universität wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Impact of checkpoint inhibitor therapies on the immunogenicity status of patients participating in Roche clinical immunotherapy studies“

verfasst von / submitted by

Péter Hunyadi, BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2023 / Vienna, 2023

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 066 875
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Bioinformatik
Betreut von / Supervisor:	Univ.-Prof. Mag. Dr. Thomas Rattei

Abstract

Background

Checkpoint Inhibitor (CPI)-based immunotherapies which target the CTLA-4 or PD-1/PD-L1 pathway have achieved impressive success in the treatment of different cancer types. Testing novel cancer immunotherapies in early clinical (Phase 1) studies face the situation that a large number of patients today have been previously exposed to those CPI-based therapies. Little is known about the impact of past CPI-treatments on a patient's response to novel anti-cancer therapeutic molecules. Understanding such impact is important to recruit the "right" patients for testing novel cancer immunotherapies in early clinical studies. Studying the immunogenicity of patients against the studied molecules is therefore important as this property is linked with the effectiveness of the applied immunotherapies.

To this end, data has been collected from a large virtual cohort of over 1800 patients with more than 20 different cancer types who participated in one of 16 Phase 1 immunotherapy clinical trials at Roche.

Results

In this thesis paper I present the descriptive analysis of this patient population, enrichment of immunophenotype information with the help of gene signatures and thorough analysis of treatment history. Afterwards, multiple correspondence analysis (MCA) takes place on relevant properties to subset the samples and finally correlation analysis is performed on these subsets. Possible relationships between aspects of the treatment history and immunophenotype/ROPRO are discussed and, considering these, some recommendations are provided for possible work in the future that could extend the observations made here and and can support the decision on patient eligibility criteria for oncologic immunotherapy Phase 1 clinical trials.

Keywords: Checkpoint Inhibitors, CD8 Immunophenotype, ROPRO, Multiple Correspondence Analysis

Zusammenfassung

Hintergrund

Auf Checkpoint-Inhibitoren (CPI) basierende Immuntherapien, die auf die CTLA-4 - oder PD-1/PD-L1 Signalwege abzielen, haben beeindruckende Erfolge bei der Behandlung verschiedener Krebsarten geleistet. Die Situation vom Testen neuartiger Krebsimmuntherapien in frühen klinischen Studien (Phase-1) ist so, dass eine große Anzahl von Patienten schon zuvor mit CPI-basierten Therapien behandelt wurde. Es ist aber nicht wohl bekannt, wie frühere CPI-Behandlungen die Reaktion eines Patienten auf neuartige anti-Krebs therapeutische Moleküle beeinflussen. Es wäre bedeutend, um diese Auswirkungen zu verstehen, die „richtige“ Patienten für die Erprobung neuartiger Krebsimmuntherapien in frühen klinischen Studien zu finden. Die Untersuchung der Immunogenität von Patienten gegen die studierte Moleküle ist daher wichtig, da diese Eigenschaft mit der Wirksamkeit der angewandten Immuntherapien verknüpft ist.

Zu diesem Zweck wurden Daten aus einer großen virtuellen Kohorte von über 1800 Patienten mit mehr als 20 verschiedenen Krebsarten gesammelt, die an einer von 16 klinischen Phase-1-Studien zur Immuntherapie bei Roche teilgenommen haben.

Ergebnisse

In dieser Arbeit stelle ich die deskriptive Analyse dieser Patientenpopulation, die Anreicherung von Immunphänotyp Informationen mit Hilfe von Gensignaturen und eine gründliche Analyse der Behandlungsgeschichte vor. Danach findet eine Mehrfache Korrespondenzanalyse (MCA) für relevante Eigenschaften statt, um die Proben zu unterteilen, und schließlich wird eine Korrelationsanalyse für diese Untergruppen durchgeführt. Potenzielle Zusammenhänge zwischen Aspekten der Behandlungsgeschichte und Immunphänotyp/ROPRO werden diskutiert und während ich sie betrachte, gebe ich einige Empfehlungen für mögliche zukünftige Arbeit, die die hier gemachten Beobachtungen erweitern und die Entscheidung über Auswahlkriterien der Patienten für onkologische Immuntherapie Phase-1-Studien unterstützen können.

Schlüsselwörter: Checkpoint-Inhibitoren, CD8 Immunphänotyp, ROPRO, Mehrfache Korrespondenzanalyse

Contents

Introduction

Checkpoint Inhibitor (CPI)-based immunotherapies which target the CTLA4 or PD1/PD-L1 pathway have achieved impressive success in the treatment of different cancer types. The number of patients with cancer who receive immune checkpoint-based therapies is rapidly increasing following a growing number of approved treatments.

Testing novel cancer immunotherapies in early clinical (Phase 1) studies face the situation that a large number of patients today have been previously exposed to those CPI-based therapies. Little is known about the impact of past CPI-treatments on a patient's response to novel therapeutic molecules against cancer[1]. Understanding such impact is important to recruit the “right” patients for testing novel cancer immunotherapies in early clinical studies. Studying the immunogenicity of patients against the studied molecules is therefore important as this property is linked with the effectiveness of the applied immunotherapies.

To this end, data has been collected from a large virtual cohort of over 1800 patients with more than 20 different cancer types who participated in one of 16 still ongoing Phase 1 immunotherapy clinical trials at Roche. The data has been integrated, harmonized and merged into the Roche Oncology Trial Dataset (ROTD) and includes medical history, concomitant medication, lab parameters, and vital signs information as well as RNA-seq data derived from tumor biopsies.

The main goal of the project is to learn about the effect of previous CPI on the immunogenicity status of patients. To this end, the ROTD dataset will be investigated, as it holds sufficient information to examine the effect of CPI, because a large portion of patients have been previously treated with CPI, often with multiple different ones. However, information about the response to the studied molecules is not available, therefore the project relies on cancer immunophenotypes[2] as an estimator of immunogenicity and a prognostic score of overall survival, ROPRO[3]. A comprehensive descriptive analysis will be performed to be able to focus on the relevant properties and utilizing the available information to find relationships with the two predictors.

Background

Checkpoint Inhibition Therapies

Immune Checkpoint Inhibitors (CPI) are a group of cancer immunotherapeutic agents, which gained enormous interest in the past decade for their spectacular success against certain cancer types for some patients. They exhibit their function by blocking receptor-ligand interactions between T cells and other immune cells or tumor cells, which would downregulate T cell response. By strengthening the patient's own immune system, a more precise, longer-lasting and for the body less exhausting attack can be elicited against the cancer cells. However this approach is not without challenges, which will be also briefly explained in the following chapter.

Current CPI strategies are based on inhibiting the CTLA-4 and PD-1 immune checkpoints. The inhibition of these cell surface molecules with antibodies has already been studied in the late 1990s as means for upregulating immune response against tumors in mice models. The first molecule of this methodology to conclude phase 3 trials in 2010, ipilimumab, an antibody targeting CTLA-4, showed that it would highly improve median overall survival (OS) for late-stage melanoma [4] and was rapidly approved by the United States Food and Drug Administration (FDA). Not long after, monoclonal antibodies to block another immune checkpoint, PD-1 (pembrolizumab and nivolumab) or its ligand, PD-L1 (atezolizumab and durvalumab) were also approved as 1st and 2nd line treatments against multiple malignancies. These demonstrated higher response rates and fewer side effects than anti-CTLA-4.

The most notable success of CPI is its ability to elicit an unprecedented durable clinical response. Long-term remission (5+ years), even after discontinuing medications, is well documented in cases of melanoma, especially after the disappearance of all visible metastases. However, it soon became clear that other cancer types do not respond quite as well and the majority of patients do not respond (primary resistance), or relapse after a certain time (acquired resistance). There has been up to 60% non-responders in certain tumor types [5] and the incidence of acquired resistance rises, as immunotherapy becomes more widely utilized, even in generally well-responding cancer such as advanced melanoma, where after treatment, approximately one-fourth to one-third relapses [6]. Immune-related adverse events (irAEs) can also complicate the therapy, some of which can be even life-threatening. The enhancement of general T cell function is likely to damage healthy tissue as well as tumor cells, therefore it is of utmost importance to identify patients with a high chance for adverse effects to be able to appropriately treat them [7].

To tackle the problems of CPI, a more holistic view of cancer is required, where the disease is considered as an interaction between tumor cells and the immune system. This relationship is incredibly complex and multifactorial and our current understanding of it is still not sufficient. This is especially

the case as different tumor types (indications) react very differently to treatment as cancer specific pathways govern treatment response [8].

Some key factors of the so-called “cancer immunogram” have been established, however the impact of them varies from patient to patient and a lot of yet unknown factors still exist [9]. These immune parameters include: the amount of T cells infiltrating the tumor and how much they express immune checkpoint inhibitors, the tumor microenvironment (TME) and tumor mutational burden (TMB). The latter was considered a negative factor before, but in terms of CPI, higher TMB forces the tumor cells to present more neoantigens which give the immune system different attacking points against the cancer. Overall, the need for effective biomarker-based selection of patients has been expressed, but the number of viable predictive biomarkers is limited[10,11].

In recent years, the combination of different CPIs[12] and CPIs with other treatments, such as low-dose radiotherapy [13], have been shown to increase response rate and OS compared to CPI monotherapies and have garnered more solid proof with consecutive, long-term trials. Currently CPIs are also investigated to be used as adjuvants in other indications. In the next few years, CPI is going to be used in new combinations with therapeutic agents and they will target other subsets of cells of the immune system and they will be tested across all areas of oncology. In-depth immune monitoring will allow new hypotheses for the application of CPI which will be verified in appropriate preclinical models [11].

Patient cohort and cancer treatment history

In order to be able to conduct large-scale descriptive analysis, data has been collected and merged into a large virtual cohort, the Roche Oncology Trial Dataset (ROTD). It holds information about over 1800 patients with more than 20 different cancer types who participated in one of 16 Phase 1 immunotherapy clinical trials at Roche. The data has been integrated, harmonized and includes medical history, concomitant medication, lab parameters, and vital signs information as well as RNA-seq data derived from tumor biopsies. However, information about the response to the studied molecules is not available, as the studies are still ongoing/were ongoing during the time of the project.

The prior cancer medication (cancer treatment history) is available in detail for each patient in the data. It has been long recognised that previous treatments can affect tumor immunogenicity [14] and studying the effect of previously administered medication is in the focus of this project as well, especially the effect of previous CPI treatments.

The history of treatments can be represented on a timeline with the therapies having a start and end date (figure A). But changing treatments rapidly or combining different treatment types is a common strategy in oncology, therefore to consider the entire history, it needs to be simplified.

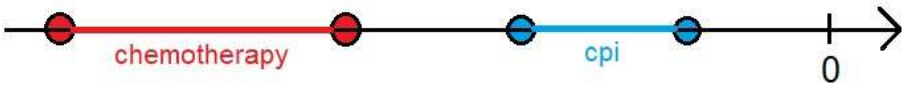

Figure A: Example of visual representation of patient history, treatments have an start and end point, 0 is the start of trial (day 0)

One way to simplify the treatment history is with the concept of Lines of Therapy (LoT). Depending on certain rules, like the ones determined by Hess *et al* [15], a treatment after another can belong to the same LoT or constitute a new one. Once the LoTs are determined, the number, order and length of LoTs enable the extraction of useful information about the patient's treatment history.

RNA-Seq

RNA-Seq is a next-generation sequencing (NGS) technology that allows the profiling of the entire transcriptome (mRNA, tRNA, etc) and the quantification of all RNA sequences within a sample. It captures a snapshot of the transcriptome at a certain moment and grants insight into the function of genes and process of transcription in a way DNA sequencing is not able to. This is a very useful technique to investigate the expression levels (through the quantity of mRNA) of not only a couple, but all of the genes. To investigate the effect of a certain condition on gene expression, differential expression analysis can be performed by comparing the RNA content of two samples, one with the condition, one without, to figure out which genes are down- or upregulated, which are turned off completely [16,17].

Immunogenicity - Predictors

Immunogenicity status of a patient here is understood as the ability of a patient's immune system to react to a stimulus/CPI treatment. Accurately, it can only be measured in hindsight, after the administration of medical agents, for example CPI therapies. As this work focuses on data and results from phase 1 trials, another metric is required to survey the impact of probationary CPI treatments other than the obvious response data that would involve the clinically categorized response to the studied molecules after the first phase of the trials. As a consequence, two available properties are considered instead, as good predictors of the patients' response to the undertaken trial: CD8 immunophenotype of the tumor and a calculated value of a prognostic model of overall survival, ROPRO (Real wOrld PROgnostic score).

CD8 Immunophenotype

In 1998, Naito *et al* established CD8+ tumor-infiltrating lymphocyte (TIL) localization patterns as a prognostic factor for colorectal cancer [18]. They showed that cytotoxic T-cells infiltrating the deeper layers of the tumor, into the so-called parenchyma, correlates with patient survival. Inversely, TILs only being recruited to the solid tumor's outer layer, the stroma, or completely excluded, means a worse prognosis. From this classification came the definition for CD8 Immunophenotype, an indicator for the immune system's ability to respond to CPI therapy [19] . The 3 localization patterns have later been termed immune-inflamed, immune-excluded and immune-deserted [2], but sometimes only considered as binary predictors (immune-rich/poor)[20]. Each of them are associated with specific underlying biological mechanisms: in the inflamed phenotype, CD8+ cells are present in the tumor parenchyma around tumor cells and cells of the inflamed tumor microenvironment (TME) express PD-L1 cell surface receptors[19]. In the excluded phenotype, immune cells are restricted to the tumor stroma and are absent from the parenchyma. The TGFbeta signaling pathway was shown to promote T cell exclusion [21] and as such inhibiting it can be a viable strategy to boost the effectiveness of CPI. In the desert phenotype, T cells are absent from both tumor layers and so, they are mostly unresponsive to CPI treatments [19]. However, it needs to be mentioned that in numerous cancer indications which do not respond as well to CPI, an immune-poor tumor is a better indicator of OS and chemotherapy is still the gold standard treatment [22].

To determine the Immphe, one needs to measure the abundance of tumor invasive immune cell types, however this does not have a gold standard method. In clinical practice the most reliable and recommended method is based on immunohistochemical (IHC) staining of tissue sections followed by calculation of the immune cell density [23], [24], which is a costly and cumbersome method requiring lab expertise. In the ROTD dataset, tumor biopsy was taken around the study start and IHC staining was conducted for a large subset of tumors.

When immune staining is not an option, another way to infer the Immphe is by conducting bulk RNA sequencing on the tumor biopsy and analyzing it with gene expression signatures [8]. While these signatures are immune cell specific, they can be used across all cancer types, but optimal signatures may be established for individual indications, like done by Behring et al for ductal breast cancer [20].

ROPRO

Real-world prognostic score (ROPRO) is a multivariate predictive risk model, taking several properties of real-world patients (clinical, demographic, etc.) into account to calculate a score which can be used to prognose the future survival of patients suffering from cancer. Based on the Flatiron Health database (an Electronic Health Record-derived de-identified database), which consists of a large amount of

patient's info, 27 factors have been selected by a machine learning algorithm, which combined, with certain weights, best prognose patient OS.
The variables contribute independently and homogeneously to OS, across all types of cancer.
Becker *et al* have shown that ROPRO has a high prognostic relevance in clinical Phase I and III trials. [3] Which is why its use for the population of ROTD is promising. The ROPRO score was able to be calculated for all patients within the cohort by scientists at Roche, as all factors necessary for the score were available.

Project goal

Another concern with CPIs is that their long-term impact on the immune system and physiology of patients are not well documented and studied [1], the possible combined effect of multiple, different CPI treatments even less so. However, it would be highly important to have knowledge about the immunogenicity status of the participants of immunotherapy oncology trials before the start of the trial, as it would make it easier to recruit the right patients for these trials.
The main goal of the project is to learn about the effect of previous CPI on the immunogenicity status of patients. To this end, the ROTD dataset will be investigated, as it holds sufficient information to examine the effect of CPI, because a large portion of patients have been previously treated with CPI, often with multiple different ones. A comprehensive descriptive analysis will be performed to be able to focus on the relevant properties and utilizing the available information to find relationships with the two predictors.

Materials and methods

Data acquisition

All of the used data was collected by Roche during phase 1 clinical trials, they are controlled and they were harmonized, anonymised, organized into separate tables. These tables are currently only accessible internally until the clinical trial results are published.

The RNA-Seq data was already preprocessed, the mapped read counts measured, normalized and log-transformed.

Inferring immunophenotype from RNA-Seq data

Signature scores were calculated with two different tools, one with the R package BioQC v1.21.2 [25], the other with GSVA v1.38.2 [26]; both of these packages are part of the Bioconductor project [27]. The GSVA enrichment score was calculated with the 'gsva' function, with default parameters, while the BioQC score is obtained from the wmwTest function with the valType='r' argument.

Preliminary accuracy measurement (to compare signatures) was conducted with a logistic regression model ('glm' base R function with family=binomial(link='logit')) fit on randomly selected 80% of samples (training set) with immunophenotype (Immphe) indicated (502) with the model "Immphe ~ GSVA" and 'predict' function was used on the remainder 20% (testing set) to classify into the binary classes (Inflamed/Non-inflamed), these were compared against the ground truth values. To get a mean accuracy, the random selection of the training set and subsequent steps were repeated fifty times and from the achieved accuracy values the mean was calculated.

