POLITECNICO
MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

Executive Summary of the Thesis

Statistical assessment of Radiomics role in prognosis of patients with Intrahepatic Cholangiocarcinoma

Laurea Magistrale in Mathematical Engineering - Ingegneria Matematica

Author: Noemi Rossi

Advisor: Prof. Francesca Ieva

Academic year: 2020-2021

1. Introduction

Intrahepatic Cholangiocarcinoma (IHC) is an aggressive disease that affects the liver. Its incidence is increasing over last decades and diagnosis is difficult at early stages, due to IHC complicated biology. Five-years survival rate ranges from 25% to 40%. The main treatment is surgery, chemotherapy has a limited effectiveness and an optimal strategy for patients with resectable IHC is not well characterized. Moreover, prognostic factors are still debated, robust biomarker are lacking and precision medicine approach with an adequate non-invasive preoperative assessment of tumor biology and prognosis is still not available [1].
In recent years, an approach of non-invasive image-based tissue analyses, namely Radiomics, has emerged [2]. Radiomics is able to capture additional information not currently considered by clinicians in making prognosis, and more specifically, it recognizes patterns which are related to clinical feature. In this way relevant data are mined from any image modalities, such as computed tomography (CT), and these data can be used jointly to traditional clinical information, to model clinical outcomes, for evidence-based clinical decision support. In literature concerning IHC, it has been demonstrated that models including radiomic features outperform traditional ones, predicting pathology data and patients' outcome with high accuracy.
In this thesis we aim to build robust models that are capable to predict pathology data and survival response in patients with IHC, making use of the information provided by patients' clinical history and radiomics. The data analysed are collected by Humanitas University and are related to patients that have undergone a liver resection for IHC. These data are embedded in a multicentre study that aims to understand the importance and the role of radiomics in predicting targeted outcomes. At first, we aim to investigate whether radiomics of both tumour (core) and peritumoral area (margin) contribute to improve performances. Afterwards, as we are provided with radiomic data covering three phases of Computed Tomography scan (the technique used for collecting the diagnostic images) we are interested in analyse the importance of each of these three phases, understanding if they carry the same information or if each one gives its own contribution in prediction. Moreover, in performing these tasks, the multicentre nature of the data present in this study needs to be

considered, discovering if there are differences among centres.

2. IHC Dataset

The data analysed in this thesis are provided by Department of Hepatobiliary and General Surgery of Humanitas Clinical and Research Center. The information provided concerns patients that have undergone a liver resection for IHC confirmed at final pathology, from 2009 to 2019. The patients comply strictly inclusion/exclusion criteria and come from 6 different hospitals. The number of collected samples in the dataset is 261. Each row of the dataset corresponds to a patient which is associated with the hospital of origin and for which information concerning clinical and radiomic characteristics was recorded. Variables used to describe the patients are divided into:

- **Clinical Variables:** they correspond to information that is known to clinicians without the use of radiomics. They are categorized in Preoperative features, that correspond to details that are available prior to the curative surgery, and in Postoperative features, which are derived from information obtained from histological pathological samples after surgery.
- **Radiomic Variables:** they correspond to the quantitative features mined from the diagnostic images, that represent how voxels (pixels) intensity values are distributed in the target area. Radiomic information is extracted from the three difference phases of the CT scan, namely Portal, Arterial and Late, from both the main tumor zone (Core) and surrounding peritumoral area (Margin). For each of these phases, for both core and margin, 50 radiomic variables are collected, producing 300 covariates in total.

To prepare the data for the analysis, dealing with the large number of radiomic features and the small sample size, we have imputed the missing values of clinical input covariate with Multiple Imputation and we have deleted the highly correlated radiomic covariates (with 0.85 as threshold).
The outcomes that are analysed within this work are the following clinical covariates:

- **Microscopic Vascular Invasion (MVI)**: it is a binary variable that assess the presence or not of microscopic vascular invasion.
- **Grading**: it is a categorical variable with values {1, 2, 3} that describes the aggressiveness of the tumor. We are interested in the two classes classification 1-2 vs 3.
- **Overall Survival (OS)**: OS represents the overall survival time (in days) of the patient. It is associated with a categorical variable that indicates the censoring state of the observation.
- **Recurrence-Free Survival (RFS)**: RFS represents the relapse-free time (in days) of the patient. It is associated with a categorical variable that indicates the censoring state of the observation.

