UiO : Department of Biosciences
University of Oslo

A Study of *Ex Vivo* Drug Screen Survival Predictions in AML Patients for Clinical Application

Andrea M. Brodersen

2020

Abstract

Acute Myeloid Leukemia (AML) is a heterogeneous malignancy involving the clonal expansion of myeloid progenitor cells (blasts) in the bone marrow and peripheral blood. Most AML patients are treated with intensive chemotherapy, but despite initially high complete response rates, many patients relapse and die from their disease. New methods are required to better stratify patients in different treatment groups. A model system that can potentially identify personalized treatments from both novel and approved drugs is Ex vivo drug screens. However, little is known about the level of the predictivity of drug screen outcomes for clinical efficacy, and the preservation of cancer dependencies to *ex vivo* cultures. In this study, we used L2 regularized Cox regression to analyze *ex vivo* drug profiling data of 349 drugs form an *ex vivo* drug screen covering 55 AML patients who were subjected to standard 7+3 chemotherapy to predict patient survival.

Strikingly, while the WHO-AML classification system showed some predictive value for the prognosis of AML patients (out of currently available clinical prediction methods), ex vivo drug screens proved to be superior at predicting survival compared to mutations, karyotype, age, gender and WHO class combined. Surprisingly, we also see that the doctor's prognosis based on the ELN-NET system did not have predictive values. This suggests a potential use of *ex vivo* drug sensitivity data in cancer prognostics and forecasting. Additionally, analyzing the regression coefficients (log-hazard ratios) we could determine the direct association between drug sensitivities in patient samples and survival outcomes. As a proof of principle, *ex vivo* response to the standard treatment drug, Daunorubicin, was associated with favourable prognosis. This reveals a direct linear relation between response to a drug *ex vivo* and probability of survival in the clinic. We also discovered 33 drugs that were significantly associated with survival; we believe that some of these drugs have treatment potential as alternatives to the standard treatment or as second-line treatments. Furthermore, drug target data was projected using UMAP in order to reveal a functional enrichment of the protein network targeted by the 33 identified drugs. Based on the drug-target network we found a clear separation of high-risk and low-risk clusters, which we believe can provide valuable insight into the underlying mechanisms of cellular dependencies that drive AML cancer progression. entailing possibilities for drug development. Finally, we have discovered a negative. In conclusion, drug sensitivity screening is a superior method for predicting AML patient survival, and we envision that *ex vivo* drug screening can be implemented in the clinic as a tool for clinical decision-making and patient stratification.

Acknowledgements

I have had the pleasure of conducting my master's project at the Department of Molecular Cell Biology at the Institute of Cancer Research at OUS (Radium Hospital). This study would not have been possible without the help of our collaborators at the Department of Haematology at Oslo University Hospital led by head physician Yngvar Fløisand, PhD.

Firstly, I express my deepest gratitude to my main supervisor Aram Nikolai Andersen for learning under your guidance. I thank you for your incredible patience, persistence and interest in teaching me computation from the bottom. I am very grateful for your willingness to help me with all difficulties along the way and answering my endless questions, regardless of their complexity. Your belief in me has made an everlasting impression and substantially strengthened my academic confidence, I am forever in gratitude.

Secondly, I would like to express my gratitude to my co-supervisor, Professor Jorrit Enserink, for allowing me the opportunity to work in your group and your continued belief and interest in this project. I am forever grateful for your guidance and feedback.

Further, I would like to thank my co-supervisor Pilar Ayuda Durán, PhD, for originally taking me under your wing as a master's student, as well as your open-mindedness in regards to the progress of the project. This project would not have been possible without your hard work on the *ex vivo* drug screen. I would also like to extend my gratitude to Ignacio Llorente García, PhD, for your continuous help and your good advice. Further also wish to thank the rest of the Enserink lab for advice, support and feedback during this time.

Aside from my academic supervisors I also wish to express my deepest gratitude to the patients who have been a part of this study. I hope, and believe, your participation and selflessness can bring good for those diagnosed with acute myeloid leukemia in the years to come. I also wish to thank the representatives of the Norwegian Blood Cancer Association for your personal cancer stories and your words of encouragement regarding the project. It also made an immense impact on me as a reminder that the data I work with are actual people and not just numbers.

I wish to thank my friends for their constant support and care during this time. I especially wish to thank Linda Håkensbakken and Jeanne Corrales for the joy and support you have provided in the lab. Your guidance during the entire project has meant a lot to me. And to Hannah Brodin, my flat-mate, for her continuous support and providing an eager ear during the last six months.

I would like to thank my family for their continuous support, encouragement and interest regarding this project. My dog, Indi, also deserves an acknowledgement for giving me an excuse to go out and clear my head during stressing times. Finally, my loving parents, who have believed in me throughout this process. Your backing throughout my academic endeavours will be forever appreciated. I also appreciate all the effort and care you have put in to making this experience as comfortable as possible.

Andrea M. Brodersen, December 2020

Contents

List of Figures

List of Tables

CHAPTER 1

Introduction

1.1 Background

1.1.1 Cancer and cancer diversity

Cancer is one of the most researched diseases in our time as it affects so many people, yet there is room for massive advancements to be made, especially in terms of treatment. As many as one in two men and one in three women will develop cancer during their lifetime. Additionally, one in five, unrelated to gender, will die from cancer [1]. As so many people are affected by cancer, finding new and more effective treatments is a central topic in biological sciences and medicine. Media often talk about "the cure of cancer" but the picture is much more complicated than what the phrase entails. Cancer is used as a bag term for a large number of diseases. The National Cancer Institute (NCI) defines cancer as: "A term for diseases in which abnormal cells divide without control and can invade nearby tissues. Cancer cells can also spread to other parts of the body through the blood and lymph systems" [2].

What makes cancer such a diverse group of diseases is its variability, not only amongst the various cancer types but also within specific cancers or even a single tumour. At its core, the various cancers are named after the tissue the cancer cells resemble and organ from which the cancer originates [3]. Additionally, cancers are very heterogeneous even within specific cancer types as cancer development is driven by random mutations. Cells with mutations in oncogenes or tumour suppressors will be subjected to natural selection to make the cell, and its clones, grow and adapt to the environment. This selection process can result in cancers with similar-looking phenotypes, but with different underlying driving mechanisms. Furthermore, these genetic differences can, in turn, give rise to variation in transcription, epigenetic regulation, further mutation burden and so on.

1.1.2 Targeted therapy and personalized medicine

As cancer is such a complex disease the variability both within and between cancer types causes an enormous challenge in terms of treatment. Finding treatments that work across the spectrum of individuality is a hurdle to overcome. Treatment sensitivity may vary greatly between individual patients. Additionally, within-tumour heterogeneity may result in one specific treatment not being enough to kill all cancer cells and treatment-resistant variants may be selected for.

Over the last few decades, patients diagnosed with the same type of cancer usually receive a standard treatment. Many standard regimens entail radiation therapy, surgery or administration of unspecific and cytotoxic chemotherapies that kill both cancer cells and healthy tissue. The use of standard treatments is problematic for several reasons. Firstly, because the current way of diagnosing cancers does not take patient individuality into account, treatment efficacy is highly unpredictable. Secondly, standard treatments using chemotherapeutics can often result in severe side effects.

The problems with the standard treatment regimen have shifted the focus of patient treatment towards a more tailored system. Tailored treatment systems entail targeted therapy, which aims to selectively inhibit the mechanisms that drive a specific cancer type. A form of tailored treatment selection is personalized medicine which is based on finding relevant markers and deciding on

treatment according to said markers for individual patients [4]. Finding markers is done through an extensive diagnosis using screens to identify mutations and chromosomal abnormalities, and sometimes immunophenotyping or transcriptional profiling [5]. Consequently, the treatment will consist of highly specific drugs predicted to work for the individual patient.

An example of targeted therapy can be found in breast cancer where a mutation screen can identify a mutation in the BRCA genes (1 and 2) which is involved in DNA-repair. If BRCA is mutated, homologous recombination is impaired, severely increasing the risk of breast cancer [6]. A common treatment option is PARP inhibitors are common. These disable double stranded break repair, another means of DNA repair. The combination of BRCA mutation and PARP inhibitors with radiation-induced DNA damage will result in the death of cancer cells as they cannot repair themselves, whereas the healthy cells with functioning BRCA can survive through homologous recombination [7]. Another example of targeted therapy can be found in chronic myelogenous leukemia where patients with the Philadelphia chromosome's key mutation, BCR-ABL, can be treated using Imatinib. BCR-ABL is a fusion protein caused by recombination between chromosomes 9 and 22, resulting in a constitutively active tyrosine kinase that leads to increased proliferation, genetic instability and disruption of the cell cycle. Imatinib targets the tyrosine kinase domain in ABL, but also other tyrosine kinases such as c-kit and PDGF-R resulting in reduced tyrosine kinase activity [8].

Despite the fact that targeted therapy has been revolutionary for treatment of some cancers, there are still challenges related to the heterogeneity between individuals in addition to the development of resistance to targeted therapy. Furthermore, finding effective treatments predicted from specific single-markers for most cancers has not been trivial [9]. A more realistic strategy has been to collect larger amounts of data for each patient in order to get a more complete profile of the individual patient that can help predict the appropriate treatment option.

1.1.3 Drug screens

Shifting the current treatment system towards personalized medicine is both tedious and expensive. Nevertheless, more cost and time-efficient methods exist. One such method is high-throughput drug screens, where the need for intermediate biomarkers is bypassed by going straight to testing treatment options on acquired patient cells. How the patient cells react to the drugs (drug responses) will then help make informed decisions on what optional treatments are suitable for individual patients.

As tactics for personalized medicine are becoming more sophisticated and efficient, so are the methods for analyzing them. Machine learning is becoming more widespread in biological medicine. Machine learning a term describing the use of computers to detect patterns in data based on complex mathematical and statistical methods [10]. Machine learning is becoming a powerful tool, in cancer research particularly, as it can aid in areas like patient stratification and increasing accuracy for treatment prediction is showing promise for making great advances in personalized medicine

1.2 Acute Myeloid Leukemia

Acute Myeloid Leukemia (AML) is a blood cancer affecting the myeloid lineage of cells marked in orange in figure 1.1. The cancerous cells will limit the production of normal cells, causing malfunctioning granulocytes and a lack of platelets and red blood cells [11, 12]. AML occurs in both children and adults. However, the average onset of the disease is at 70 years of age in Norway. Additionally, the average incidence of AML in Norway from 2011-2015 was 129 people per year with a five-year survival rate ranging from 90% to 10% depending on underlying conditions, age and severity of the disease [11]. AML usually arises as a *de novo* disease spontaneously, but can also occur as a consequence of previous anti-cancer therapy [13].

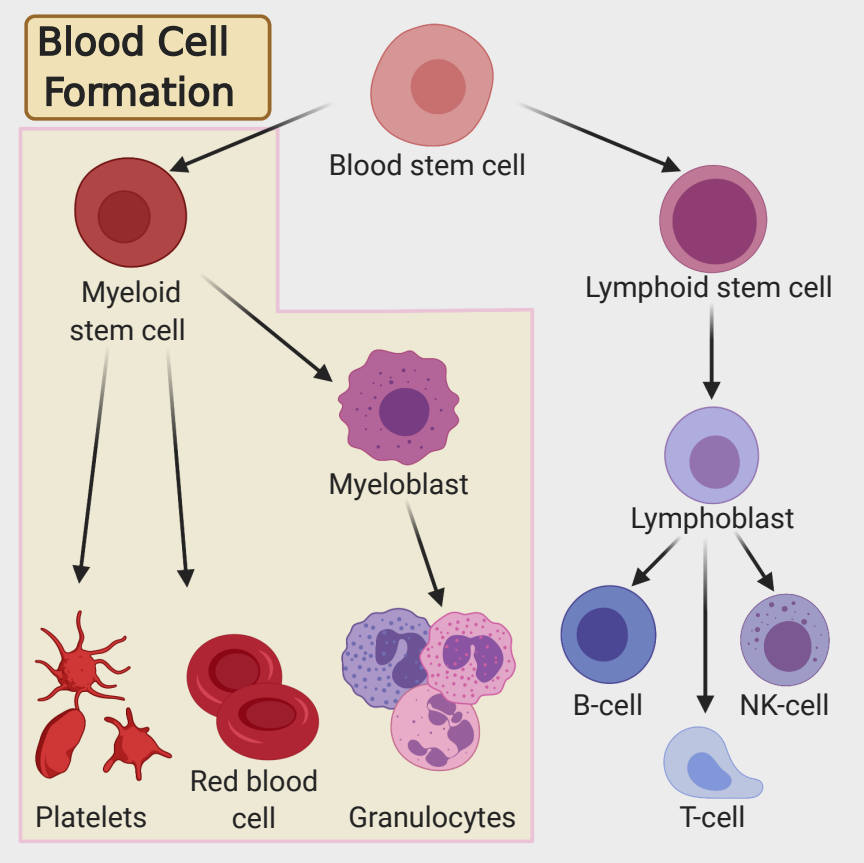


Figure 1.1: **AML development:** An overview of the development of blood cells from a blood stem cell with the cells affected in Acute Myeloid Leukemia marked in the orange frame. Created with BioRender.

AML is highly associated with large chromosomal rearrangements or aberrations. Apart from the karyotypic abnormalities, AML is a highly heterogeneous disease genetically, with a large number of mutations identified. Only a few mutations are found in more than 25% of AML patients, including nucleophosmin (NPM1), Fms-related tyrosine kinase 3 (FLT3), and DNA methyltransferase 3A (DNMT3A) [14]. Table 1.1 shows an overview of key genetic abnormalities found in AML and their incidence [15]. This overview illustrates the vast heterogeneity and complexity of AML.

Table 1.1: **Heterogeneity of AML:** An overview of key genetic subgroups detected in AML, and their incidence, illustrating the complex heterogeneity of AML. The table is based on a study conducted by Papaemmunuli et al in 2016 containing 1540 patients [15].

AML genetic subtype	**Incidence**
NPM1 mutation	27%
Mutated chromatin, genes involved in RNA splicing or both	18%
TP53 mutations, chromosomal aneuploidy or both	13%
Driver mutations but no other class-defining lesions (FLT3, DNMT3A)	11%
Inversion(16)(p13.1q22) or t(16;16)(p13.1;q22) resulting in CBFB-MYH11	5%
biallelic CBPA mutations	4%
t(15;17)(q22;q12); PML-RARA	4%
t(8;21)(q22;q22) RUNX1-RUNX1T1	4%
MLL fusion genes t(x;11)(x;q23)	3%
inv(3)(q21q26.2) or t(3;3)(q21;q26.2); GATA2, MECOM (EVI1)	1%
IDH2 mutations	1%
t(6;9)(p23;q34); DEK-NUP214	1%
no detectable driver mutations	4%

1.3 Treatment of AML

Although certain systems have been implemented to stratify AML patients into risk groups, the treatment still remains largely the same. The common practice in Norway is providing a standard treatment to all patients regardless of their risk group or underlying genetic background. The standard treatment in Norway is a combination treatment regimen consisting of a 100-200 mg/m^2 Cytarabine intravenously around the clock for 7 days followed by a three-day treatment of 60mg/m^2 Daunorubicin (see figure 1.2). This treatment regimen is repeated over six months with a three to four-week rest in between treatments [16].

Cytarabine is an analogue to cytosine and acts as an antineoplastic agent by inhibiting DNA synthesis [17]. Daunorubicin is an antibiotic and cytotoxic agent that inhibits DNA replication, DNA repair, RNA synthesis and protein synthesis [18]. As both agents used in the standard treatment of AML have a very broad mechanism of action, they can also be considered toxic agents as they target both cancerous and healthy cells.

There has also been a practice of additional treatment of bone marrow transplantation. The bone marrow transplantation can be either heterologous (from another person, preferably unrelated) or autologous (healthy hematopoetic stem cells are purified from the patient themselves). The bone marrow transplant can only take place if the patient is under 70 years of age and does not have any underlying health condition that might affect the treatment, and is still associated with high mortality [19]. Both treatments also require close monitoring as well as supportive treatments of supplementary red blood cells and platelets in addition to treatment of infections [19].

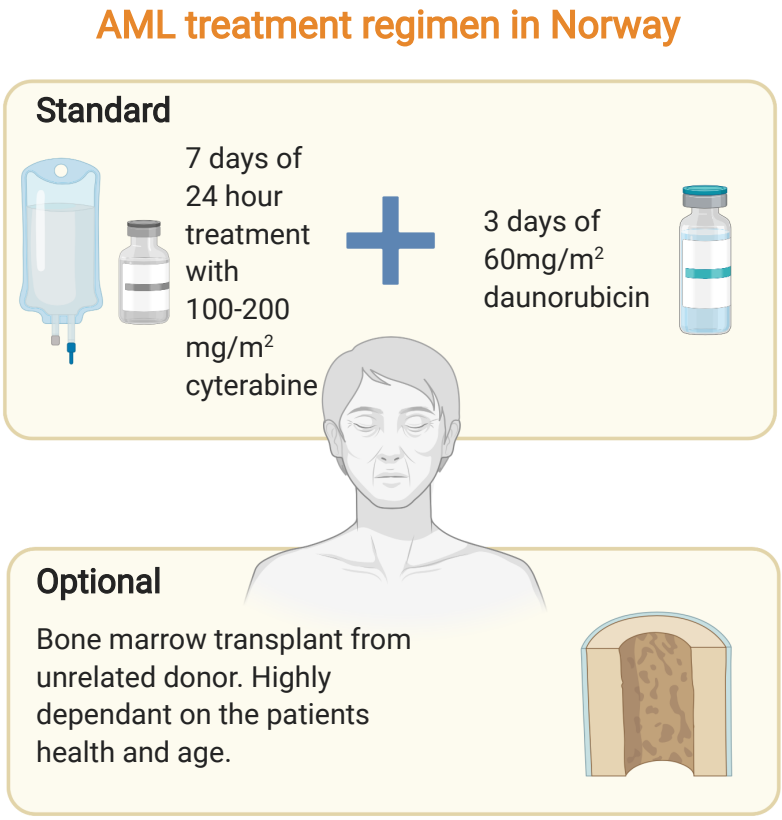


Figure 1.2: **Treatment of AML:** A simplified overview of the treatment regimen of AML in Norway. The standard treatment is repeated over 6 months in four-week intervals. Created with BioRender.

A special subgroup of AML, acute promyelocytic leukemia (APL), recieves an alternative treatment to AML. APL patients recieve All-trans-retinoic acid (ATRA, tretinoin, or Vesanoid) or arsenic trioxide (ATO, Trisenox). APL needs a different therapy, as the blast cells affected by APL contain proteins that when released into the blood stream cause severe clotting, and is thereby not suited for cytotoxic chemotherapy [20].

In cases of relapse post treatment there have been some experimental use of second line drugs like the Bcl-2 inhibitors Navitoclax and Venetoclax. Both drugs can work in some cases, but have severe side effects including neutropenia (low white blood cell count) and thrombocytopenia (low platelet count). Navitoclax is closely related with thrombocytopenia and is thereby dose restricted, Venetoclax on the other hand is showing more promise but is not commonly used in the clinic for AML [21]. There are also clinical trials using targeted therapy for specific genetic markers like Gemtuzumab ozogamicin for mutations in CD33, Ivosidenib for IDH1 gene mutation and Enasidenib if the AML cells have an IDH2 gene mutation [22]. Furthermore, clinical trials have been conducted in Norway using immunotherapy [23]. Regardless of the second line and alternative treatments, most AML patients still receive the standard treatment of Cyterabine and Daunorubicine.

1.4 Problems Associated With Standard Treatment

With the standard treatment mentioned above the average five year survival of AML patients is roughly 50% [11]. Nevertheless, most patients are put on the standard treatment regimen in hopes that it will cause them to go into remission. Regardless, the standard combination therapy has a long list of side affects including[24]:

1. hair loss
2. fatigue
3. dry mouth
4. loss of appetite
5. low red and white blood cell counts

These side effects can take a toll on all patients, especially the elderly or the ones with underlying health conditions [19]. Furthermore, the low cell counts of red and white blood cells are often accompanied with a high increase in risk of infections in addition to other side effects. Worryingly, the current treatment strategy can cause death, either directly from the severe the side effects, or as a secondary effect from infects manifesting in the absence of white blood cells[24].

As mentioned in section 1.1, the standard treatment is based on superficial similarity between patients and not the actual drivers of the cancer in the individual patients. Additionally, the current treatment regimen uses unspecific cytotoxic drugs, killing both healthy and cancerous cells. As the diversity in the genetics of AML patients is so vast, the response to the standard treatment is also highly variable. Although there are screening methods to find markers (see section 1.5) that can aid in assessing the potential success of the standard therapy, or finding another therapeutic option, the current practice still involves a certain element of guesswork. As there is uncertainty related to the standard treatment, elderly patients and patients with other severe health conditions are often put on palliative care as a last resort [19].

Even tough patients are fit for the standard treatment, it is not only extremely demanding for the body, but is also mentally challenging. The current treatment combination requires long hospital stays, both during treatment and often for observation in relation to the side effects post treatment. The amount of hospitalization limits the patient's social contact with family and friends, which can be isolating and lonely for any individual. However, the current treatment regimen attempts to allow patients to have as much time at home as possible due to the intervals of treatment [19]. Regardless, putting all patients thorough this physical and mental strain that could increase probability of death is not optimal. There must be a way to determine at an earlier stage whether patients should undergo the standard treatment regimen, so that the patients who do not have effect from the treatment, can be spared from undergoing this demanding process.

1.5 Approaches to Diagnosing Patients with Higher Precision

The common practice for diagnosing AML in Norway is mainly based on a blood test and a bone marrow biopsy. The blood test is used to determine the cell count of the various blood cells in addition to platelets and whether immature blood cells are present in the blood [25]. The bone marrow is used to do a haematological study in order to determine the blast count. The AML diagnosis will be given in cases where the blast count is more than 20 % (with the exception of AML with t(15;17), t(8;21), inv(16) or t(16;16)) [26]. In later years the bone marrow biopsy has also been used for genetic tests in order to determine how the patient will respond to the standard treatment.

As haematological and morphological studies of blood or bone marrow can offer challenges to correctly define what type of leukemia is present in a patient, there has been an increasing use of other prognostic tools to more accurately diagnose patients [25]. These tools allow for more detailed information about the cancer based on genetics, immunology and so on. It is also worth mentioning that the current development of diagnosing cancer is leaning towards separating the cancer into

smaller and more specific cohorts in relation to biomarkers, like specific insertions/deletions, single nucleotide variants of chromosomal translocations. [27]. Concurrently, there is a growing challenge with the increasing amount of information to extract which markers are relevant for treatment choice. This more narrow classification approach reflects the current trend in drug development, where drugs with highly specific targets are in focus. Concurrently, the increasing amount of information poses a growing challenge when it comes to extracting which markers are relevant for treatment choice. Döhner et al has written a review on the various diagnosis techniques used in AML [26]. The remainder of this section will discuss a few of these techniques in short in correlation with Döhner's descriptions.