For the multiclass classification model, multinomial regression, 'glmnet' engine was used with multinom_reg() function from Tidymodels v0.1.3[28] . For the k-nearest neighbor (KNN) model, nearest_neighbor() function, with the 'kknn' engine was used instead. The model formula in every case was using Bioqc r-score and GSVA enrichment score (GSVA) as combined predictors of the Immphe. To obtain optimal model parameters, training was performed on 75% of samples with Immphe indicated (628) and tested on the rest with random resampling fifty times. With the resulting parameters the model was fit on all samples with known Immphe and the 'predict' function was used to infer the classification on the samples (366) with signature scores but unknown Immphe.

For the binary classification, a logistic regression model ('glm' base R function with family=binomial(link='logit')) was fit on samples with Immphe indicated (628) with the model "Immphe ~ BIOQC + GSVA" and the 'predict' function was used to classify the samples with signature scores but unknown Immphe.

Treatment history analysis

As a first step, unique medications were grouped into 4 higher-level treatment groups (chemotherapies, immunotherapies, targeted therapies, radiotherapy) with the help of a translation table (Supplementary Table S1).

"Lines of Therapy" was a key definition used in the treatment history part of the project and it was established according to the following rules (based on a document from Flatiron company):

1. If two treatments, next to each other, are from different treatment groups, the more recent constitutes a new Line of Therapy (LoT).
2. Two treatments next to each other, from the same group, can only be of the same LoT if the end day of the earlier treatment and the start day of the latter treatment are less than 120 days apart (except after Rituximab, in which case, within 270 days another Rituximab treatment is still the same LoT).

When considering the last day before study start, extreme outliers were removed with the use of 'identify_outliers()' function from rstatix v0.6.0, by discarding the samples from the 'is.extreme' column.

Multiple Correspondence Analysis in R

For Multiple Correspondence Analysis, the 'FactoMineR' v2.3[29] package's 'MCA' function was utilized and to simply visualize the results, 'fviz_mca_biplot' and 'fviz_mca_var' functions from the 'factoextra' v1.0.7[30] package was used.

Sorting into subsets

At this point I decided to subset all samples to a smaller subset with the most information: the samples with both ROPRO values and known or inferred Immphe were considered for analysis. Then within this subset, subsets of different cancer indications were compared, based on their ROPRO and Immphe distributions.

The distribution of ROPRO was compared with two-sample Wilcoxon tests (wilcox.test()). The frequency distributions of Immphe against cancer indications were placed into 2-by-2 contingency tables to be adequate for testing with Chi-square tests (chisq.test()). Significance was assumed with p-value <= 0,05.

Correlation analysis

Details about the investigated properties of the whole population in this part: For Immphe, Treatment Length (Trlen), (existence or absence of) Intermittent Therapy (Inter) and the Type of Last Received Treatment (Last_treat) the count of samples in their subcategories are used. In the case of ROPRO,

either the distribution of values or the number of samples from the highest and lowest deciles (sometimes all 10 deciles) within the respective subgroups are compared.

Visualizations: ROPRO distributions for the groups of a property were visualized. Ratio of the samples in a decile and all samples within its respective group were calculated for each decile and shown.

$$Ratio\ X(in\ feature\ of\ interest) = \frac{Sample\ count\ in\ decile\ X}{Sum\ of\ sample\ count\ of\ all\ \ deciles}$$

Additionally, Last Treatment groups were separated by Immphe and displayed with bar plots.

Tests: Comparing the distributions of ROPRO between two groups was done with Wilcoxon rank-sum tests (W score and p-values are reported). The counts within the highest and lowest ROPRO deciles of 2 or more subgroups are placed within a contingency table and the (ratio of) ROPRO counts are compared for the groups with a chi-square test (X-squared score and p-values are reported). Similarly, contingency tables were set up for binary immunophenotype to compare groups; the chi-square test was used again to show significant differences between the groups. Significance in every case was assumed with p-value <= 0,05.

Results

1. Descriptive Analysis of Population

As mentioned in the introduction, my first task was to get to know the patient cohort better and get familiar with as many facets of the dataset as possible. I hereby present a descriptive analysis of some properties of the population that are of interest in the first place, distribution of sex and age, types of cancer, experience with CPI and available information about Immphe and ROPRO. This initial exploration of the data set is beneficial for performing more analysis steps down the road.

First of all, I considered basic traits as sex and age. In the population there are 1050 males, 812 females and sex is unknown for 2 patients. The age distribution for each sex can be seen in figure 1.1, the mean age for females is 56,87, while for males it is a bit higher 60,62. The female histogram has a heavier left tail, because in the age range 34-42 more females are participating in the trials than males.

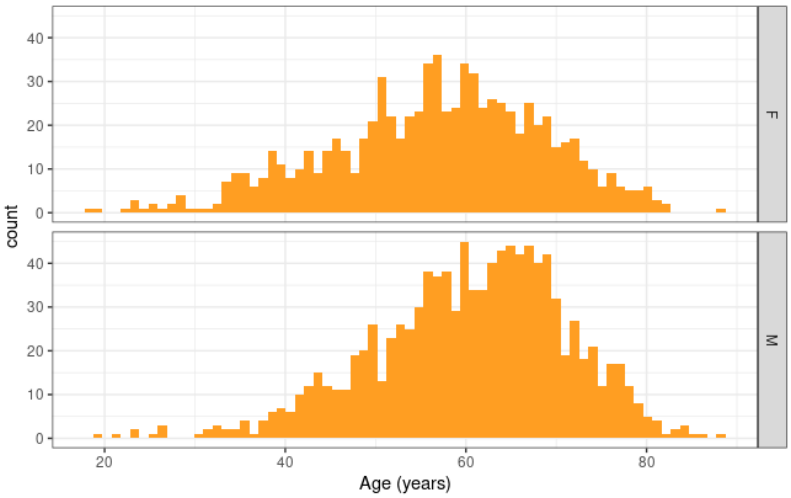

Figure 1.1: Age distribution of the whole cohort faceted by sex (F = female, M = male).

Let us take a look at the prevalence of CPI experience in the cohort. Out of the 1864 samples, 453 (24,3%) have received at least one CPI treatment before the start of the study. As CPI treatments are not equally issued for all cancer types, large differences can be expected in CPI experience per indication. Table 1.1 shows the 'top indications' around the top of the table as it is in descending order of total samples and the number of CPI experienced/naive samples.

Indication	CPI Naive	CPI Experienced	Total
NSCLC	121	197	318
Renal Cell Carcinoma	166	33	199
Colorectal Cancer	189	3	192
unclear or rare indication	136	26	162
HNSC	97	51	148
Breast Cancer	131	4	135
Urothelial Bladder Cancer	106	21	127
Melanoma	23	101	124
Ovarian Cancer	70	3	73
PAAD	70	0	70
Lymphoma	55	8	63
Esophageal Cancer	57	0	57
Cervical Carcinoma	41	2	43
Gastroesophageal Cancer	36	1	37
Sarcoma	35	0	35
SCLC	29	0	29
Mesothelioma	18	1	19
Uveal Melanoma	9	2	11
Cholangiocarcinoma	10	0	10
Small Intestinal Neoplasm	8	0	8
Prostate Cancer	4	0	4

Table 1.1: The number of cpi experienced/naive patients by every indication, descending order by total count.

As all 27 factors, which contribute to the ROPRO score, are available in the dataset, ROPRO was calculated by Roche for 1863 out of 1864 patients and also included in the dataset. The distribution of all the values can be seen in figure 1.2. From the model design close to normal distribution is expected for a larger population, which is well reflected in the figure. The ROPRO deciles, which is the binning of the continuous variable into 10 equal pools, will be of interest, as a significant difference between the OS of samples within the highest and lowest deciles has been exhibited within the original publication [3], so it is going to be important to investigate this for the ROTD as well.

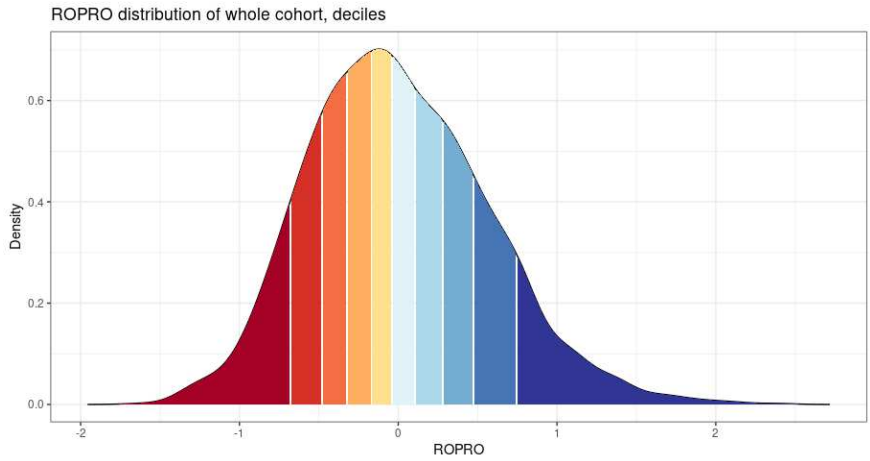

Figure 1.2: Distribution of all ROPRO values, deciles (each 10% of samples) shown with different colors.

The CD8 Immunophenotype has also been determined and recorded from tumor biopsy in a subset of patients, this is also available in the dataset: 707 samples (37,9%) with indicated CD8 immunophenotype (so it is unknown for 1154 samples), 324 samples (17,4%) with desert, 168 (9%) excluded, 215 (11,5%) inflamed Immphe (figure 1.3). Desert is the most common, but luckily all three Immphe are fairly well represented in the data.

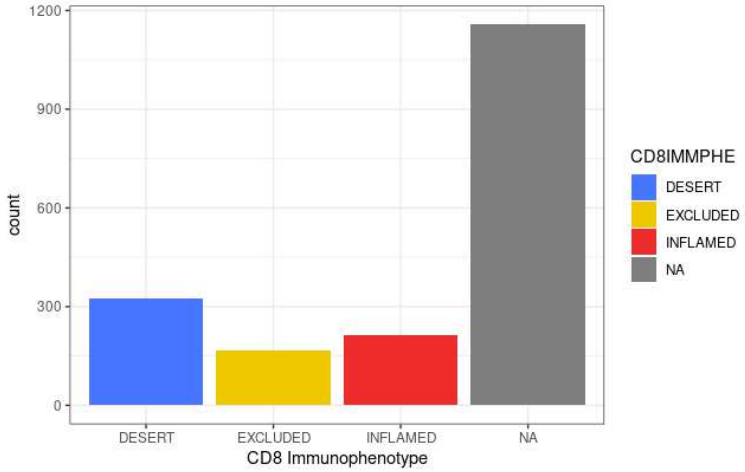

Figure 1.3: Total number of samples in each Immphe group

As most of the population does not actually have their Immphe indicated, it is important for my work to extend the Immphe information by inferring it from gene signatures, using tumor RNA-Seq data, if possible. Therefore, I looked at how many patients have RNA-Seq information, which do not have Immphe. The sample numbers are reported in figure 1.4. Most of the samples with Immphe also have RNA-Seq, but there are quite a lot of samples with RNA-Seq data, but unknown Immphe. This made it ideal to pursue inferring Immphe with the help of RNA-Seq data and gene signatures.

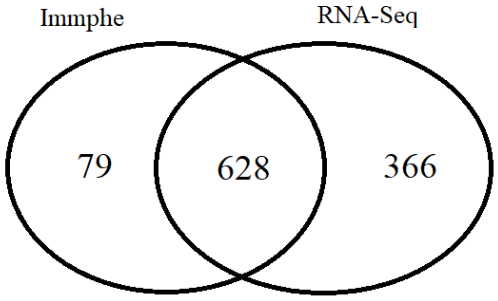

Figure 1.4: Showing the overlap of samples with available immunophenotype and RNA-Seq information.

There is quite an extensive patient treatment history information in the data set, which allows to show general treatment paths, extraction and analysis of relevant features. This is a crucial step as CPI experience is the focal point of the project, therefore this analysis step requires a broader presentation and discussion. These will be handled in the 'Analysis of Treatment History' and 'Discussion' chapters.

2. Enrichment of immunophenotype data

Within the dataset, 707/1864 samples had an immunophenotype already determined. It was feasible for the project to enrich the Immphe information and infer it for as many patients as possible, in order to be able to use this extended information for analysis down the line.

As previously determined, RNA-Seq data is available for 366 samples which lack a predetermined immunophenotype; inferring Immphe for these samples from RNA-Seq would expand Immphe information from 707 to 1073 samples (+51,34%). In order to do this, methods have to be assessed and the best ones chosen for the data and use case.

The task consists of two steps:

(1) Choose a suitable signature that can be used to classify Immphe from the RNA-Seq data.

(2) Perform multiple classification methods and choose the right one for the data by contrasting accuracy.

2.1 Comparing signature scores

For this step, 14 gene signatures were selected which are specific to immune cells (13 T cell and/or interferon gamma (IFNg) signatures, 1 natural killer cell (NK) signature) and are adequately documented (the last part of signature names refer to the main author: Atezo [31], Li19 [32], Danaher17 [33], Tirosh16 [34], Zhang18 [35], SadeFeldman18 [36], Wu20 [37], Ayers17 [38]). The task is to select one which could be best suited for classifying the Immphe for the samples with absent value with the help of the ROTD. To this end, the GSVA scores of each signature are going to be utilized. This score is a summarization of the expression counts of the genes of the signature by the R package GSVA (BioQC is an alternative tool that is also going to be utilized later). It is calculated for every sample with every signature with RNA-Seq, which means 14 x 994 signature scores. As a first line of analysis I compare the correlation between GSVA profiles, I computed the correlation matrix from their values (figure 2.1).

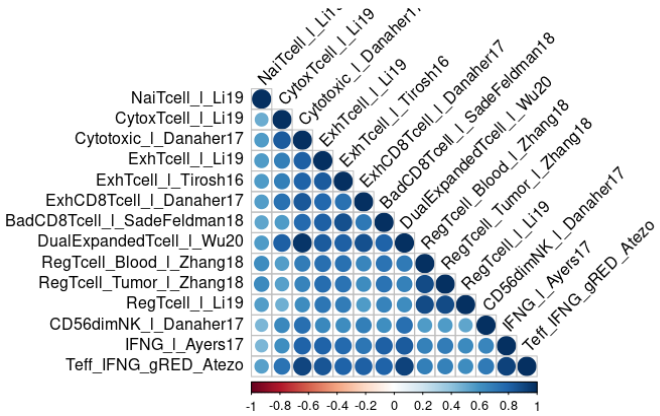

Figure 2.1: Correlation matrix of GSVA scores of 14 gene signatures

It is notable that the scores correlate with each other substantially, the 'CD56dimNK' NK cell signature and 'NaiTcell' naive T cell signature less so with all the others. However, considering the differences between the resulting Pearson's correlation coefficients, I can expect non-equal ability of the signatures to infer the Immphe.

Preliminary accuracy measurement: Now using only the GSVA scores and a 'preliminary' classification method I compared the accuracy of signatures. This showed that the 'Teff_IFNG_gRED_Atezo' signature outperforms all other signatures, reaching 0,7795 (SD=0,0427, SE=0,006), shown in figure 2.2. This signature corresponds to characteristic effector T-cell genes and also holds interferon-γ (CD8A, GZMA, GZMB, IFNγ, EOMES, CXCL9, CXCL10, and TBX21).
This signature is going to be used throughout the next steps.

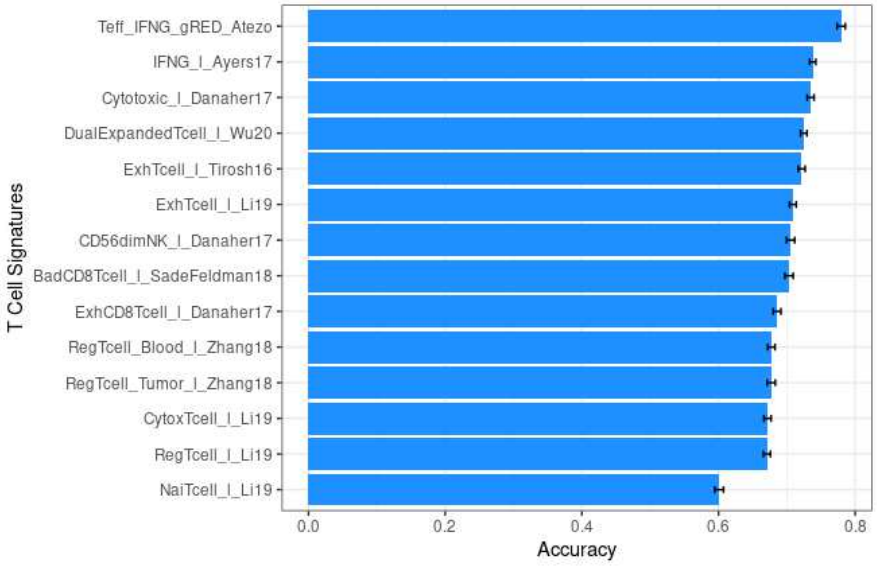

Figure 2.2: Accuracy (mean and SE) comparison of 14 gene signatures

Showing relationship of signature scores and Immphe to highlight trend:
Another set of signature scores was calculated with a different tool, BIOQC. The relationship of signature scores across all indications and Immphe is shown in figure 2.3 and figure 2.4. A clear trend is visible with both signature sets; desert and inflamed seem more separated for GSVA.