3. Pipeline of the Work

Given the nature of the outcomes, i.e. binary variables for pathology outcomes and time-to-event data for survival outcomes, we have to face two different problems:

- **Classification:** To classify pathology data, Logistic Regression (LR) and Mixed Effects Models (MEMs) are employed. With LR we select the best models for describing the outcomes, trying several variable selection techniques. With MEMs we take into account the multilevel nature of the data in the best models determined.
- **Survival Analysis:** To predict survival response we use Cox-Proportional Hazards Models (Cox-PH) and Shared Frailty Models. With Cox-PH models we select the best models for describing the outcomes. With Shared Frailty models we take into account the multilevel nature of the data in the best models determined.

Initially, we focus only on the Portal phase of the CT scan, as this is the main phase on which decisions are made and as we have data for all patients. To understand the impact that radiomics may have on the predictive ability of the model, we test the models with different sets of input covariates: we start considering the clinical covariates alone, adding progressively radiomic covariates belonging to the tumor and to the peritumoral margin area.
Afterwards, multiview aspect of the data is exploited. Considering all radiomic data, the number of patients included in the analysis decreases

to 190, due to the presence of missing records. Since the sample size is small and the number of features is very large as we take into account all radiomics data, we reduced the dimensionality of the dataset using two Multiview approaches, i.e. Multiview Canonical Correlation Analysis (MCCA) and Kernel Multiview Canonical Correlation Analysis (KMCCA) [3]. With these techniques we are able to decrease the number of the feature in the dataset considering all radiomic phases jointly to optimize the process. Classification and Survival Analysis are performed on the Multiview reduced dataset.

4. Application and Results

4.1. Classification of Pathology Data

Separately for each outcome, we apply Logistic Regression to clinical and radiomics covariates belonging to core and margin of the Portal phase, applying various methods of variable selection, namely Forward Selection, Backward Selection, Stepwise Selection, Ridge Regression, Lasso Regression and Principal Components Regression. To analyse the importance of the radiomic features in classification, three scenarios of grouped covariates are considered, following a clinical rationale:

- Clinical
- Clinical + Portal(Core)
- Clinical + Portal(Core+Margin)

Logistic Regression is used to identify the best model for each of the above cases, not taking into account multicentre aspect of the data. Then, with covariates identified in each of the best models, centre effect is analysed fitting Mixed Effects Models. To select the best model we examined the following performance metrics: accuracy, specificity, sensitivity, precision, precision-recall AUC and ROC AUC. To identify the covariates retained by the various variable selection techniques, the models have been trained on the entire dataset and the performances tested on the whole dataset are considered. However, in order to have a more realistic estimation of the performances, two different techniques of cross-validation are used to test the models:

- **Method 1 of Cross-Validation:** usually Stratified K-fold Cross-Validation with k=50.
- **Method 2 of Cross-Validation:** the data are split into a training set (80%) and a test set (20%) stratifying the outcome. The validation procedure was repeated 100 times over 100 different samples. The performances for each metric produced on each individual sample was collected in a dataset of 100 sample for later use.

The best models are selected looking at performances in cross-validation. For both MVI and Grading, after a thorough selection process, we identified the best models: they are the one obtained applying Logistic Regression on Clinical+Portal(Core+Margin) features with Backward Selection. Fitting MEMs with the covariates obtained in the best models, from values of VPC, that are respectively 19.91% and 26.62% for MVI and Grading, we conclude that the effect of the grouping factor is strongly present in the data, as it explains a part of the variability in the data.

In order to have a powerful way of demonstrating the added value of radiomics and margin, than simply consider that the best models selected are the one that have in input Clinical+Portal(Core+Margin) features, the performances on the test set of the various samples of the cross-validation 2 are exploited. We have collected the performances on each single test set sample producing a new dataset with 100 rows that correspond to a sample and a column for each performance matrix, 6 in total. This dataset is produced for every MEM that is fitted and it is used to perform *Permutation Tests* on the mean. Permutation tests are used by us to test, for each performance metric, whether there is statistical evidence to say that, by including radiomics and margin, the average of a given metric is higher than in the case where less information is considered. Formally, we will do the following one-sided tests for the means:

- H_0 : Clinical $\geq$ Core vs
  H_1 : Clinical $<$ Core
- H_0 : Clinical $\geq$ Core + Margin vs
  H_1 : Clinical $<$ Core + Margin
- H_0 : Core$\geq$ Core + Margin vs
  H_1 : Core $<$ Core + Margin

The results are reported in Table 1.