1.5.1 Immunophenotyping

Immunophenotyoing of acute leukemia is used not only for assessing the correct diagnosis, but also to assess the patients into risk groups (see section 1.6 in addition to measuring the residual disease post treatment. Immunophenotyping uses flow cytometry with a large group of selected markers (usually clusters of differentiation markers) in order to measure and detect a large variety of physical and chemical properties of the AML cells. Immunophenotyping can be used to separate three types of acute leukemias: AML, ALL and acute promyelocytic leukemia (APL). Furthermore, immunophenotyping can be used to detect subgroups within all three classes, in AML these include with AML minimal differentiation and AML with granulocytic or monocytic differentiation [5].

1.5.2 Cytogenetics

Cytogentics is the study of chromosomes and is thereby used in AML diagnosing to determine if there are chromosomal abnormalities present in the bone marrow, which in turn can affect prognosis. For instance cytogenetics is often used to determine whether the cancer has a complex karyotype or a monosomal karyotype, as these variants are associated with a very poor prognosis [28] . Additionally, fluorescence in situ hybridization (FISH) can be used to further study the cytogentics on a molecular level to detect specific fusion genes, translocations and other rearrangements. FISH is especially useful for detecting minimal residual disease after treatment [28].

1.5.3 Molecular genetics

In certain cases molecular genetics techniques, such as reverse transcriptase–polymerase chain reaction (RT-PCR), can be used to identify mutations associated with high-risk prognosis in order to determine whether patients should receive palliative care instead of unspecific chemotherapy. Examples of such high-risk genes are NPM1, CEBPA, and FLT3. As these genes are associated with high risk, there is ongoing research to find effective inhibitors for them. For instance there is extensive research to find and improve FLT3 inhibitors for AML treatment [29].

Standard procedure for diagnosis of AML in Norway involves all the above mentioned tests. Additional tests might be performed depending on underlying health risks. The diagnosis is usually set two to five days after the medical examination has commenced [30].

1.6 Assessment of AML

Regardless of the large heterogeneity in genetic AML profiles, certain mutations and karyotypical aberrations are components used to assess patients into risk groups through classification systems. These risk classification systems can be used in order to determine whether the patient is fit for standard treatment or whether the patient has a subtype of AML that will require an alternative treatment. There are three main assessment systems: the French, American and British (FAB) classification system, its successor, the World Health Organization (WHO) classification system and the 2017 European LeukemiaNet (ELN) system. However, it is important to note that the WHO system is more commonly used as a diagnostics tool. There are other classification systems for risk

assessment of AML as well, however these are not relevant for this thesis and will therefore not be discussed.

1.6.1 The FAB classification system

The FAB classification system was developed in the 1970 as a joint effort between seven French, American and British haematologists [31]. Due to its time of invention the FAB classification technique is largely based on the morphology of cells studied under the microscope after staining. The system was originally based on six subgroups but with the advances of immunological markers two new subgroups were added [32]. The current FAB system can be seen in 1.2 [33].

Table 1.2: **FAB classification system:** The names of the seven subtypes of AML in the FAB classification system alongside their names and association with prognosis compared with the average in AML[34].

FAB subtype	Name	Prognosis compared to AML average
M0	Undefferentiated acute myeloblastic leukemia	Worse
M1	Acute myloblastic leukemia with minimal maturation	Average
M2	Acute myloblastic leukemia with maturation	Better
M3	Acute Promyelocytic leukemia	Best
M4	Acute myelomonocytic leukemia	Average
M4 eos	Acute myelomonocytic leukemia with eosinophilia	Better
M5	Acute monocytic leukemia	Average
M6	Acute erythroid leukemia	Worse
M7	Acute megakaryoblastic leukemia	Worse

The FAB system is used for assessing risk in patients to determine whether they should have standard treatment or whether they should be put on palliative care. However, it also aids in identification on subgroups who should receive alternative treatments[35]. An example of this is the subgroup M3 or APL (which was discussed in section 1.3). APL patients have the best prognosis of all AML subtypes if detected early, but prognosis can severely worsen with delayed detection [20].

1.6.2 The WHO classification system

Like the FAB classification system, the WHO system aims at dividing AML into subgroups based on morphology. However, as the WHO system was developed in 2001 and later revised in both 2009 and 2016, it takes into account other factors that we know affect prognosis today, such as mutations, genetic and chromosomal abnormalities[36]. The WHO system is much more detailed than the FAB system. It is based on dividing AML into large classes with highly specific subgroups. The subgroups in the WHO system can contain a few mutations, or very specific chromosomal translocations. The four main classes of the WHO system are listed in table 1.3 [37] . Each main group has multiple subgroups based on genetic abnormalities [34]

Table 1.3: **The WHO classification system:** An overview of the four main WHO classes for AML, their characteristics and their incidence [34, 37].

WHO class	Characteristics	Incidence
1	AML with recurrent genetic abnormalities	30-40%
2	AML with multi-lineage dysplasia	10-15%
3	AML and myelodysplastic syndrome, therapy related	5-10%
4	AML, not otherwise categorised	30-40%

After assessing the specific subgroup of leukemia, these classes can further be used to determine the prognosis for the individual patients. The prognosis set by the risk classification systems will, alongside factors such as age and underlying health conditions, in turn decide whether the patient is suited to undergoing treatment or if he/she should be put on palliative care [34].

1.6.3 The ELN system

The European LeukemiaNet, an international expert panel, attempted to make a classification system of AML based on findings from using previous classification systems such as WHO and FAB. The ELN system was developed first in 2010 and later revised in 2017. As opposed to the WHO and FAB systems, the ELN-system has very clear guidelines for how genetic abnormalities should be associated with risk [38]. However, the system does not take into account co-occurring genetic abnormalities.

1.7 *Ex Vivo* Drug Screens

An *ex vivo* drug screen is a tool that aims to identify novel or approved drugs with patient-specific anti-cancer properties for possible personalized treatment. It is most commonly used in myelodysplastic syndromes (MDSs) and AML due to the heterogeneity of these cancers [39].*Ex vivo* drug screens are based on isolating cells from patient material before adding the cells in well plates. The cells are dispensed in equal amounts in each well, and drugs will be added in a set concentration range. The cells will be counted in order to determine changes in growth in response to the drug (see figure 1.3 and chapter 2 for further explanation).

Drug sensitivity can be measured in multiple ways, however the common practice is using Hill equations to fit a dose response curve to the data. Followingly, a variety of metrics can be used to determine drug sensitivity: inhibitory concentration 50 (IC50), effective concentration 50 (EC50) and area under the curve (AUC) are the most commonly used (see figure 1.3).

IC50 describes the half maximal inhibitory concentration, EC50 describes the half maximal effective concentration and AUC describes the entire area under the fitted dose response curve [40]. The different drug sensitivity metrics describe different qualities related to drug sensitivity and which metric to use depends on what you desire to study.

IC50 and EC50 both describe differences in sensitivity based on the dosage, but they do not explain the effect a drug has on cell growth. Additionally, neither IC50 nor EC50 for responses with differing maximal inhibitory effects. However, if drug responses share the same potency across patient samples both the IC50 and EC50 can be used to compare the change in patient sensitivity. Further, the IC50 has little validity for describing responses that have a lower potency and it can only be used in cases where responses go below 50% [40].

AUC, on the other hand, will detect both changes in doses sensitivity and maximal potency for a

drug. Additionally, AUC as a measure of drug-sensitivity has been reported to yield a higher degree of reproducibility in screens when panels of drugs and cell lines are compared across independent studies. Furthermore, it has been reported that when comparing AUCs across studies using different concentration ranges it is beneficial to adjust the integration interval [41].

Additionally, it is also common to use the differential integral between drug responses from healthy cells (or donors) and drug responses from cancer cells (or patients) to adjust the AUCs for non-cancer specific drug toxicity. This is sometimes referred to as selective drug sensitivity score [42]. In a seminal paper on ex vivo drug sensitivity screening of AML by Pemovska et al. they also reported to adjust the integration range to cover 10-100% of maximal inhibition [42].

Ex vivo drug screens are not commonly practised in the clinic, but can often be a component of clinical trials as a precision medicine approach. In the trial setting they can be utilized to determine which drug is best suited for a patient, or to study the effects of drugs in development on patient cells based on the cells' responses. In addition to patient cells, healthy donor cells are also included in the *ex vivo* drug screen as reference point to determine whether the drugs are toxic or cancer specific. Despite its usage in clinical trials and in drug exploration, little is known about the extent of information from an *ex vivo* drug screen that can be extrapolated for clinical application.

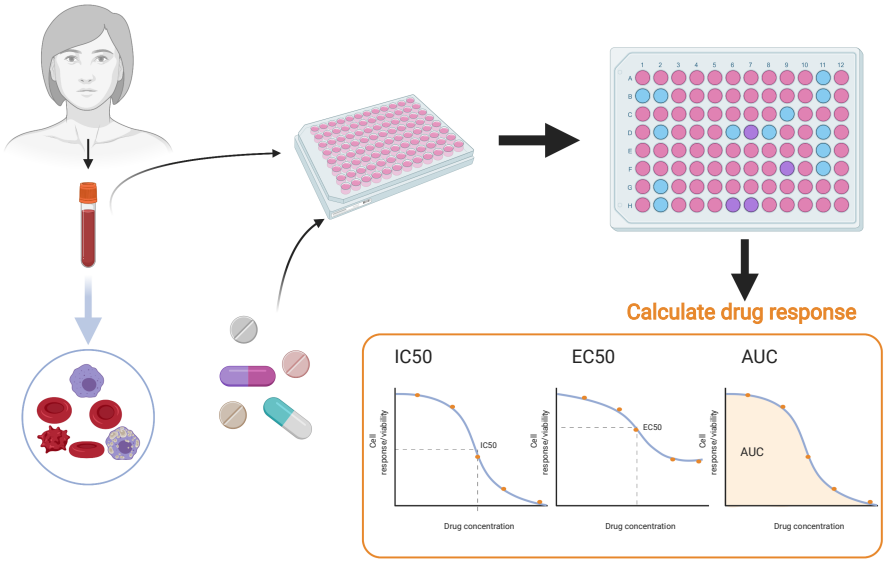


Figure 1.3: ***Ex vivo* drug screen:**A simplified illustration of an *ex vivo* drug screen. In AML, mono nuclear cells are isolated from patient tissue (either bone marrow or blood). These cells are a mixture of healthy and cancerous cells as indicated by the blue ring. The cells will be transferred into a well plate in with an equal amount of cells in each well. A large variety of drugs in a range of concentrations can subsequently be added to the wells of the plate. The plate will thereafter be analyzed by performing a cell count in each well to asses the cell response, to the added drugs. The drug responses can be evaluated in various ways, the lower right square shows examples using IC50, EC50 and AUC. Created with BioRender.

1.8 The Current Tactic for Drug Development

Historically there have been great advancements in cancer therapy, from the early days where the only treatment was surgery, to the discovery of radiation therapy, moving towards chemotherapeutic compounds to present day where immune therapy and precision medicine is the main focus[43].

The goal of precision medicine is to take into account individual variability in relevant cancer markers, in order to accurately predict which treatment options will work. The aim of this treatment tailoring is to create drugs that target mutations specific to subgroups identified on the basis of "common heterogeneity". The process of studying drug effects on the premise of the genetic background of subgroups is referred to as pharmacogenomics [9].

The focus in drug development is on highly specific drugs targeted to a few mutations in relation to a disease. Additionally, there is an increasing focus on developing drugs towards conditional cancer dependencies, entailing that drugs are developed against targets ,or factors, that are conditionally important for growth and viability of the cancer cells [4]. Further more, there is an increasing use of immunotherapy in AML, for instance using monoclonal antibodies, Alirocumab and Evolocumab, developed for treatment of diseases caused by the gene PCSK9 [44]. The current trend is thereby moving away form the previous use of chemotherapy where unspecific drugs kill both healthy and cancerous cells.

1.9 Applications of Machine Learning in Cancer Biology

Machine learning is a term describing the use of statistics and computer algorithms to learn patterns in data or finding predictive relations between sets of observations [10]. Machine learning is typically categorized into two groups, supervised and unsupervised learning, see figure 1.4.

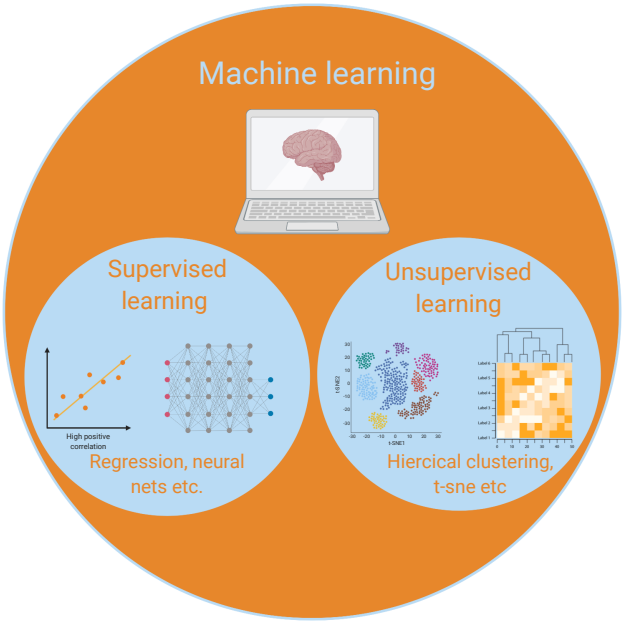


Figure 1.4: **Machine Learning:**An illustration of how machine learning encompasses both supervised and unsupervised learning with examples of techniques for both subgroups. Created with BioRender.

Supervised learning is concerned with optimizing, or training, models for maximizing prediction of measurable outcomes or target variables from other sets of variables. Techniques for performing supervised learning include linear and logistic regression, decision trees, support vector machines and neural networks (deep learning). Supervised learning can for instance use neural networks for image classification [45], tree models for diagnosing COVID19 based on chest X-rays[46] or Cox regression for survival predictions in breast cancer patients [47].

Unsupervised learning on the other hand, involves finding low-dimensional patterns or differentiable classes in high-dimensional data through the use of various clustering techniques, such as hierarchical clustering, k-means clustering or principal component analysis (PCA). Unsupervised learning can for example be used to determine aggressive phenotypes of prostate cancer using hierarchical clustering [48] or to determine genetic ancestry based on single nucleotide polymorphisms (SNPs) by PCA [49].

The application of machine learning in cancer research is gaining momentum due to the increased focus on personalized medicine in addition to the advantages of machine learning in clinical forecasting and patient diagnostics. Machine learning in clinical cancer research is mainly focused on two key issues [50]:

1. Prediction: Forecasting survival, prognostics, and treatment selection

2. Inferences on disease relevant biological mechanisms, often to aid further drug development

An example of forecasting is the use of Cox regression to predict survival from genomic aberrations in chronic lymphocyte leukemia [51]. An example of inference is doing whole genome sequencing on blood to determine mutations associated with development of haematological cancers using the Genome Analysis Tool-Kit, GATK, a set of bioinformatic tools for analyzing high-throughput sequencing [52].

1.9.1 Forecasting

Forecasting is a term used in machine learning referring to predicting future outcomes based on data explaining the past (hence weather forecast and not weather prediction). In oncology forecasting has a wide array of application areas but the main focus is on forecasting survival, relapse and treatment response. The main goal is to eliminate guesswork in the clinic to relieve patients of unnecessary treatment, hospitalization and tests.

Forecasting can use a wide array of machine learning techniques, but the main challenge is choosing the right one. One of the unique challenges of performing prediction and inference from biological data is the large inequality between numbers of variables measured over number of samples. For instance, various omics measurements, such as RNA sequencing or genomic profiling on a given patient can range in the thousands of individual makers. However, the number of patient samples in any given study is generally low. Thus, finding a generalizable pattern or subset of predictive covariates is an inherent challenge.

1.9.2 Inference

In addition to predicting treatment outcome and selection of appropriate therapy, a fundamental concern of machine learning in cancer precision-medicine is to derive and understand the predictive biological relations or mechanisms to aid the development of new drugs/therapies (as seen in the lower part of figure 1.5). For any given model used in cancer medicine, interpretability is critical, such that inferences can be made about which variables are important for a prediction and what they mean.

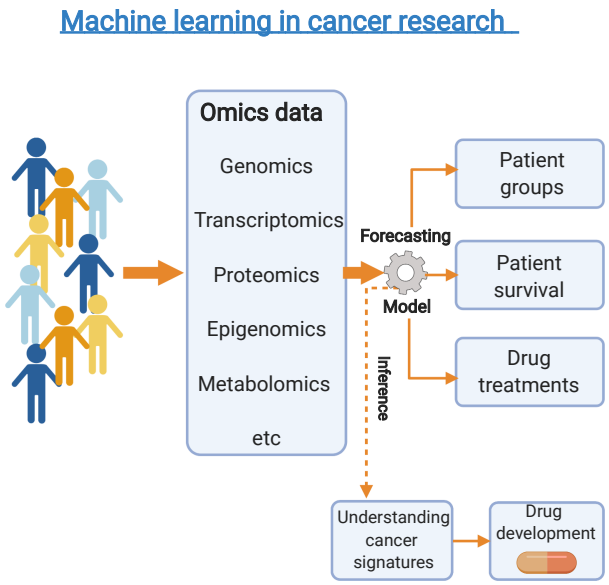


Figure 1.5: **Machine Learning in cancer research:**An example of the applications of machine learning on omics data from cancer patients separated into forecasting and inference. Created with BioRender.

Finding relevant drug-sensitivity related patterns in cancer data (genomics, transcriptomics etc.) is an inherent challenge due to the high degree of variability and presence of confounders in cancer, in combination with the commonly low sample size. The advent of high-throughput drug-screens is overstepping this challenge by directly characterizing drug sensitivity profiles of individual cancer samples, and thereby skipping identification of markers as well as expensive tests. However the problem of a large number of variables in relation to a reduced sample size still holds true for more direct methods like high-throughput drug screens.

1.9.3 Problems associated with machine learning on biological data

There is a problem in cancer biology in terms of finding generalizable patterns or subsets of predictive covariates on biological data due to the variable-to-sample size problem. Increasingly complex models, containing high numbers of variables, are more prone to what is known as over-fitting, which is the condition where a model has high predictive accuracy on its training data, but low accuracy on new data. This is related to something known as the bias-variance trade-off. The bias-variance trade-off is the phenomenon of when the complexity of a model increases, the bias will be reduced and the variance will increase. Both bias and variance contribute to prediction error, however low-bias high-complexity models can over-fit the few sample points that exists in a low sample size context and thus will have high sensitivity to sampling variance. Such models lack generality and thus have low predictive accuracy on new observations (1.6). Selecting a less complex model, like linear regression, tend to be more predictive when the sample variance is low.

In addition to choosing high-bias models, removing or reducing non-predictive covariates, a process known as variable selection, is also a common strategy to counter over-fitting. A widely used method to reduce complexity is by penalizing the total magnitude or number of correlations that can be trained in the model during the optimization procedure. This process is known as regularization.

However, the most important aid to maximize predictivity is to have as much data as possible so that the model can be trained on a sample, which represents the true diversity within possible future observations, as best possible. The data acquisition is the main bottleneck in machine learning, especially in biology due to the cost and time it takes to produce it as well as limited access to patient samples.

Finally, a disadvantage of high-complexity models is their lack interoperability and reduced ability to make inference. For example a high complexity model like a neural net with high numbers of non.linear relations through the so-called hidden layer can often be challenging to interpret. Advantages of linear models include the direct interptretability of the regression coefficients.

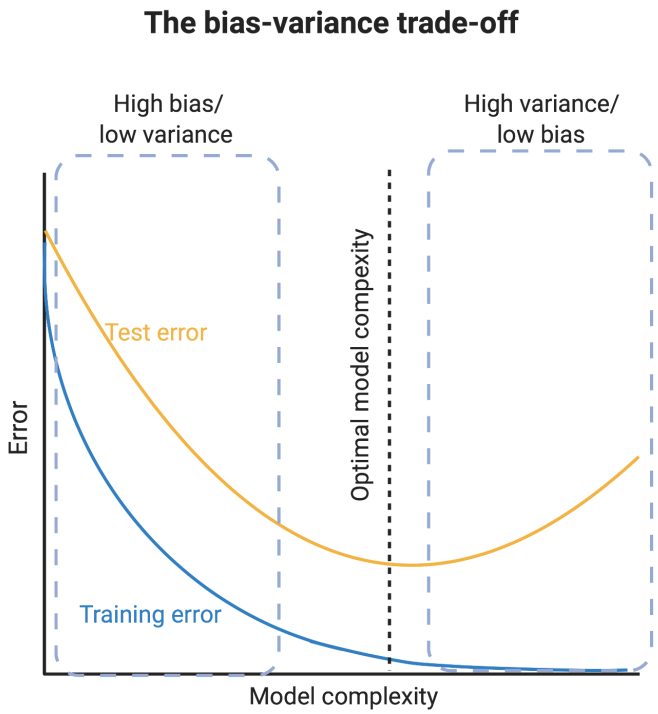


Figure 1.6: **An illustration of the bias-variance trade-off:** A models complexity (x-axsis) is highly associated with bias and variance. Both bias an variance contribute to prediction error (y-axis) on training and test data. Less complex models are associated with high bias and low variance as simpler generalizations within the training data are made. The prediction error in both the training (blue) and test (orange) data tend to be high for low complexity models. More complex models on the other hand will be associated with low bias and high variance. High complexity models tend to be prone to overfitting on the training data and do poorly on new observations (test data). The key is to find a model that has the correct complexity to minimize the error for the test data in order to optimize prediction, here indicated by the stapled line. Created with BioRender.

1.10 Scope of the Study

This study aims to answer the following questions:

- What information can be extrapolated from an *ex vivo* drug screen for clinical application? And to what extent can it be used to forecast patient survival?
- Can an *ex vivo* response correlate with a clinical response to a drug?
- Can information from *ex vivo* drug screens be used to stratify patients into treatment groups?
- Can we make inferences about the mechanisms of AML disease progression based on clinically predictive information in *ex vivo* drug screens?

The experiments and analyses were conducted at the Departement of Molecular Cell Biology in Jorrit Enserink's group, Cancer Molecular Medicine, at the Norwegian Radium Hospital.

1.11 Outline

The master's thesis has been organized in the following manner:

Chapter 2 provides an in depth explanation of the methods and materials used for experiments and analysis.

Chapter 3 presents the results from the analyses of the *ex vivo* drug screen and clinical data as well as the results from exploration of drug target networks.

Chapter 4 discusses the results that are presented in Chapter 3 with consideration of prior studies.

Chapter 5 Presents the conclusions drawn based on the obtained results. Furthermore, future perspectives of the project are presented.