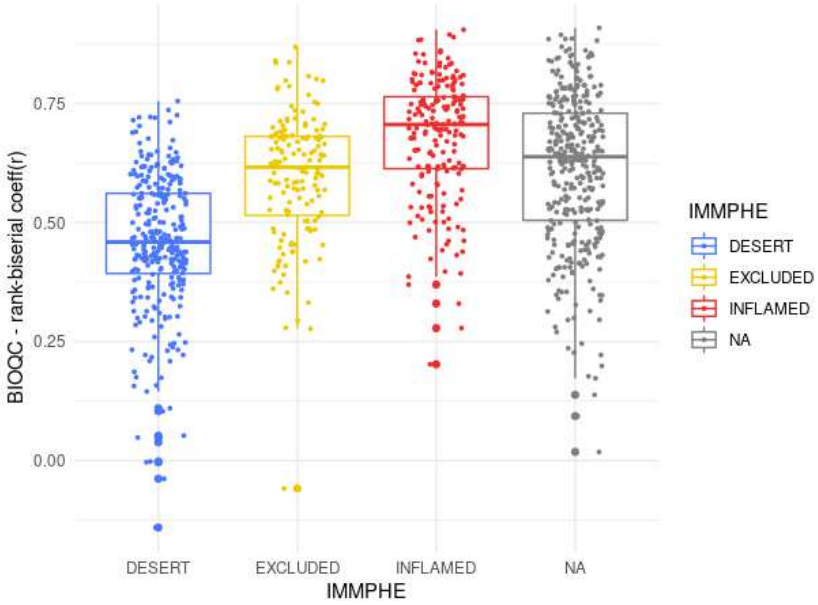

Figure 2.3: BIOQC scores from the best signature (Teff_IFNg) by quartiles, separated by classes of Immphe

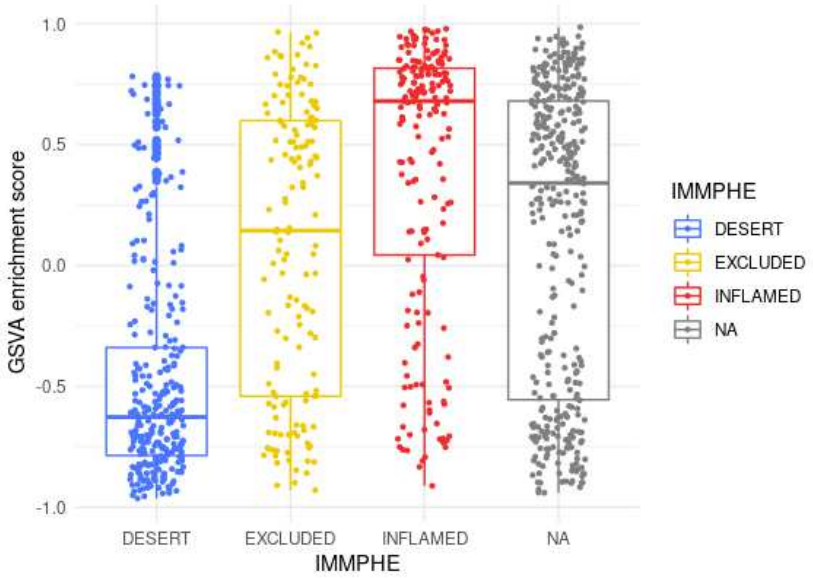

Figure 2.4: GSVA scores of the best signature (Teff_IFNg) by quartiles, separated by classes of Immphe

GSVA scores seem to imply a clearer trend than the BIOQC scores, but the trend follows the same line with both. Next, the GSVA landscape was separated by cancer indications (figure 2.5), which shows the same trend for most indications with enough samples. Melanoma and Bladder Cancer are notably different. For the most sample-rich indication, NSCLC, the inflamed phenotype is well separated from the other two, but desert and excluded are very similarly distributed.

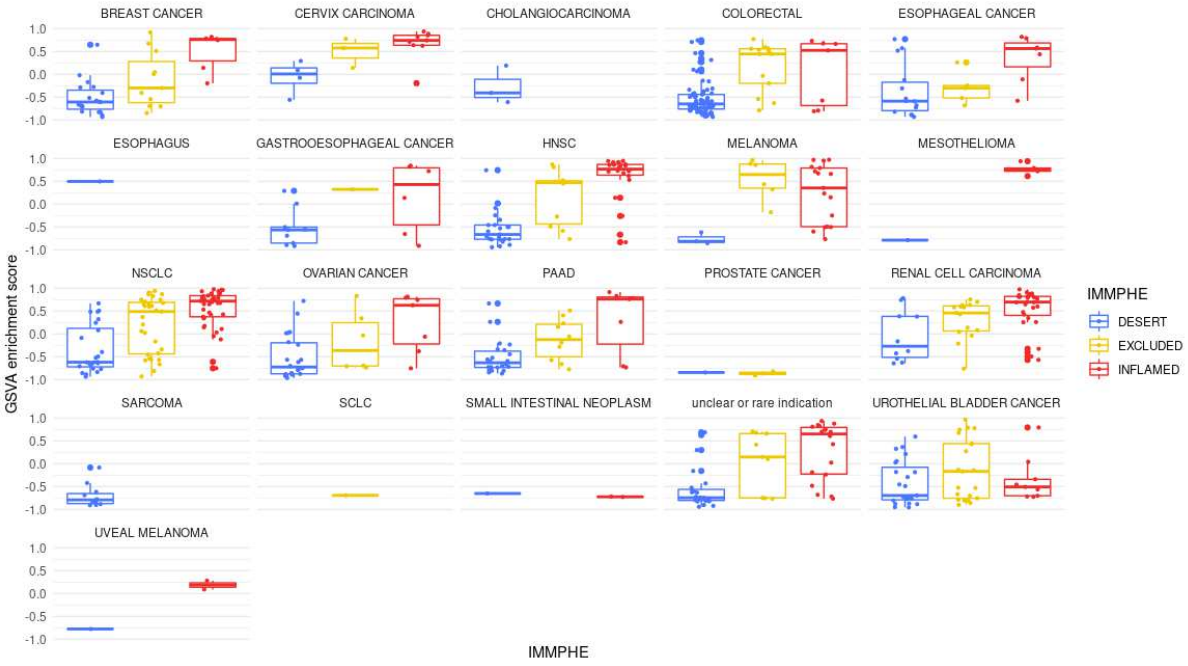

Figure 2.5: GSVA ES trend per all indications

2.2 Comparing accuracy of classification methods

Considering the visible trend in the signature scores and Immphe, separating its classes based on the scores is the next step. There are multiple classification methods that can be used with the signature scores computed earlier. I present the methods, apply and compare them to be able to choose the method that yields the most accurate results and is the most useful for the following analysis steps. I will use multiclass classification and a two-step binary classification to reach sufficient accuracy for all three classes. Afterwards, I will also conduct a binary classification for two Immphe classes (Inflamed and Non-inflamed = Excluded + Desert) and discuss if that is a more suitable method in the light of results for all three classification methods.

Tertiary classification

For the multiclass classification, a logistic model and K-nearest neighbor (KNN) models are being compared. The performance of these models is reported in table 2.1. The logistic model showed higher accuracy and AUC, but also a bit lower sensitivity than KNN.

model	metric	mean	standard error
Logistic model (penalty=0,00001)	accuracy	0,638	0,0235
	AUC-ROC	0,746	0,0156
	sensitivity	0,546	0,0152
K-nearest neighbor (neighbors=15)	accuracy	0,628	0,0169
	AUC-ROC	0,713	0,0184
	sensitivity	0,560	0,0204

Table 2.1: Comparing performance of multiclass classification models

After the performance test, the samples without a label were classified with the logistic model and the new class labels were visualized within the two-dimensional BIOQC - GSVA signature score feature space (figure 2.6). This method leaves out almost all Excluded predictions, suggesting that Excluded and Desert are not reasonably distinguishable which is in line with our knowledge about the biological similarity of the two Immphe classes. Considering this and the fairly low accuracy of multiclass classification, it was decided to roll back the method and to only separate Inflamed from Excluded and Desert (together Non-inflamed) samples with highest accuracy possible, thereby performing a binary classification.

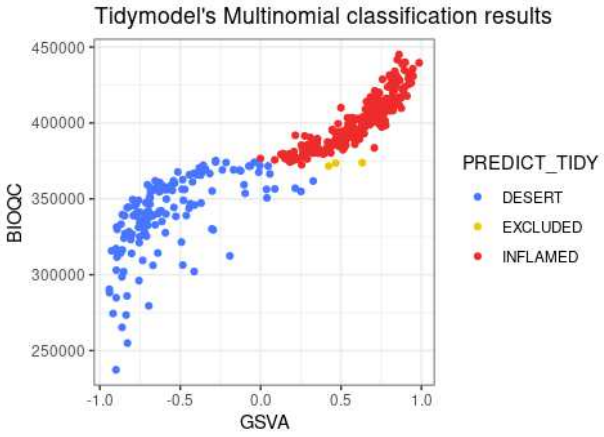

Figure 2.6: Multinomial classification results of samples without label in GSVA and BIOQC value space, predicted classes with color

For this last binary classification step the signature score is needed which can result in the highest accuracy, if they are even different. The mean accuracy reached is the highest when using both signatures, although the difference from BIOQC alone is minimal. However, for the final classification 'GSVA+BIOQC' is used and it can be seen that the accuracy achieved with it is higher than for the preliminary classification. The resulting Immphe labels are merged with the previously existing labels (which were converted to inflamed-noninflamed).

Signature score	Mean ± SD
GSVA	0,7515 ± 0,03369
BIOQC	0,7943 ± 0,03456
GSVA+BIOQC	0,7985 ± 0,03175

Table 2.2: Accuracies reached by using BIOQC and/or GSVA signatures during the binary classification

3. Analysis of treatment history

The treatment history is a focal point of this thesis work and it is important to take a closer look at what factors are available and of interest, and what scientific assumptions can be made to determine the effect of received treatments on the immunogenicity status of patients.

First, to get a general view of the medication history of the whole population, the Lines of Therapy (LoT) were established based on our ruleset (in Materials and Methods). A Sankey plot shows the patient groups moving between LoTs of different treatment types (figure 3.1). The relevant treatments were grouped into 4 groups (Supplementary Materials table 1.) which are used here and only the first 4 LoTs are being shown. Further lines are cropped because the number of patients in them is low (<20% for LoT5) and no useful information can be extracted. The highest LoT count within the cohort is 13.

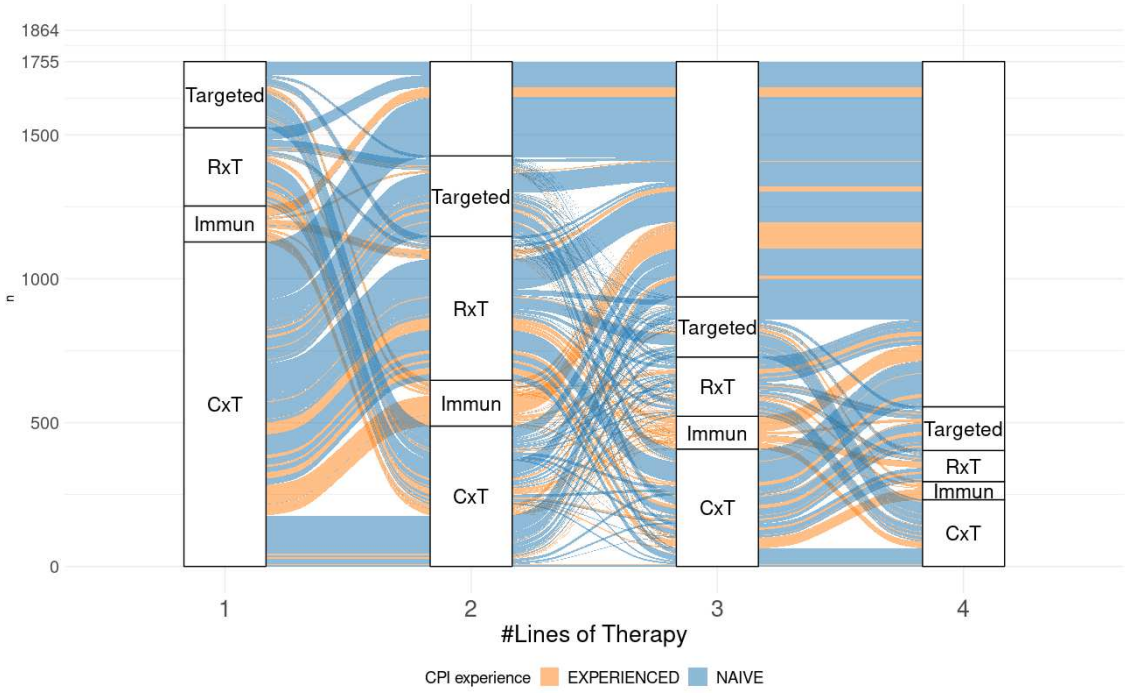

Figure 3.1: The first 4 LoTs from relevant treatment types, strands flowing from treatment to treatment are patients (groups); the higher limit of y axis, 1864, is the total number of patients in the cohort. CPI experienced groups are colored with orange and CPI naives with blue. (CxT: chemotherapy, RxT: radiotherapy, Targeted: targeted therapies, Immun: immunotherapies, empty columns indicate that LoTs stopped for the current patient)

Chemotherapy (CxT) is the main first line treatment by a large margin, very few patients participate in a trial after just a single LoT, if that is not CxT. Almost a third of all patients received CxT first and radiotherapy (RxT) right after, this is a major treatment profile in our cohort. Targeted therapy, second after CxT, is also a very common treatment path for these patients (~15%). The most common number

of LoTs is 2, followed closely by 1 and 3 LoTs. Regarding CPI experience, more than half of the CPI experienced patients started on CxT, a large number of them received Immunotherapy right after. And most of the other CPI exp. started on Immunotherapy already. These are probably patients with cancer indications which on average react well to CPI treatment.

The distribution of LoT counts for each cancer indication was visualized next (figure 3.2). Based on cancer type, treatment profiles are quite different and this can be seen in the LoT distributions as well. Breast, ovarian and colorectal cancers have the longest right tail (large amounts of LoTs on average) which correspond to their low cure and survival rates. As these cancers do not respond particularly well to CPI, participating in CPI trials can only be seen as a last resort, this is also seen from the plot. With renal cell carcinoma (RCC) (and uveal melanoma although very low sample size) a lot of patients did not receive any prior treatment from the four considered types. The large discrepancy between breast cancer and RCC is separately shown in figure 3.2B.

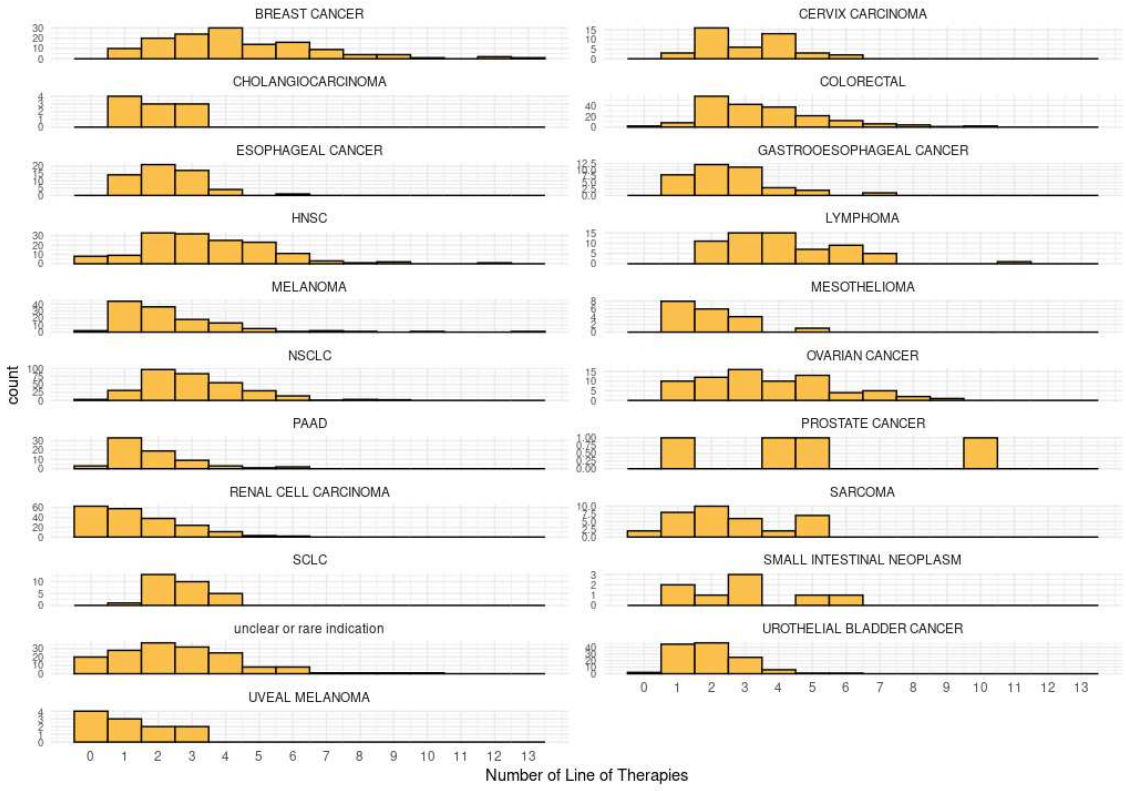

Figure 3.2A: Amount of LoT patients received in each indication group

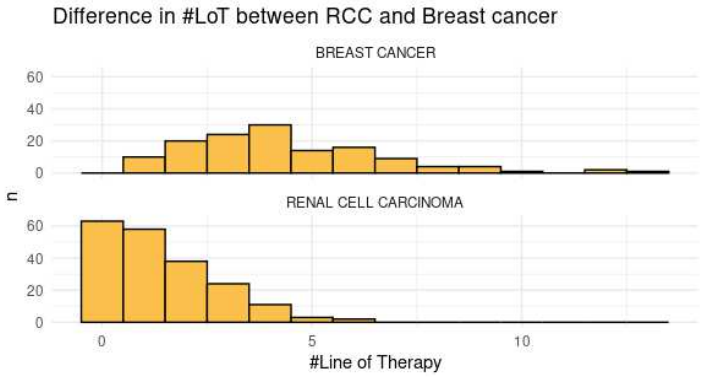

Figure 3.2B: Histograms of LoT amount distribution in the two most different groups, Renal Cell Carcinoma and Breast Cancer.