From the results of these tests it can be seen that radiomics brings added value to the predictive performances of the models they are inserted.

Table 1: P-values of Permutation Tests applied to values of performances obtained in Cross-validation 2 method while classifying MVI and Grading with MEMs

	Clinical vs Core		**Clinical vs Core+Margin**		**Core vs Core+Margin**	
	MVI	GRADING	MVI	GRADING	MVI	GRADING
ACCURACY	0.0221	0.0068	0.0001	<0.0001	0.0301	0.0004
SPECIFICITY	0.0408	0.006	0.0009	<0.0001	0.0935	0.1111
SENSITIVITY	0.0544	0.0125	0.0001	<0.0001	0.0110	<0.0001
PRECISION	0.0734	0.0005	0.0473	<0.0001	0.3610	0.6422
PR AUC	0.0043	<0.0001	< 0.0001	<0.0001	0.1126	0.0013
ROC AUC	0.0066	0.4055	< 0.0001	<0.0001	0.0549	<0.0001

In particular, these data highlight the additional information that the part of radiomics associated with the area surrounding the tumour brings. Moreover, in this analysis of pathology data classification, both for MVI and Grading, there is a strong centre effect. Therefore, the variable recording the hospital of origin of the patients is crucial in explaining part of the variability of the data.

4.2. Survival Analysis

In the case of this study, in the first part, Survival Analysis is applied only to radiomic covariate belonging to Portal phase, jointly with clinical ones. In these analyses we want to understand the radiomics role, and thus the importance of tumor (Core) and peritumoral area (Margin) in predicting outcome. In addition, in order to understand whether radiomics can provide adequate non-invasive preoperative assessment, we decided to consider preoperative and postoperative clinical covariates separately, following a clinical rationale. To answer these questions, in order to study the prognostic impact of the radiomic features on the responses, six scenarios of grouped covariates are analysed:

- Preoperative
- Preoperative + Portal(Core)
- Preoperative + Portal(Core+Margin)
- Postoperative
- Postoperative + Portal(Core)
- Postoperative + Portal(Core+Margin)

For each of the above cases, not taking into account the multicentre nature of the data Cox Proportional Hazard models are fitted jointly with Stepwise Selection, in order to find the best model. With the covariates given by the latter, a Shared Frailty model is used in to analyse the grouping effect. To select the best model Concordance index (C-Index) is used as evaluation metric, since it validates the predictive ability of a survival model. The model selected is the one with the highest C-index.
For both OS and RFS, after a thorough selection process, we found out that the best models are the one obtained applying Cox-PH model to Postoperative+Portal(Core+Margin) features. Fitting Shared Frailty with the covariates obtained in the best models, from p-values of the Commongen-Andersen test of heterogeneity, that are respectively 0.654 and 0.865 for OS and RFS, we conclude that there is no statistical evidence to say that the grouping factor is significant in survival analysis. To understand the importance of radiomics in the prediction of the survival response, Table 2 is produced to have a unified view of the results. From Table 2 different conclusions can be drawn. First, it can be seen that, by adding radiomic covariates to clinical ones, the C-index value increases in both cases. Therefore, radiomics gives added predictive value in Survival Analysis. Secondly, it can be noticed that, by adding Margin covariates to clinical and Core, the C-index rises. Hence, the features belonging to Margin are important for predictive purposes, as they give added value

Table 2: Summary of Cox-PH models results with C-Index

	OS	**RFS**
PREOP	0.682	0.66
PREOP+CORE	0.713	0.668
PREOP+CORE+MARGIN	0.752	0.710
POSTOP	0.755	0.677
POSTOP+CORE	0.766	0.716
POSTOP+CORE+MARGIN	0.797	0.733

compared to Core variables. Lastly, C-Indexes in Preop+Core+Margin and Postop+Core cases are comparable. This not only testifies to the additional value of the Margin, as already observed in the previous case; but shows how preoperative information integrated with radiomics can achieve similar performance to the postoperative case. Thus, an adequate non-invasive preoperative assessment is possible taking into consideration both tumour and margin radiomics information.