Appendix A contains supplementary methods

Appendix B contains supplementary results

Bibliography.

CHAPTER 2

Materials and Methods

The following chapter describes methods and materials utilized in the study. All data preparation and the drug screen were performed by Pilar Ayuda-Durán, PhD before my time in the lab commenced. The experimental methodology is briefly described in section 2.1.2. Every analysis from 2.1.3 and onward have been executed by me using the statistical programming environment R (version 3.6.2 GUI 1.70 el capitan build) and RStudio (version 1.2.5033). At the end of this chapter there is a table containing information about the R packages used, their version and for which method they were used.

2.1 Data Preparation

2.1.1 Obtaining data

To answer the aims of this study the two following datasets were used:

1. Drug Screen data previously generated in the Enserink lab testing 349 drugs from the Selleck Anti-cancer Compound Library (see A) on 119 AML patients. This dataset will henceforth be referred to as the DS-dataset.
2. Clinical data covering survival, gender, mutations and age as well as other clinical variables (see B) were generated at the department of Hematology at Oslo university hospital. The Clinical data covered 210 patients. The clinical data will henceforth be referred to as the clinical dataset.

In the DS-dataset the patients included are all AML patients who have been diagnosed and treated at the Department of Hematology at Oslo University Hospital in Norway. All 119 patients have had bone marrow samples collected after signing a written informed consent form. The ethical review board (REK Sør-øst 2015-1012) has approved the study in accordance with the Declaration of Helsinki [53].

The clinical dataset was obtained in cooperation with Yngvar Fløisland, PhD head physician at the department of Hematology at Oslo university hospital. All clinical data was anonymized prior to leaving the clinic according to Norwegian laws concerning general data protection regulation (GDPR) and the Norwegian patient registry regulation (NPR) [54, 55].

2.1.2 Clinical samples and Drug screen

The bone marrow derived blast cells were purified from the patient samples by gradient centrifugation using Lymphoprepa ™(Stemcell)before culturing in Mononuclear Cell Medium (PromoCell) that

had been supplemented with 100 units/ml Penicillin and 100 µg/mL streptomycin. Followingly, 10,000 cells per 25 µL were dispensed in a 384-well plate (Greiner Bio-One) by a Multi Drop Combi peristaltic dispenser (Thermo Scientific) (if cell counts were low, fewer cells were used without a noticeable change in the outcomes of the experiment). Subsequently, the cells were treated with compounds from the Selleck Anti-cancer Compound Library consisting of 349 compounds with anti-cancer properties (see A). Each compound was dissolved in DMSO before being dispensed in the 384-well microplate using an Echo 550 (Labcyte Inc.) at the Biotechnology Center of Oslo. The compounds were dispensed in five concentrations covering a 10,000-fold concentration range from 1 to 10,000 nM. The only exception was VPS, which was dispensed at 100 mmol/L to 1 mmol/Ll. DMSO was added as a negative control and BzCl served as a positive control. The plates (seven per patient) were then incubated for 72 hours at 37°C in a 5% CO2, humidified environment. CellTiter-Glo illuminecent assay (Promega) was used to assess cell responses with an EnVision 2104 Multilable plate reader (Perkin Elmer). The final readouts were measured in counts per second (CPS) [56].

2.1.3 Outlier identification and removal

In order to ensure the quality of the results, patients with a median log10(CPS) under 7.5 were removed from the DS-dataset. After removal of outliers, 55 patients overlapped between the DS-dataset and the clinical dataset.

2.1.4 Normalization to BzCL and DMSO

The CPS of CellTiter-Glo illuminecent assay (Promega) is a quantitative measure of the relative ATP levels in a given well and serves as an alternate measure for the number of viable cells. To measure the drug responses as a relative change in cell growth, all CPS values were normalized to the median CPS of the positive control (BzCL) and the negative control (DMSO) using the following formula:

$$Response_{Normalized} = \frac{CPS - median(CPS_{BzCL})}{median(CPS_{DMSO}) - median(CPS_{BzCL})} \tag{2.1}$$

All negative response values were adjusted to the lowest positive value in each plate.

2.1.5 AUC calculation, log-transformation and standardization

The drug sensitivity was measured as the area under the curve (AUC) of the dose responses for every drug-patient combination using the function below:

$$AUC = \sum_{d=2}^{m} \frac{Response_d(log_{10}(C_d) - log_{10}(C_{d-1}))}{log_{10}(C_m) - log_{10}(C_1)}, \tag{2.2}$$

where C_d indicates a specific concentration of a drug, ranging from the lowest concentration $d = 1$ to the highest concentration $d = m$.

The AUCs were transformed to fold drug sensitivity, $-log_2(AUC)$, such that a value of -1 would be a halved population and a value of 1 would be a doubled population over the integration interval of the AUC.

Both drug sensitivty scores, AUCs and -log2 transformed AUCs were also standardized over the 55 patient samples using the following formula:

$$Z_i = \frac{X_i - mean(X)}{sd(X)}, \tag{2.3}$$

where X_i is the input drug sensitivity score for patient i, and $mean(X)$ and $sd(X)$ is the mean and standard deviation of a drug sensitivity for a given drug over the patient population. Thus, Z_i represents

a standardized drug sensitivity score for patient i, where the Z-scores for each drug are distributed with zero mean and standard deviation of one.

2.2 Regression and Work With Regression Coefficients

2.2.1 Cox proportional-hazard model

To predict patient survival we used the Cox proportional-hazard model, also known as Cox regression, on the different drug sensitivity measures and the clinical data. Cox regression is a form of linear regression that can investigative the effect a magnitude of variables has on a timed event. In our case the event is death. The survival times were computed as the time from first registered visiting date in the clinic, until the registered date of death. The last recorded visiting date was used for censoring surviving patients.

The Cox regression uses a hazard function, $h(t|X_i)$, to model how change in a baseline risk for death at time t, $h_0(t)$, changes as a function of some covariates X_i for patient i,

$$h(t \mid X_i) = h_0(t)exp(\beta_1 X_{i,1} + ... + \beta_p X_{i,p}) = h_0(t)exp(\beta \cdot X_i) = h_0(t)\theta_i. \tag{2.4}$$

Here β are the regression coefficients that represent a change in log-hazard ratio as function of a drug sensitivities X. A positive coefficient indicates that sensitivity to a specific drug is negatively associated with survival by increasing the likelihood of dying. Likewise a negative coefficient indicates the opposite, and zero indicates no relation between a variable and survival.

Given the model and input variables X_i the likelihood of death occurring for patient i at its true time of death T_i, over the set of patients (indexed by j), who are still alive and still in the study (not censored), is given by

$$L_i(\beta) = \frac{h(T_i \mid X_i)}{\sum_{j:T_j \geq T_i} h(T_i \mid X_j)} = \frac{h_0(T_i)\theta_i}{\sum_{j:T_j \geq T_i} h_0(T_i)\theta_j} = \frac{\theta_i}{\sum_{j:T_j \geq T_i} \theta_j}. \tag{2.5}$$

If the hazard-ratio $\theta_i > 1$, the likelihood of death for patient i will be greater than the baseline probability.

2.2.2 Univariate Cox regression on clinical data

To test for survival association within clinical variables we used univariate Cox regression. Mutation and chromosomal abnormalities were encoded as binary variables. Prognosis, WHO-class and FAB-class was represented as binary dummy variables. Age was represented as birth year and gender was represented binary where the baseline, zero, represented male and one represented female.

2.2.3 Model testing and concordance index

In order to objectively compare different models we split the data into a training set and a test set. The training set is designated for training the models. The test set is used to assess the predictive accuracy of the models on unobserved data. The test set approach gives a clear indication of how the models will perform in prognostic forecasting on new patient data. The data was randomly separated into a training set of 45 patients and a test set of 10 patients. The proportions of alive and dead patients were maintained from the original dataset, such that the test set always contained four and six patients respectively. This randomized splitting into training and test data was performed 200 times for each model to test predictive accuracy on data with different sample compositions.

The predictive accuracy on test data was scored with a concordance index (C-index). The C-index measures the fraction of patient pairs in the test data where the patient with lower survival time is also predicted with higher risk. The C-index is calculated using the equation

$$\text{C-Index} = \frac{\sum_{i_j} 1_{T_j<T_i} \times 1_{\theta_j>\theta_i} \times \delta_j}{\sum_{i_j} 1_{T_j<T_i} \times \delta_j}, \tag{2.6}$$

where i and j are patients, T is the survival time, θ is the hazard-ratio associated with the patients, and δ_j is a binary variable where $\delta_j = 1$ indicating that patient j died. The C-index will thereby be an output between zero and one. A model that perfectly predicts the outcome for all patients will provide an output equal to one. A model that incorrectly predicts the outcome for all patients will have an output equal to zero. A C-index of 0.5 is thereby equal to random guessing.

2.2.4 Training

To train the Cox model, the partial likelihood function is maximized with respect to the regression coefficients:

$$\arg\max_{\beta} L(\beta) = \prod_{i=1}^{n} L_i(\beta), \tag{2.7}$$

which is equivalent to minimizing the negative log-partial likelihood function

$$\arg\min_{\beta} -l(\beta) = -\sum_{i=1}^{n} (\beta \cdot X_i - log(\sum_{j:Y_j \geq Y_i} exp(\beta \cdot X_j))). \tag{2.8}$$

2.2.5 Regularization and cross-validation

In datasets with a high number of covariates p compared to sample size n, overfitting the model to the training data is a high risk. This can be countered by restricting the total magnitude, or number of coefficients, that can be trained, by penalizing the sum magnitude of β-values during the optimization process, thus constraining the maximum achievable model complexity. This is called regularization and works by adding a penalty term to the negative log-likelihood function that is minimized. If the penalty is on a sum of squared β-values, it is called L2-regularization, and the following function is minimized:

$$\arg\min_{\beta} -l(\beta) + \lambda \sum_{k=1}^{p} \beta_k^2. \tag{2.9}$$

If the penalty is a sum of absolute β-values, it is called L1-regularization, and the following function is minimized:

$$\arg\min_{\beta} -l(\beta) + \lambda \sum_{k=1}^{p} |\beta_k|. \tag{2.10}$$

The constraining of the model training is determined by size of the λ-parameter, which is selected to maximize prediction accuracy in unobserved data. Linear regression with L2-regularization is known as Ridge regression, and linear regression with L1-regularization is known as Lasso. The fundamental difference between the two methods is that Lasso will shrink coefficients of non-predictive covariates to zero, and thus works as a variable selection technique. On the other hand Ridge regression only shrinks the magnitude of the coefficients towards zero, and tends to maintain better predictive accuracy compared to Lasso. Lasso will usually shrink the model to a number of covariates less than the sample size n, regardless of whether more covariates have predictive relations to the target variable.

An intermediate method to Ridge regression and Lasso, called elastic-net uses, both the L1 and the L2-penalty, which are weighted with an α-parameter:

$$\arg\min_{\beta} -l(\beta) + \lambda((1-\alpha)/2 \sum_{k=1}^{p} \beta_k^2 + \alpha \sum_{k=1}^{p} |\beta_k|). \tag{2.11}$$

Here $\alpha = 1$ results in Lasso, and $\alpha = 0$ results in Ridge regression. Elastic-net, with α-values between 0 and 1, also shrinks coefficients to zero, but tends to retain more non-zero coefficients compared to Lasso, and thus is a softer variable selection method.

To optimize the regularized regression for prediction in unobserved data, the λ-parameter needs to be optimized for prediction of out-of-sample data (non-training data). This was done by testing several λ-parameters and selecting the best one, using leave-one-out cross-validation. In this procedure one patient is retained from training dataset, and thereby the the model optimization, such that the model is trained on 44 patients. Followingly, the prediction accuracy is evaluated on the one retained patient. This splitting is iterated through the whole dataset, and the average prediction accuracy is computed for that particular λ-parameter. The prediction accuracy was evaluated by the partial-likelihood deviance

$$Dev_{\lambda}^{i} = -2(l_i(\hat{\beta}^{(-i)}(\lambda)) - l^{-i}(\hat{\beta}^{(-i)}(\lambda))), \tag{2.12}$$

where $\hat{\beta}^{(-i)}(\lambda)$ are the model coefficients trained on the training data minus data from patient i, using penalty λ. $l_i(\hat{\beta}^{(-i)}(\lambda))$ is the log-partial likelihood for the left out patient i, and $l^{-i}(\hat{\beta}^{(-i)}(\lambda))$ is the maximized log-partial likelihood for the model. The λ-parameter, which yields the lowest $Dev_{\lambda} = \frac{1}{n}\sum_{i=1}^{n} Dev_{\lambda}^{i}$, is selected.

2.2.6 Variable importance

To evaluate importance of each variable (drug sensitivity) on forecasting survival, models with one variable (drug) taken out ($p-1$), were tested and scored using the C-index. The difference in mean C-index from the complete model (p) was computed to score the change in predictive accuracy upon retaining a drug.

2.2.7 Bootstrapping and t-test of regression coefficients

To statistically evaluate the associations between clinical outcomes and ex vivo drug sensitivities, we estimated standard errors for the model coefficients through Ridge regression on bootstrapped training data samples. Bootstrapping is a method which uses random sampling with replacement in order to simulate sampling variability from the original data set. Regression coefficients, representing the log-hazard ratio change for each drug sensitivity, were estimated from $B = 200$ bootstrapped training datasets, of which the means and standard errors were computed. Statistical significance in association between ex vivo drug sensitivity for drug d and survival was assessed through the t-statistic

$$t_{\beta_d} \frac{\bar{\beta_d}}{s.e.(\beta_d)}, \tag{2.13}$$

where $\bar{\beta_d} = 1/B\sum_{b=1}^{B}\beta_{d,b}$ indicates the mean and $s.e.(\beta_d) = 1/B\sqrt{\sum_{b=1}^{B}(\beta_{d,b} - \bar{\beta_d})^2}$ is the standard error. A t-statistic above 1.96 or below -1.96 was considered statistically significant.

2.2.8 Quantile analysis

Patient drug sensitivity distributions were stratified into sensitivity quantiles of 25%, 50% and 75%. The quantile would hold the top percentage of patients with the strongest responses to a given drug. Based on the drug sensitivity cut-offs Kaplan-Meyer survival analyses were performed for the two groups.

2.3 Hierarchical Clustering

Hierarchical clustering was done on the $-log_2(AUC)$ drug sensitivity matrix to study the relation between the drugs as well as the patients. Euclidean distance measure and Ward-D2 clustering was used. The Hierarchical clustering was visualized in a heatmap. The coefficients (log-hazard ratio

change) of significantly associated drugs were indicated along the rows and mutations, whereas gender and survival status were indicated along the columns.

To weigh the clustering by the forecasting potential of ex vivo drug sensitivities, the same clustering was performed on a log-hazard ratio matrix, where the $-log_2(AUC)$ data was multiplied with the Ridge estimated coefficients for each drug ($Xj,d\beta_d$, where Xj,d is the $-log_2(AUC)$ for drug d and patient j and β_d is the regression coefficient).

2.4 Generation of Drug Target Networks

Functional enrichment of drug targets and cell processes in the significantly associated drug sets was assessed through projection and clustering of drug target networks. All experimentally verified drug-drug target interactions were collected from the Drug Target Commons database, in addition to those reported in the Selleck library (see A).

2.4.1 Drug Target Commons Database

The Drug Target Commons (DTC) database is a community driven web platform improving the consensus on drug-target interactions. DTC contains information about compounds, targets and bio-activities uploaded by a magnitude of labs globally. DTC covers:

1. over seven million compounds
2. over 13 thousand targets
3. Over 14 million bio-activities

2.4.2 STRING database

To map and cluster the enriched drug-targets based on functional relations we used interaction data from the STRING database. The STRING database is a collection of protein-protein interactions, both known and predicted. The interactions cover both direct/physical interactions and indirect/functional interactions. The database covers over 24 million proteins across more than 5000 different organisms. The data collected in the STRING database stems from:

1. High-throughput experiments
2. Genomic Context Predictions
3. Conserved Co-expression
4. Automated text mining
5. Previous knowledge collected in primary databases

2.5 UMAP of string data

All drug-targets were represented in a binary matrix with their reported string interactions, and projected in two dimensions into functionally related clusters using Uniform Manifold Approximation and Projection (UMAP). UMAP is an algorithm used to do dimension reduction, similarly to techniques such as t-SNE and PCA. UMAP is based on manifold learning techniques, meaning that it is a non-linear dimension reduction technique.

We used the Pearson correlation coefficient as a distance metric and tested different number of neighbours-parameters for the clustering analysis. Number of neighbours is used to determine what

class a variable should be included in by looking at the nearest number of neighbours and determining their clusters in order to cluster the variable in question[57] .

2.6 GO-analysis

Gene Ontology enrichment analysis (GO-analysis) was used to assess the optimal functional clustering in UMAP in the distinct k-means clusters in the UMAP space. K clusters for the kmeans were set to k=15 with approximately 50 drug targets in each cluster.

GO-analysis is used to study the enrichment in a set of genes. It can also be used to indicate annotations on the gene set. The annotations are:

1. Molecular function (MF)
2. Biological Process (BP)
3. Cellular compartment (CC)

To select the optimal UMAP clustering, the UMAPs were scored by the number of enriched GO terms and average $-log(p-value)$ for the enrichment (using a Fisher's exact test). To annotate the drug target network, the most significant GO terms from CC, BP and MF for each respective cluster were selected.

Drug localization in the drug target networks was computed by the Euclidean distance form all targets related to a drug. Drugs were binary color-coded based on clinical association with drug sensitivity (-1 representing association with negative prognosis, and +1 representing association with positive prognosis). The Drug targets were color-coded based on the average binary clinical association from all of the drugs reported to inhibit said target.

2.7 Packages in R

The table below is an overview over which packages have been used in R to conduct the methods mentioned above. The table also provides information about where to find the citation for the packages.

Table 2.1: **R packages:** An overview of packages used in this project, their version, what section they have been utilized for and their citation

R-package	Verison	Used in section	Citation
umap	0.2.6.0	2.4	[58]
org.Hs.eg.db	3.10.0	2.4	[59]
clusterProfiler	3.14.3	2.4	[60]
ChemmineR	3.38.0	2.4	[61]
STRINGdb	2.2.0	2.4	[62]
glmnet	3.0-2	2.2	[63]
survival	3.1-8	2.2	[64]
survminer	0.4.6	2.2	[65]
heatmap3	1.1.7	2.2	[66]

CHAPTER 3

Results

3.1 Assessment of Drug Screen Predictivity

3.1.1 Overview and pattern conservation

In order to assess the predictive potential of *ex vivo* drug screens in forecasting survival of AML patients, a drug screen was performed on bone marrow and peripheral blood samples from a cohort of newly diagnosed AML patients. All patients received the standard treatment of Daunorubicin and Cytarabine and were followed up over a five-year period, during which disease progression and treatment response was recorded.

The study comprised 210 patients with clinical metadata, mutation mapping and karyotype aberrations collected. Additionally, regular clinical assessment of AML through cytogenetic classification using the WHO and FAB systems was included in the dataset for each patient.

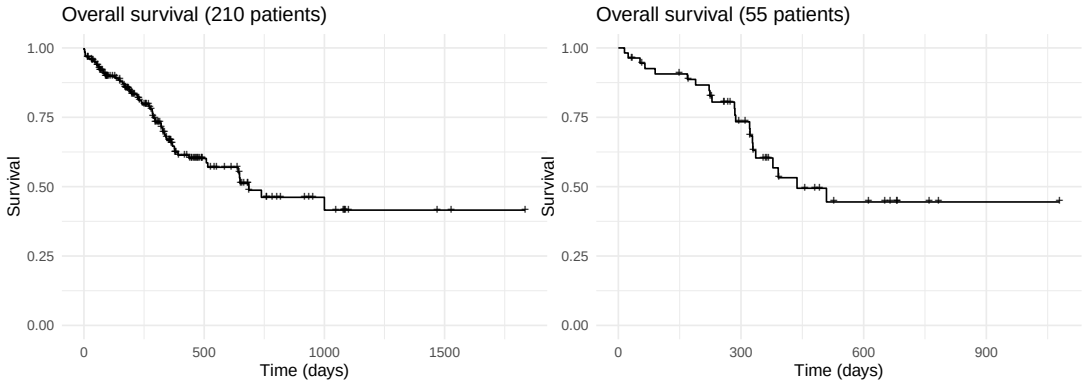
The drug screen was performed on samples from 119 patients with a drug library of 349 anticancer drugs from the Selleck Anti-Cancer Compound Library (see appendix B). Each drug was tested at five different concentrations ranging from 1 to 10,000 nM. The drug screen was performed using an ATP chemiluminescent assay yielding counts per second (CPS) value representing viable cell density after 72 hours of drug treatment.

A total of five outliers were removed from the DS dataset on the basis that their median log_10 CPS was under 7.5. The overlap between the DS data and the clinical data resulted in a dataset consisting of 55 patients with 385 variables: CPS for each drug at all five concentrations in the *ex vivo* drug screen for all 349 drugs as well as the 36 variables from the clinical dataset, in addition to survival status and time recordings from the patient cohort.

To confirm that the combined dataset with a reduced population had maintained a distribution which was representative of the overall distribution of the original clinical dataset, Kaplan-Meier plots were generated to investigate the two populations. The larger population of 210 patients from the clinical dataset (figure 3.1b) exhibited an estimated overall survival of about 41% with the patient longest included in the study recorded at 1835 days. The reduced population in the combined dataset (figure 3.1c), used for the remainder of the study, was recorded to have an estimated survival of roughly 43%, with the patient longest included in the study recorded to be 1079 days. The reduced dataset thereby showed a distribution sufficiently conserving the patterns seen in the original dataset.

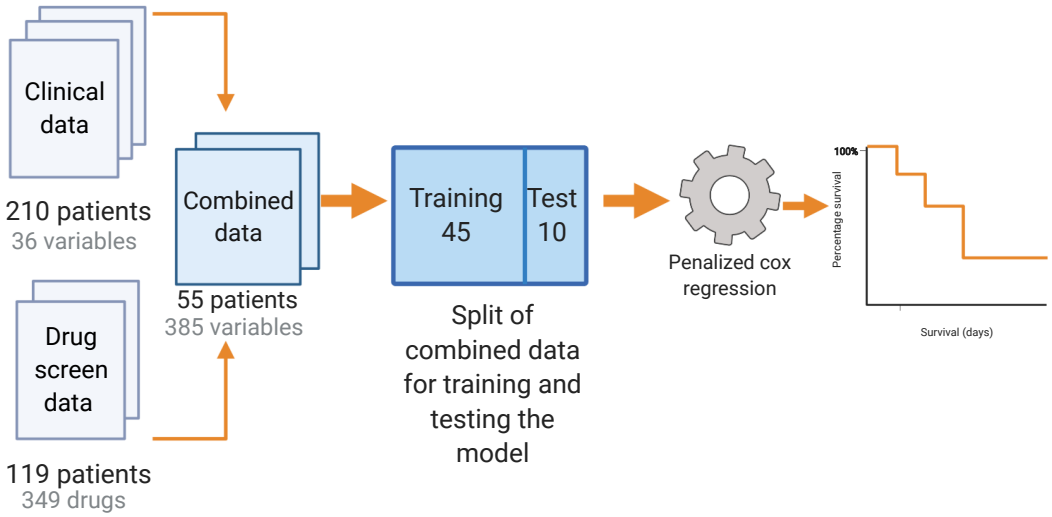
Ensuingly, the CPS values from the drug screens were normalized in accordance with the positive and negative plate controls according to formula 2.1, resulting in a drug response representing the relative change in growth. In order to score drug sensitivity, AUCs were computed from the normalized dose-response curves for all 349 drugs used in the *ex vivo* drug screen. The AUC's were also -log_2 transformed and standardized for individual patients. Cox regression was then used in

order to predict survival from the different variables. In cases where the number of variables exceeded the number of patients, regularization was used to counter overfitting and find the minimal predictive patterns (see figure 3.1a for a schematic overview).



