	38508	treosulfan	Ingredient	Standard
903643	10753	tretinoin	Ingredient	Standard
35604205	1747556	venetoclax	Ingredient	Standard
1308290	11202	vincristine	Ingredient	Standard

Table S4. List of concepts to describe procedures.

ID	Code	Name	Class	Standard Concept Caption	Domain	Vocabulary
4120445	234336002	Hemopoietic stem cell transplant	Procedure	Standard	Procedure	SNOMED
4081380	278257006	Peripheral blood stem cell graft	Procedure	Standard	Procedure	SNOMED
4028623	23719005	Transplantation of bone marrow	Procedure	Standard	Procedure	SNOMED
4144157	425983008	Autologous peripheral blood stem cell transplant	Procedure	Standard	Procedure	SNOMED
4242257	58390007	Allogeneic bone marrow transplantation	Procedure	Standard	Procedure	SNOMED
4143404	425843001	Allogeneic peripheral blood stem cell transplant	Procedure	Standard	Procedure	SNOMED
4125488	234335003	Allogeneic unrelated bone marrow transplant	Procedure	Standard	Procedure	SNOMED
4125487	234334004	Allogeneic related bone marrow transplant	Procedure	Standard	Procedure	SNOMED
46271079	709115004	Transplantation of autologous hematopoietic stem cell	Procedure	Standard	Procedure	SNOMED
4240337	58776007	Autologous bone marrow transplant	Procedure	Standard	Procedure	SNOMED
4144882	427423003	Allograft of bone marrow from matched unrelated donor	Procedure	Standard	Procedure	SNOMED
4142405	426425001	Allograft of bone marrow from sibling donor	Procedure	Standard	Procedure	SNOMED
4059885	19944001	Autologous bone marrow transplant without purging	Procedure	Standard	Procedure	SNOMED
40486968	446253009	Allogeneic bone marrow transplantation with purging	Procedure	Standard	Procedure	SNOMED
40492289	445757003	Allogeneic bone marrow transplantation without purging	Procedure	Standard	Procedure	SNOMED
44790154	228311000000103	Allograft of bone marrow from haploidentical donor	Procedure	Standard	Procedure	SNOMED
4186582	46280001	Autologous bone marrow transplant with purging	Procedure	Standard	Procedure	SNOMED
44793170	378841000000102	Allograft of bone marrow from unmatched unrelated donor	Procedure	Standard	Procedure	SNOMED
4145532	427615008	Allograft of cord blood to bone marrow	Procedure	Standard	Procedure	SNOMED
4083057	278256002	Cord cell transfusion	Procedure	Standard	Procedure	SNOMED

ID	Code	Name	Class	Standard Concept Caption	Domain	Vocabulary
40484034	442557006	Grafting of bone marrow using allograft from unmatched unrelated donor	Procedure	Standard	Procedure	SNOMED
4139690	425563003	Grafting of cord blood to bone marrow	Procedure	Standard	Procedure	SNOMED
37152106	1172516002	High-dose chemotherapy with stem cell transplant	Procedure	Standard	Procedure	SNOMED
4125486	234333005	Imperfect T-cell depleted allogeneic bone marrow graft	Procedure	Standard	Procedure	SNOMED
4121104	234331007	Syngeneic bone marrow transplant	Procedure	Standard	Procedure	SNOMED
44783964	698075004	Syngeneic peripheral blood stem cell transplantation	Procedure	Standard	Procedure	SNOMED
4122920	234332000	T-cell depleted allogeneic bone marrow graft	Procedure	Standard	Procedure	SNOMED
1243851	174241000112100	Transplantation of allogeneic hematopoietic stem cell	Procedure	Standard	Procedure	SNOMED
46272967	711429001	Transplantation of autologous progenitor cell	Procedure	Standard	Procedure	SNOMED

Table S5. List of concepts to describe measurement (blood).