The end date of last treatment before trial start are compared per indication (Figure 3.3A-B). Removing extreme outliers clearly removes a large amount of variance for multiple indications. Afterwards, the difference does not seem to be significant among most indications, only between the ones on the low and high end.

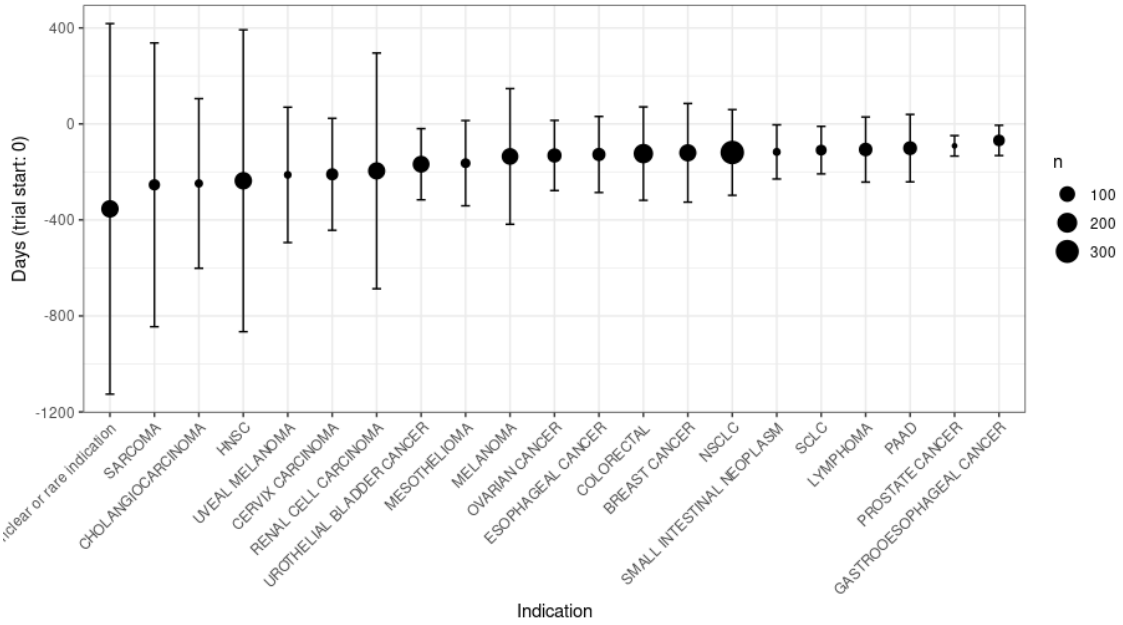

Figure 3.3A: Differences in the end day of last treatment before trial start (Day 0) per indication (Mean plus standard deviation)

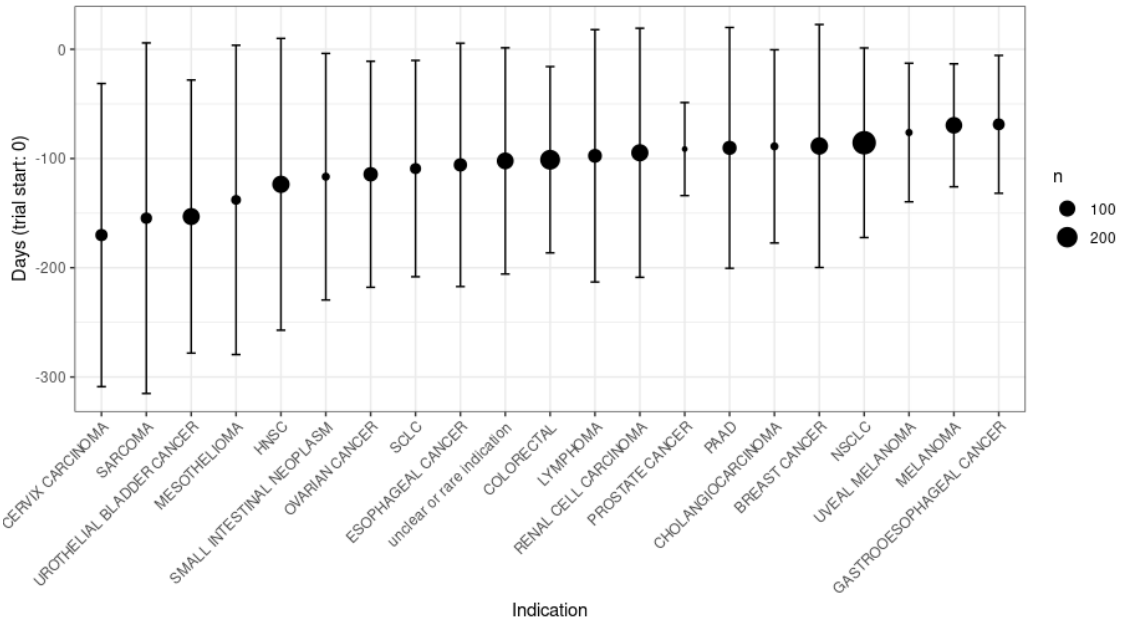

Figure 3.3B: Differences in the end day of last treatment before trial start (Day 0) per indication, extreme outliers removed. (Mean plus standard deviation)

Finally the last day of treatment before the start of the study is shown again, this time separated by treatment type (Figure 3.4). No huge difference in median, similarly to faceting per indication in figure 3.3. The fact that CPI seems overall closer to trial start coincides with NSCLC and melanoma being on average closer with the other view as well.

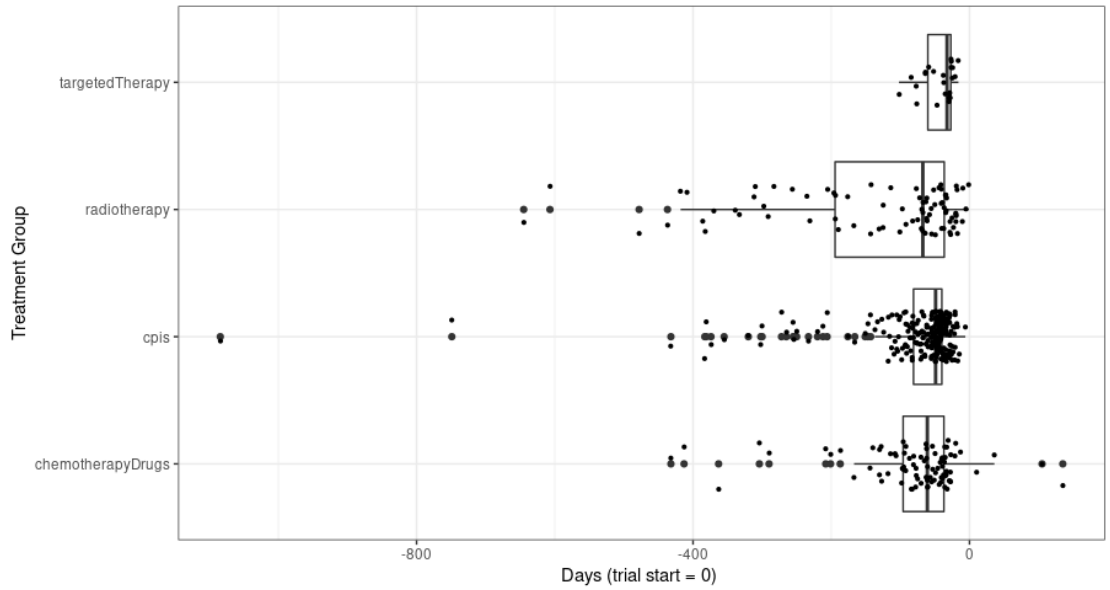

Figure 3.4: Last treatment end dates, grouped by treatment group

Focusing on the CPI experienced subset

The goal of the project is to investigate the effect of previously taken CPI upon patient immunogenicity, therefore it is exceptionally important to focus on the CPI experienced subset and so the next comparisons highlight the previous medications of this patient group, especially the CPI treatments. Properties discussed: LoTs, CPI treatment lengths, intermittent therapy, last day of treatment before trial start.

A similar Sankey plot can be shown for this subset like the one created for the whole population (figure 3.5). It can be noted from the plot that the last treatment is most commonly CPI and even if the last treatment is from other three therapies, the second to last is most likely a CPI LoT. Thus, a combination of two non-CPI therapies (eg. chemotherapy and radiotherapy) right before study start is not common in CPI-experienced patients, instead two distinct treatment profiles can be distinguished: intermittent (last treatment before trial is non-CPI) and not intermittent (last treatment CPI).

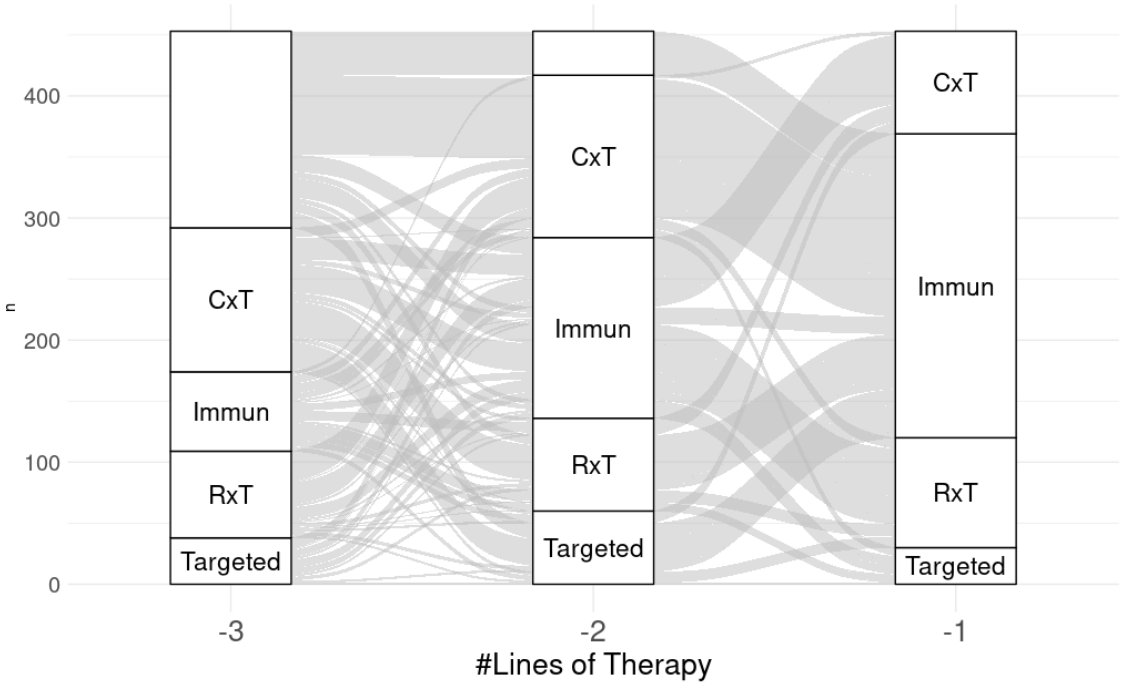

Figure 3.5: The last 3 LoTs of CPI-experienced patients from relevant treatment types, strands flowing from treatment to treatment are patients (groups). The higher limit of y axis, 453, is the number of CPI-experienced patients. -1 on the x axis means the last LoT, -3 is two LoTs before that. (CxT: chemotherapy, RxT: radiotherapy, Targeted: targeted therapies, Immun: immunotherapies, empty columns indicate no previous LoTs for the current patient)

In figure 3.6 is the distribution of the lengths of CPI treatments. The median is 150 but the average is lower, the most common length is around 60-70 days.

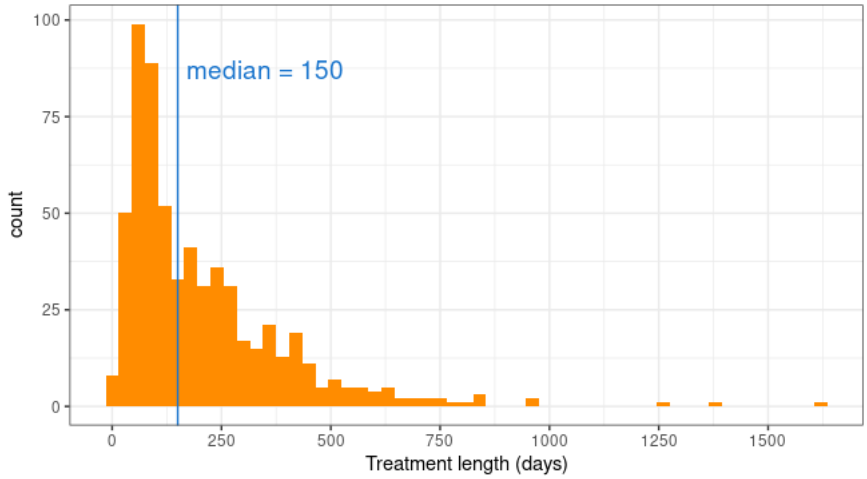

Figure 3.6: Histogram of the length of all CPI treatments with median line

Next, I looked at an important property that was established during the project, the presence or absence of intermittent therapies. Here, intermittent treatments mean one or more non-CPI LoT between the last CPI LoT and trial start. Its importance comes from the fact that the largest influencing treatment on the trial is believed to be the last previously received LoT.

The duration of the last received CPI treatments can be seen in figure 3.7, split between the group which received additional treatment(s) afterwards while the others did not (these groups are fairly equal in patient numbers too). No major differences can be noticed in this regard.

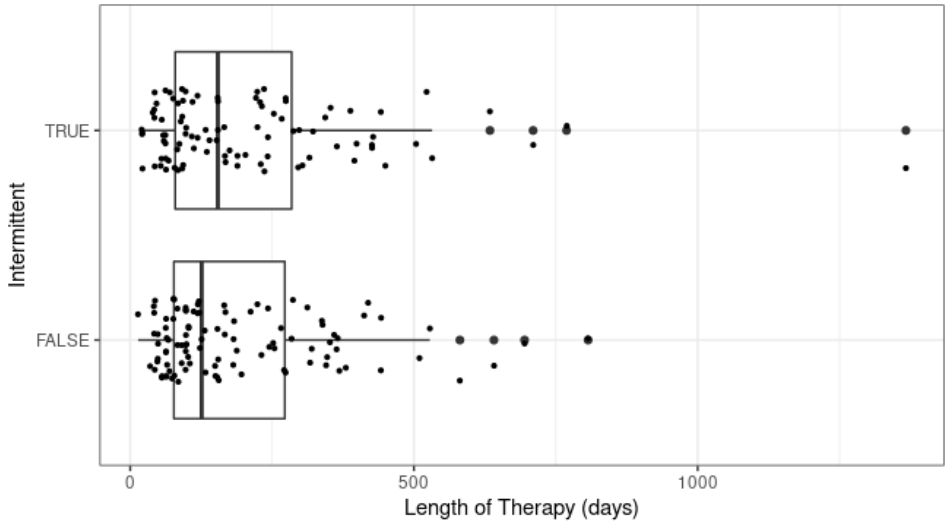

Figure 3.7: Lengths of last CPI treatments, split into two groups: no intermittent group = last treatment is a CPI treatment; intermittent group = last treatment is something other than CPI tr.

Finally, I looked at the ending day of the last treatment before the trial started (figure 3.8), in this case, the end date of the 'no intermittent' group is a lot closer to trial start on average than the 'intermittent' group, but the last day for half of patients in the 'intermittent' group is under 150 days, which is considered recent. This is important because based on this it can be assumed that those treatments would impact the immunogenicity of patients during the trial.

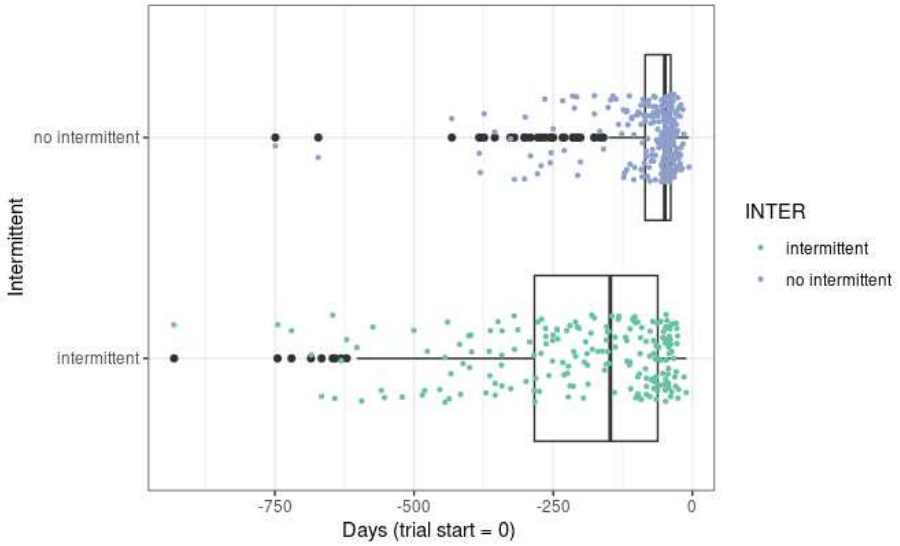

Figure 3.8: Last treatment day distributions, grouped into two groups: no intermittent group = last treatment is a cpi treatment; intermittent group = last treatment is something other than cpi tr.