(a)

(b)



(c)

Figure 3.1: **Patient survival:** ***(a)*** Overall survival from all 210 patients with clinically available data involved in the study. The estimated overall survival is at roughly 41%, and the longest survival time was recorded to be 1835 days. Ticks indicate last follow-up date. ***(b)*** Overall survival in the 55 patients with overlap in the two datasets. The estimated overall survival is roughly 43%, and the longest recorded survival was 1079 days. Ticks indicate last follow-up date. ***(c)*** A schematic overview of the process of assessing accuracy of AML patient survival prediction from *ex vivo* drug screens and clinical data using penalized Cox regression.

3.1.2 Survival association of clinical variables

In order to investigate the association of the individual covariates in the clinical dataset with patient survival, we performed univariate Cox regression for all available variables. The regression was done on the data including all 210 patients to increase statistical coverage. All covariates except birth year were represented as binary variables, such that the reported regression coefficients indicated the log-hazard ratio for the indicated variable, while for birth year it indicates the log-hazard ratio change per year. A positive log-hazard ratio represents an increased probability of dying over the baseline risk. To assess statistical significance, t-statistics were computed from the log-hazard ratios for all clinical variables (figure 3.2.

The analysis showed that only three variables tested in the clinic had a statistical significance associated with patient survival: WHO class 4, deletion in chromosome 20 and FAB class 4-5. All three variables were associated with an increase in risk and thereby indicated a poor prognosis. Three variables showed a weak tendency towards a negative association, or good prognosis: translocation of chromosomes 9 and 11, FLT3TKD and year of birth. However, none of these variables were significant. Additionally, for gender, with male representing the baseline risk, there was a slight, but non-significant, tendency towards an increased risk for females.

Surprisingly, the doctors' prognoses, which are based on cytogenetic data according to the ELN system [38], all had non-significant associations with patient survival (figure 3.2).

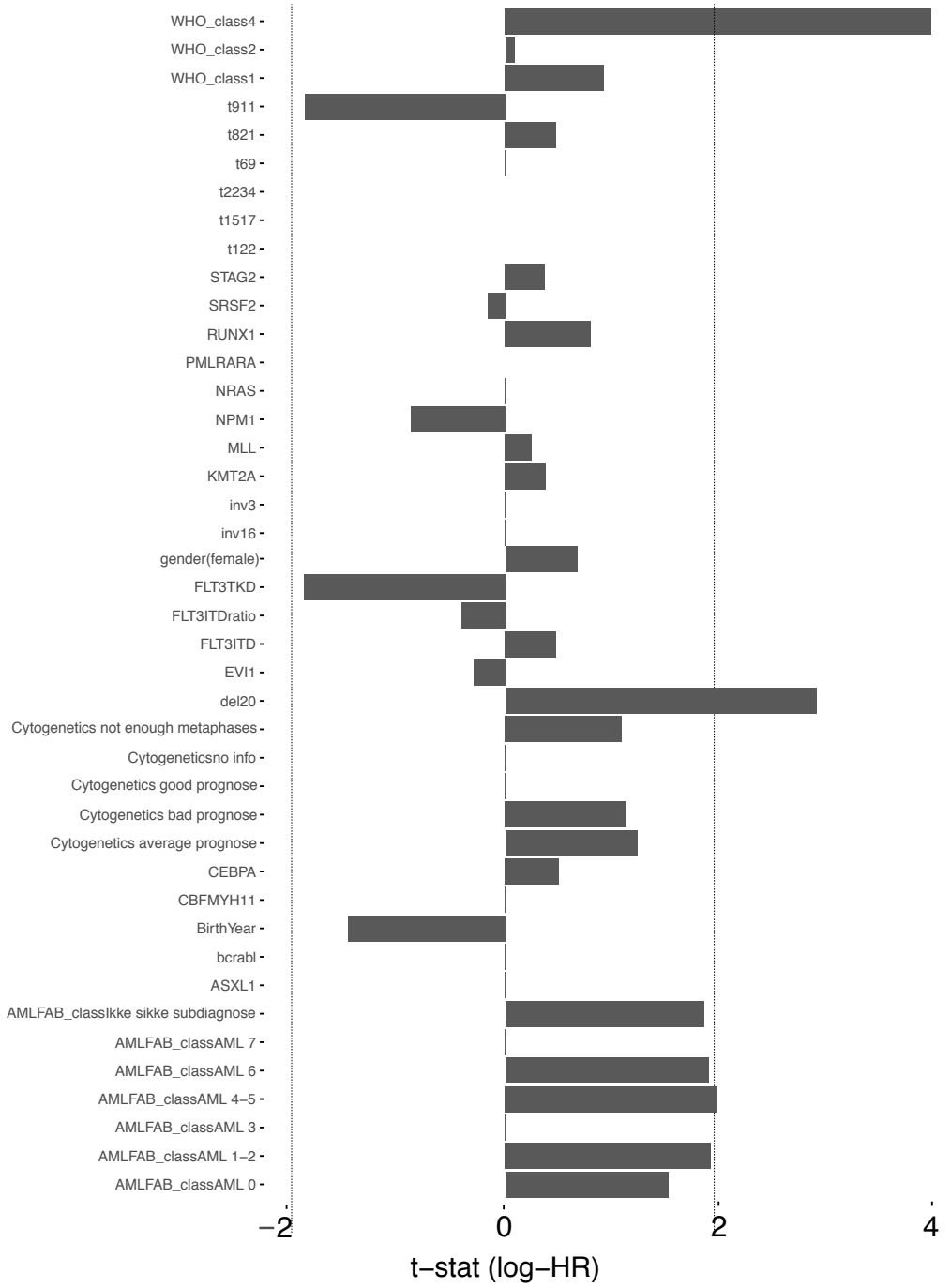


Figure 3.2: **Statistical association of clinically available data and patient survival:** Associations are reported as t-statistics of log-hazard ratios (log-HR) where the direction indicates association with risk. Significance thresholds 1.96 and -1.96 are indicated by the dotted lines for positive and negative associations, respectively.

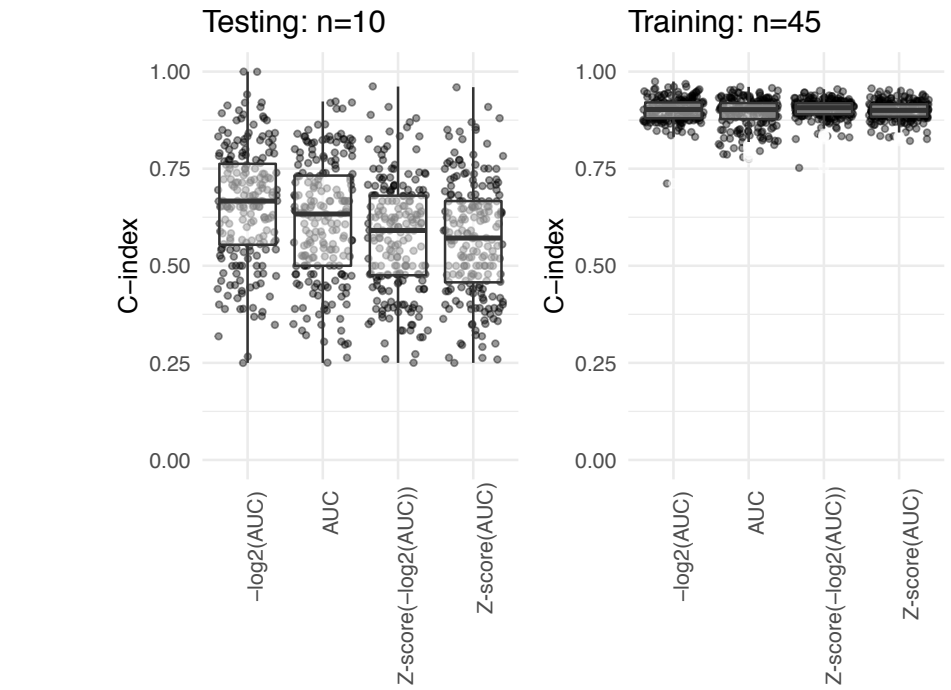
3.1.3 Survival prediction from *ex vivo* drug screens

To optimize survival prediction from *ex vivo* drug screens, we tested multiple drug sensitivity measures: AUC and $-log_2$-transformed AUC. To predict patient survival, we used a multivariate Cox regression model for all 349 drugs used in the *ex vivo* drug screen for the 55 overlapping patients. Penalized regression (Ridge, Lasso or elastic-net) was performed because of the low sample size relative to the number of drugs. In order to test the model on unobserved data, we randomly split the dataset into 45 patients for training and 10 patients for testing. The proportions between alive and dead in the training and testing sets were conserved concerning the clinical data with 6 alive and 4 deceased patients in every test set. The random splitting of the datasets and testing of the model were performed for 200 iterations. The models were scored on their predictive accuracy, using a concordance index (C-index) that measures the fractions of correctly matched predicted risk and survival time for pairs of patients (figure 3.3a.

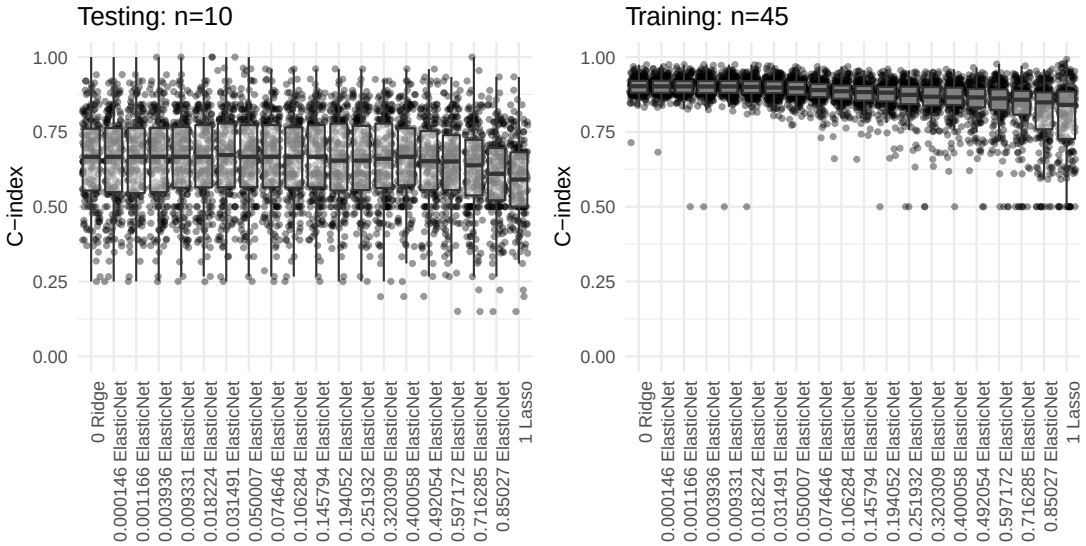
Using Ridge regularized Cox regression, we found that all methods showed high predictive accuracy in the training set with a median C-index above 0.9. However, on test data the $-log_2$-transformed AUCs had the highest predictive accuracy with a median C-index just below 0.7. Furthermore, standardization significantly lowered the predictive accuracy of both drug sensitivity metrics. The difference in predictive accuracy between the test and training data indicated substantial over-fitting to the training data.

After deciding to -log2 transform the AUCs, we proceeded to study different regularization techniques to overcome the problem of a large number of variables to a small number of samples resulting in overfitting (the model is significantly more adapt to make predictions on the training data than the test data seen in figure 3.3a). We therefore performed the same testing and training regimen on the DS-dataset while testing different regularization techniques, by modulating the α-parameter that scales between L2-penalty (Ridge) and L1-penalty (Lasso) (see section 2.2.5. An α-value of 0 equals Ridge regression, whereas an α-value of 1 equals Lasso regression. All α-values between 0 and 1 are called elastic-net. The prediction accuracy was then evaluated using the C-index (figure 3.3b.

We found that regularization using Lasso, or elastic-nets with stronger L1-penalties, caused a reduction in predictive accuracy. Furthermore, there was little difference in the median C-index of Ridge and elastic-nets with lower α-values.



(a)



(b)

Figure 3.3: **Model optimization:** ***(a)***Prediction accuracy was assessed for various drug sensitivity measures by obtaining a median C-index in the testing and training datasets. ***(b)*** Prediction accuracy was assessed for regularization methods by plotting α-parameters between 0 and 1 for the training and testing datasets and measuring their respective median C-indexes.

3.1.4 Comparing survival prediction across datasets

To study the predictive accuracy of the *ex vivo* drug screen data compared to clinically available data, we used the optimized Cox model with Ridge regression to predict survival from the clinical data, the drug sensitivity data ($-log_2$-transformed AUCs) and the combined dataset containing both clinical variables and drug sensitivity data. We followed the same testing and training regimen as described above in section 3.1.3 (figure3.4.

We saw that all three models were capable of predicting survival in the training set with scores over 0.5. However, in terms of predicting survival in the test set, the clinical dataset had no predictive power with a score below 0.5, which is even less accurate than random guessing. On the other hand, the two data sets containing drug sensitivities from the *Ex vivo* drug screen show predictive power on foreign data with scores just below 0.70.

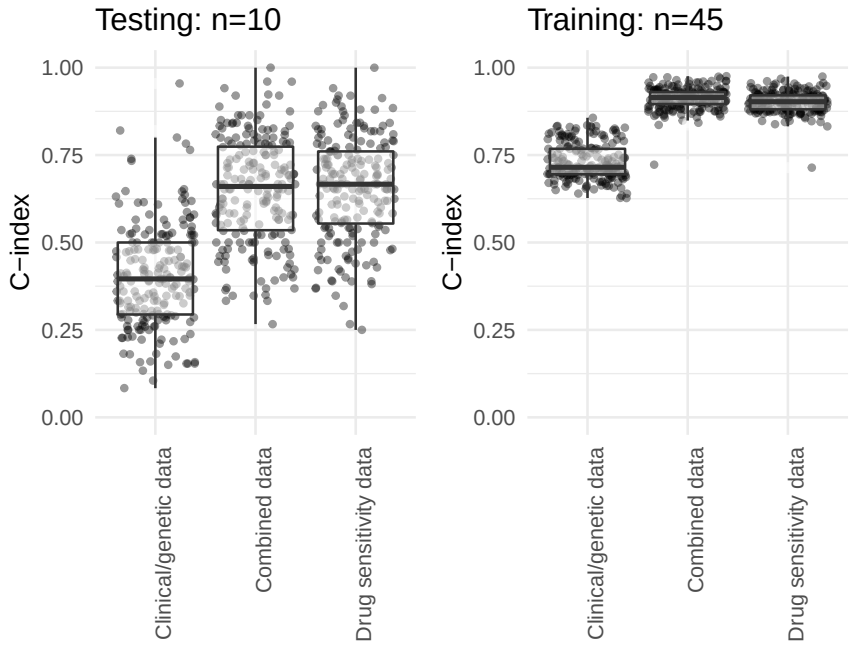


Figure 3.4: **Survival prediction compared for Cox regression models trained on clinical data, drug sensitivity data and combined data:**. Prediction accuracy is scored with C-index for 200 randomized stratifications into test data (left) and training data (right).

3.2 Identifying Predictive Drugs

In order to estimate the smallest number of drugs needed to make solid survival predictions, we used the property of L1-regularization as a variable selection technique, and how elastic-net with higher leverage on the L1-penalty (α-values closer to one) will increase the number of coefficients forced to zero. Based on the C-index in figure 3.3b we see that the predictivity drops in the test data set with an α-value of 0.40, indicating that penalizing with a stronger α removed variables that are important for predicting survival. We found that the median number of non-zero coefficients associated with an α-value of 0.40, is 20 drugs. This means that in order to predict survival based on data from an *ex vivo* drug screen the smallest number of drugs is on average 20 drugs, given the library and sample size at hand.

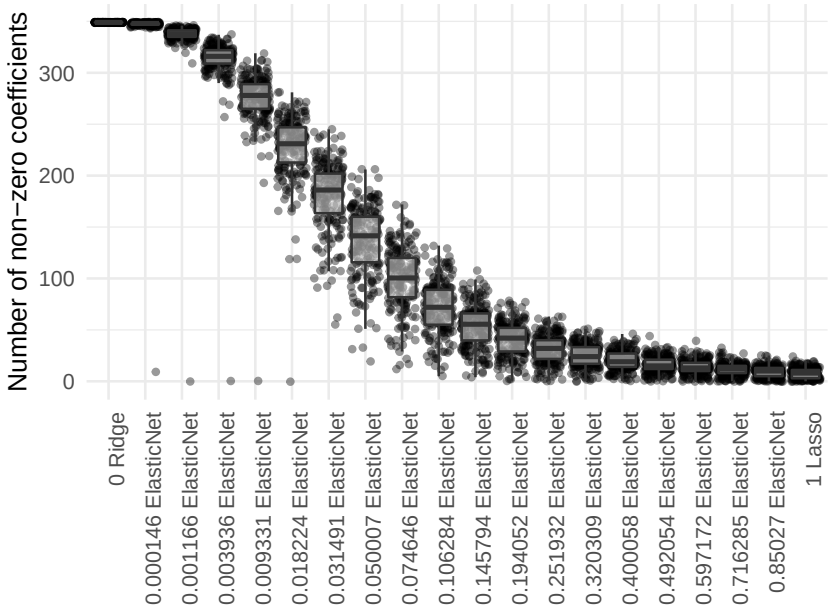


Figure 3.5: **Identifying the smallest number of predictive drugs:** Number of non-zero coefficients in models trained as a function of α-scaling between L1 and L2 regularization.

Further, we wanted to determine which drugs had a statistically significant association with patient survival out of the 349 drugs, meaning that we wanted to find the drugs associated with predicting either increased risk (bad prognosis) or lower risk (good prognosis). To do this, we performed Ridge regression on bootstrapped patient data and computed mean and standard errors for the regression coefficients (log-hazard ratio for drug sensitivity change). Statistical significance was concluded for drugs with a t-statistic greater or smaller than 1.96 or -1.96, respectively(figure 3.6).

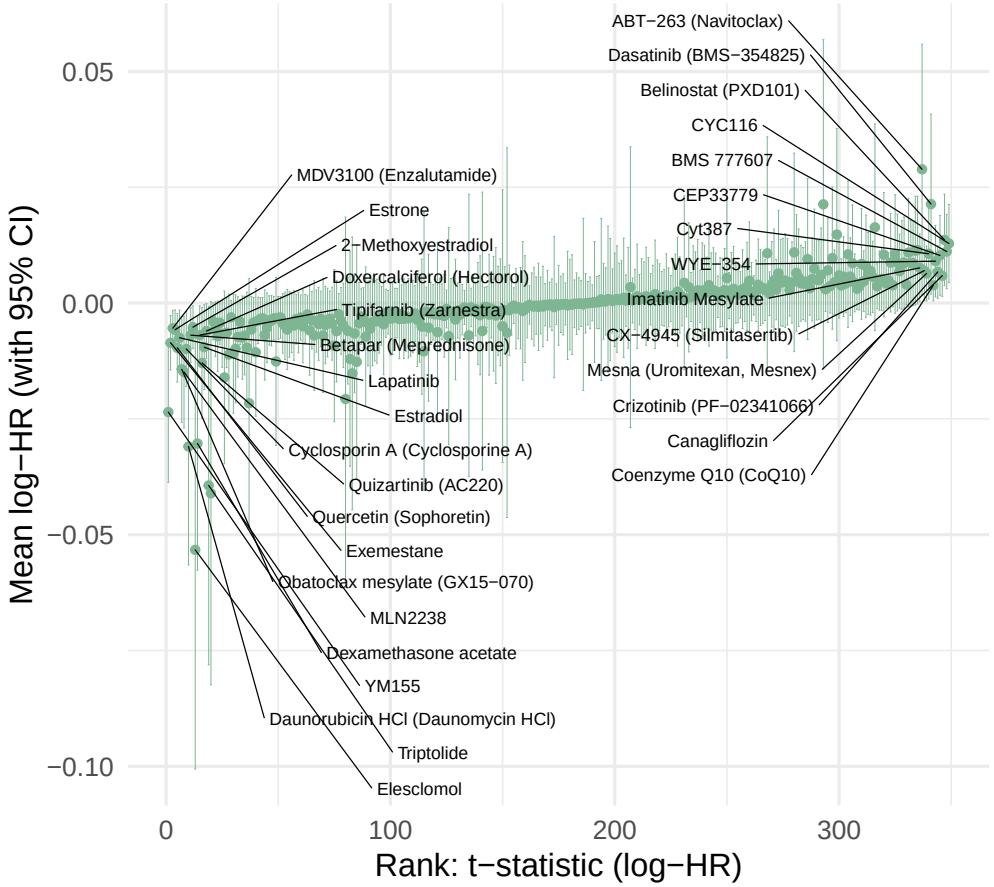


Figure 3.6: **Identifying drug sensitivities with statistically significant associations with survival:** Mean log-hazard ratio with a 95% confidence interval per unit of change drug sensitivity plotted against the drugs ranked by a t-statistic in relation to their log-hazard ratio. 33 drugs with a statistically significant association are indicated. The upper right corner indicates drug sensitivities associated with poor prognosis. The lower left corner indicates drug sensitivities associated with good prognosis.

We found 33 drugs with a statistically significant association with patient survival (figure 3.6). 14 of these drugs had drug sensitivities associated with an increase in risk, and thereby poor prognosis. 19 drugs had drug sensitivities associated with reduced risk, and thereby a good prognosis. It is however important to note that the prognosis in relation to these drugs is based on all patients received the standard of Cytarabine and Daunorubicine in the clinic.

As a proof of principle, we find Daunorubicine in the cluster associated with reduced risk (figure 3.6, lower left), indicating that a good response to this drug *ex vivo* predicts a favourable response to the standard treatment in the clinic. Further, we find the drug ABT-263 (Navitoclax) in the opposite group (figure 3.6, upper right). Venetoclax, a molecular analogue of Navitoclax, is a drug that has been used in the clinic in cases where patients do not respond to the standard treatment, or for patients with high risk [67].