ID	Code	Name	Class	Standard Concept Caption	Vocabulary
37208514	1106091000000103	Basophil count in blood	Observable Entity	Standard	SNOMED
37399328	1022601000000101	Blast cell count	Observable Entity	Standard	SNOMED
4212803	391426006	Bone marrow aspirate examination	Procedure	Standard	SNOMED
4043761	167907006	Bone marrow examination	Procedure	Standard	SNOMED
37208633	1107391000000104	Eosinophil count in blood	Observable Entity	Standard	SNOMED
37393848	1022441000000101	FBC - full blood count	Observable Entity	Standard	SNOMED
4094758	250220000	Hemoglobin finding	Clinical Finding	Standard	SNOMED
4098519	250644007	LDH blood measurement	Procedure	Standard	SNOMED
37394163	1027691000000105	Leucocyte count	Observable Entity	Standard	SNOMED
4148615	30630007	Neutrophil count	Procedure	Standard	SNOMED
37208699	1108071000000101	Neutrophil count in blood	Observable Entity	Standard	SNOMED

ID	Code	Name	Class	Standard Concept Caption	Vocabulary
37399327	1022461000000100	Nucleated red blood cell count	Observable Entity	Standard	SNOMED
37398606	1015501000000103	Percentage basophils	Observable Entity	Standard	SNOMED
37399255	1015561000000104	Percentage eosinophils	Observable Entity	Standard	SNOMED
37398607	1015541000000100	Percentage granulocytes	Observable Entity	Standard	SNOMED
37398605	1015471000000105	Percentage neutrophils	Observable Entity	Standard	SNOMED
37393325	1015551000000102	Percentage nucleated red blood cells	Observable Entity	Standard	SNOMED
37393631	1019541000000102	Percentage reticulocyte count	Observable Entity	Standard	SNOMED
4267147	61928009	Platelet count	Procedure	Standard	SNOMED
37208696	1108041000000107	Platelet count in blood	Observable Entity	Standard	SNOMED

Table S6. List of concepts to describe WHO/ECOG performance status.

ID	Code	Name	Class	Standard Concept Caption
37174738	1230349005	Assessment using WHO performance status scale	Procedure	Non-Standard
37538797	C102116	ECOG	Observable Entity	Non-Standard
4308014	423740007	ECOG performance status	Observable Entity	Non-Standard
4175026	425389002	ECOG performance status - grade 0	Clinical Finding	Standard
4173614	422512005	ECOG performance status - grade 1	Clinical Finding	Standard
4172043	422894000	ECOG performance status - grade 2	Clinical Finding	Standard
4174241	423053003	ECOG performance status - grade 3	Clinical Finding	Standard
4174251	423237006	ECOG performance status - grade 4	Clinical Finding	Standard
4173456	423409001	ECOG performance status - grade 5	Clinical Finding	Standard
36303470	89262-0	ECOG Performance Status [Interpretation]	Clinical Observation	Standard
4308794	424122007	ECOG performance status finding	Clinical Finding	Non-Standard
1030477	LP72504-1	ECOG performance status grade	LOINC Component	Non-Standard
36305299	89246-3	ECOG Performance Status panel	Clinical Observation	Standard

ID	Code	Name	Class	Standard Concept Caption
36305384	89247-1	ECOG Performance Status score	Clinical Observation	Standard
4167763	273437007	ECOG scale for physical assessment	Staging / Scales	Standard
4162589	373802001	WHO performance status finding	Clinical Finding	Standard
4190931	373803006	WHO performance status grade 0	Clinical Finding	Standard
45464997	13CJ.00	WHO performance status grade 0	Read	Non-Standard
45438222	13CK.00	WHO performance status grade 1	Read	Non-Standard
4161577	373804000	WHO performance status grade 1	Clinical Finding	Standard
4161578	373805004	WHO performance status grade 2	Clinical Finding	Standard
45454796	13CL.00	WHO performance status grade 2	Read	Non-Standard
45514942	13CM.00	WHO performance status grade 3	Read	Non-Standard
4162590	373806003	WHO performance status grade 3	Clinical Finding	Standard
4161579	373807007	WHO performance status grade 4	Clinical Finding	Standard
45428297	13CN.00	WHO performance status grade 4	Read	Non-Standard
4162588	373801008	WHO performance status scale	Staging / Scales	Standard
37157581	1230348002	WHO performance status score	Observable Entity	Non-Standard

WHO/ECOG=World Health Organisation/Eastern Cooperative Oncology Group

Table S7. List of concepts to describe genetic mutations.