4. Multiple Correspondence Analysis

In order to check associations among properties previously discussed, without having to do it pairwise, MCA (Multiple Correspondence Analysis), a version of PCA, can be performed, which takes categorical variables and localizes them in a two-dimensional space, where features closer to each other are assumed to be more related. After this step, observed similarities can be investigated further, separately, by checking back to the original data.

Here, first I take into consideration all patients in the cohort and then subset to better show the impact of properties that only involve a group of samples. This is indeed necessary as there are few properties that all patients possess (e.g. ROPRO).

4.1 All samples with ROPRO (1853)

When taking the entirety of the population into account, the properties shown in table 4.1 can be considered, namely tumor indication, CPI experience and ROPRO deciles (the binning of the continuous ROPRO variable). However, as Immphe is not available for all patients, binary immunophenotype becomes a tertiary variable with 'unknown Immphe' being a third categorical variable.

Short name	**Description**	**Values**
INDICAT	*indication of primary tumor*	TRUE / FALSE
CPI_EXP	cpi treatment experience status	naive / experienced
ROPRO_DECILE	10 deciles of ROPRO values (ordered by overall population)	quantile 1-10
IMMPHE_BIN	binary immunophenotype	inflamed / non-inflamed / unknown Immphe

Table 4.1: All features (4) and their descriptions included in the MCA of subset A

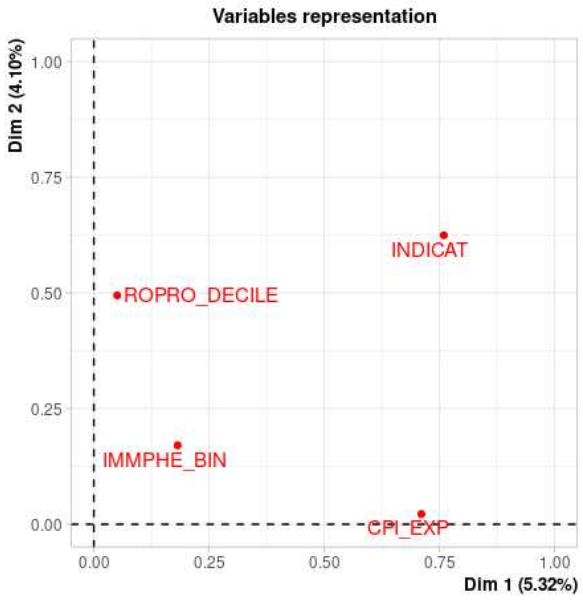

Figure 4.1: Variables of subset A in the MCA's two dimensional space

In a 'biplot' (figure 4.1), the overall features can be placed and the following can be noted: The variance is not well explained by the first two dimensions (5,3+4,1=9,4%) which is consistent with the observation that features with high amounts of possible values (categories) result in a high number of groups of samples based on category intersections, which spreads the individuals apart. Here, both the indication and the ROPRO decile columns have numerous categories. Indication discriminates on both dimensions, which makes sense as it is dependent on all other variables. ROPRO and CPI-experience status are located directly on the two perpendicular axes. This suggests close to no relationship between them.

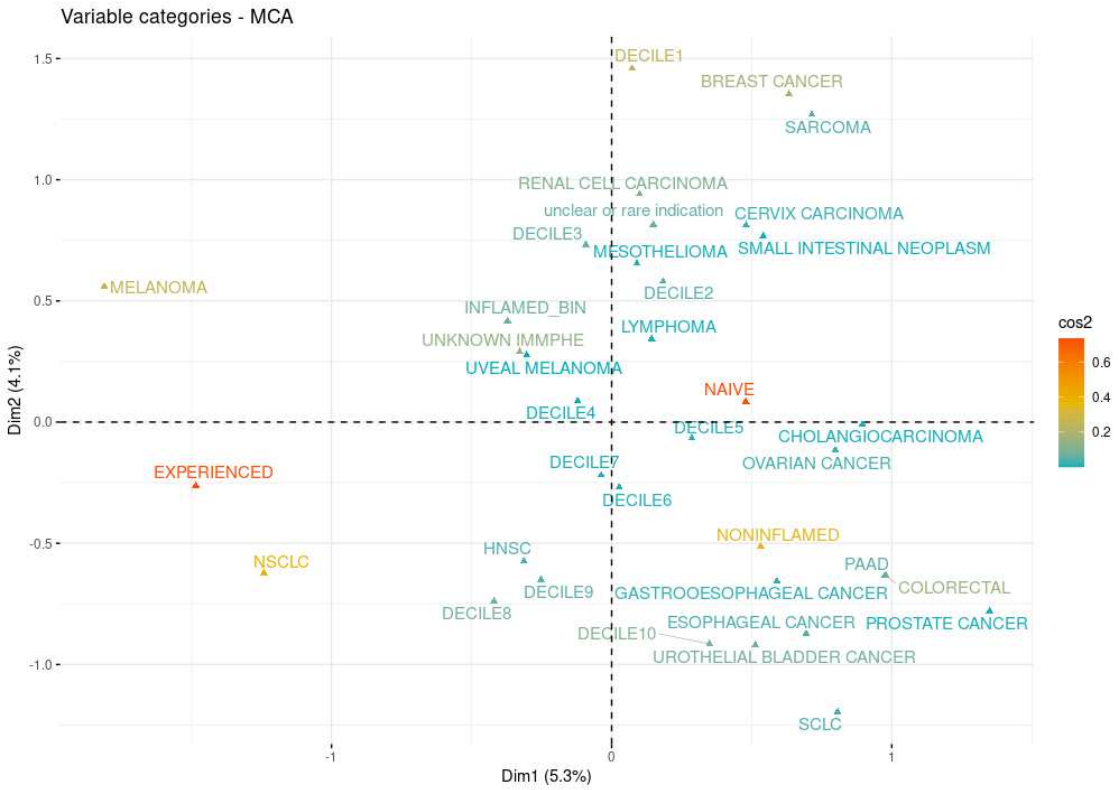

Figure 4.2: Variable categories of subset A in the MCA's two dimensional space

In a 'variable plot' on the other hand, all of the categories are presented. Cos2 means the sum of explained variance by the first two dimensions. If it is low, the variable categories may not be well represented in this 2D space. The ROPRO deciles have quite low values as well as most of the indication types; this makes their evaluation less clear.

In figure 4.2 it can be seen that the y axis corresponds well to the ROPRO deciles. The CPI-experience property doesn't comply with the deciles on the y axis, instead it sits on the two ends of the x axis. Therefore it can be assumed that it doesn't correlate well to that property. However, the CPI experienced category is surrounded by lung cancer and melanoma, both of which have high amounts of experienced samples. Some indications are close to high deciles (indicating lower overall survival), on the other hand, breast cancer and renal cell carcinoma are close to low deciles. The immunophenotype group 'noninflamed' is also closer to high deciles than inflamed, which suggests a worse prognosis for it. And finally, indications are all over the figure, which reinforces the observation that they inflict immense variance upon all other properties.

4.2 All CPI-experienced samples (452)

For the second view, only the CPI-experienced samples are taken into account as this is the group I am most curious about. Its features and categories are summarized in table 4.2. From this step onward

during the MCA analysis, the ROPRO category column is changed from deciles to quantiles, because of the consideration for decreasing sample size within the deciles.

Short name	Description	Values
INTER	was there intermittent treatment between last CPI and study start	TRUE / FALSE
LAST_DRUG_TYPE	type of last CPI treatment by inhibition target	CTLA-4, PD-1, PD-1L
ROPRO_QUAN	4 quantiles of ROPRO values (ordered by overall population)	quantile 1-4
IMMPHE_BIN	binary immunophenotype	inflamed / non-inflamed
TRLEN3	length of last CPI treatment	shortmid: <12month, long: >12m

Table 4.2: The 5 features and their descriptions included in the MCA of subset B

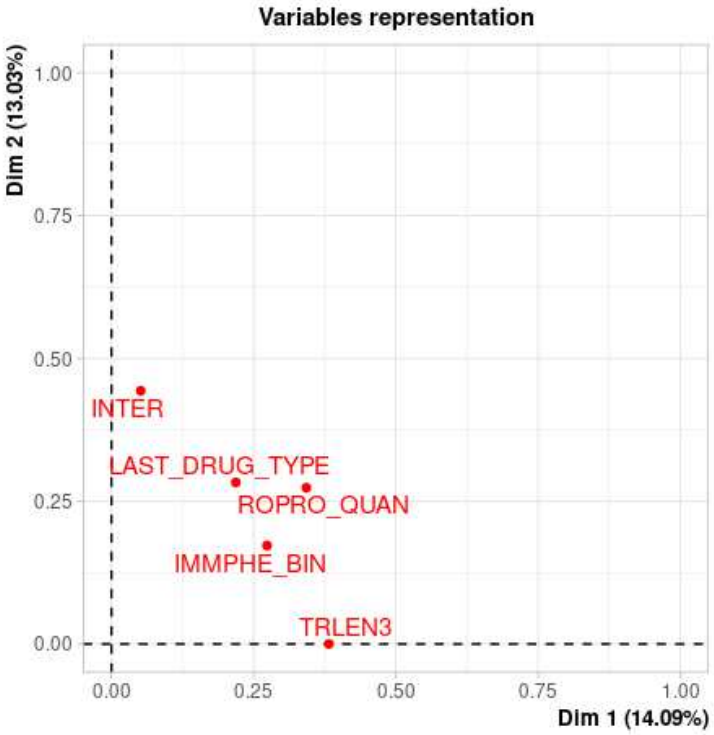

Figure 4.3: Variables of subset B in the MCA's two dimensional space

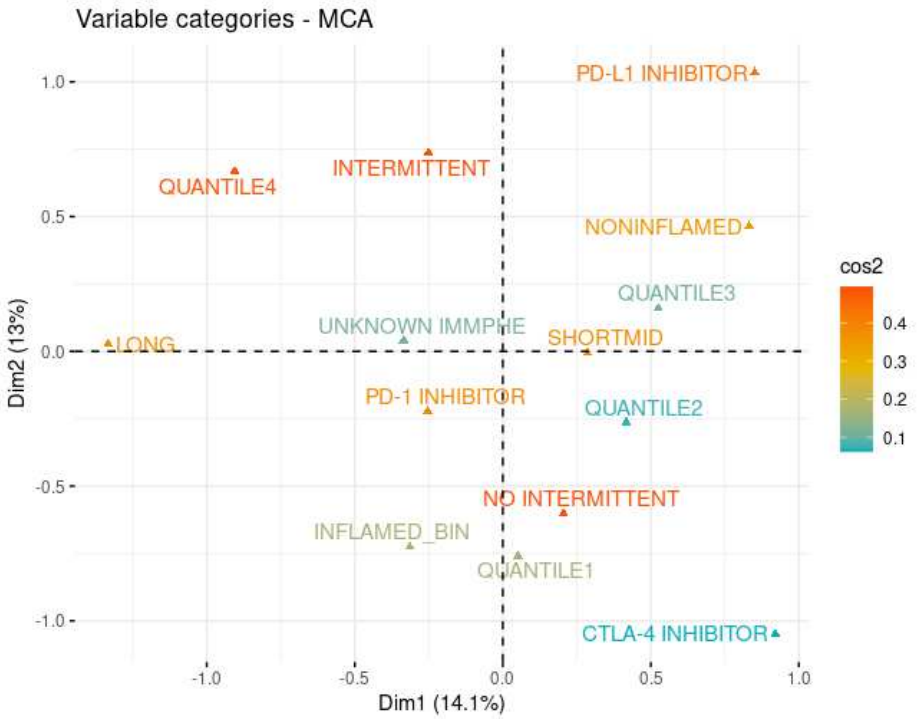

Figure 4.4: Variable categories of subset B in the MCA's two dimensional space

In figure 4.3, 'intermittent' and 'treatment length' are located on perpendicular axes, which implies their low degree of relationship.

The categories within figure 4.4 reveal some interesting points: Treatment length is located on a different axis from most other variables. It could be assumed that it will not have a clear correlation with other variables. Quantile 4 of ROPRO (worst prognosis) is further away from the other quantiles, most similar to long treatment length and 'intermittent=TRUE'. Quantile 1 (best prognosis), inflamed immunophenotype and 'intermittent=FALSE' are close together as indicating similarity and showing the "good" properties.

'Last drug type' could play a role towards immunogenicity, as in figure 4.3, it is close to both ROPRO and immunophenotype. However, it needs to be kept in mind that a very small amount received CTLA-4 inhibitor (Ipilimumab). As it sits near the edge of the coordinate system, it discriminates well between patients.

4.3 Intersection of CPI experience and known immunophenotype (223)

For this view the same features are investigated as for the subset in 4.2, detailed in table 4.2. By taking a look at the biplot (figure 4.5), the variance is better explained by the first two dimensions for this subset (15,99% + 15,18% = 31,17%). Clear separation can be noticed between two groups of properties

and association inside the groups can be better seen. However, in section 4.2, ROPRO was close to both 'last drug type' and immunophenotype; this is not the case here.

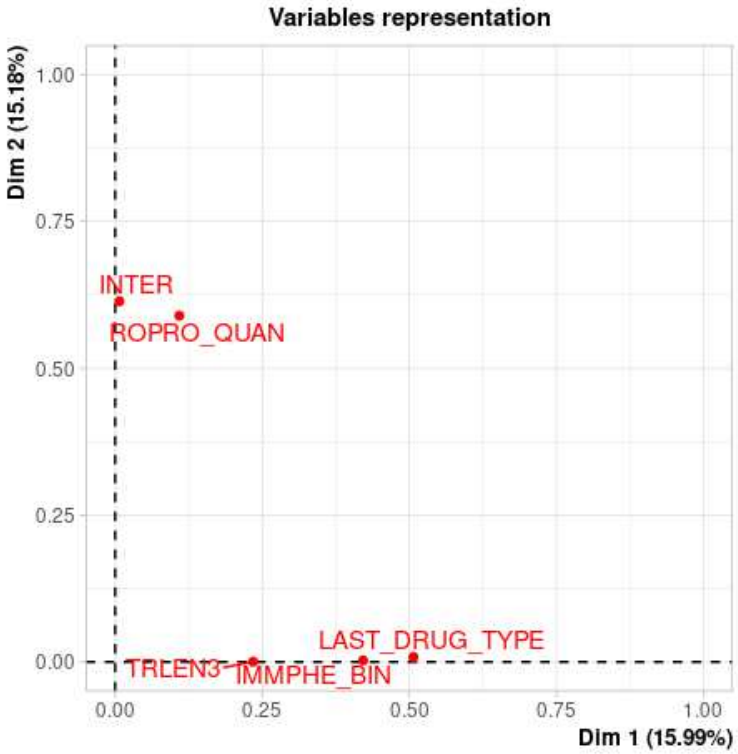

Figure 4.5: Variables of subset C in the MCA's two dimensional space

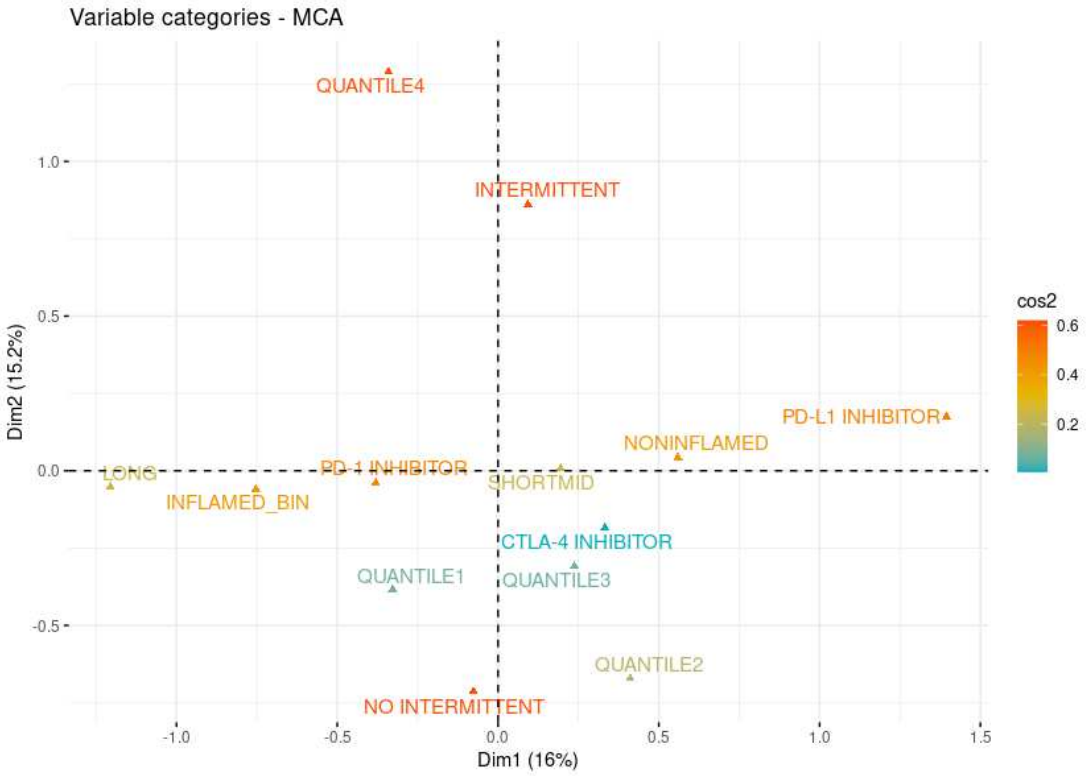

Figure 4.6: Variable categories of subset C in the MCA's two dimensional space

The categories of 'Treatment length' and 'intermittent' line up well on the same axis with Immphe and the quantiles of ROPRO, respectively (figure 4.6).