Further, we wanted to study which drugs affected predictivity the most. Thus, using the same training and testing regimen as described in section 3.1.3, we scored the prediction accuracy on test data for models where one drug was left out. The difference in the mean C-index compared to the model with all drugs included was calculated to score C-index loss for each drug (figure 3.7).

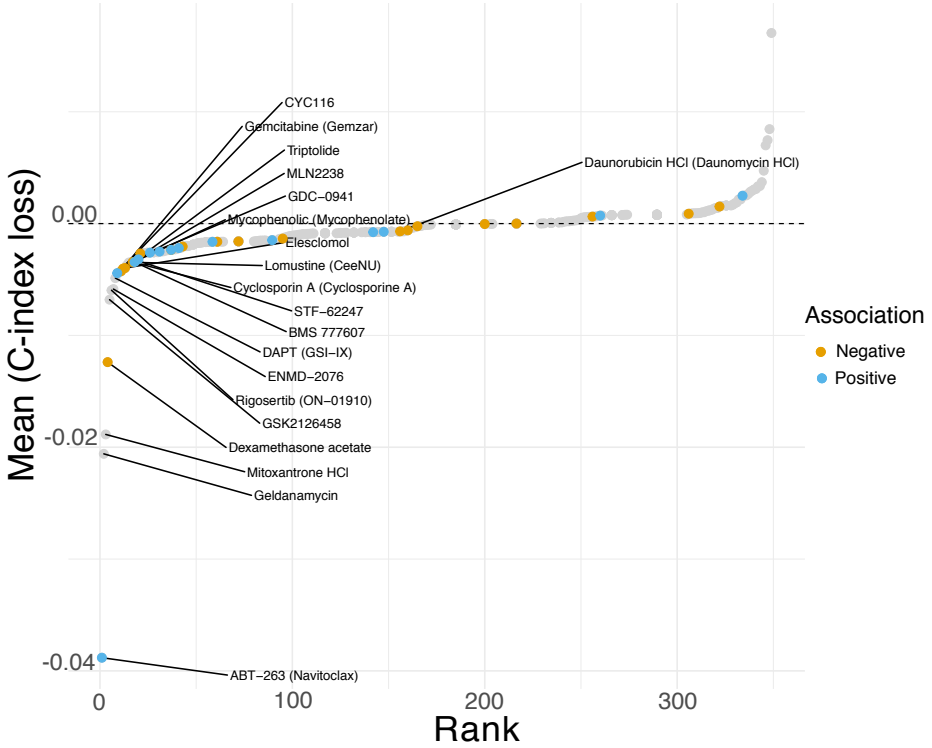


Figure 3.7: **Variable importance for survival predictions:** A waterfall plot of the mean C-index loss for models leaving out one drug relative to the full model. The top 19 drugs and Daunorubicin are indicated. The color-coding highlights the drug sensitivities identified with a statistically significant association with risk. Blue indicates a positive association and orange indicates a negative association.

We found that very few of the 33 drugs with significant association (indicated in either blue or orange, depending or their direction of affecting risk) had a substantial effect on predictivity on their own when left out of the model. However, we could observe a weak displacement in predictivity reduction for the 33 drugs. This is indicative of a high degree of multicollinearity between the drugs in the dataset, but the absence of strong correspondence could also be due to low sample size. The one exception we identified was Navitoclax, resulting in a predictivity loss of 4%. This effect on predictivity suggests that the drug sensitivity profile of Navitoclax contains unique information for forecasting patient survival.

To verify the identified drugs, we plotted the dose-response curves for all 55 patients and colour-coded them according to survival. A selection of six of the 33 drugs is shown in figure 3.8. Three of the plots correspond to drugs associated with higher log-hazard ratios: WYE-354, ABT-263 (Navitoclax) and Imatinib. Three plots correspond to drugs associated with lower log hazard ratios: Daunorubicin, Lapatinib and Estrone. For some of the drugs, like WYE-354 and Daunorubicin, we see clear differences in the responses for alive and dead patients. On the other hand, some drugs had a less prominent difference in the distribution of survival, like Imatinib.

Further, we wanted to study how the change in survival was associated with drug sensitivity cut-offs for the six drugs (figure 3.8). Thus, we stratified the distributions of drug sensitivities for the 55 patients into quantiles (25%, 50% and 75%) of sensitive and non-sensitive patient groups. Subsequently, we studied the survival of the two groups by plotting Kaplan-Meyer estimators of the percentage of patient survival over time.

We found that some drugs, like ABT-263(Navitoclax), show clear separation in survival for the 25% and 75% quantiles of sensitivities and less clear separation of survival in the 50% quantile. The separation in drug sensitivity to Navitoclax across the population could indicate a varying dependency on Bcl-2. Surprisingly, Estradiol also showed clear separations for the 25% and 75% quantiles, but a less clear separation at the 50% quantile. These groups of survival separation based on sensitivity can be material for further studies. Imatinib and Lapatinib, on the other hand, showed a more prominent separation in survival as we included more weak responses. This indicates that weaker responses to these drugs have a higher association with survival than the strongest responses.

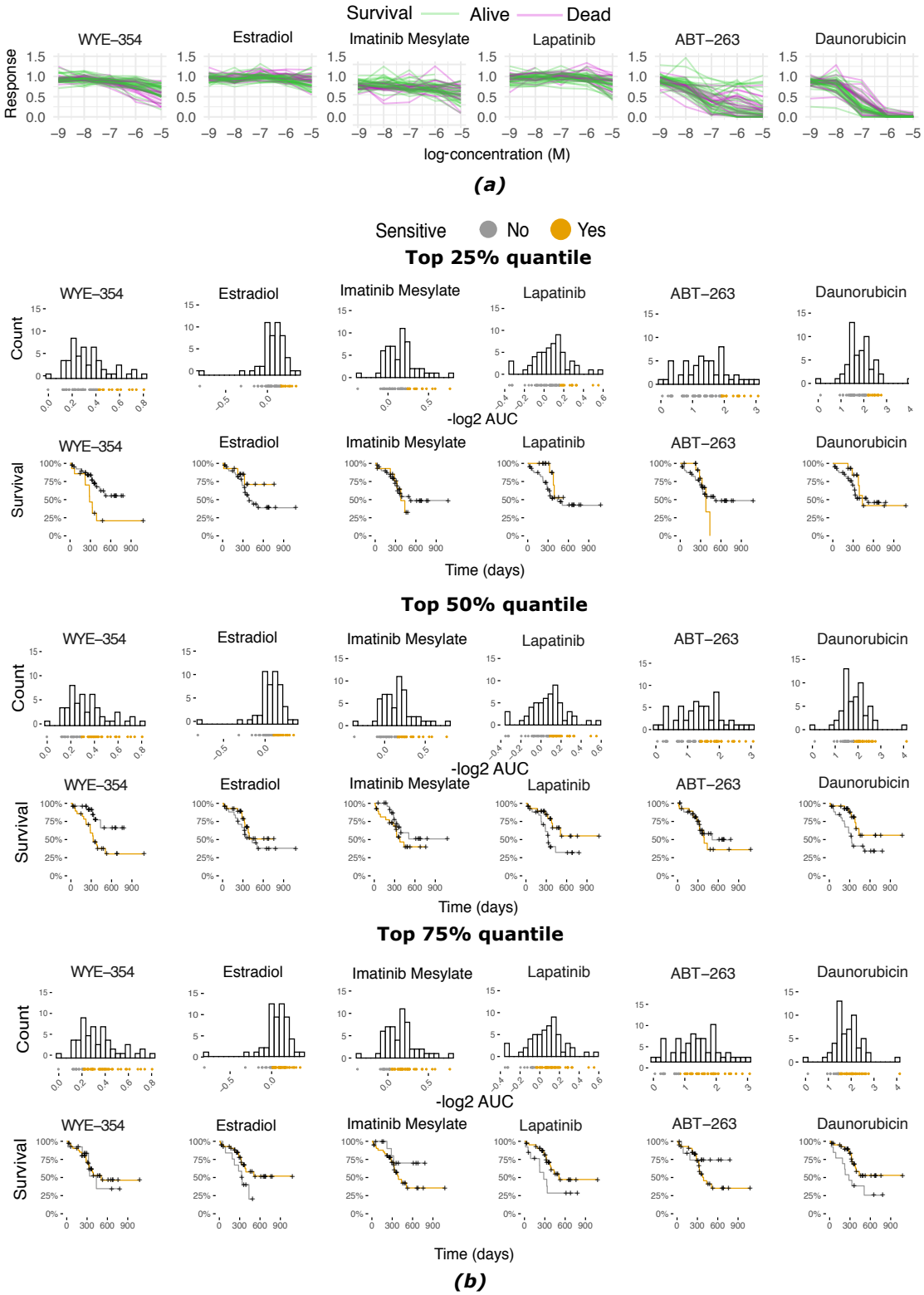


Figure 3.8: **Drug sensitivity separation:** ***(a)*** Dose response curves of six of the 33 dugs associated with log hazard ratios. The responses have been colour coded based on survival. Green indicates alive and purple indicates dead) ***(b)*** Histograms and survival plots for the six drugs are separated by quantiles of 25%. The orange line indicates patients deemed sensitive and the gray line indicates patients deemed as non-sensitive. Ticks indicate censoring.

3.3 Patient Clustering

Ensuingly, we wanted to study the association of drug sensitivity profiles in relation to potential patient groups. We performed hierarchical Ward-D2 clustering on the drug sensitivities with the Euclidean distance measure. To weight the clustering with the forecasting potential of the drug sensitivities, we multiplied the $-log_2$ AUC with the Ridge regression coefficients for each drug. This resulted in an array where each element represented the log-hazard ratio for each patient-drug combination. The sum of each patient column thereby resulted in the total log-hazard ratio, or risk, for a given patient (figure 3.9).

Firstly, we found that the clustering results in a clear stratification of drugs into groups based on the risk-weighted drug sensitivity profiles.

On top of the heatmap, we find drugs that are associated with strong drug-responses. Some of the drugs in this group have previously been associated with increased log hazard ratios, for instance, ABT-263 (Navitoclax) and WYE-354. In the bottom of the heatmap we find drugs with reduced log-hazard ratios. In the middle of the group with reduced log-hazard ratios we find some strong responses, such as for Daunorubicin, which is part of the standard treatment. In the middle of the heatmap we find a group of drugs with mixed drug sensitivities flanked by drugs with neutral sensitivity.

Secondly, in the rightmost patient cluster, it becomes evident that the survival clustering of patients is highly dependent on a balance in dug sensitivities that are positively and negatively associated with survival. Deceased patients tend to have strong responses to drugs associated with increasing log-hazard ratios, and weaker responses to drugs associated with lower log-hazard ratios. For patients that are alive, on the other hand, the opposite applies.

Furthermore, we found that the clustering of the patients relates to survival to some extent. We see alive and dead patients clustered together respectively. To some extent, we also see certain patterns for clustering of prognosis, WHO class and gender. On the other hand, for the mutations and chromosomal abnormalities, it is evident that the coverage is suboptimal even after removing the variables with the lowest coverages. However, we see that certain mutations group together, like NPM1 and FLT3ITD.

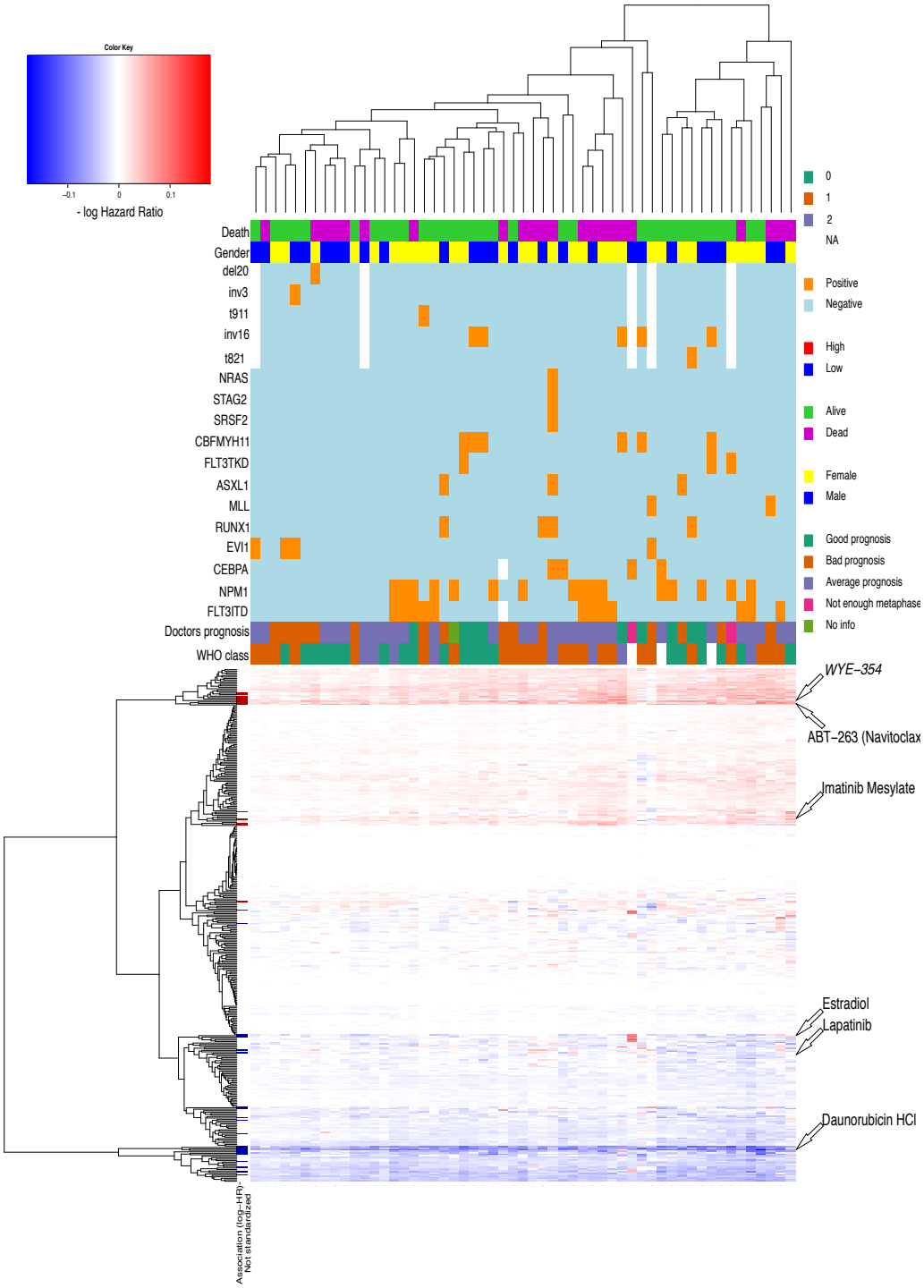


Figure 3.9: **Association of dug sensitivity profiles and patient groups:** hierarchical clustering of risk weighted dose responses are visualized as a heatmap. Clinical variables with sufficient coverage are indicated with coloured labels below the clustered patients on the X-axis. The 33 drugs with association to survival have been colour coded accordingly next to the total clustered drugs on the Y-axis.The six drugs uesd as examples in figure 3.8 are indicated with arrows on the right side of the heatmap

3.4 Identification of Survival Associated Protein Networks in AML Based on Drug-Target Data

In light of the large coverage of drug targets in the Selleck drug library, we wanted to investigate if we could generate a drug-target network for AML to better understand the underlying cellular mechanisms of disease progression and survival. An overview of the process is shown in figure 3.10.

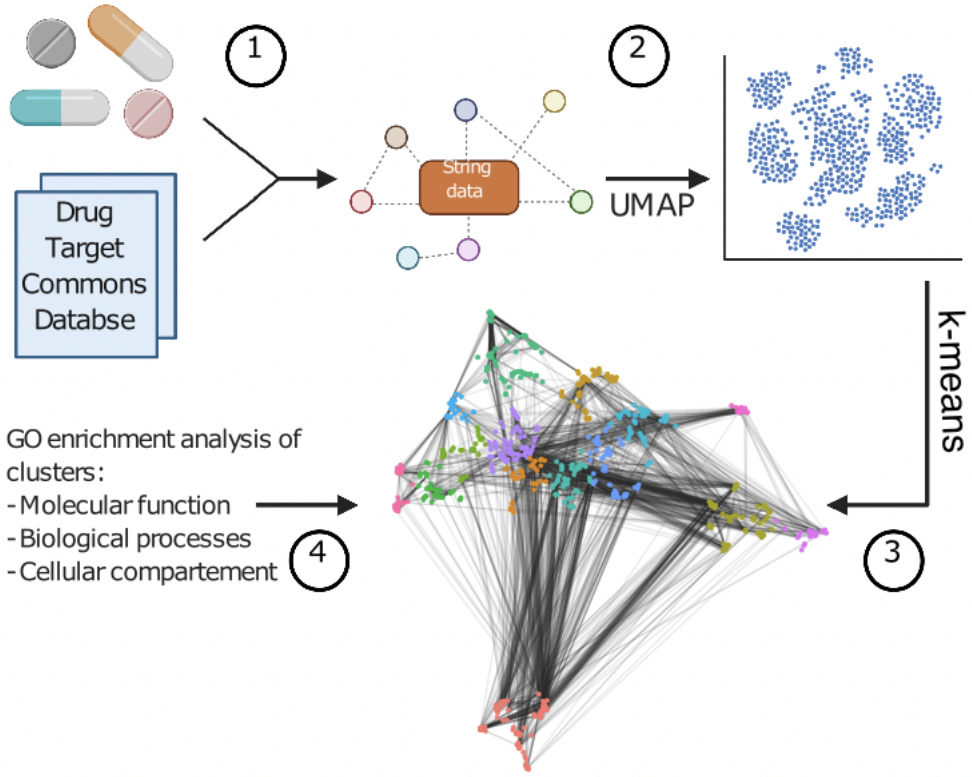


Figure 3.10: **Creating a survival associated protein network in AML based on drug target data:** ***(1)*** Drug targets were collected from the Selleck library and the DTC database for the 33 drugs identified as statistically significantly associated with survival. ***(2)*** String data for interactions between all drug targets was collected and projected using UMAP. ***(3)*** k-means clustering was performed to identify drug-target clusters. ***(4)*** GO-analysis was done to assess the optimal number of clusters from the k-means analysis and to asses GO-terms to each cluster of drug targets.

We began using the 33 drugs identified with a statistically significant association with survival, and gathered information about their drug targets form the Selleck Anti Cancer Compound Library (B). Additionally, we used information from the Drug Target Commons (DTC) database to supplement with additional known drug targets for all 33 drugs (point 1 in figure 3.10).

After collecting all the drug targets for the 33 drugs, we used the STRING database to gather information about known and predicted functional protein interactions for our collected drug targets. The STRING database covers both direct, or physical, interactions and indirect, or functional, interactions, such as participation in similar pathways. After collecting the interaction data, we generated a binary matrix representing all drug targets and their reported string interactions. To visualize the network, we used Uniform Manifold Approximation and Projection (UMAP) (point 2 in

figure 3.10). Functional drug-target clusters in the UMAP were identified with k-means clustering and analyzed for gene ontology (GO) enrichment (point 3 in figure 3.10).

The GO-analysis was used to functionally characterize the individual clusters of drug targets based on molecular function (MF), biological process (BP) and cellular compartment (CC). The GO analysis was also used to determine the ideal UMAP clustering based on the number of enriched GO terms and the average $-log(p-value)$ using a Fisher's exact test for the enrichment (point 4 in figure 3.10).

After generating the UMAP with the optimized number of clusters, we computed the position of the drugs based on the average distance to its related targets (3.11). The drugs were binary colour coded according to how the drugs associated with survival. The interaction between drug targets and drugs was colour coded in accordance with the survival association of the drugs.

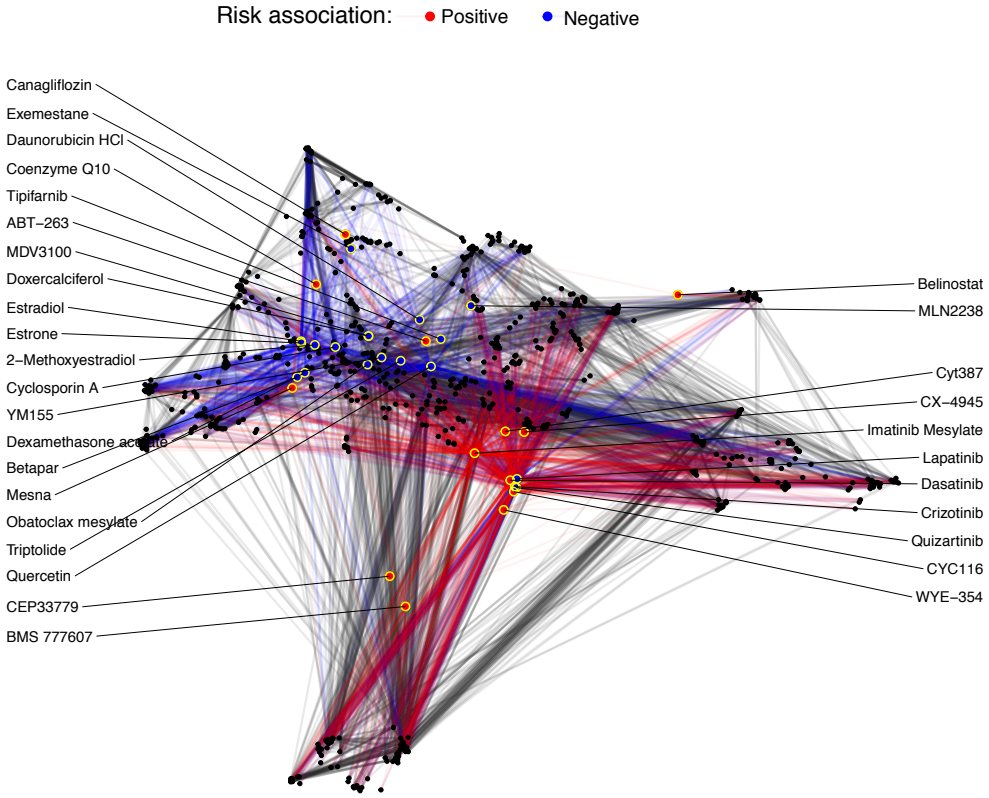


Figure 3.11: **Interaction between drugs and targets in an AML protein network:** A UMAP of an AML drug-target network generated based on the 33 drugs identified as statistically significantly associated with survival. Drugs have been mapped onto the network based on the average distance to all its related targets. The drugs are coloured based on association with survival. Red drugs are associated with high log-hazard ratios and blue drugs are associated with low log-hazard ratios. Lines are drawn from each drug to all its targets and coloured in relation to the drug.

