

RADIOMICA PER LA PREDIZIONE DELLA RISPOSTA ALLA CHEMIOTERAPIA NEL TUMORE DELLA MAMMELLA

PREMIO ANGELO DRIGO

Dr. Giulio Benetti – 7 Novembre 2024

Relatore: Prof. Marcello Lunardon

Correlatore: Dr. Carlo Cavedon

INDICE

INTRO

2. Il tumore alla mammella
3. La chemioterapia neoadiuvante

ASPETTI TECNICI

4. La risonanza magnetica
5. L'ispezione visiva
6. La radiomica
7. Il progetto
8. Il processo di estrazione
9. Dalla regressione lineare a LASSO
10. La selezione delle features e la LOOCV

11. Le variabili cliniche e istologiche
12. I gruppi di covariate
13. Il confronto tra tre classificatori

14. Lo studio prospettico
15. BIRRa
16. Qt Creator + Python
17. La finestra principale
18. PyInstaller
19. Risultati della validazione
20. Confronto con revisione sistematica
21. Conclusioni

RISULTATI

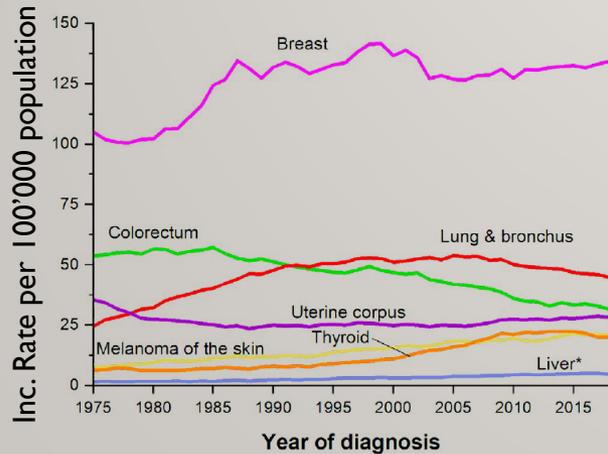
VALIDAZIONE

2 IL TUMORE ALLA MAMMELLA

Estimated New Cases

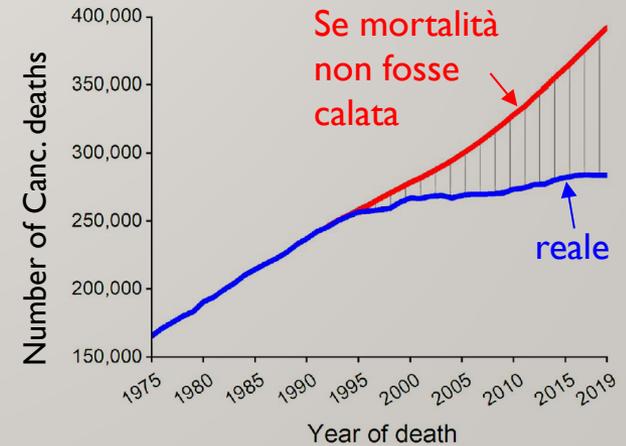
			Males	Females		
Prostate	268,490	27%		Breast	287,850	31%
Lung & bronchus	117,910	12%		Lung & bronchus	118,830	13%
Colon & rectum	80,690	8%		Colon & rectum	70,340	8%
Urinary bladder	61,700	6%		Uterine corpus	65,950	7%
Melanoma of the skin	57,180	6%		Melanoma of the skin	42,600	5%
Kidney & renal pelvis	50,290	5%		Non-Hodgkin lymphoma	36,350	4%
Non-Hodgkin lymphoma	44,120	4%		Thyroid	31,940	3%
Oral cavity & pharynx	38,700	4%		Pancreas	29,240	3%
Leukemia	35,810	4%		Kidney & renal pelvis	28,710	3%
Pancreas	32,970	3%		Leukemia	24,840	3%
All Sites	983,160	100%		All Sites	934,870	100%

Dati USA



- Incidenza: 2.1 Mdonne/y
- Terza causa di neoplasia
- Prima nel sesso femminile

- Incidenza: +0.3% / y
- Mortalità: -0.8% / y



3 LA CHEMIOTERAPIA NEOADIUVANTE (NAC)

- Carcinomi mammari in stadio localmente avanzato per i quali l'intervento chirurgico non è indicato come primo approccio terapeutico, poiché non sarebbe radicale
- Tumori mammari operabili se potenziale vantaggio acquisito dalla sua somministrazione



- Riduzione del volume tumorale (downstaging)
- Controllo precoce di metastasi a distanza
- Ridurre il rischio di recidiva
- Vantaggio in termini di sopravvivenza libera da malattia (EFS) e sopravvivenza globale (OS)

neutropenia febbrile, cardiotossicità, tossicità polmonare, neurotossicità, malignità ematologica, ...