The three views all paint a different picture of the data, which is caused by considering different subsets of the samples and the properties of the MCA method itself. However, some variables consistently show correlation with each other (for example Immphe with 'Treatment Length'), which relationships need to be further evaluated. It can be also noted that indication does not correspond to any other trait and is the source of immense variance (Part 4.1, figure 4.1-2). The aim is going to be to mitigate this variance for follow-up analysis.

5. Final subsetting per cancer indication

Based on the previous MCA analysis I was able to see that the cancer indication is a key factor which influences variance of the ROPRO value distribution. Additionally, from the descriptive analysis it is known that there is a large discrepancy in the ratio of CPI experienced/naive patients per indication (table 1.1). To reduce this variance and better represent the underlying effects, the aim is to subset the whole cohort and analyze its properties without the interfering influence of the cancer type.

It would be possible to select just one or two large indications and check the data within a single group of cancer types. However, it would be of interest to combine similar (in terms of the two predictive variables) subsets in order to increase the sample sizes which can be investigated.

To this end, I compare the ROPRO distributions and Immphe ratio of each indication to the two main CPI experienced indications (NSCLC and Melanoma) and the condition for merging two indications is for them to not significantly deviate in either aspects.

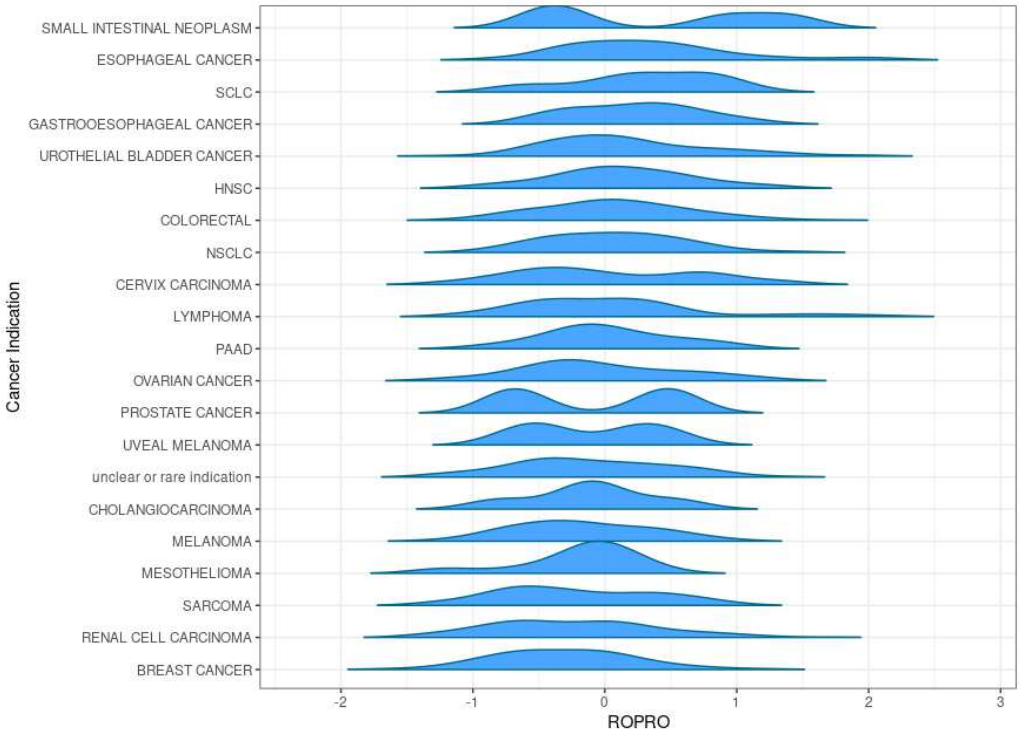

Figure 5.1: The distribution of ROPRO values of 1073 samples per indication (descending order by group mean).

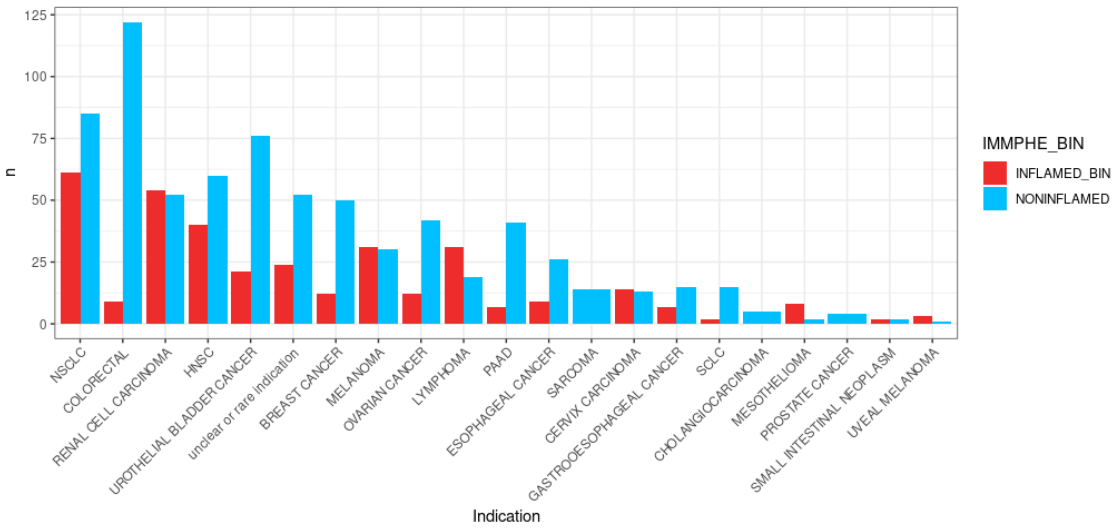

Figure 5.2: Sample count in the two Immphe classes shown for all indications

Figures 5.1-2 show that both the ROPRO distributions and the Immphe ratios are widely varying across indications. Also, it is of importance to choose indications with sufficient CPI-experienced samples. Therefore only the following indications are considered: NSCLC, HSNC, Urothelial Bladder Cancer, Renal Cell Carcinoma, Lymphoma, Melanoma (based on Table 1.1).

Indication 1	Indication 2	ROPRO Distribution	Immphe Ratio
NSCLC	HSNC	W = 6695, p-value = 0,2702	X-squared = 0,021594, df = 1, p-value = 0,8832
	Urothelial Bladder Cancer	W = 6640, p-value = 0,4117	X-squared = 9,683, df = 1, p-value = 0,00186
	Renal Cell Carcinoma	W = 10087, p-value = 3,932e-05	X-squared = 1,7252, df = 1, p-value = 0,189
	Lymphoma	W = 4032, p-value = 0,2705	X-squared = 5,3286, df = 1, p-value = 0,02098
	Melanoma	W = 5599, p-value = 0,003551	X-squared = 1,0811, df = 1, p-value = 0,2985

Table 5.1: Comparing five indications with lower sample size to NSCLC, in terms of ROPRO distribution and Immphe Ratio.

From the table 5.1 can be seen that NSCLC and HNSC do not differ significantly in either metric, so they can be joined. There is no other indication that fits into this group based on both metrics. Lymphoma is non-significant in ROPRO and only slightly significant in Immphe, so it could be considered as a candidate, but to remain strict with the conditions, it will not be joined to the lung cancer group. After combining samples from NSCLC and HNSC it totals at 246 samples (as seen in Table 1.1). Melanoma is the indication with the second highest CPI-experienced sample count, so it would be very important to analyze it. However, because of its emphasized status in immune checkpoint therapies, which means a much higher number of CPI-experienced patients than in other indications, it makes sense to keep Melanoma as its own individual group (61 samples).

Lung cancer and Melanoma groups will be the main two subsets for the final analysis, but two more indications, Renal Cell Carcinoma (106 samples) and Urothelial bladder cancer (97 samples) can be examined as well, as their sample counts are sufficiently large.

6. Correlation analysis of predictors of overall survival

In the final and pivotal chapter of the thesis, I assess the results of all previously explained steps to combine them in an analysis which contrasts categories determined from the treatment history with Immphe and ROPRO profiles of the proposed groups. Additionally, I compare ROPRO to Immphe distributions, in order to confirm expected correlation between them.

The analysis is to be performed on four subsets of different cancer indications, which were determined in the previous section called 'Sorting into subsets'. The groups are compared both visually and with the help of statistical tests, to investigate the possible influence of treatments received before the start of the trial on the immunogenicity of the patients.

First, the relationship of ROPRO and immunophenotype is examined. The other investigated properties are all aspects of the treatment history of the involved patients. These were selected and their importance described in earlier chapters (MCA, treatment history). Namely, they are the binary variable 'presence or absence of intermittent therapy between the last CPI treatment and study start' or intermittent therapy, the length of last CPI treatment in days and the type of last received treatment from the four main treatment groups (cpi, chemotherapy, radiotherapy and targeted therapy).

6.1a Relationship of ROPRO and Immphe

In the MCA chapter some correspondence can already be noticed between ROPRO and Immphe but it can also be assumed that these two features do not correlate too well (particularly from analysis C of said chapter). Here this relationship was investigated further. In figure 6.1, a trend can be seen in the ROPRO distributions where the median of the inflamed group is lower (which means better OS) than of the non-inflamed group, however this difference is not significant with a Wilcoxon rank-sum test (W = 8035, p-value = 0,1947).

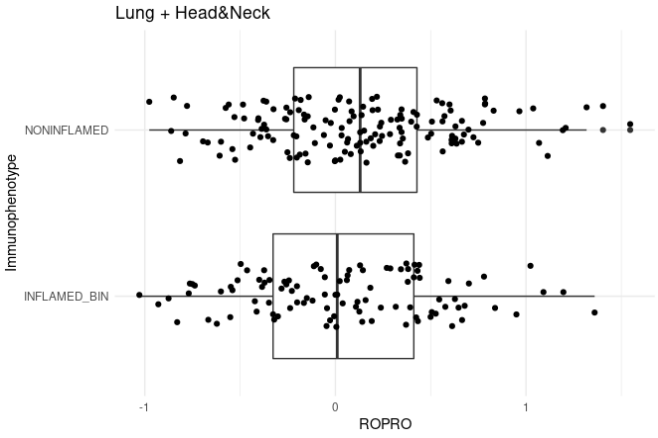

Figure 6.1: ROPRO values separated by the two Immphe groups, shown as points and as Box plots (NSCLC / HNSC subset)

Sample counts per ROPRO decile can also be considered, to easier check the tails of the distributions (Figure 6.2-3). This way it can be seen that in the lowest decile there are more inflamed samples and in the highest more non-inflamed (Table 6.1). However this difference is not significant with the chi-square test (X-squared = 0,936, p-value = 0,333).

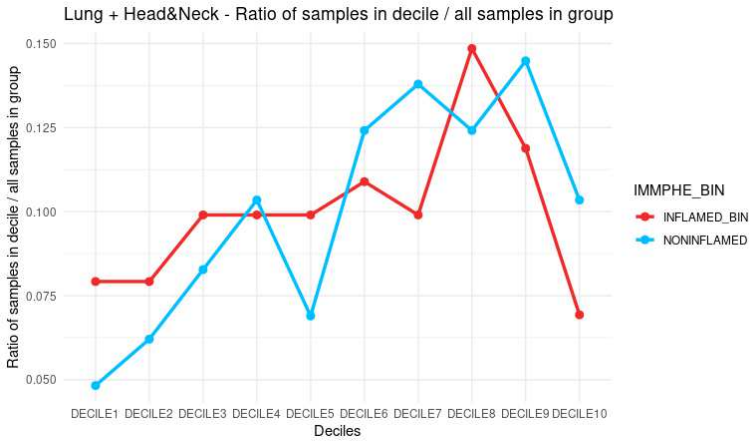

Figure 6.2: Ratio of samples in each ROPRO decile divided by sample count in the respective Immphe group (NSCLC / HNSC subset)

	Decile 1	Decile 10
Inflamed	8	7
Non-inflamed	7	15

Table 6.1: Number of samples in ROPRO decile 1&10 per Immphe classes

Between the lowest and highest quartiles the same difference does not show up even in sample numbers (figure 6.3, table 6.2) and it is not at all significant (X-squared = 0,1, p-value = 0,752).

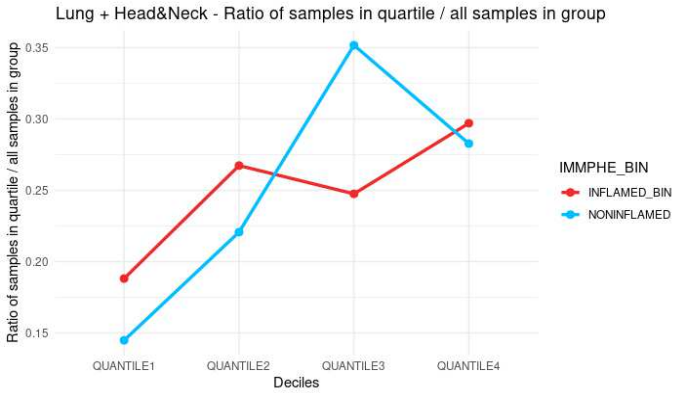

Figure 6.3: Ratio of samples in each ROPRO quartile divided by sample count in the respective Immphe group (NSCLC / HNSC subset)

	Quartile 1	Quartile 4
Inflamed	19	30
Non-inflamed	21	41

Table 6.2: Number of samples in ROPRO quartile 1&4 per Immphe classes

6.1b Contrast with Melanoma shows opposite trend

Very interestingly, the Melanoma group shows a clear trend too, in regards to the relationship of ROPRO and Immphe, however the correlation is opposite to the trend in the NSCLC / HNSC subset as well as in the overall population. When comparing the difference within the ROPRO distributions of the two Immphe classes (figure 6.4), it is significant (W = 285, p-value = 0,0089).

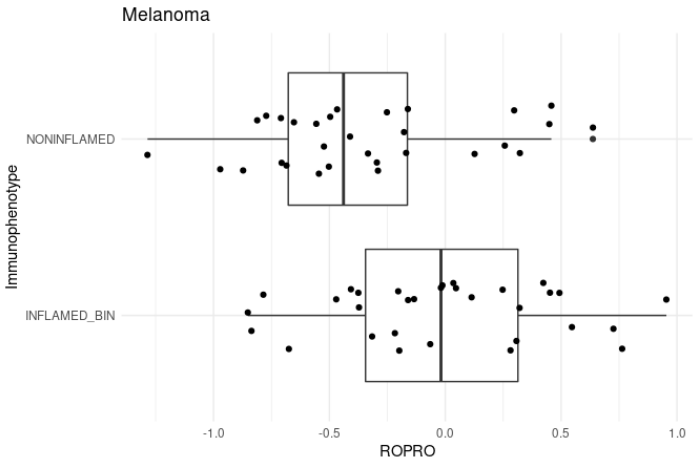

Figure 6.4: ROPRO values separated by the two Immphe groups, shown as points and as Box plots (Melanoma subset)

When comparing the tails of the distributions (figure 6.5-6), both with the deciles and quartiles it shows that decile 1 is associated with Non-inflamed IP and decile 10 is with Inflamed, the same is true with quartile 1 and 4. The difference between the deciles is not significant in table 6.3 (X-squared = 1,333, p-value = 0,248) probably because of the low sample size, but between quartiles it is in table 6.4 (X-squared = 4,545, p-value = 0,033).

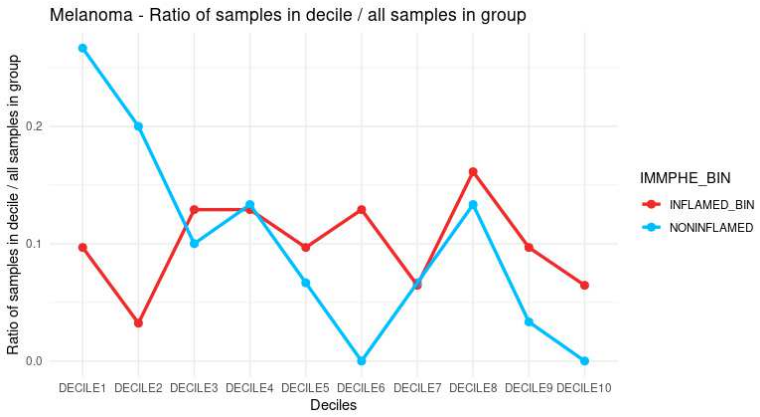

Figure 6.5: Ratio of samples in each ROPRO decile divided by sample count in the respective Immphe group (Melanoma subset)

	Decile 1	Decile 10
Inflamed	3	2
Non-inflamed	8	0

Table 6.3: Number of samples in ROPRO decile 1&10 per Immphe classes for Melanoma

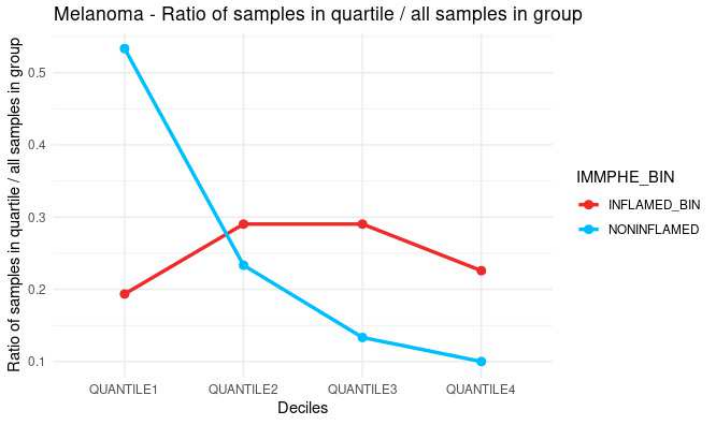

Figure 6.6: Ratio of samples in each ROPRO quartile divided by sample count in the respective Immphe group (Melanoma subset)

	Quartile 1	Quartile 4
Inflamed	6	7
Non-inflamed	16	3

Table 6.4: Number of samples in ROPRO quartile 1&4 per Immphe classes for Melanoma

6.2 Treatment length

The length of the previous CPI treatment is also a property that is assumed to be related to changes in immunogenicity indirectly, as it may indicate if the patient was already resistant to CPI treatments, acquired resistance during treatment or responded well to the treatment without becoming resistant. The ROPRO distributions of the three categories (short < 42 days, 42 days < middle < 1 year, long > 1 year) seem to be very similar (W = 1221, p-value = 0,5194). After summarizing the middle and short classes, because of the low sample size of short, in figure 6.7 can be also seen that no large differences show even based on CPI experience (naive group same distribution).