To our surprise, we see a clear functional separation of drugs and drug targets associated with high risk and low risk. We see a cluster of drug to drug-target interactions in the lower right of the UMAP

associated with decreased survival when receiving the standard treatment. On the other hand, in the upper left corner, we find drugs and drug targets associated with increased survival when receiving the standard treatment.

As a consequence of the clear separation of risk association in the drug-target network, we wanted to study if there were biological similarities within the 15 drug target clusters in light of their risk association. We therefore associated each cluster of the UMAP with GO terms. We included the top most significant GO terms for all three GO classes for each cluster. All drug targets were then colour coded based on the clinical association with drug sensitivity based on an average of all the drugs reported to inhibit said target (figure 3.12).

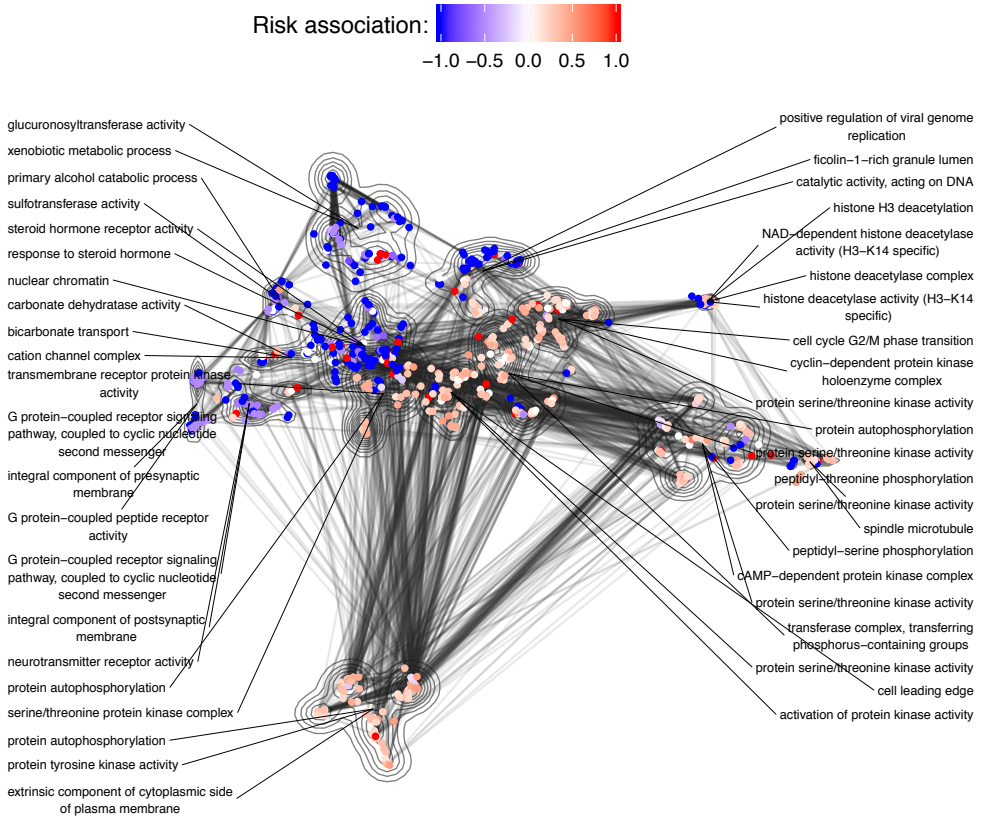


Figure 3.12: **Identifying drug target network clusters with GO-terms:** Top significant GO-terms were enriched for each drug-target cluster in the the AML protein network generated in figure 3.11. The GO-terms cover molecular function, biological process and cellular compartment. Drug targets are colour-coded based on the average clinical association with drug sensitivity based on all drugs inhibiting the given target.

We discovered that each cluster of drug targets was mainly associated with either increased risk or reduced risk if the targets in the cluster were inhibited. The clusters, associated with a reduction in risk if targets are inhibited (in blue, see figure 3.12) are enriched for pathways related to hormone regulation, regulation of foreign DNA and chromatin modifications. For example, the two top clusters are associated with xenobiotic metabolic processes and regulation of viral genomes.

Further, we find that the clusters dominated by drug targets associated with an increase in risk, if inhibited show, a saturation of kinase activity. Specifically, activity related to tyrosine, threonine and serine kinases is highly present in these high-risk clusters. For example, we see that the lowest cluster is associated relatively strongly with increased risk upon inhibition and contains targets related to protein tyrosine kinase activity. This may reveal an increased dependence on tyrosine kinase signaling in standard treatment-resistant AML.

3.5 Experimental Limitations

Some parts of this project included experimental limitations which are listed below.

- This study is largely based on a population of 55 patients as a result of the overlap between the clinical dataset and the drug screen dataset. This resulted in poor coverage of some clinical variables and especially affected the clustering analysis in terms of mutations and chromosomal abnormalities.
- All patients included in this project have received the same standard treatment, regardless of the findings in this study and previous studies using the same data. This has limited the investigation of survival to be in the light of a singular treatment, and we have no indicators of how changing the treatment in accordance to our findings would alter survival.

CHAPTER 4

Discussion

4.1 Assessment of Drug Screen Predictivity

Acute myeloid leukaemia is a highly heterogeneous disease, but a standard treatment is still the current practice. Even though risk assessment systems are in place, there are currently no alternative treatment options commonly used in Norwegian clinics, and it remains a challenge for clinicians to determine which patients are eligible for the standard treatment resulting in a survival rate of 40% (figure 3.1c). As a means of hurdling this problem, we studied how *ex vivo* drug screens could aid in predicting survival using machine learning.

In the first part of this study, we tested if *ex vivo* drug screens contained information that could be used to forecast patient survival. Indeed, we found that *ex vivo* drug screens have the ability to predict survival using a dataset of only 55 patients. We see that *ex vivo* drug screens are better at predicting survival than the current variables used to aid in setting a prognosis in the clinic. It is however important to note that, when performing the multivariate Cox regression, the clinical dataset used in this study only covers 55 patients of sparsely represented binary variables. Regardless of this, we see that even with a larger dataset of 210 patients there are still only three variables that have a significant association with survival when using a univariate Cox regression. However, a larger dataset with larger variable coverage across patients might elevate the predictive ability of clinical data both in the univariate and multivariate Cox models.

To our surprise, the doctor's prognosis does not have a statistically significant association, regardless of the severity of the prognosis. Furthermore, the good prognosis, even with an incidence of 43 out of 210, does not hold any predictive power at all. This is indicative of the need for a new, or improved, system to assess the prognosis that is less reliant on genetic markers. It is, however, noteworthy that the final prognoses are not only based on mutations and chromosomal abnormalities, but the fact that doctors take other factors into consideration, like underlying disease and previous medical history.

We see that there is overfitting of the training data for all our multivariate models, as observed by discrepancies in C-indexes for the training and testing predictions. However, we have used a model linear regression with regularization to maintain minimal complexity to battle this problem. Additionally, the test set of 10 patients is representative of adding new patients to the study in the future. A group of 10 patients might be close to a realistic supplementary patient group due to the low incidence of AML. Furthermore, when we tested the model on 10 patients, it resulted in a training set of only 45 patients, which is a highly limited dataset for use in machine learning. The best way to minimize over-fitting and maximizing the generalization in a model, is to increase the amount of data to get a more complete picture of the true variance in a larger population.

Further, we have determined that the optimal method for representing *ex vivo* drug sensitivities is by using a $-log_2$ transformation of the AUCs to maximize survival predictions on out of sample

data. We believe that the $-log_2$ transformation is better because it results in a scenario where halved and doubled populations are weighted equally, whereas untransformed AUCs represent an average relative cell density over the integration range and will be exponentially distributed as a function of variations in cell growth rate. Thus, the transformation increases the likelihood of more normally distributed data as opposed to the non-transformed AUCs.

These findings indicate that there exists a linear relationship between the $-log_2$ transformed AUCs and the log hazard-ratio. The AUC can be seen as an average response value over the integration range of a drug response curve. Thus, the AUC represents a relative growth of cells in culture in response to a drug, given the integration range. This, in turn, makes the transformed AUCs representative of a differential growth rate, as can be seen by:

$$(n(t,drug)/n(t,control))_{AUC} = e^{tg_{drug|AUC} - tg_{control}} = e^{t\Delta g_{drug|AUC}}, \tag{4.1}$$

where t is time and g represents the growth rate under the drug treatment and DMSO control respectively. The log hazard ratios, on the other hand, portray the differential decay rate of patient populations in the clinic. Combining these ideas mathematically has resulted in the following equation:

$$log\frac{h(t,AUC_{drug})}{h(t)} = -\beta \times log(AUC_{drug}) = -\beta \times t\Delta g_{drug|AUC} \tag{4.2}$$

This suggests that there is a direct negative linear relationship between the degree of inhibition of a cancer cell proliferation in culture and the probability of survival of a patient in the clinic when exposed to the same drug, as observed for Daunorubicin.

In this study we have not performed Hill curve fitting of the dose responses, and rather performed a discrete integration under the empirical response-curve between the lowest and highest dose in log_{10}-scale. This is in order to avoid curve-fitting bias or artifacts of non-Hill type responses or drug responses with insufficient data points to support a full curve fit. Because, we normalize on the maximal dose range, such that the integration interval is in one unit, the AUC becomes a weighted average for the drug response. This value is sometimes referred to as drug sensitivity score [42]. Furthermore, we have not used healthy donor cells as a reference point for drug toxicity. In some studies the differential integral between the response-curves of patient samples and healthy donors is used [42]. In our study this would simply shift the average of the drug sensitivity distribution and would have no effect on the variational relation between drug sensitivity and survival for individual drugs. However, for analytical purposes in our quantile analysis, a reference point for cancer unspecific toxic responses would be valuable.

Additionally, we also found that Ridge regression or elastic-net with more weight on L2-regularization, is the best way of penalizing the *ex vivo* drug screen data. We believe the L2-penalty is superior, as it is maintains more variables in the analysis, whereas larger α-values in elastic net or Lasso will force a disproportionate number of coefficients to zero. The fact that this hard variable selection reduces predictivity, suggests a high degree of correlation (or collinearity) between variables. This is supported by our finding that the predictivity loss when removing a single drug is minuscule. Using Ridge, where no coefficients are forced to zero, therefore includes all variables potentially leading the model to detect patterns that are lost with a harder variable selection.

4.2 Identifying Predictive Drugs

After determining that *ex vivo* drug screens could predict survival, we found that a minimum of 20 drugs was needed to maintain the predictive accuracy, using the respective α-value of 0.4. However, this amount is based on an average for all 200 Cox regressions performed on random samples of 45, and thus the number is likely higher.

We have also found a set of 33 drugs with a significant association with survival. Interestingly, we find Daunorubicin in the group associated with reduced risk. This finding is coherent with the fact that all patients have received a treatment where Daunorubicine is one of the components. An *ex vivo* response is thereby related to the patients having a positive effect on the treatment. In the same group we find Quizertinib, an inhibitor of FLT3, which is commonly mutated in AML. In the group associated with elevated hazard ratios, and thereby an increased risk, we find the Bcl-2 inhibitor Navitoclax. This corresponds well with the fact that a molecular homologue of Navitoclax, Venetovlax, is being used as a second-line treatment in clinical trials if patients do not respond to Daunorubicin [67]. A response to Navitoclax ex vivo could thereby be indicative of a poor response to the standard treatment. Additionally, we found that Navitoclax was the drug that affected predictivity the most when left out of the analysis. As Bcl-2 is an anti-apoptotic protein, this finding indicates that resistance to standard treatment in AML is related to an increased dependency on anti-apoptotic mechanisms.

Interestingly, we also find another Bcl-2 protein family inhibitor, Obatoclax, in the group of drugs associated with reduced risk. This could be an indication that the Bcl-2 family of proteins can influence treatment resistance in different ways, or that there are some off-target effects to either of the two drugs. Interestingly, Obatoclax is also included in clinical trials for AML treatment [68]. Furthermore, we found a group of multiple estrogen regulators (Estradiol, Estrone and Exemestane), both steroid hormones and estrogen inhibitors. A previous study performed by Roma and Spanguolo proposes to use Estrogen receptors (ER) as a new biomarker for deciding on AML treatment [69]. They further report that activation of ERα is in line with a good response to standard therapy and that ERβ suppresses cancer cell growth [69]. This is also in line with our observations of Estradiol in the quantile analysis, where we see that strong responders have an increased chance of survival.

From the quantile analysis we also believe that there can be found treatment clues for AML. For instance, we see that those who strongly respond to the mTOR inhibitor, WYE-354 and Navitoclax have the most severe prognosis. This could entail a possible combination treatment of the two as an alternative to standard treatment for these individuals. Interestingly, we see that the strongest responses to Daunorubicine are not associated with a better outcome compared to the more weak responses. This may indicate that if patients are very responsive to the chemotherapeutic, it will be too strong to treat the patients and yield severe side effects, as both healthy cells and cancerous cells will die.

When studying the effects the individual drugs had on predictivity, we surprisingly found that removing Daunorubicin from the analysis had a minimal effect on predictivity. We believe this is caused by a great deal of multicollinearity in the data, where the response to Daunorubicin can be explained from the response patterns of the other drugs. Furthermore, the observation that Navitoclax has the strongest reduction in C-index loss, at 4%, indicates that Navitoclax embeds a unique drug sensitivity pattern that is not represented by other drugs in the Selleck library.

Additionally, in light of the quantile analysis, we see that survival differences for many drugs are dependant on a patient stratification based on strong or weak responses. This could potentially explain why leaving out single drugs has a small effect on predictivity due to sparse representation of high-leverage patient responses under the sampling scheme used to test prediction accuracy. Intriguingly, we also see that some drugs increase predictivity when they are left out of the regression analysis. We believe these are drugs that do not have an association with survival and only contribute noise to the analysis. Furthermore, we find a group of drugs that have largely reduced predictivity when left out, but are not present in the identified group of 33 drugs with statistically significant association to survival. We suppose these drugs are the ones with a clear average change in the log-hazard ratios, but with a wide 95% confidence interval that crosses the zero line, and thus represent high-variance drug sensitivity distributions that are highly sensitive to sub-sampling with the small patient number present in this study.

4.3 Patient Clustering

From the clustering analysis performed, it becomes apparent that there is a significant number of patients who have no detected genetic marker and die. This indicates the need for better markers, or better yet, a new system that will allow patients to get the correct treatment and that can assess prognosis. In cases like these, preforming an *ex vivo* drug screen could be the answer both to predict survival and to find a possible alternative treatment. Additionally, extensive genetic analyses can be expensive and time-consuming, further indicating the need for a modernized, efficient and effective alternative.

Furthermore, we see that 30 of the 33 identified drugs are placed in areas with high log-hazard ratio weighted drug sensitivities. However, there are multiple drugs in these areas that have not been identified, due to a lack of statistical association with survival, which could be interesting to study further. The last 3 drugs identified are placed in a space of relatively weak log-hazard ratio weighted responses. This could be due to the drugs having a small association with survival.

4.4 Identification of Protein Networks in AML Based on Drug Target Data

Upon identifying a relation between *ex vivo* drug sensitivity and survival in the clinic, we decided that we would further investigate whether we could use the underlying information from the drug screens to understand the underlying mechanisms of AML survival through the identified cellular dependencies. Interestingly, we see two distinct groups of drug target clusters and target interactions that are associated with a positive or negative change in log-hazard, respectively.

When we mapped the drugs onto the network itself, we saw that the drug sensitivities associated with increased hazard and a poor prognosis when receiving the standard treatment, are located in the lower right of the map. The clusters associated with their targets are associated with protein kinase activity, specifically serine/threonine and tyrosine kinases. This correlates well with a review provided by Ling et al in 2017. They report an increase in the development of protein kinase inhibitors as developing drugs for AML treatment, with multiple drugs being in phase III of clinical trials [70].

Furthermore, in the group with low hazard and good prognosis (marked in blue in figure 3.12) we found targets associated with foreign entities, like xenobiotic metabolic processes and viral genome regulation. This is in line with the standard treatment as Daunorubicin is an anthracycline antibiotic. Additionally, we see multiple clusters identified to be involved in epigenetic modification. In the centre of figure 3.12, we find drugs associated with chromatin regulation, and in the upper right corner of the blue region we find drugs associated with histone H3 deacetylation. There are multiple papers published about the role of histone H3 deacetylation of AML and its induction of the t(8:21) translocation and its repression of RUNX1. However, why this subtype is associated with a good response to standard treatment remains unclear. Additionally, as we see clear clusters of drug targets, we believe that it might potentially be a future strategy to target multiple drug-targets within a cluster, rather than a highly specific target. This could potentially reduce the rate of developing resistance and be a more robust option for treatment.

However, within groups associated with a risk in one direction or the other, we often find drugs that are associated with the opposite hazard level. This could be attributed to the fact that a drug highly associated with hazard in one particular direction could have targets (or off-target effects) in a group associated with the opposite hazard. It could also represent a protein with a negative regulatory role of the surrounding drug-targets in the cluster.

It is, however, important to note that this drug-target network has been created solely based on the 33 previously identified drugs. We hope that increasing the number of drugs can provide a more

detailed picture of how AML is driven. In the future, we aim to use all 349 drugs included in the study to generate a map with more coverage to reveal new drug clusters and their association with patient survival.

We believe that the drug-target network can help understand the underlying mechanisms of AML based on drug-to-drug target interactions as well as functional interactions between drug targets. Additionally, we envision that the use of drug-target networks can become a powerful tool for future drug development by identifying novel cellular dependencies that are associated with survival and disease progression.

CHAPTER 5

Conclusion

There is a pressing need for improving the way we assess risk in AML patients, and to open up to alternatives in order to the current standard treatment to increase patient survival. Thus, being able to predict whether a patient will respond to treatment, is a necessity. Tactics like *ex vivo* drug screens are used in clinical trials to determine how patient cells respond to specific drugs *ex vivo*, and potentially to find treatment alternatives. In this master's thesis we aimed to address whether *ex vivo* drug screens could forecast patient survival, and whether there was more information in these high-throughput screens that could have clinical applications.

We have found that *ex vivo* drug screens have the ability to predict patient survival better than the current markers, risk assessments and prognoses used in the clinic. We have optimized predictivity through an extensive study of various data handling methods resulting in a $-log2$ transformation of AUCs for dose responses. Based on this finding, we have mathematically determined that cancer cell fitness *ex vivo* exhibits a negative linear relationship with patient fitness in the clinic when all patients are subjected to the same treatment.

Further, we have identified 33 drugs that have an unambiguous association with survival that can potentially be used in the clinic to help aid in setting a prognosis and to give indicators of treatment options. Furthermore, we have determined that Navitoclax is the drug that has the largest effect on the predictivity of patient survival, underlining its importance as a second-line treatment option for AML patients.

Finally, we sought to determine whether there was more underlying information in an *ex vivo* drug screen that could be of clinical importance. Our research on drug-drug target interactions has resulted in an extensive drug target network that can help shed light on the complicated mechanisms driving AML. Furthermore, we believe there is great potential in this drug-target network to find and determine targets for future drug development related to AML.

5.0.1 Future studies

Patient population and treatment

As described in previous sections, the number of patients in this study affects the coverage of clinical variables and prediction accuracy. It would be beneficial to run our analysis on a larger patient population to determine how this affects survival predictivity, specifically in light of clinical variables. Additionally, it would be interesting to conduct a similar analysis on a patient population that has received an alternative to the standard treatment.

Screen dimension and regularization

Further, it would be of interest to use the elastic net with the optimal α-value in the Cox model to see how variable selection will affect the survival predictions. Additionally, it would be interesting to study to what degree we can shrink the dimensions of the screen to find the lowest number of drugs needed to maintain prediction accuracy across different patient sub-samples. This could save economic costs associated with individual patient screening. With more patient data we can determine the causes for the discrepancy between C-index loss and statistical association for single drugs.

Dose-response sensitivity in relation to survival

Extensive studies on the quantile analysis would also be of interest to determine if underlying factors behind the differential survival association to weak or strong drug responses.

Generating a larger drug-target network

Generating a larger drug-target network, using all 349 drugs, is also of great interest, as we hope to gain a deeper understanding of the mechanisms regulating AML disease progression in the clinic.

Appendices

APPENDIX A

Supplementary Tables

A.1 Selleck Anti-cancer Compound Library

The following pages include a table of the complete Selleck Anti-cancer Compound Library. The table includes names of compounds/drugs, the CAS number and drug-target(s).