4.3. Multiview Dimensionality Reduction

In this Section we consider, for both core and margin, all three radiomic phases to perform classification and survival analysis, in order to explore the contribution that each of the phases can make to prediction. We aim to understand whether including each phase is decisive for modelling the outcomes.

Portal, Arterial and Late phases of the CT scan can be considered as a multiple view representations of the tumor and its surrounding area. Thus, *Multiview Learning* techniques are natural candidates to decrease the number of input covariates in the models, properly accounting for the multiview aspect of the data. In order to analyse the effect of considering all radiomics information and multiview modelling on the prediction performances, four settings of different covariates are considered:

- Only Portal Covariates: this case is the baseline case adopted to see if it is worth considering the three radiomic phases.
- All Radiomics covariates simply concatenated: this case is adopted in order to be able to understand the difference made by using a Multiview approach.
- MCCA: MCCA is performed on both core and margin of all radiomic phases, repeating the process twice for both areas separately, to jointly reduce the dimensionality of each view to 10 components, producing a total of 60 features.
- KMCCA: same as in MCCA but components are computed using a Gaussian Kernel.

For each of these cases, always jointly considering clinical information, the results and performances of both Classification (with Backward Selection) and Survival Analysis (with Stepwise Selection) are analysed. Regarding Classification, we use Permutation tests to test, for each performance metric, whether there is statistical evidence to say that by including all radiomics information, with multiview approach or not, the average of a given metric is higher than in the case where only Portal information is considered. The result for MVI and Grading are reported in Figures 1 and 2. It can be concluded

Figure 1: Permutation Tests results for MVI MEMs considering multiview aspect of radiomics

ACCURACY		PRECISION	
H0	P-value	H0	P-value
Portal>KMCCA	<0.0001	Portal>KMCCA	<0.0001
Portal>MCCA	<0.0001	Portal>MCCA	0.0031
Portal>Radiomica	<0.0001	Portal>Radiomica	<0.0001
Radiomica=KMCCA	0.51	Radiomica=KMCCA	0.805
Radiomica>MCCA	0.9988	Radiomica>MCCA	0.9958
SPECIFICITY		**PR AUC**	
H0	P-value	H0	P-value
Portal>KMCCA	<0.0001	Portal>KMCCA	<0.0001
Portal>MCCA	0.0032	Portal>MCCA	0.004
Portal>Radiomica	<0.0001	Portal>Radiomica	<0.0001
Radiomica=KMCCA	0.757	Radiomica=KMCCA	0.337
Radiomica>MCCA	0.9839	Radiomica>MCCA	0.995
SENSITIVITY		**ROC AUC**	
H0	P-value	H0	P-value
Portal>KMCCA	<0.0001	Portal>KMCCA	<0.0001
Portal>MCCA	0.0003	Portal>MCCA	0.0003
Portal>Radiomica	<0.0001	Portal>Radiomica	<0.0001
Radiomica=KMCCA	0.515	Radiomica=KMCCA	0.09
Radiomica=MCCA	0.145	Radiomica>MCCA	0.9986

Figure 2: Permutation Tests results for Grading MEMs considering multiview aspect of radiomics

ACCURACY		PRECISION	
H0	P-value	H0	P-value
Portal>KMCCA	<0.0001	Portal>KMCCA	<0.0001
Portal=MCCA	0.228	Portal=MCCA	0.325
Portal=Radiomica	0.492	Portal>Radiomica	0.155
Radiomica>KMCCA	<0.0001	Radiomica>KMCCA	<0.0001
Radiomica>MCCA	0.0374	Radiomica>MCCA	0.0089
SPECIFICITY		**PR AUC**	
H0	P-value	H0	P-value
Portal>KMCCA	<0.0001	Portal>KMCCA	<0.0001
Portal=MCCA	0.453	Portal>MCCA	0.0008
Portal=Radiomica	0.418	Portal>Radiomica	0.0491
Radiomica>KMCCA	<0.0001	Radiomica>KMCCA	<0.0001
Radiomica>MCCA	0.0576	Radiomica>MCCA	0.0562
SENSITIVITY		**ROC AUC**	
H0	P-value	H0	P-value
Portal>KMCCA	0.0012	Portal>KMCCA	<0.0001
Portal>MCCA	0.0952	Portal>MCCA	0.0575
Portal=Radiomica	0.781	Portal=Radiomica	0.282
Radiomica>KMCCA	0.0089	Radiomica>KMCCA	<0.0001
Radiomica=MCCA	0.349	Radiomica>MCCA	0.0058