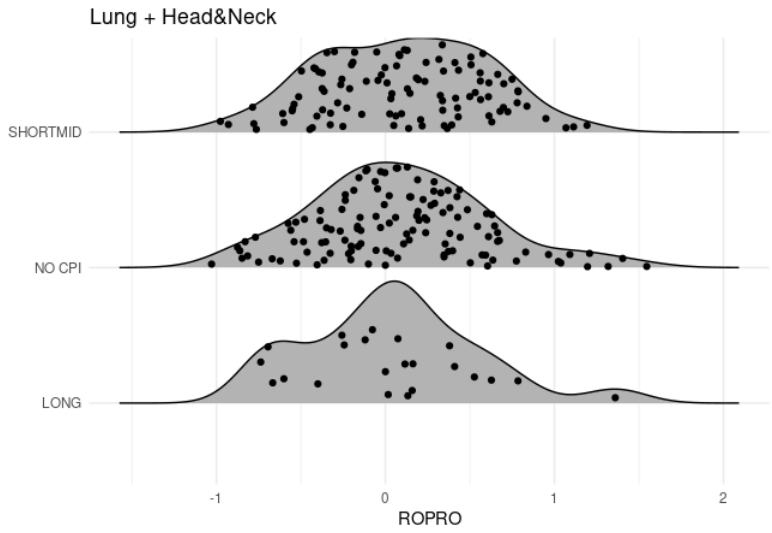

Figure 6.7: ROPRO distribution faceted by length of last cpi treatment

On the other hand, when looking at them, the contingency table for Immphe and treatment lengths categories show clear differences (table 6.5). Even after summarizing short and middle (table 6.6), the sample sizes are not balanced (only 20% of samples are in the long category) and the difference is on the brink of being significant (X-squared = 3,7896, p-value = 0,0516).

Immphe\Tr. Length	Long Duration	Middle Duration	Short Duration
Inflamed	14	37	2
Non-inflamed	8	57	6

Table 6.5: Contingency table of Immphe classes by treatment length groups (three groups)

Immphe\Tr. Length	Long Duration	Short+Middle Duration
Inflamed	14	39
Non-inflamed	8	63

Table 6.6: Contingency table of Immphe classes by treatment length groups (two groups)

6.3 Intermittent Therapies

One of the most important questions of this project was whether the influence of intermittent therapies could be identified. Considering ROPRO values of the first subset, it seems to have an influence on it, as some difference is visible in the median and quartiles (figure 6.8), statistically this difference is significant (W = 2339, p-value = 0,03639).

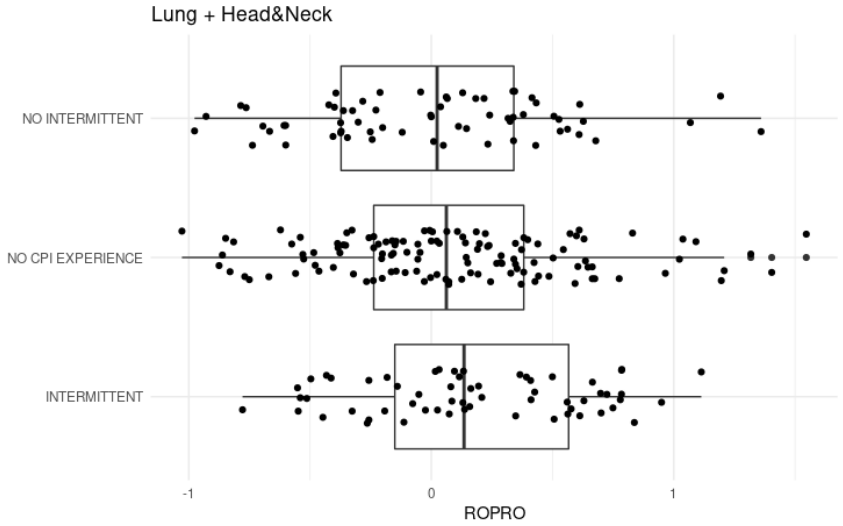

Figure 6.8: ROPRO value distribution separated by intermittent therapy status

Now looking at the deciles and quartiles (figure 6.9-10), the difference in the tails is not as clear but still visible. The p-value for deciles is very close to being significant (X-squared = 3,7403, p-value = 0,0531) for samples in table 6.7 and for quartiles this effect mostly diminishes in table 6.8 and the significance is no longer noticeable (X-squared = 1,0639, p-value = 0,3023).

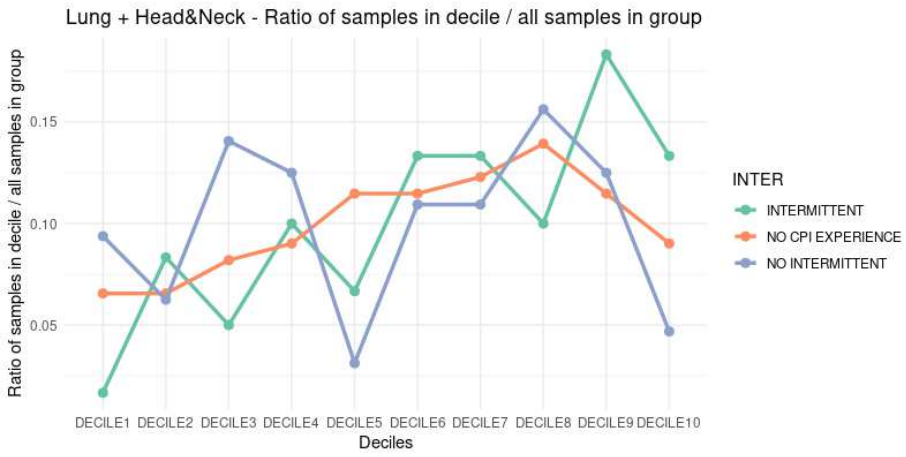

Figure 6.9: Ratio of samples in each ROPRO decile divided by sample count in the respective intermittent group

	Intermittent	No Intermittent
Decile1	1	6
Decile10	8	3

Table 6.7: Contingency table of samples in the lowest and highest deciles from the two intermittent categories

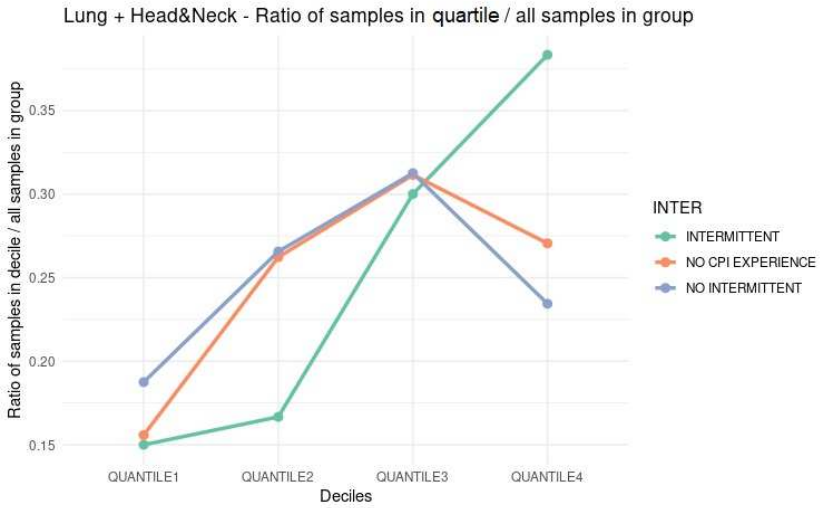

Figure 6.10: Ratio of samples in each ROPRO quartile divided by sample count in the respective intermittent group

	Intermittent	No Intermittent
Quantile1	9	12
Quantile4	23	15

Table 6.8: Contingency table of samples in the lowest and highest quantiles from the two intermittent categories

On the side of Immphe, some disparity can be noticed between the intermittent and no intermittent groups, even though right over the significance threshold (X-squared = 3,493, p-value = 0,0616). Within intermittent, non-inflamed Immphe is more common, which is consistent with the putative relationships of non-inflamed and intermittent groups with higher ROPRO.

	Intermittent	No Intermittent
Inflamed	20	33
Non-inflamed	40	31

Table 6.9: Contingency table of samples in the two Immphe classes from the two intermittent categories

6.4 The type of the last treatment before study

From the investigation of intermittent treatments in the previous part, the question logically arises: does it matter which type of therapy is received last? This time the CPI naive patients are also incorporated, as the intermittent therapy is a term only relevant for CPI experienced patients, but most patients received some type of treatment from the four main considered types.

In figure 6.11, the previous trend can be noticed of CPI experienced patients entering trials after a CPI treatment have lower ROPRO than experienced patients receiving another type of treatment last. After breaking down intermittent therapies into types, the difference of distributions was tested in 3 comparisons: chemotherapy for CPI exp. to CPI treatment (only existent in CPI exp. column) (W = 1347, p-value = 0,09561), chemotherapy for CPI exp. to radiotherapy for CPI exp. (W = 400, p-value = 0,4254) and radiotherapy for CPI exp. to CPI treatment (W = 939, p-value = 0,0206). Targeted therapy was not tested because of the low sample sizes.

Radiotherapy last (RxT - Experienced boxplot) is significantly different from CPI last (Immun - Experienced boxplot), but chemotherapy is not. For naives, their distributions for each treatment almost perfectly match experienced patients with CPI last.

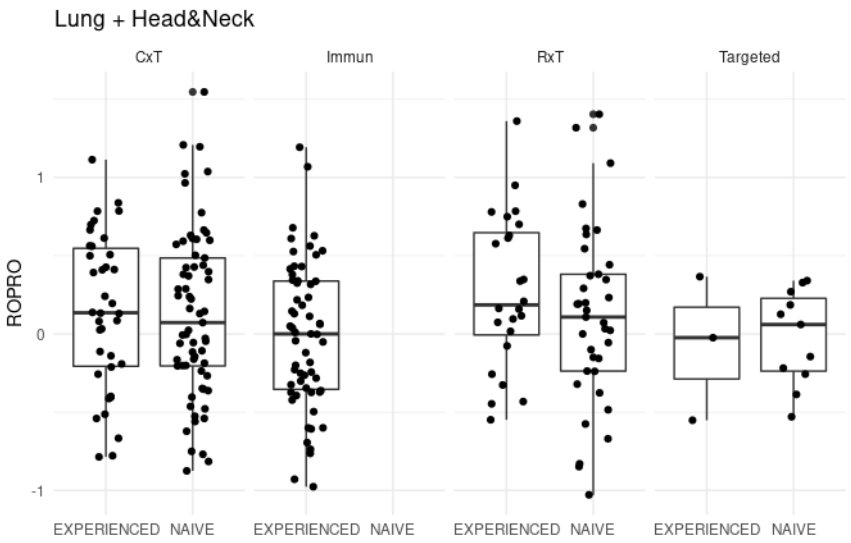

Figure 6.11: Distribution of ROPRO separated by last received treatment type and CPI experience status

When showing Immphe groups in a similar fashion, differences become apparent (figure 6.12–13). When chemotherapy or radiotherapy is the last line of therapy, non–inflamed phenotypes are more common for both CPI experienced and naive patients. A ≥ 1 inflamed/non–inflamed ratio can only be found for samples with CPI last or without relevant treatment.

On the CPI experienced side (figure 6.12) the CxT and CPI groups show a significant difference in inflamed to non-inflamed ratio (I-NI ratio) (X-squared = 5.4846, df = 1, p-value = 0.01918) as opposed to non-significant differences between CPI to RxT and CxT to RxT. However, there is a noticeable difference in this regard between CxT and RxT, more than in the CPI naive group. When looking at the CPI naives (figure 6.13), the I-NI ratio is very similar within each major treatment type (CxT, RxT, targeted) to the respective CPI experienced I-NI ratio.

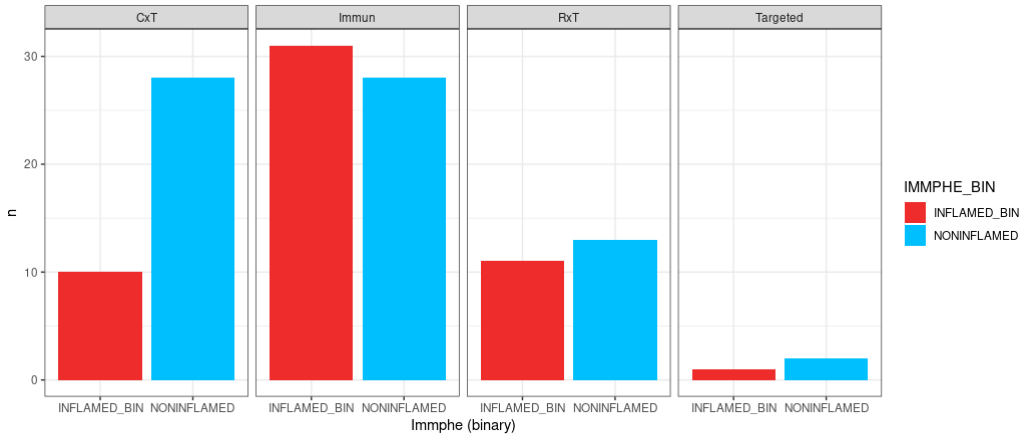

Figure 6.12: Total sample counts within Immphe classes faceted by last received treatment type; only CPI experienced patients

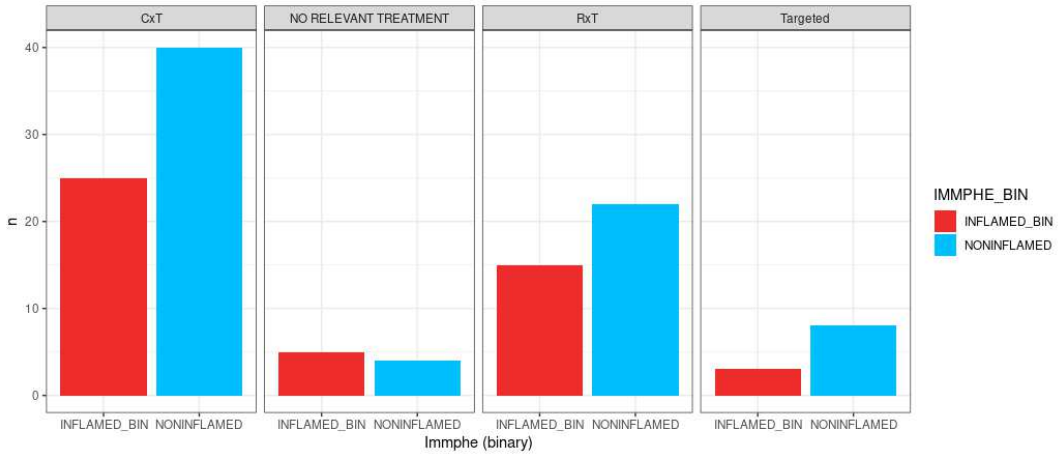

Figure 6.13: Total sample counts within Immphe classes faceted by last received treatment type; only CPI naive patients

I tested the difference of Immphe ratio between the following groups of CPI experienced patients (figure 6.12): CxT to CPI from table 6.10 (X-squared = 5,485, p-value = 0,01918), RxT to CPI from table 6.11 (X-squared = 0,09743, p-value = 0,7549) and CxT to RxT from table 6.12 (X-squared = 1,706, p-value = 0,1915) and only the difference between CxT and CPI is significant.

	Chemotherapy	CPI
Inflamed	10	31
Noninflamed	28	28

Table 6.10: Contingency table samples in Immphe classes in CxT compared to CPI treatment group (CPI experienced patients)

	CPI	Radiotherapy
Inflamed	31	11
Noninflamed	28	13

Table 6.11: Contingency table samples in Immphe classes in RxT compared to CPI treatment group (CPI experienced patients)

	Chemotherapy	Radiotherapy
Inflamed	10	11
Noninflamed	28	13

Table 6.12: Contingency table samples in Immphe classes in CxT compared to RxT treatment group (CPI experienced patients)

Discussion

Chapter 1-5:

Throughout the Results the analysis steps were laid out in the following order: Immphe was inferred for a large subset of patients, then MCA analysis was conducted which also used these inferred data points, subsetting data before the final analysis took place and finally the connection between treatment history and the predictors was investigated. Some of the decisions throughout the work will be discussed and points of observation will be better explained in this chapter.