	CAS Number	Catalo	Target
(-)-Epigallocatechin gallate	989-51-5	S2250	
17-DMAG HCl (Alvespimycin)	467214-21-7	S1142	HSP (e.g. HSP90)
17-AAG (Tanespimycin)	75747-14-7	S1141	HSP (e.g. HSP90)
2-Methoxyestradiol	362-07-2	S1233	HIF
3-Methyladenine	5142-23-4	S2767	PI3K
A-769662	844499-71-4	S2697	AMPK activator
Abiraterone (CB-7598)	154229-19-3	S1123	P450 (e.g. CYP17)
Abitrexate (Methotrexate)	59-05-2	S1210	DHFR
ABT-263 (Navitoclax)	923564-51-6	S1001	Bcl-2
ABT-737	852808-04-9	S1002	Bcl-2
ABT-751	141430-65-1	S1165	Microtubules
ABT-888 (Veliparib)	912444-00-9	S1004	PARP
Adrucil (Fluorouracil)	51-21-8	S1209	DNA/RNA Synthesis
AEE788 (NVP-AEE788)	497839-62-0	S1486	EGFR
Afatinib (BIBW2992)	439081-18-2	S1011	EGFR
AG14361	328543-09-5	S2178	PARP
Altretamine (Hexalen)	645-05-6	S1278	Potential alkylating agent
AMG 900	945595-80-2	S2719	Aurora Kinase
Aminoglutethimide (Cytadren)	125-84-8	S1672	Aromatase
Amuvatinib (MP-470)	850879-09-3	S1244	c-Kit, PDGFR, FLT3
Anagrelide HCl	58579-51-4	S3172	PDE
Anastrozole	120511-73-1	S1188	Aromatase
Andarine (GTX-007)	401900-40-1	S1140	Androgen Receptor
APO866 (FK866)	658084-64-1	S2799	NMPRTase
Aprepitant (MK-0869)	170729-80-3	S1189	Neurokinin receptor
AR-42 (HDAC-42)	935881-37-1	S2244	HDAC
AT-406	1071992-99-8	S2754	E3 Ligase
AT7519	844442-38-2	S1524	CDK
AT9283	896466-04-9	S1134	Aurora Kinase
AUY922 (NVP-AUY922)	747412-49-3	S1069	HSP (e.g. HSP90)
Axitinib	319460-85-0	S1005	c-Kit, BCR-Abl-T315I
AZ 3146	1124329-14-1	S2731	Mps1
AZ628	878739-06-1	S2746	Raf
Azacitidine (Vidaza)	320-67-2	S1782	DNA damaging agent
Azathioprine (Azasan, Imuran)	446-86-6	S1721	Purine synthesis inhibitor
AZD6244 (Selumetinib)	606143-52-6	S1008	MEK
AZD7762	860352-01-8	S1532	CHK1/2
AZD8055	1009298-09-2	S1555	mTOR
Barasertib (AZD1152-HQPA)	722544-51-6	S1147	Aurora Kinase
BAY 11-7082 (BAY 11-7821)	19542-67-7	S2913	E2 conjugating
Belinostat (PXD101)	414864-00-9	S1085	HDAC
Bendamustine HCL	3543-75-7	S1212	Alkylating agent
Betapar (Meprednisone)	1247-42-3	S1689	Glucocorticoid
Bexarotene	153559-49-0	S2098	Retinoid receptor activator
BEZ235 (NVP-BEZ235)	915019-65-7	S1009	PI3K/mTOR/ATM/ATR
BI 2536	755038-02-9	S1109	PLK
BI6727 (Volasertib)	755038-65-4	S2235	PLK
BIBF1120 (Vargatef)	656247-17-5	S1010	FGFR
BIBR 1532	321674-73-1	S1186	Telomerase
Bicalutamide (Casodex)	90357-06-5	S1190	Androgen Receptor
BIIB021	848695-25-0	S1175	HSP (e.g. HSP90)
BIRB 796 (Doramapimod)	285983-48-4	S1574	p38 MAPK

BKM120 (NVP-BKM120)	944396-07-0	S2247	PI3K
Bleomycin sulfate	9041-93-4	S1214	DNA/RNA Synthesis
BMS 777607	1025720-94-8	S1561	c-MET/Axl/Ron/Tyro3
BMS 794833	1174046-72-0	S2201	c-Met
BMS-599626 (AC480)	714971-09-2	S1056	EGFR
Bortezomib (Velcade)	179324-69-7	S1013	Proteasome
Bosutinib (SKI-606)	380843-75-4	S1014	Src/Abl
Brivanib (BMS-540215)	649735-46-6	S1084	FGFR
BTZ043 racemate	957217-65-1	S1097	Antimycobacterial
Busulfan (Myleran, Busulfex)	55-98-1	S1692	Alkylating agent
BX-795	702675-74-9	S1274	IκB/IKK
CAL-101 (GS-1101)	870281-82-6	S2226	PI3K
Canagliflozin	842133-18-0	S2760	SGLT
Capecitabine (Xeloda)	154361-50-9	S1156	DNA/RNA Synthesis
Carboplatin	41575-94-4	S1215	DNA/RNA Synthesis
Carmofur	61422-45-5	S1289	DNA/RNA Synthesis
Cediranib (AZD2171)	288383-20-0	S1017	VEGFR
Celecoxib	169590-42-5	S1261	COX-2
CEP33779	1257704-57-6	S2806	JAK2
CH5132799	1007207-67-1	S2699	PI3K/mTOR
CHIR-99021 (CT99021) HCl	252917-06-9 (free base)	S2924	GSK-3
Chrysophanic acid (Chrysophanol)	481-74-3	S2406	EGFR
CI-1040 (PD184352)	212631-79-3	S1020	MEK
Cisplatin	15663-27-1	S1166	DNA/RNA Synthesis
Cladribine	4291-63-8	S1199	DNA/RNA Synthesis
Clafen (Cyclophosphamide)	50-18-0	S1217	DNA/RNA Synthesis
Clofarabine	123318-82-1	S1218	DNA/RNA Synthesis
Coenzyme Q10 (CoQ10)	303-98-0	S2398	Ubiquinone
CP-466722	1080622-86-1	S2245	ATM/ATR
Crenolanib (CP-868596)	670220-88-9	S2730	PDGFR
Crizotinib (PF-02341066)	877399-52-5	S1068	c-Met/ALK
CUDC-101	1012054-59-9	S1194	EGFR
CX-4945 (Silmitasertib)	1009820-21-6	S2248	CK2
CYC116	693228-63-6	S1171	Aurora Kinase
Cyclopamine	4449-51-8	S1146	Hedgehog
Cyclophosphamide monohydrate	6055-19-2	S2057	Alkylating agent
Cyclosporin A	59865-13-3	S2286	Calcineurin
Cyt387	1056634-68-4	S2219	JAK
Cytarabine	147-94-4	S1648	DNA/RNA Synthesis
Dacarbazine (DTIC-Dome)	4342-03-4	S1221	DNA/RNA Synthesis
Dacomitinib (PF299804,PF-00299804	1110813-31-4	S2727	ErbB
Dalcetrapib (JTT-705)	211513-37-0	S2772	CETP
Danusertib (PHA-739358)	827318-97-8	S1107	Aurora Kinase
Dapagliflozin	461432-26-8	S1548	SGLT
DAPT (GSI-IX)	208255-80-5	S2215	Beta Amyloid
Dasatinib (BMS-354825)	302962-49-8	S1021	Bcr-Abl
Daunorubicin HCl (Daunomycin HCl)	23541-50-6	S3035	Topoisomerase II
DCC-2036 (Rebastinib)	1020172-07-9	S2634	Bcr-Abl
Decitabine	2353-33-5	S1200	DNA Methyltransferase
Deforolimus (Ridaforolimus)	572924-54-0	S1022	mTOR
Desmethyl Erlotinib (CP-473420)	183321-86-0	S2826	EGFR
Dexamethasone	50-02-2	S1322	Anti-inflammatory
Dexamethasone acetate	1177-87-3	S3124	Anti-inflammatory

D-glutamine	5959-95-5	S1893	
Dimesna	16208-51-8	S1201	Uroprotectant
Disulfiram (Antabuse)	97-77-8	S1680	Acetaldehyde dehydrogenase
DMXAA (ASA404)	117570-53-3	S1537	VDA
Docetaxel (Taxotere)	114977-28-5	S1148	Microtubule stabilizer
Dorzolamide HCl	130693-82-2	S1375	Carbonic Anhydrase
Dovitinib (TKI-258)	405169-16-6	S1018	c-Kit
Doxercalciferol (Hectorol)	54573-75-0	S1467	Vitamin D
Doxorubicin (Adriamycin)	25316-40-9	S1208	Topoisomerase II
E7080 (Lenvatinib)	417716-92-8	S1164	VEGFR
Elesclomol	488832-69-5	S1052	HSP (inductor of HSP70)
ENMD-2076	934353-76-1	S1181	Aurora Kinase, Flt3/4
Entinostat (MS-275, SNDX-275)	209783-80-2	S1053	HDAC
Enzastaurin (LY317615)	170364-57-5	S1055	PKC
Epirubicin HCl	56390-09-1	S1223	Topoisomerase
Epothilone A	152044-53-6	S1297	Microtubule stabilizer
Epothilone B (EPO906)	152044-54-7	S1364	Microtubule stabilizer
Erlotinib HCl	183319-69-9	S1023	HER1/EGFR
Estradiol	50-28-2	S1709	Estrogen receptor
Estrone	53-16-7	S1665	Estrogen/progestogen Receptor
Etoposide (VP-16)	33419-42-0	S1225	Topoisomerase II
Everolimus (RAD001)	159351-69-6	S1120	mTOR
Evista (Raloxifene HCl)	82640-04-8	S1227	Estrogen/progestogen Receptor
EX 527	49843-98-3	S1541	Sirtuin
Exemestane	107868-30-4	S1196	Aromatase
Ezetimibe (Zetia)	163222-33-1	S1655	Cholesterol absorption
Febuxostat (Uloric)	144060-53-7	S1547	Xanthine oxidase
Fingolimod (FTY720)	162359-56-0	S5002	S1P receptor, PP2A activator
Flavopiridol (Alvocidib) HCl	131740-09-5	S2679	CDK
Floxuridine	50-91-9	S1299	DNA/RNA Synthesis
Fludarabine (Fludara)	21679-14-1	S1491	DNA/RNA Synthesis
Fludarabine Phosphate (Fludara)	75607-67-9	S1229	DNA/RNA Synthesis
Flutamide (Eulexin)	13311-84-7	S1908	Androgen receptor
Fluvastatin sodium (Lescol)	93957-55-2	S1909	HMG-CoA Reductase
Formestane	566-48-3	S2208	Steroidal aromatase
Ftorafur	17902-23-7	S1300	DNA/RNA Synthesis
Fulvestrant (Faslodex)	129453-61-8	S1191	Estrogen receptor antagonist
Ganetespib (STA-9090)	888216-25-9	S1159	HSP (e.g. HSP90)
GDC-0879	905281-76-7	S1104	Raf
GDC-0941	957054-30-7	S1065	PI3K
Gefitinib (Iressa)	184475-35-2	S1025	EGFR
Geldanamycin	30562-34-6	S2713	HSP90
Gemcitabine (Gemzar)	95058-81-4	S1714	DNA synthesis
Gemcitabine HCl (Gemzar)	122111-03-9	S1149	DNA synthesis
Gossypol	303-45-7	S2303	Dehydrogenase inhibitor
GSK1120212 (Trametinib)	871700-17-3	S2673	MEK
GSK1904529A	1089283-49-7	S1093	IGF-1R
GSK2126458	1086062-66-9	S2658	mTOR
GSK461364	929095-18-1	S2193	PLK
GSK690693	937174-76-0	S1113	Akt
GW3965 HCl	405911-17-3	S2630	Liver X Receptor
GW4064	278779-30-9	S2782	FXR agonist
Hydrocortisone (Cortisol)	50-23-7	S1696	Anti-inflammatory

Hydroxyurea (Cytodrox)	127-07-1	S1896	DNA synthesis
IC-87114	371242-69-2	S1268	PI3K
Idarubicin HCl	57852-57-0	S1228	Topoisomerase II
Ifosfamide	3778-73-2	S1302	DNA/RNA Synthesis
Imatinib (Gleevec)	152459-95-5	S2475	PDGFR
Imatinib Mesylate	220127-57-1	S1026	Bcr-Abl
Imiquimod	99011-02-6	S1211	Toll-like receptor 7 agonist
Iniparib (BSI-201)	160003-66-7	S1087	PARP
INK 128 (MLN0128)	1224844-38-5	S2811	mTOR
Irinotecan	97682-44-5	S1198	Topoisomerase I
Irinotecan HCl Trihydrate (Campto)	136572-09-3	S2217	Topoisomerase I
Isotretinoin	4759-48-2	S1379	Possible RAR and RXR agonist
Ispinesib (SB-715992)	336113-53-2	S1452	Kinesin
Itraconazole (Sporanox)	84625-61-6	S2476	Antifungal agent
JNJ 26854165 (Serdemetan)	881202-45-5	S1172	E3 Ligase
JNJ-26481585	875320-29-9	S1096	HDAC
JNJ-38877605	943540-75-8	S1114	c-Met
JNJ-7706621	443797-96-4	S1249	CDK/Aurora Kinase
Ku-0063794	938440-64-3	S1226	mTOR
KU-55933	587871-26-9	S1092	ATM/ATR
KU-60019	925701-49-1	S1570	ATM/ATR
KX2-391	897016-82-9	S2700	Src
Lapatinib	231277-92-2	S2111	EGFR
Lapatinib Ditosylate (Tykerb)	388082-77-7	S1028	EGFR
LDE225 (NVP-LDE225, Erismodegib)	956697-53-3	S2151	Smoothened
LDN193189	1062368-24-4	S2618	Alk2/3
Lenalidomide (Revlimid)	191732-72-6	S1029	TNF-alpha
Letrozole	112809-51-5	S1235	Aromatase
Leucovorin Calcium	6035-45-6	S1236	Folic acid analog
Linifanib (ABT-869)	796967-16-3	S1003	KDR, CSF-1R, FLT1/3, PDGFR
Linsitinib (OSI-906)	867160-71-2	S1091	IGF-1R
Lomustine (CeeNU)	13010-47-4	S1840	Alkylating agent
Lonidamine	50264-69-2	S2610	Glycolysis
LY2109761	700874-71-1	S2704	TGF-beta/Smad
LY2157299	700874-72-2	S2230	TGF-beta/Smad
LY2228820	862507-23-1	S1494	p38 MAPK
LY2603618 (IC-83)	911222-45-2	S2626	CHK1
LY294002	154447-36-6	S1105	PIK3K
Maraviroc	376348-65-1	S2003	CCR5
Masitinib (AB1010)	790299-79-5	S1064	c-Kit
MDV3100 (Enzalutamide)	915087-33-1	S1250	Androgen Receptor
Medroxyprogesterone acetate	71-58-9	S2567	Estrogen/progestogen Receptor
Megestrol Acetate	595-33-5	S1304	Androgen Receptor
Mercaptopurine	50-44-2	S1305	DNA/RNA Synthesis
Mesna (Uromitexan, Mesnex)	19767-45-4	S1735	Chemotherapy adjuvant
Mifepristone (Mifeprex)	84371-65-3	S2606	Estrogen/progestogen Receptor
Mitoxantrone HCl	70476-82-3	S2485	Folic acid antagonist
MK-0752	471905-41-6	S2660	Beta Amyloid
MK-1775	955365-80-7	S1525	Wee1
MK-2206 2HCl	1032350-13-2	S1078	Akt
MLN2238	1072833-77-2	S2180	Proteasome
MLN8237 (Alisertib)	1028486-01-2	S1133	Aurora Kinase
MLN9708	1201902-80-8	S2181	Proteasome

Mocetinostat (MGCD0103)	726169-73-9	S1122	HDAC
Motesanib Diphosphate (AMG-706)	857876-30-3	S1032	VEGF/PDGF/c-Kit/RET
Mycophenolate mofetil (CellCept)	128794-94-5	S1501	Inosine monophosphate dehydr
Mycophenolic acid (Mycophenolate)	24280-93-1	S2487	Inosine monophosphate dehydr
Nelarabine (Arranon)	121032-29-9	S1213	DNA/RNA Synthesis
Neratinib (HKI-272)	698387-09-6	S2150	HER2/EGFR
Nilotinib (AMN-107)	641571-10-0	S1033	Bcr-Abl
Nocodazole	31430-18-9	S2775	Microtubule inhibitor
NU7441 (KU-57788)	503468-95-9	S2638	DNA-PK
Nutlin-3	890090-75-2	S1061	E3 Ligase
NVP-BSK805 2HCl	1092499-93-8 (free base)	S2686	JAK
Obatoclax mesylate (GX15-070)	803712-79-0	S1057	BCL2
Olaparib (AZD2281)	763113-22-0	S1060	PARP
OSI-420	183320-51-6	S2205	EGFR
OSI-930	728033-96-3	S1220	c-Kit/FLT/KDR/c-Raf/Lck
Ostarine (MK-2866)	841205-47-8	S1174	Androgen Receptor
Oxaliplatin (Eloxatin)	61825-94-3	S1224	DNA/RNA Synthesis
PAC-1	315183-21-2	S2738	Caspase
Paclitaxel (Taxol)	33069-62-4	S1150	Microtubule stabilizer
Palomid 529	914913-88-5	S2238	PI3K/Akt/mTOR
Pamidronate Disodium	57248-88-1	S1311	Prevents osteoporosis
Pazopanib	444731-52-6	S3012	c-Kit
PCI-24781	783355-60-2	S1090	HDAC
PCI-32765 (Ibrutinib)	936563-96-1	S2680	BTK
PD 0332991 (Palbociclib) HCl	827022-32-2	S1116	CDK
PD0325901	391210-10-9	S1036	MEK
PD153035 HCl	183322-45-4	S1079	EGFR
PD173074	219580-11-7	S1264	FGFR
Pelitinib (EKB-569)	257933-82-7	S1392	EGFR
Pemetrexed (Alimta)	150399-23-8	S1135	DHFR
PF 573228	869288-64-2	S2013	FAK
PF-03814735	942487-16-3	S2725	Aurora Kinase
PF-04217903	956905-27-4	S1094	c-Met
PF-3845	1196109-52-0	S2666	FAAH
PF-562271	717907-75-0	S2890	FAK
PH-797804	586379-66-0	S2726	p38 MAPK
PHA-665752	477575-56-7	S1070	c-Met
PHA-793887	718630-59-2	S1487	CDK
Phloretin (Dihydronaringenin)	60-82-2	S2342	SGLT1/2
PI-103	371935-74-9	S1038	DNA-PK/PI3K/mTOR
PIK-75	372196-77-5	S1205	PI3K/DNA-PK
PIK-90	677338-12-4	S1187	PI3K
PIK-93	593960-11-3	S1489	PI3K/PI4K
Pioglitazone (Actos)	111025-46-8	S2590	PPAR activator
Pomalidomide	19171-19-8	S1567	TNF-alpha
Ponatinib (AP24534)	943319-70-8	S1490	Bcr-Abl
Prednisone (Adasone)	53-03-2	S1622	Immunosuppressant
Procarbazine HCl (Matulane)	366-70-1	S1995	DNA/RNA Synthesis
Quercetin (Sophoretin)	117-39-5	S2391	Possible CYP2C8 inhibitor
Quizartinib (AC220)	950769-58-1	S1526	FLT3
R406 (free base)	841290-80-0	S1533	Syk
R935788 (Fostamatinib disodium, R7	1025687-58-4	S2206	Syk
Raltitrexed (Tomudex)	112887-68-0	S1192	DNA/RNA Synthesis

Ranolazine (Ranexa)	95635-55-5	S1799	Calcium levels
Rapamycin (Sirolimus)	53123-88-9	S1039	mTOR
Regorafenib (BAY 73-4506)	755037-03-7	S1178	VEGFR/c-Kit/RET/Raf-1
RG108	48208-26-0	S2821	DNA Methyltransferase
Rigosertib (ON-01910)	1225497-78-8	S1362	PLK
Roscovitine (Seliciclib, CYC202)	186692-46-6	S1153	CDK
Rosiglitazone (Avandia)	122320-73-4	S2556	PPAR
Rucaparib (AG-014699 , PF-0136733	459868-92-9	S1098	PARP
Ruxolitinib (INCB018424)	941678-49-5	S1378	JAK
Salinomycin (Procoxacin)	53003-10-4	S2352	Antibiotic
Saracatinib (AZD0530)	379231-04-6	S1006	Src/Bcr-Abl
SB 203580	152121-47-6	S1076	p38 MAPK
SB 216763	280744-09-4	S1075	GSK-3
SB 431542	301836-41-9	S1067	ALK
SB 525334	356559-20-1	S1476	ALK
SB 743921	940929-33-9	S2182	Kinesin
SB590885	405554-55-4	S2220	B-Raf
SB939 (Pracinostat)	929016-96-6	S1515	HDAC
SGI-1776 free base	1025065-69-3	S2198	Pim
SGX-523	1022150-57-7	S1112	c-Met
Simvastatin (Zocor)	79902-63-9	S1792	HMG-CoA reductase
Sirtinol	410536-97-9	S2804	Sirtuin
SNS-032 (BMS-387032)	345627-80-7	S1145	CDK
SNS-314 Mesylate	1146618-41-8	S1154	Aurora Kinase
Sodium butyrate	156-54-7	S1999	HDACs
Sorafenib (Nexavar)	475207-59-1	S1040	Raf-1/B/Raf/VEGFR
Sotrastaurin (AEB071)	425637-18-9	S2791	PKC
SRT1720	1001645-58-4	S1129	Sirtuin
S-Ruxolitinib	941685-37-6	S2902	JAK
STF-62247	315702-99-9	S1041	May activate autophagy
Streptozotocin (Zanosar)	18883-66-4	S1312	DNA synthesis
SU11274	658084-23-2	S1080	c-Met
Sunitinib Malate (Sutent)	341031-54-7	S1042	PDGFR/VEGFR/c-Kit
TAE684 (NVP-TAE684)	761439-42-3	S1108	ALK
TAK-733	1035555-63-5	S2617	MEK
TAME	901-47-3	S2225	APC
Tamoxifen Citrate (Nolvadex)	54965-24-1	S1972	Strogen receptor
Tandutinib (MLN518)	387867-13-2	S1043	FLT3/c-Kit/PDGFR
Telatinib (BAY 57-9352)	332012-40-5	S2231	FLT3/c-Kit/PDGFR
Temozolomide	85622-93-1	S1237	Alkylating agent
Temsirolimus (Torisel)	162635-04-3	S1044	mTOR
Teniposide (Vumon)	29767-20-2	S1787	DNA damaging agent
TG101348 (SAR302503)	936091-26-8	S2736	JAK
Thalidomide	50-35-1	S1193	immunomodulatory agent
Tie2 kinase inhibitor	948557-43-5	S1577	Tie-2
Tipifarnib (Zarnestra)	192185-72-1	S1453	Farnesyltransferase
Tivozanib (AV-951)	475108-18-0	S1207	VEGFR1/2/3; c-Kit; PDGFR
Tofacitinib (CP-690550, Tasocitinib)	477600-75-2	S2789	JAK
Topotecan HCl	119413-54-6	S1231	Topoisomerase I
Toremifene Citrate (Fareston, Acapo	89778-27-8	S1776	Estrogen/progestogen Receptor
Torin 1	1222998-36-8	S2827	mTOR
Torin 2	1223001-51-1	S2817	mTOR, ATM/ATR
Tosedostat (CHR2797)	238750-77-1	S1522	Aminopeptidase

TPCA-1	507475-17-4	S2824	IκB/IKK
Tretinoin (Aberela)	302-79-4	S1653	Retinoic acid
Triamcinolone Acetonide	76-25-5	S1628	Anti-inflammatory
Trichostatin A (TSA)	58880-19-6	S1045	HDAC
Triciribine (Triciribine phosphate)	35943-35-2	S1117	Akt
Triptolide	38748-32-2	S3604	Immunosuppressive
TW-37	877877-35-5	S1121	Bcl-2
Ubenimex (Bestatin)	58970-76-6	S1591	Protease inhibitor
Valproic acid sodium salt (Sodium v	1069-66-5	S1168	HDACs
Vandetanib (Zactima)	443913-73-3	S1046	VEGFR
Vatalanib 2HCl (PTK787)	212141-51-0	S1101	VEGFR2/KDR/Flt-1/c-Kit
Vemurafenib (PLX4032)	918504-65-1	S1267	B-Raf
Vinblastine	57-22-7	S1248	Microtubule inhibitor
Vincristine	2068-78-2	S1241	Microtubule inhibitor
Vinpocetine (Cavinton)	42971-09-5	S2110	Sodium Channel
Vismodegib (GDC-0449)	879085-55-9	S1082	Hedgehog
Vorinostat (SAHA)	149647-78-9	S1047	HDACs
VX-680 (MK-0457, Tozasertib)	639089-54-6	S1048	Aurora Kinase
WAY-362450	629664-81-9	S2694	FXR
WP1130	856243-80-6	S2243	DUB inhibitor
WYE-354	1062169-56-5	S1266	mTOR
WZ4002	1213269-23-8	S1173	EGFR
XAV-939	284028-89-3	S1180	Wnt/beta-catenin
XL147	956958-53-5	S1118	PI3K
XL765 (SAR245409)	1349796-36-6	S1523	mTOR
Y-27632 2HCl	129830-38-2	S1049	ROCK
YM155	781661-94-7	S1130	E3 Ligase IAP
YM201636	371942-69-7	S1219	PIP5K (Pikfyve)
YO-01027	209984-56-5	S2711	Gamma-secretase
Zibotentan (ZD4054)	186497-07-4	S1456	Endothelin Receptor
Zileuton	111406-87-2	S1443	5-lipoxygenase
ZSTK474	475110-96-4	S1072	PI3K

APPENDIX B

Supplementary data

The code for reproduction of the results was developed in R, version 3.6.2 and is available at a GitHub repository (https://github.com/Andreabrodersen/A-study-of-ex-vivo-drug-screen-survival-predictions-in-AML-patients-for-clinical-application.git).