ID	Code	Name	Class	Standard Concept Caption
37312067	788740009	Acute myeloid leukemia with biallelic CEBPA mutation	Morph Abnormality	Standard
42511604	9878/3	Acute myeloid leukemia with biallelic mutation of CEBPA	ICDO Histology	Standard
45771384	703819004	Acute myeloid leukemia with CEBPA mutation	Morph Abnormality	Standard
35622696	764855007	Acute myeloid leukemia with CEBPA somatic mutations	Disorder	Standard
778218	AMLCEBPA	AML with Biallelic Mutations of CEBPA (AMLCEBPA)	Condition	Non-Standard
35948202	1833	CEBPA (CCAAT enhancer binding protein alpha) gene variant measurement	Genetic Variation	Standard

ID	Code	Name	Class	Standard Concept Caption
40778879	LP115339-6	CEBPA gene	LOINC Component	Non-Standard
37052855	LP227934-9	CEBPA gene full mutation analysis	LOINC Component	Non-Standard
44816923	75032-3	CEBPA gene full mutation analysis in Bone marrow by Sequencing	Lab Test	Standard
40766729	64012-8	CEBPA gene mutations found [Identifier] in Blood or Tissue by Molecular genetics method Nominal	Lab Test	Standard
37044509	LP227935-6	CEBPA gene targeted mutation analysis	LOINC Component	Non-Standard
3173762	911752951000004103	CEBPA sequence variant identified in excised malignant neoplasm	Observable Entity	Standard
35965641	33749	CEBPA:D63N	Variant	Non-Standard
35975078	33746	CEBPA:R288P	Variant	Non-Standard
36011491	CEBPA:Truncating Mutations	CEBPA:Truncating Mutations	Variant	Non-Standard
35952780	18674	DDX41 (DEAD-box helicase 41) gene variant measurement	Genetic Variation	Standard
19566863	OMOP5478787	DDX41 protein: Substitution in position 396 of I replaced by T measurement	Gene Protein Variant	Standard
19598455	OMOP5511266	DDX41 protein: Substitution in position 525 of R replaced by H measurement	Gene Protein Variant	Standard
37160954	1172625004	DDX41-related hematologic malignancy predisposition syndrome	Disorder	Standard
35967129	31204	DDX41:I396T	Variant	Non-Standard
35975314	31192	DDX41:R525H	Variant	Non-Standard
42542255	1891	FLT3 Acute myeloid leukemia	Condition	Non-Standard
37049191	LP228449-7	FLT3 gene internal tandem duplication	LOINC Component	Non-Standard
37058355	LP310331-6	FLT3 gene p.Asp835 mutations	LOINC Component	Non-Standard
37025196	LP228450-5	FLT3 gene targeted mutation analysis	LOINC Component	Non-Standard

ID	Code	Name	Class	Standard Concept Caption
36012809	FLT3:D835A	FLT3:D835A	Variant	Non-Standard
36012810	FLT3:D835del	FLT3:D835del	Variant	Non-Standard
36012811	FLT3:D835E	FLT3:D835E	Variant	Non-Standard
36012812	FLT3:D835H	FLT3:D835H	Variant	Non-Standard
36012813	FLT3:D835N	FLT3:D835N	Variant	Non-Standard
36012814	FLT3:D835V	FLT3:D835V	Variant	Non-Standard
36012815	FLT3:D835Y	FLT3:D835Y	Variant	Non-Standard
36012816	FLT3:D839G	FLT3:D839G	Variant	Non-Standard
36012826	FLT3:Internal tandem duplication	FLT3:Internal tandem duplication	Variant	Non-Standard
36012838	FLT3:Tyrosine kinase domain mutation	FLT3:Tyrosine kinase domain mutation	Variant	Non-Standard
37116868	733888001	Glioblastoma isocitrate dehydrogenase 1 mutation	Morph Abnormality	Standard
725078	81120	IDH1 (isocitrate dehydrogenase 1 [NADP+], soluble) (eg, glioma), common variants (eg, R132H, R132C)	CPT4	Standard
1260674	LP438656-3	IDH1 and IDH2 gene	LOINC Component	Non-Standard
36034349	LP420044-2	IDH1 and IDH2 genes targeted mutation analysis	LOINC Component	Non-Standard
19643639	OMOP5499694	IDH1 exon 4 mutation measurement	Gene DNA Variant	Standard
36034117	LP265705-6	IDH1 gene	LOINC Component	Non-Standard
1762366	LP432138-8	IDH1 gene exon 4	LOINC Component	Non-Standard
1992609	LP431494-6	IDH1 gene targeted mutation analysis	LOINC Component	Non-Standard
3168003	911752301000004101	IDH1 sequence variant identified in excised malignant neoplasm	Observable Entity	Standard
36013313	IDH1:A134D	IDH1:A134D	Variant	Non-Standard
36013315	IDH1:G123R	IDH1:G123R	Variant	Non-Standard