The first step towards the final correlation analysis was the inference of Immphe. This was feasible because of the availability of RNA-Seq data for a large subset of patients in the ROTD cohort which did not possess determined Immphe. This method was able to expand the available Immphe information by 51%. Analysis, using two different methods (BIOQC, GSVA), showed high correlation of T-cell and other immune signatures with Immunophenotype as expected and so increases confidence in using inferred Immphe as a meaningful biomarker. Multiclass and binary classification were both applied to classify samples based on Immphe signature into desert, inflamed and excluded and inflamed, non-inflamed, respectively. Binary classification with Immphe (inflamed and non-inflamed categories) yielded 79,85% accuracy as opposed to 63,8% with a tertiary classification. Based on the observation that desert and excluded are hard to distinguish computationally, which lines up with the known traits of the phenotypes, it seemed more informative to use two categories (Inflamed vs. non-inflamed) rather than the original three classes (inflamed, excluded, desert).

Afterwards the treatment history of patients was explored, with a focus on the CPI experienced patient population. As CPIs are used more extensively in first line settings, this is the subset that will be more and more prevalent in future studies. In the treatment history analysis, I pivoted to consider the last treatment(s) before the trial, as these are the most recent, they may have the largest lingering effect on the immunogenicity.

By evaluating the views of MCA, two connections can be suspected between considered properties, intermittent therapy to ROPRO and treatment length to Immphe seem to be related, but as they were only visually inspected, their relationships need further evaluation.

The largest observed differences within the dataset during MCA and in the number of CPI experienced patients (compared to naives) are based on indication. This makes sense as different cancer types have different severity and are in general treated differently. So the number of LoTs, most common prior treatments vary, which influence the immunogenicity more so than other factors. Considering this heterogeneity in the patient population, the two main predictors of immunogenicity have been contrasted across indications so that similar cancer types in terms of ROPRO and Immphe can be merged to form larger subgroups for the last analysis step. NSCLC is the largest indication and it fits

together very well with HSNC, which also has numerous samples, and by combining these two, a reasonably homogenous subset of the whole cohort can be carved out. Three other indications are also selected which do not match with other indications, but they can provide a different perspective as common cancers and enough samples are available from them: Melanoma, the best responding cancer type to CPI, Renal Cell Carcinoma and Urothelial bladder cancer.

At the sorting part all samples without Immphe are left behind. This limitation ensures that for all considered samples both predictors of OS are present, as the two metrics capture different facets of patient immunogenicity.

Chapter 6:

When investigating a general trend it was seen earlier that Inflamed Immphe and low ROPRO corresponded to better OS, separately from each other. But when testing the correlation between ROPRO and Immunophenotype (chapter 6.1), the correlation is relatively little. This trend is even turned around (Pearson r becomes negative) in the case of Melanoma: Inflamed corresponds to worse ROPRO. These findings reflect on the fact that Immphe and ROPRO cover different aspects of patient wellbeing and overall survival and reinforce the proposal that both should be considered separately.

The length of the last CPI treatment (chapter 6.2) seems to be related to the Immphe status of patients, group with long treatment (>1 year) has more samples of the inflamed category, which is expected for patients who are taking a CPI medicine for a longer cycle, as this phenotype pretty much is a prerequisite for positive response to therapy. However, after such a long exposure to the same treatment, it is possible that acquired CPI resistance develops [39]. This resistance mechanism could answer why ROPRO is not significantly different between the groups. Even though T-cells are infiltrating the tumor, therefore contributing to an Inflamed phenotype, they are no longer able to effectively combat cancer.

In chapter 6.3 the relationship of intermittent therapy and ROPRO was investigated, where correlation was visible between the presence of an intermittent treatment and worse overall survival. Requiring such a treatment is likely dependent on some form of resistance to CPI. I argue these patients are not suitable for participating in trials of checkpoint inhibitor drugs.

Finally, through the examination of the relationship of the type of last treatment to Immphe & ROPRO (chapter 6.4), two significant differences were found between groups, both for the CPI experienced subset. First, the ROPRO distribution of the group with CPI as last treatment differs from the group with RxT (1) and the Immphe profile of CPI groups differs from CxT (2) (but not CPI v. RxT as for ROPRO).

Additionally, there are a lot less noninflamed samples with chemotherapy (CxT) as last treatment rather than radiotherapy (RxT). The reason can be that there is only one line of therapy after CPI in most cases in the experienced group (known from chapter3) and few combinations of CxT-RxT (like in naives), “sudden” medication with CxT often correspond to worse prognosis than RxT before trial, Immphe mirrors that.

Methods:

To infer immunophenotype, signature scores calculated by GSVA and BioQC, both available on the Bioconductor platform and classification was done with K-nearest neighbor (KNN) and logistic regression, widely used classifiers. As the data contains multiple categorical variables and the continuous ones can be easily transformed into categorical, Multiple Correspondence Analysis (MCA), similar to Principal Component Analysis, was a natural choice to explore and visualize the structure behind the data. For the final analysis, the relationship of each investigated feature with each of the two predictors was inspected separately, by looking for significant differences between subsets, through statistical tests (Wilcoxon rank-sum and Chi-square tests as applicable).

Limitations:

The present thesis work can naturally be extended in future work. To approximate the immunogenicity status, the immunophenotype was used as only a categorical variable (partly inferred from RNA-Seq). While it is a very useful property, as tumor RNA-Seq data was available for a number of patients anyway, investigating gene expression as a whole could have potentially provided a more comprehensive aspect of immune status. Available features of patients could have been investigated in other, more rigorous ways: batch effect removal could have been applied to remove potentially unwanted effects, this could have been used in chapter 5 as well instead of the manual subsetting. Instead of exploring the correlation of features and predictors one-by-one, a combined model could have been established which would have also measured the strength of effect of the distinct properties on the immunogenicity.

These issues can be sorted out in future projects, especially, when response data from current or recently finished clinical trials becomes available.

References

1. Robert C. A decade of immune-checkpoint inhibitors in cancer therapy. In: Nature Publishing Group UK [Internet]. 30 Jul 2020 [cited 21 Apr 2022]. doi:10.1038/s41467-020-17670-y

2. Chen DS, Mellman I. Elements of cancer immunity and the cancer–immune set point. Nature. 2017;541: 321–330.

3. Becker T, Weberpals J, Jegg AM, So WV, Fischer A, Weisser M, et al. An enhanced prognostic score for overall survival of patients with cancer derived from a large real-world cohort. Ann Oncol. 2020;31: 1561–1568.

4. Improved Survival with Ipilimumab in Patients with Metastatic Melanoma. New England Journal of Medicine. 2010. pp. 1290–1290. doi:10.1056/nejmx100063

5. Sharma P, Hu-Lieskovan S, Wargo JA, Ribas A. Primary, Adaptive, and Acquired Resistance to Cancer Immunotherapy. Cell. 2017;168: 707–723.

6. Zaretsky JM, Garcia-Diaz A, Shin DS, Escuin-Ordinas H, Hugo W, Hu-Lieskovan S, et al. Mutations Associated with Acquired Resistance to PD-1 Blockade in Melanoma. N Engl J Med. 2016;375: 819–829.

7. Havel JJ, Chowell D, Chan TA. The evolving landscape of biomarkers for checkpoint inhibitor immunotherapy. Nat Rev Cancer. 2019;19: 133–150.

8. Li B, Severson E, Pignon J-C, Zhao H, Li T, Novak J, et al. Comprehensive analyses of tumor immunity: implications for cancer immunotherapy. Genome Biol. 2016;17: 1–16.

9. Blank CU, Haanen JB, Ribas A, Schumacher TN. CANCER IMMUNOLOGY. The “cancer immunogram.” Science. 2016;352: 658–660.

10. Mechanism and potential predictive biomarkers of immune checkpoint inhibitors in NSCLC. Biomed Pharmacother. 2020;127: 109996.

11. Sharma P, Siddiqui BA, Anandhan S, Yadav SS, Subudhi SK, Gao J, et al. The Next Decade of Immune Checkpoint Therapy. Cancer Discov. 2021;11: 838–857.

12. Larkin J, Chiarion-Sileni V, Gonzalez R, Grob J-J, Rutkowski P, Lao CD, et al. Five-Year Survival with Combined Nivolumab and Ipilimumab in Advanced Melanoma. N Engl J Med. 2019;381: 1535–1546.

13. Herrera FG, Ronet C, Ochoa de Olza M, Barras D, Crespo I, Andreatta M, et al. Low-Dose Radiotherapy Reverses Tumor Immune Desertification and Resistance to Immunotherapy. Cancer Discov. 2022;12: 108–133.

14. Blankenstein T, Coulie PG, Gilboa E, Jaffee EM. The determinants of tumour immunogenicity. Nature Reviews Cancer. 2012. pp. 307–313. doi:10.1038/nrc3246

15. Hess LM, Li X, Wu Y, Goodloe RJ, Cui ZL. Defining treatment regimens and lines of therapy using real-world data in oncology. Future Oncol. 2021;17: 1865–1877.

16. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet. 2009;10: 57–63.

17. Ozsolak F, Milos PM. RNA sequencing: advances, challenges and opportunities. Nat Rev Genet.

2011;12: 87–98.

18. Naito Y, Saito K, Shiiba K, Ohuchi A, Saigenji K, Nagura H, et al. CD8+ T cells infiltrated within cancer cell nests as a prognostic factor in human colorectal cancer. Cancer Res. 1998;58: 3491–3494.

19. Herbst RS, Soria J-C, Kowanetz M, Fine GD, Hamid O, Gordon MS, et al. Predictive correlates of response to the anti-PD-L1 antibody MPDL3280A in cancer patients. Nature. 2014;515: 563–567.

20. Behring M, Ye Y, Elkholy A, Bajpai P, Agarwal S, Kim H-G, et al. Immunophenotype-associated gene signature in ductal breast tumors varies by receptor subtype, but the expression of individual signature genes remains consistent. Cancer Med. 2021;10: 5712–5720.

21. Mariathasan S, Turley SJ, Nickles D, Castiglioni A, Yuen K, Wang Y, et al. TGFβ attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. Nature. 2018;554: 544–548.

22. Ali HR, Chlon L, Pharoah PDP, Markowetz F, Caldas C. Patterns of Immune Infiltration in Breast Cancer and Their Clinical Implications: A Gene-Expression-Based Retrospective Study. PLoS Med. 2016;13: e1002194.

23. Feldmeyer L, Hudgens CW, Ray-Lyons G, Nagarajan P, Aung PP, Curry JL, et al. Density, Distribution, and Composition of Immune Infiltrates Correlate with Survival in Merkel Cell Carcinoma. Clin Cancer Res. 2016;22: 5553–5563.

24. Desmedt C, Zoppoli G, Gundem G, Pruneri G, Larsimont D, Fornili M, et al. Genomic Characterization of Primary Invasive Lobular Breast Cancer. J Clin Oncol. 2016;34: 1872–1881.

25. Zhang JD, Hatje K, Sturm G, Broger C, Ebeling M, Burtin M, et al. Detect tissue heterogeneity in gene expression data with BioQC. BMC Genomics. 2017;18: 277.

26. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics. 2013;14: 7.

27. Huber W, Carey VJ, Gentleman R, Anders S, Carlson M, Carvalho BS, et al. Orchestrating high-throughput genomic analysis with Bioconductor. Nat Methods. 2015;12: 115–121.

28. Website. Available: Kuhn M, Wickham H (2020). Tidymodels: a collection of packages for modeling and machine learning using tidyverse principles.. https://www.tidymodels.org.

29. Lê S, Josse J, Husson F. FactoMineR: AnRPackage for Multivariate Analysis. Journal of Statistical Software. 2008. doi:10.18637/jss.v025.i01

30. Extract and Visualize the Results of Multivariate Data Analyses [R package factoextra version 1.0.7]. 2020 [cited 6 Sep 2021]. Available: https://CRAN.R-project.org/package=factoextra

31. Fehrenbacher L, Spira A, Ballinger M, Kowanetz M, Vansteenkiste J, Mazieres J, et al. Atezolizumab versus docetaxel for patients with previously treated non-small-cell lung cancer (POPLAR): a multicentre, open-label, phase 2 randomised controlled trial. Lancet. 2016;387: 1837–1846.

32. Li H, van der Leun AM, Yofe I, Lubling Y, Gelbard-Solodkin D, van Akkooi ACJ, et al. Dysfunctional CD8 T Cells Form a Proliferative, Dynamically Regulated Compartment within Human Melanoma. Cell. 2019;176: 775–789.e18.

33. Danaher P, Warren S, Dennis L, D'Amico L, White A, Disis ML, et al. Gene expression markers

of Tumor Infiltrating Leukocytes. J Immunother Cancer. 2017;5: 18.

34. Tirosh I, Izar B, Prakadan SM, Wadsworth MH 2nd, Treacy D, Trombetta JJ, et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. Science. 2016;352: 189–196.

35. Zhang L, Yu X, Zheng L, Zhang Y, Li Y, Fang Q, et al. Lineage tracking reveals dynamic relationships of T cells in colorectal cancer. Nature. 2018;564: 268–272.

36. Sade-Feldman M, Yizhak K, Bjorgaard SL, Ray JP, de Boer CG, Jenkins RW, et al. Defining T Cell States Associated with Response to Checkpoint Immunotherapy in Melanoma. Cell. 2018;175: 998–1013.e20.

37. Wu M, Li X, Zhang T, Liu Z, Zhao Y. Identification of a Nine-Gene Signature and Establishment of a Prognostic Nomogram Predicting Overall Survival of Pancreatic Cancer. Front Oncol. 2019;9: 996.

38. Ayers M, Lunceford J, Nebozhyn M, Murphy E, Loboda A, Kaufman DR, et al. IFN-γ-related mRNA profile predicts clinical response to PD-1 blockade. J Clin Invest. 2017;127: 2930–2940.

39. Schoenfeld AJ, Hellmann MD. Acquired Resistance to Immune Checkpoint Inhibitors. Cancer Cell. 2020;37: 443–455.

Supplementary material

Treatment group	Therapies or medicines corresponding to group
CPIs	"Ipilimumab","Atezolizumab", "Avelumab", "Durvalumab", "Nivolumab", "Pembrolizumab", "Cemiplimab"
Chemotherapy	"Cyclophosphamide", "Mechlorethamine", "Chlorambucil","Melphalan", "Dacarbazine", "Nitrosoureas", "Temozolomide", "Thiotepa","Bendamustine","Procarbazine","Busulfan", "Daunorubicin","Doxorubicin","Epirubicin","Idarubicin","Mitoxantrone","Valrubicin", "Doxorubicin Pegylated Liposomal", "Paclitaxel", "Docetaxel", "Abraxane", "Taxotere", "Vorinostat", "Romidepsin","Panobinostat", "Irinotecan","Topotecan", "Etoposide", "Teniposide", "Tafluposide", "Azacitidine", "Azathioprine", "Capecitabine", "Cytarabine", "Doxifluridine", "Fluorouracil", "Gemcitabine", "Hydroxyurea", "Mercaptopurine", "Methotrexate", "Tioguanine", "Thioguanine","Decitabine","Cladribine", "Bleomycin","Actinomycin","Mitomycin", "Carboplatin", "Cisplatin", "Oxaliplatin", "Tretinoin" ,"Alitretinoin", "Bexarotene", "Vinblastine", "Vincristine", "Vindesine", "Vinorelbine", "Paclitaxel Protein-Bound", "Irinotecan Liposomal","Ifosfamide","Lomustine", "Eribulin", "Ixabepilone", "Cabazitaxel","Pemetrexed", "Trifluridine/Tipiracil","Arsenic","Cytarabine Liposomal", "Selinexor", "Fludarabine","Talazoparib"
Targeted therapy	"Megestrol", "Anastrozole", "Exemestane", "Fulvestrant", "Letrozole", "Leuprolide","Tamoxifen", "Bicalutamide","Abiraterone","Medroxyprogesterone","Degarelix","Triptorelin","Goserelin","Triptorelin", "Ibrutinib", "Binimetinib", "Encorafenib", "Cabozantinib", "Enzalutamide", "Trametinib", "Cobimetinib", "Dasatinib", "Dabrafenib", "Imatinib", "Regorafenib", "Sorafenib", "Axitinib", "Pazopanib", "Niraparib", "Lenvatinib", "Olaparib", "Afatinib", "Ceritinib", "Osimertinib", "Alectinib", "Crizotinib", "Gefitinib", "Erlotinib", "Loreatinib", "Palbociclib", "Vemurafenib", "Vismodegib", "Lorlatinib", "Abemaciclib", "Sunitinib", "Temsirolimus", "Everolimus", "Brigatinib", "Neratinib","Nilotinib","Bosutinib","Ruxolitinib","Idelalisib","Pertuzumab", "Ramucirumab", "Necitumumab", "Bevacizumab", "Bevacizumab-Awwb", "Cetuximab", "Panitumumab", "Trastuzumab", "Ado-Trastuzumab Emtansine", "Olaratumab","Elotuzumab", "Rituximab","Rituximab/Hyaluronidase","Isatuximab-Irfc","Alemtuzumab", "Rituximab-Abbs","Ziv-Aflibercept","Brentuximab Vedotin","Obinutuzumab" "Bortezomib", "Ixazomib", "Carfilzomib", "Daratumumab","Isatuximab", "Venetoclax"

Table S1: Grouping therapies/medicines into higher-level groups