Bibliography

[1] *Lifetime Risk of Developing or Dying From Cancer*. English. The American Cancer Society. Jan. 13, 2020. URL: https://www.cancer.org/cancer/cancer-basics/lifetime-probability-of-developing-or-dying-from-cancer.html (visited on 10/18/2020).

[2] *NCI dictionary of canver terms - cancer*. English. National Cancer Institute. URL: https://www.cancer.gov/publications/dictionaries/cancer-terms/def/cancer (visited on 10/18/2020).

[3] *Cancer Classification*. English. The American Cancer Society. URL: https://training.seer.cancer.gov/disease/categories/classification.html (visited on 10/18/2020).

[4] Lin, A. and Sheltzer, J. "Discovering and validating cancer genetic dependencies: approaches and pitfalls". In: *Nature Reviews Genetics* 21 (2020), pp. 671–682. URL: https://doi.org/10.1038/s41576-020-0247-7.

[5] Béné, M., Nebe, T., and Bettelheim, P. e. a. "Immunophenotyping of acute leukemia and lymphoproliferative disorders: a consensus proposal of the European LeukemiaNet Work Package 10". In: *Leukemia* 25 (Jan. 2011), pp. 567–574. URL: https://www.nature.com/articles/leu2010312#citeas.

[6] Roy, R., Chun, J., and Powell, S. N. "BRCA1 and BRCA2: different roles in a common pathway of genome protection". In: *Nature reviews. Cancer* 12 (1 2011), pp. 68–78. URL: https://doi.org/10.1038/nrc3181.

[7] Tutt, A. et al. "Oral poly(ADP-ribose) polymerase inhibitor olaparib in patients with BRCA1 or BRCA2 mutations and advanced breast cancer: a proof-of-concept trial". In: *The Lancet* 376.9737 (2010), pp. 235–244. URL: http://www.sciencedirect.com/science/article/pii/S0140673610608926.

[8] Hare, T. et al. "In vitro Activity of Bcr-Abl Inhibitors AMN107 and BMS-354825 against Clinically Relevant Imatinib-Resistant Abl Kinase Domain Mutants". In: *Cancer Research* 65.11 (2005), pp. 4500–4505. eprint: https://cancerres.aacrjournals.org/content/65/11/4500.full.pdf. URL: https://cancerres.aacrjournals.org/content/65/11/4500.

[9] Vogenberg, F. R., Isaacson Barash, C., and Pursel, M. "Personalized medicine: part 1: evolution and development into theranostics". English. In: *P and T : a peer-reviewed journal for formulary management* 35.10 (Oct. 2010), pp. 560–576. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2957753/ (visited on 10/30/2020).

[10] Hao, K. "What is machine learning". English. In: *MIT Technology Review* (Nov. 17, 2018). URL: https://www.technologyreview.com/2018/11/17/103781/what-is-machine-learning-we-drew-you-another-flowchart/ (visited on 10/29/2020).

[11] *Akutt myelogen leukemi (AML)*. Norwegian. Norsk Helseinformatikk (Norwegian health informatics). June 12, 2020. URL: https://nhi.no/sykdommer/barn/kreftsykdommer/akutt-myelogen-leukemi-oversikt/?page=2 (visited on 10/20/2020).

[12] Lowenberg, B., Downing, J. R., and Burnett, A. "Acute Myeloid Leukemia". In: *New England Journal of Medicine* 341.14 (1999). PMID: 10502596, pp. 1051–1062. eprint: https://doi.org/10.1056/NEJM199909303411407. URL: https://doi.org/10.1056/NEJM199909303411407.

[13] De Kouchkovsky, I. and Abdul-Hay, M. "Acute myeloid leukemia: a comprehensive review and 2016 update". In: *Blood cancer journal* 7.6 (June 2016). PMID: 27367478. eprint: https://doi.org/10.1038/bcj.2016.50. URL: https://doi.org/10.1038/bcj.2016.50.

[14] DiNardo, C. D. and Cortes, J. E. "Mutations in AML: prognostic and therapeutic implications. Hematology". In: *American Society of Hematology. Education Program, 2016* 1 (Dec. 2, 2016), pp. 348–355. eprint: https://doi.org/10.1182/asheducation-2016.1.348. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6142505/.

[15] Papaemmanuil, E. et al. "Genomic Classification and Prognosis in Acute Myeloid Leukemia". In: *New England Journal of Medicine* 374.23 (2016). PMID: 27276561, pp. 2209–2221. eprint: https://doi.org/10.1056/NEJMoa1516192. URL: https://doi.org/10.1056/NEJMoa1516192.

[16] Dombret, H. and Gardin, C. "An update of current treatments for adult acute myeloid leukemia". In: *Blood* 127.1 (Dec. 2016), pp. 53–61. eprint: https://doi.org/10.1182/blood-2015-08-604520. URL: https://doi.org/10.1182/blood-2015-08-604520.

[17] *PubChem Compound Summary for CID 6253, Cytarabine*. English. National Center for Biotechnology Information. 2020. URL: https://pubchem.ncbi.nlm.nih.gov/compound/Cytarabine (visited on 11/09/2020).

[18] *PubChem Compound Summary for CID 30323, Daunorubicin*. English. National Center for Biotechnology Information. 2020. URL: https://pubchem.ncbi.nlm.nih.gov/compound/Daunorubicin (visited on 11/09/2020).

[19] *Akutt myelogen leukemi (AML) - Behandling (Acute Myeloid Leukemia (AML) - treatment)*. Norwegian. Blodkreftforeningen (The Norwegian Blood Cancer Association). June 1, 2017. URL: https://www.blodkreftforeningen.no/akutt-myelogen-leukemi-aml/ (visited on 10/22/2020).

[20] *Treatment of Acute Promyelocytic Leukemia (APL)*. English. The American Cancer Society. Aug. 21, 2018. URL: https://www.cancer.org/cancer/acute-myeloid-leukemia/treating/m3-leukemia.html (visited on 11/15/2020).

[21] Pollyea, D. A. et al. "Venetoclax for AML: changing the treatment paradigm". In: *Blood Advances* 3.24 (Dec. 2019), pp. 4326–4335. eprint: https://ashpublications.org/bloodadvances/article-pdf/3/24/4326/1719143/advancesadv2019000937c.pdf. URL: https://doi.org/10.1182/bloodadvances.2019000937.

[22] *Typical Treatment of Acute Myeloid Leukemia (Except APL)*. English. The American Cancer Society. Jan. 13, 2020. URL: https://www.cancer.org/cancer/acute-myeloid-leukemia/treating/typical-treatment-of-aml.html (visited on 11/11/2020).

[23] *Spennende forskning i Norge på AML og ALL*. Norwegian. Blodkreftforeningen (The Norwegian Blood Cancer Association). Mar. 29, 2017. URL: https://www.blodkreftforeningen.no/2017/03/spennende-forskning-norge-pa-all-aml/ (visited on 11/11/2020).

[24] *Chemotherapy for Acute Myeloid Leukemia (AML)*. English. The American Cancer Society. Sept. 3, 2020. URL: https://www.cancer.org/cancer/acute-myeloid-leukemia/treating/chemotherapy.html (visited on 10/22/2020).

[25] Döhner, H. et al. "Diagnosis and management of acute myeloid leukemia in adults: recommendations from an international expert panel, on behalf of the European LeukemiaNet". In: *Blood* 115.3 (Jan. 2010), pp. 453–474. eprint: https://ashpublications.org/blood/article-pdf/115/3/453/1458437/zh800310000453.pdf. URL: https://doi.org/10.1182/blood-2009-07-235358.

[26] Döhner, H. et al. “Diagnosis and management of acute myeloid leukemia in adults: recommendations from an international expert panel, on behalf of the European LeukemiaNet”. In: *Blood* 115.3 (Jan. 2010), pp. 453–474. eprint: https://ashpublications.org/blood/article-pdf/115/3/453/1458437/zh800310000453.pdf. URL: https://doi.org/10.1182/blood-2009-07-235358.

[27] McClure et al. “Practical Acute Myeloid Leukemia (AML) Biomarker Testing Using Next-Generation Sequencing (NGS) Technology: A Comprehensive, Rapid, Inexpensive and Flexible Approach”. In: *Blood* 126.23 (Dec. 2015), pp. 4973–4973. URL: https://doi.org/10.1182/blood.V126.23.4973.4973.

[28] Gulley, M. L., Shea, T. C., and Fedoriw, Y. “Genetic tests to evaluate prognosis and predict therapeutic response in acute myeloid leukemia”. In: *The Journal of molecular diagnostics : JMD* 12.1 (Jan. 2010), pp. 3–16. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2797712/.

[29] Wu, M., Li, C., and Zhu, X. “FLT3 inhibitors in acute myeloid leukemia”. In: *Journal of Hematology and Oncology* 11.133 (Dec. 2018). URL: https://jhoonline.biomedcentral.com/articles/10.1186/s13045-018-0675-4#citeas.

[30] *Utredning av akutt leukemi*. Norwegian. Helsedirektroratet (The Norwegian Directorate of Health). Feb. 28, 2017. URL: https://www.helsedirektoratet.no/pakkeforlop/akutt-leukemi/utredning-av-akutt-leukemi#utredning (visited on 10/28/2020).

[31] Bennett, J. M. et al. “Proposals for the Classification of the Acute Leukaemias ”. In: *British Journal of Haematology* 33.4 (Jan. 1975), pp. 451–458. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2141.1976.tb03563.x.

[32] Segeren, C. M. and ’t Veer, M. B. van. “The FAB classification for acute myeloid leukaemia–is it outdated?” In: *The Netherlands journal of medicine* 49.3 (Jan. 1996), pp. 126–131. URL: https://doi.org/10.1016/0300-2977(96)00024-1.

[33] Mandal, A. *Acute Myeloid Leukemia Classification*. English. News-Medical. Feb. 26, 2019. URL: https://www.news-medical.net/health/Acute-Myeloid-Leukemia-Classification.aspx (visited on 11/15/2020).

[34] *Acute Myeloid Leukemia (AML) Subtypes and Prognostic Factors*. English. The American Cancer Society. Aug. 21, 2018. URL: https://www.cancer.org/cancer/acute-myeloid-leukemia/detection-diagnosis-staging/how-classified.html (visited on 10/28/2020).

[35] *Acute Myeloid Leukemia (AML) Subtypes and Prognostic Factors*. English. The American Cancer Society. Aug. 21, 2018. URL: https://www.cancer.org/cancer/acute-myeloid-leukemia/detection-diagnosis-staging/how-classified.html (visited on 11/15/2020).

[36] Arber, D. A. et al. “The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia”. In: *Blood* 127.20 (May 2016), pp. 2391–2405. eprint: https://doi.org/10.1182/blood-2016-03-643544. URL: https://doi.org/10.1182/blood-2016-03-643544.

[37] Hess, C. J. et al. “Hypermethylation of the FANCC and FANCL promoter regions in sporadic acute leukaemia”. In: *Cellular oncology : the official journal of the International Society for Cellular Oncology* 30.4 (2008), pp. 299–306.

[38] Döhner, H. et al. “Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel”. In: *Blood* 129.4 (2017), pp. 424–447. URL: https://doi.org/10.1182/blood-2016-08-733196.

[39] Spinner, M. A. et al. “Ex vivo drug screening defines novel drug sensitivity patterns for informing personalized therapy in myeloid neoplasms”. In: *Blood Advances* 4.12 (June 2020), pp. 2768–2778. eprint: https://ashpublications.org/bloodadvances/article-pdf/4/12/2768/1745752/advancesadv2020001934.pdf. URL: https://doi.org/10.1182/bloodadvances.2020001934.

[40] Brooks, E. A. et al. “Applicability of Drug Response Metrics for Cancer Studies using Biomaterials”. In: *bioRxiv* (2018). eprint: https://www.biorxiv.org/content/early/2018/09/07/408583.full.pdf. URL: https://www.biorxiv.org/content/early/2018/09/07/408583.

[41] Pozdeyev, N. et al. “Integrating heterogeneous drug sensitivity data from cancer pharmacogenomic studies”. In: *Oncotarget* 7.32 (June 2016), pp. 51619–51625. URL: https://doi.org/10.18632/oncotarget.10010.

[42] Pemovska, T. et al. “Individualized Systems Medicine Strategy to Tailor Treatments for Patients with Chemorefractory Acute Myeloid Leukemia”. In: *Cancer Discovery* 3.12 (2013), pp. 1416–1429. eprint: https://cancerdiscovery.aacrjournals.org/content/3/12/1416.full.pdf. URL: https://cancerdiscovery.aacrjournals.org/content/3/12/1416.

[43] Arruebo, M. et al. “Assessment of the evolution of cancer treatment therapies”. English. In: *Cancers* 3.3 (Sept. 2011), pp. 3279–3330. URL: https://doi.org/10.3390/cancers3033279 (visited on 10/30/2020).

[44] Dugger, S. A., Platt, A., and Goldstein, D. B. “Drug development in the era of precision medicine”. English. In: *Nature reviews. Drug discovery* 17.3 (Dec. 2017), pp. 183–196. URL: https://doi.org/10.1038/nrd.2017.226 (visited on 10/30/2020).

[45] Sharma, N., Jain, V., and Mishra, A. “An Analysis Of Convolutional Neural Networks For Image Classification”. In: *Procedia Computer Science* 132 (2018). International Conference on Computational Intelligence and Data Science, pp. 377–384. URL: http://www.sciencedirect.com/science/article/pii/S1877050918309335.

[46] Yoo, S. H. et al. “Deep Learning-Based Decision-Tree Classifier for COVID-19 Diagnosis From Chest X-ray Imaging”. In: *Frontiers in Medicine* 7 (2020), p. 427. URL: https://www.frontiersin.org/article/10.3389/fmed.2020.00427.

[47] Nematolahi, S. and Ayatollahi, S. “A comparison of breast cancer survival among young, middle-aged, and elderly patients in southern Iran using Cox and empirical Bayesian additive hazard models”. In: *Epidemiology and health* 39 (2017). URL: https://doi.org/10.4178/epih.e2017043.

[48] Opinto, G. et al. “Hierarchical clustering analysis identifies metastatic colorectal cancers patients with more aggressive phenotype”. In: *Oncotarget* 8 (50 2017), pp. 87782–87794. URL: https://doi.org/10.18632/oncotarget.21213.

[49] Sleiman, P. M. A. et al. “GWAS identifies four novel eosinophilic esophagitis loci”. In: *Nature Communicatoins* 5 (2014). URL: https://www.nature.com/articles/ncomms6593?l=center&ci=9&cp=title&cc=1801f8f3768f38ff13019b74571b72ca#citeas.

[50] Kourou, K. et al. “Machine learning applications in cancer prognosis and prediction”. In: *Computational and Structural Biotechnology Journal* 13 (2015), pp. 8–17. URL: http://www.sciencedirect.com/science/article/pii/S2001037014000464.

[51] Döhner, H. et al. “Genomic Aberrations and Survival in Chronic Lymphocytic Leukemia”. In: *New England Journal of Medicine* 343.26 (2000). PMID: 11136261, pp. 1910–1916. eprint: https://doi.org/10.1056/NEJM200012283432602. URL: https://doi.org/10.1056/NEJM200012283432602.

[52] Genovese, G. et al. “Clonal Hematopoiesis and Blood-Cancer Risk Inferred from Blood DNA Sequence”. In: *New England Journal of Medicine* 371.26 (2014). PMID: 25426838, pp. 2477–2487. eprint: https://doi.org/10.1056/NEJMoa1409405. URL: https://doi.org/10.1056/NEJMoa1409405.

[53] *WMA DECLARATION OF HELSINKI – ETHICAL PRINCIPLES FOR MEDICAL RESEARCH INVOLVING HUMAN SUBJECTS*. Wold Medical Association, July 9, 2018. URL: https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/.

[54] *Lov om behandling av personopplysninger (personopplysningsloven)*. Norwegian Ministry of Justice and Public Security, July 20, 2018. URL: https://lovdata.no/dokument/NL/lov/2018-06-15-38/KAPITTEL_1%C2%A71.

[55] *Lov om helseregistre og behandling av helseopplysninger (helseregisterloven)*. Norwegian Ministry of Health and Care Services, Jan. 1, 2015. URL: https://lovdata.no/dokument/SF/forskrift/2007-12-07-1389/KAPITTEL_1%C2%A71-5a.

[56] Durán, P. A. "Unpublished, Personal Comunication". unpublished. 2018.

[57] Subramanian, D. *A Simple Introduction to K-Nearest Neighbors Algorithm*. English. June 8, 2018. URL: https://towardsdatascience.com/a-simple-introduction-to-k-nearest-neighbors-algorithm-b3519ed98e (visited on 11/02/2020).

[58] Konopka, T. *Uniform Manifold Approximation and Projection*. R package version 0.2.6.0. 2020. URL: https://CRAN.R-project.org/package=umap.

[59] Carlson, M. *Genome wide annotation for Human*. R package version 3.10.0. 2019.

[60] G., Y. et al. "clusterProfiler: an R package for comparing biological themes among gene clusters". In: *OMICS: A Journal of Integrative Biology* 16.5 (2012). R package version 3.14.3, pp. 284–287. URL: http://online.liebertpub.com/doi/abs/10.1089/omi.2011.0118.

[61] Y., C. et al. ""ChemmineR: a compound mining framework for R"". In: *Bioinformatics* 24.15 (2008). R package version 3.38.0, pp. 1733–1734. URL: https://doi.org/10.1093/bioinformatics/btn307.

[62] Franceschini, A. *"STRING v9.1: protein-protein interaction networks, with increased coverage and integration"*. R package version 2.2.0. 2013. URL: https://bioc.ism.ac.jp/packages/2.14/bioc/html/STRINGdb.html.

[63] Friedman, J., Hastie, T., and Tibshirani, R. "Regularization Paths for Generalized Linear Models via Coordinate Descent". In: *Journal of Statistical Software* 33.1 (2010), pp. 1–22. URL: http://www.jstatsoft.org/v33/i01/.

[64] Therneau, T. M. *A Package for Survival Analysis in R*. R package version 2.37-7. 2014. URL: http://CRAN.R-project.org/package=survival.

[65] Kassambara, A. et al. *survminer: Drawing Survival Curves using ggplot2*. R package version 0.4.8. 2020. URL: https://CRAN.R-project.org/package=survminer.

[66] Zhao, S. et al. *heatmap3: An Improved Heatmap Package*. R package version 1.1.7. 2020. URL: https://CRAN.R-project.org/package=heatmap3.

[67] DiNardo, C. D. et al. "Venetoclax combined with decitabine or azacitidine in treatment-naive, elderly patients with acute myeloid leukemia". In: *Blood* 133.1 (Jan. 2019), pp. 7–17. eprint: https://ashpublications.org/blood/article-pdf/133/1/7/1551592/blood868752.pdf. URL: https://doi.org/10.1182/blood-2018-08-868752.

[68] Schimmer, A. D. et al. "A multicenter phase I/II study of obatoclax mesylate administered as a 3- or 24-hour infusion in older patients with previously untreated acute myeloid leukemia". In: *PLOS ONE* 9.10 (Jan. 2014). URL: https://doi.org/10.1371/journal.pone.0108694.

[69] Roma, A. and Spagnuolo, P. A. "Estrogen Receptors Alpha and Beta in Acute Myeloid Leukemia". In: *Cancers* 12.4 (2020), p. 907.

[70] Ling, Y. et al. "Protein Kinase Inhibitors as Therapeutic Drugs in AML: Advances and Challenges". In: *Current pharmaceutical design* 23.29 (2017), pp. 4303–4310.

[71] Copeland, B. *Artificial intelligence*. English. Aug. 11, 2020. URL: https://www.britannica.com/technology/artificial-intelligence (visited on 10/29/2020).

[72] Zwass, V. *Neural Network*. English. Jan. 9, 2020. URL: https://www.britannica.com/technology/neural-network (visited on 10/29/2020).

[73] James, G. et al. “Section 2.2.2 The bias-variance trade-off”. In: *In An introduction to statistical learning with applications in R*. New York, New York, USA: Springer, 2013. Chap. 2.2.2, pp. 33–36.

[74] *Inference*. English. Merriam-Webster. Oct. 10, 2020. URL: https://www.merriam-webster.com/dictionary/inference (visited on 10/30/2020).

[75] Ritz, C. et al. “Dose-Response Analysis Using R”. Version 3.6.0. In: *PLOS ONE* 10.e0146021 (12 2015). URL: http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0146021.

[76] Hess, C. J. et al. “Hypermethylation of the FANCC and FANCL promoter regions in sporadic acute leukaemia”. In: *Cellular oncology* 30.4 (2008), pp. 299–306. URL: https://doi.org/10.3233/clo-2008-0426.

[77] Döhner, K. and Döhner, H. “Molecular characterization of acute myeloid leukemia”. In: 93.7 (July 2008), pp. 976–982. URL: https://haematologica.org/article/view/4939.

[78] Jang, I. S. et al. “Systematic assessment of analytical methods for drug sensitivity prediction from cancer cell line data”. In: *Pacific Symposium on Biocomputing* (2014), pp. 63–74. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3995541/.

[79] Glen, S. *Z-Score: Definition, Formula and Calculation*. English. StatisticsHowTo.com. URL: https://www.statisticshowto.com/probability-and-statistics/z-score/ (visited on 11/15/2020).

[80] Therneau, T. M. *A Package for Survival Analysis in R*. R package version 3.2-7. 2020. URL: https://CRAN.R-project.org/package=survival.

[81] *The R Project for Statistical Computing*. English. Version 3.6.0. The R Foundation. Apr. 26, 2019. URL: https://www.r-project.org/about.html.

[82] *RStudio*. English. Version 1.2.1335. Apr. 8, 2019. URL: www.rstudio.com.

[83] Mishra, D. *Weight Decay == L2 Regularization?* English. Medium, Towards data science. May 9, 2020. URL: https://towardsdatascience.com/weight-decay-l2-regularization-90a9e17713cd (visited on 10/02/2020).