ID	Code	Name	Class	Standard Concept Caption
36013316	IDH1:G70D	IDH1:G70D	Variant	Non-Standard
36013317	IDH1:G97D	IDH1:G97D	Variant	Non-Standard
36013319	IDH1:I130M	IDH1:I130M	Variant	Non-Standard
36013323	IDH1:R132C	IDH1:R132C	Variant	Non-Standard
36013324	IDH1:R132G	IDH1:R132G	Variant	Non-Standard
36013325	IDH1:R132H	IDH1:R132H	Variant	Non-Standard
35974725	1231	IDH1:R132L	Variant	Non-Standard
36013326	IDH1:R132Q	IDH1:R132Q	Variant	Non-Standard
36013327	IDH1:R132S	IDH1:R132S	Variant	Non-Standard
35974728	7070	IDH1:R132V	Variant	Non-Standard
36013328	IDH1:R49C	IDH1:R49C	Variant	Non-Standard
36013329	IDH1:S280F	IDH1:S280F	Variant	Non-Standard
36013331	IDH1:Y139D	IDH1:Y139D	Variant	Non-Standard
725079	81121	IDH2 (isocitrate dehydrogenase 2 [NADP+], mitochondrial) (eg, glioma), common variants (eg, R140W, R172M)	CPT4	Standard
19643640	OMOP5390250	IDH2 exon 4 mutation measurement	Gene DNA Variant	Standard
36034462	LP265706-4	IDH2 gene	LOINC Component	Non-Standard
1762307	LP432139-6	IDH2 gene exon 4	LOINC Component	Non-Standard
1762217	LP432136-2	IDH2 gene exon 4 targeted mutation analysis	LOINC Component	Non-Standard
1992610	LP431495-3	IDH2 gene targeted mutation analysis	LOINC Component	Non-Standard
3170900	911752311000004103	IDH2 sequence variant identified in excised malignant neoplasm	Observable Entity	Standard
36013332	IDH2:I319M	IDH2:I319M	Variant	Non-Standard
36013333	IDH2:Q316E	IDH2:Q316E	Variant	Non-Standard

ID	Code	Name	Class	Standard Concept Caption
35974750	23343	IDH2:R140G	Variant	Non-Standard
35974751	23276	IDH2:R140K	Variant	Non-Standard
35974752	6286	IDH2:R140L	Variant	Non-Standard
36013334	IDH2:R140Q	IDH2:R140Q	Variant	Non-Standard
35974757	6287	IDH2:R140W	Variant	Non-Standard
36013335	IDH2:R172G	IDH2:R172G	Variant	Non-Standard
36013336	IDH2:R172K	IDH2:R172K	Variant	Non-Standard
36013337	IDH2:R172M	IDH2:R172M	Variant	Non-Standard
36013338	IDH2:R172S	IDH2:R172S	Variant	Non-Standard
35974840	6406	IDH2:R172W	Variant	Non-Standard
35975177	6407	IDH2:R377C	Variant	Non-Standard
35976187	14071	IDH2:T435M	Variant	Non-Standard
35973148	226416	NM_016222.4(DDX41):c.1187T>C (p.Ile396Thr)	Variant	Non-Standard
35976965	OMOP4997530	CEBPA Gene Mutation measurement	Gene DNA Variant	Non-Standard
35976985	OMOP4997550	FLT3 Gene Mutation measurement	Gene DNA Variant	Non-Standard
35976993	OMOP4997558	IDH1 Gene Mutation measurement	Gene DNA Variant	Non-Standard
35976994	OMOP4997559	IDH2 Gene Mutation measurement	Gene DNA Variant	Non-Standard

ANNEX V. Glossary

Additional definitions are available in the EMA Glossary of terms <https://www.ema.europa.eu/en/about-us/glossaries>.