Se la risposta alla NAC è limitata il bilancio potrebbe essere negativo

4 LA RISONANZA MAGNETICA (MRI)

- Sensibilità diagnostica (95-100%) maggiore dell'ecografia o della mammografia[1]
- Utilizzata per monitorare l'andamento della NAC

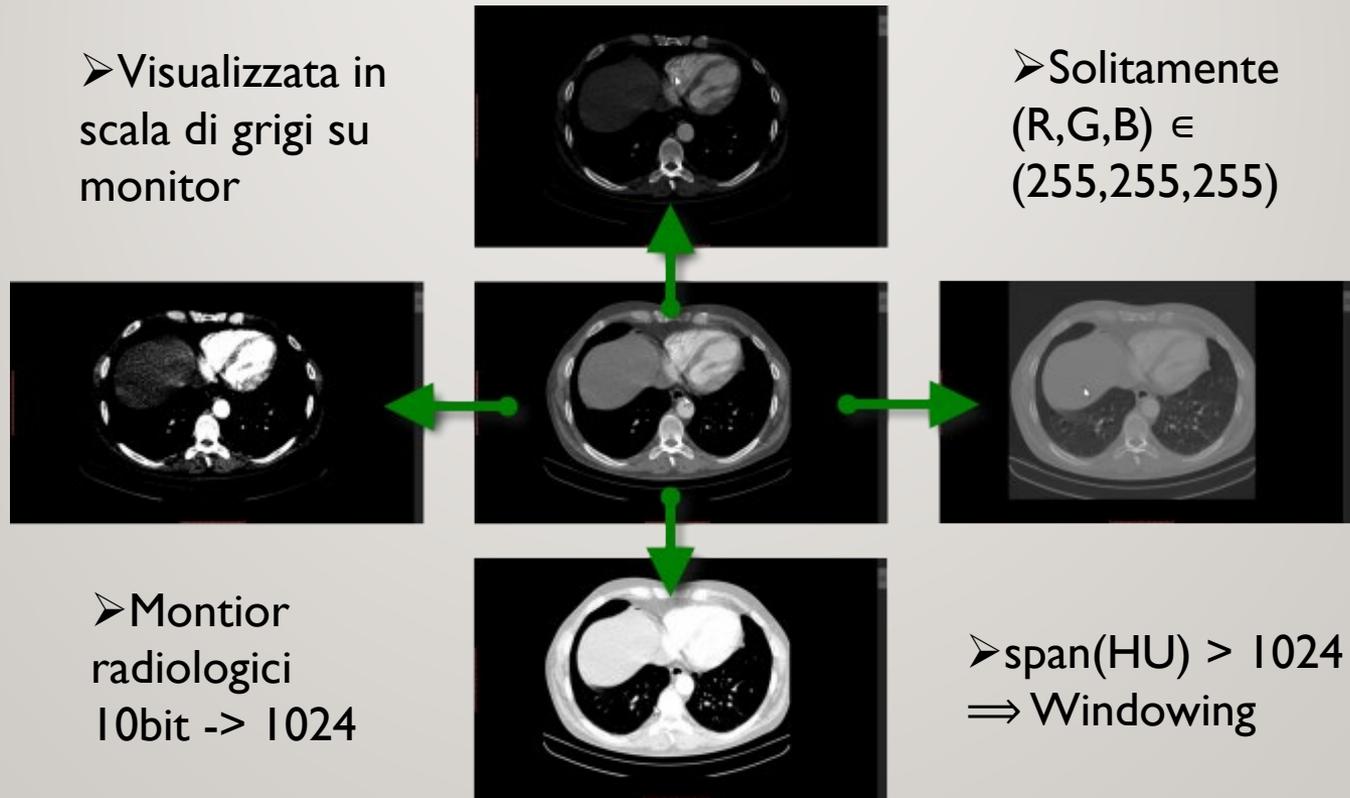
Dynamic Contrast Enhanced (DCE-MRI)

- Sfrutta mezzo di contrasto (MnC) contenente Gd (paramagnetico)
- Sequenze dinamiche (temporali) per valutare l'enhancement della lesione
- Può essere utile creare un'immagine *sub* sottraendo l'immagine pre-contrasto

Diffusion Weighted Imaging (DWI-MRI)

- Quantifica la diffusione dovuta ai movimenti microscopici delle molecole di acqua nei tessuti
- Viene quantificata tramite l'ADC (Apparent Diffusion Coefficient), solitamente valutato all'interno di una ROI (Region of Interest)

5 L'ISPEZIONE VISIVA



➤ Eye-driven \Rightarrow vedo tutto?