that considering all radiomic data is important in order to improve the predictive ability of the model. Depending on how all radiomic information is considered, whether or not the Multiview aspect of the data is taken into account, there is a greater or lesser increase in the predictive performance of the model. In the case of MVI, using

multiview techniques does not exceed the performance of the model in which all radiomics is concatenated. In the case of Grading, accounting for the multiview nature of the data using KMCCA, the model outperforms all other cases. Regarding Survival Analysis, to understand the importance of radiomics in the prediction of the survival response, Table 3 is produced to have a unified view of the result. From Table 3 it

Table 3: C-Indexes of Cox-PH models fitted for analyse the benefit of using Multiview Dimensionality Reduction techniques for radiomics

	OS	**RFS**
PORTAL PHASE	0.807	0.763
ALL RADIOMICS	0.838	0.811
KMCCA	0.838	0.729
MCCA	0.821	0.779

can be concluded that including all information supplied by the three phases of radiomics is important in order to increase the predictive performance of the model. However, in case of Survival Analysis, considering the multiview nature of the data does not increase the predictive ability of the model, which is at best equalled for OS with KMCCA, compared to the case where all radiomics is considered concatenated.

5. Conclusions

Within this work, we developed robust models capable to classify pathology data and predict survival response in patients with IHC. For what concerns Classification, we first employed Generalized Linear Models. Then, we used their random effect version to properly account for the hierarchy of the data. The same approach was followed for time-to-event data: initially we used Cox type regression models to find the best models. Afterwards, we employed Shared Frailty models. At the end of the work, in order to better consider the multiview aspect of the radiomic information available, we used MCCA and KMCCA as dimensionality reduction techniques. With these methods, we were able to decrease the number of covariates to be given as input to the model with respect to optimising the process by considering all views simultaneously. The results of the dimensionality reduction were used in Classification and Survival Analyses to understand the advantages of using all radiomic information with a multiview approach.
We highlighted the importance of considering the information provided by radiomics, in conjunction with clinical data, to have an adequate prognosis in patients with IHC. Using radiomics, predictive performances of classification and survival models improve, so that we are capable of predicting more accurately quantities that are relevant to know in order to find the proper treatment. Moreover, we discovered that both the radiomics of the core tumor zone and surrounding peritumoral area are relevant for the analysis, as together contribute in prediction of the outcomes. With regard to the three different phases of the CT scan, we showed the importance of considering them together, as they are not redundant descriptions of the same subject, but each provides added value to the analysis. How to take this information into account, whether with a multiview approach or not, depends on the modelled outcome, as there is no technique that always outperforms the others. Furthermore, we pointed out that when modelling pathology data, it is necessary to consider the grouping factor present in the data, since in MEMs the centre-related random effect is strongly present. On the other hand, for Survival Analysis, hospital grouping may not be taken into account as it is not significant in the model, probably because long-term outcomes are examined.

References

[1] Sumera Rizvi, Shahid A Khan, Christopher L Hallemeier, Robin K Kelley, and Gregory J Gores. Cholangiocarcinoma—evolving concepts and therapeutic strategies. *Nature reviews Clinical oncology*, 15(2):95–111, 2018.

[2] Robert J Gillies, Paul E Kinahan, and Hedvig Hricak. Radiomics: images are more than pictures, they are data. *Radiology*, 278(2): 563–577, 2016.

[3] Jan Rupnik and John Shawe-Taylor. Multiview canonical correlation analysis. In *Conference on data mining and data warehouses (SiKDD 2010)*, pages 1–4, 2010.