Aggregated Data

Data collected and combined from multiple sources to generate summary information, typically anonymised.

Benefit-Risk Assessment

Evaluation of the positive therapeutic effects of a medicine compared to its risks (e.g. side effects).

Common Data Model (CDM)

A standardised data structure that enables data from multiple sources to be harmonised, making analysis consistent and reproducible. DARWIN EU® utilises the OMOP CDM maintained by the OHDSI community.

Complex Studies (C3)

Studies requiring the development or customisation of specific study designs, protocols, and Statistical Analysis Plans (SAPs), with extensive collection or extraction of data. Examples include etiological studies measuring the strength and determinants of an association between an exposure and the occurrence of a health outcome in a defined population considering sources of bias, potential confounding factors, and effect modifiers.

Coordination Centre (CC)

The central hub responsible for managing and overseeing the activities within DARWIN EU®. It is based at Erasmus University Medical Center in Rotterdam, the Netherlands.

Data Access

The process of obtaining permission to use specific datasets for regulatory or scientific studies.

Data Quality Framework

A set of standards and procedures to ensure accuracy, completeness, timeliness, and consistency of data used in DARWIN EU®.

Data Source

A database or repository of structured health-related data, such as electronic health records (EHRs), insurance claims, or registries.

DARWIN EU®

The European Medicines Agency's (EMA) federated network of real-world data sources designed to generate evidence to support regulatory decision-making.

EMA (European Medicines Agency)

The regulatory body responsible for the evaluation and supervision of medicinal products in the EU, overseeing DARWIN EU®.

Evidence Generation

The process of analysing real-world data to produce scientific information that can inform healthcare or regulatory decisions.

Federated Network

A data infrastructure where data remain at their original location but can be analysed in a harmonised way across multiple partners using a common model and tools.

GDPR (General Data Protection Regulation)

The EU regulation governing the protection of personal data and privacy, crucial to how DARWIN EU® handles health data.

Health Technology Assessment (HTA)

A systematic evaluation of properties and impacts of health technology, often using DARWIN EU® data to support assessments.

Metadata

Descriptive information about a data source (e.g. its content, quality, and structure), essential for identifying relevant databases in DARWIN EU® studies.

Off-the-Shelf Studies (OTS)

Studies for which a standard protocol per study/analysis type and standardised analytics may be developed and applied or adapted, typically relating to a descriptive research question. This includes studies on disease epidemiology, for example, the estimation of the prevalence or incidence of health outcomes in defined time periods and population groups, or drug utilisation studies at the population or patient level.

OHDSI (Observational Health Data Sciences and Informatics)

An open-science collaborative community that develops tools and standards (including the OMOP CDM) to enable large-scale analytics of observational health data. OHDSI provides the technical and scientific foundation for DARWIN EU®'s analytical ecosystem.

Patient-Level Data

Data related to individuals, de-identified, used for longitudinal or detailed analyses.

OMOP (Observational Medical Outcomes Partnership)

A common data model (CDM) that standardises the structure and content of observational healthcare data, enabling systematic analysis across disparate datasets. DARWIN EU® uses the OMOP CDM to ensure interoperability and consistency in real-world evidence generation.

Real-World Data (RWD)

Data relating to individual health status or healthcare delivery that is collected from routine clinical practice rather than from randomised controlled trials.

Real-World Evidence (RWE)

Clinical evidence derived from the analysis of RWD, used to inform decisions by regulators, payers, or clinicians.

Regulatory Decision-Making

The process by which authorities like EMA assess data to authorise, monitor, or modify the use of medicines in the EU.

Routine Repeated Studies (RR)

Studies that are either Off-the-Shelf or Complex studies repeated on a regular basis, following the same protocol and study code, but with updated data and/or different data partners.

Study Protocol

A detailed plan describing how a specific real-world study will be conducted, including objectives, design, data sources, and analyses.

Very Complex Studies (C4)

Studies which cannot rely only on electronic health care databases, or which would require complex methodological work, for example, due to the occurrence of events that cannot be defined by existing diagnosis codes, including events that do not yet have a diagnosis code, where it may be necessary to combine a diagnosis code with other data such as results of laboratory investigations. These studies might require the collection of data prospectively, or the inclusion of new (not previously onboarded) data sources.