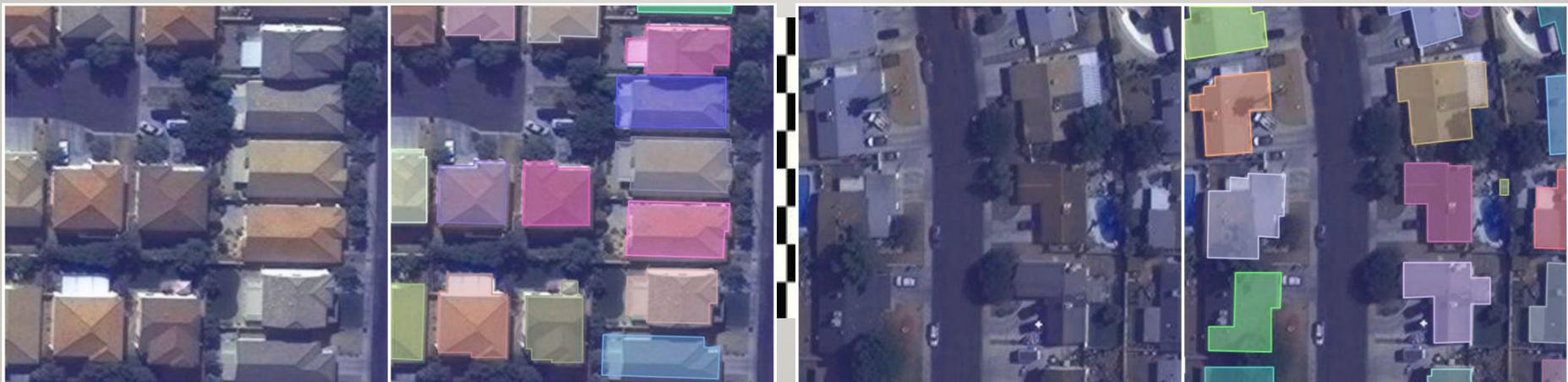
➤ USER+environment -dependent \Rightarrow riproducibile?

6 LA RADIOMICA

- Processo di analisi quantitativa dell'imaging digitale radiologico
- Si basa sull'estrazione di dati (**features**) dalle immagini mediante algoritmi
- Le features sono dati oggettivi da analizzare in termini quantitativi.

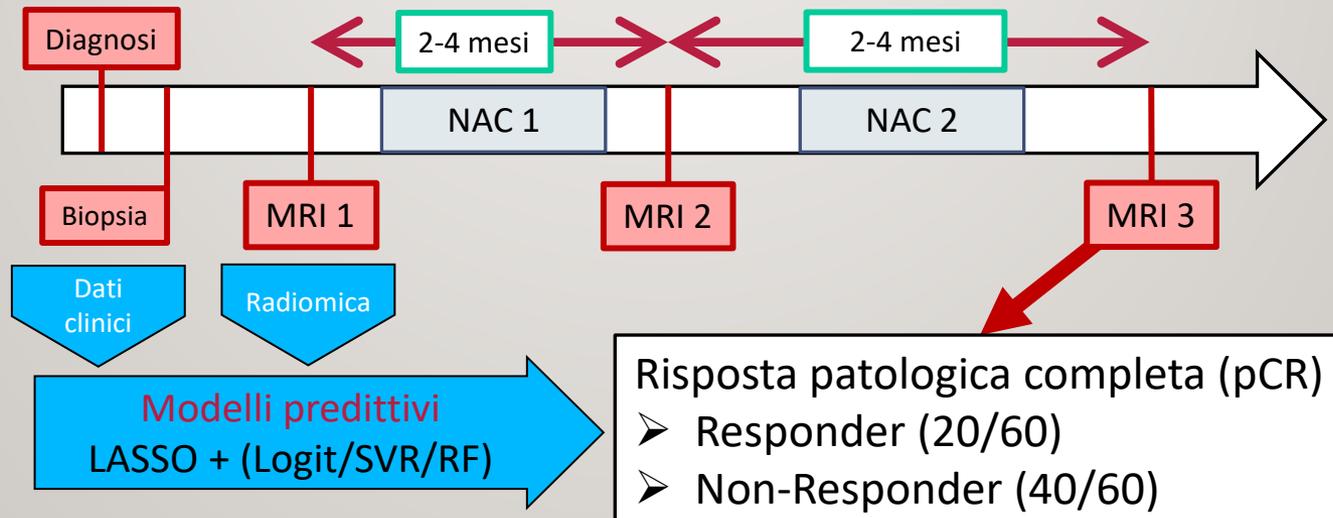
Features:

- PRIMO ORDINE: distribuzione statistica dei valori di grigio (media, curtosi, asimmetria)
- DI FORMA: analisi riferita alla geometria del VOI (Volume of Interest)
- TESSITURALI: analisi delle caratteristiche specifiche dell'ambiente in cui si trova il voxel



7 IL PROGETTO

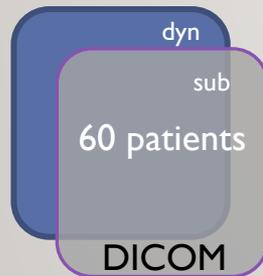
- Alcuni tumori non rispondono alla NAC, tuttavia gli effetti collaterali della terapia possono essere importanti



- **Predire** la risposta dei pazienti alla NAC prima dell'inizio del trattamento per proporre un percorso di cura personalizzato
- **Confrontare** le prestazioni di diversi modelli predittivi
- Sono stati usati pazienti sottoposti a NAC nel periodo 2016-19

8 IL PROCESSO DI ESTRAZIONE

MRI 1



Segmentazione
semi-automatica

HealthMyne
QIDS®



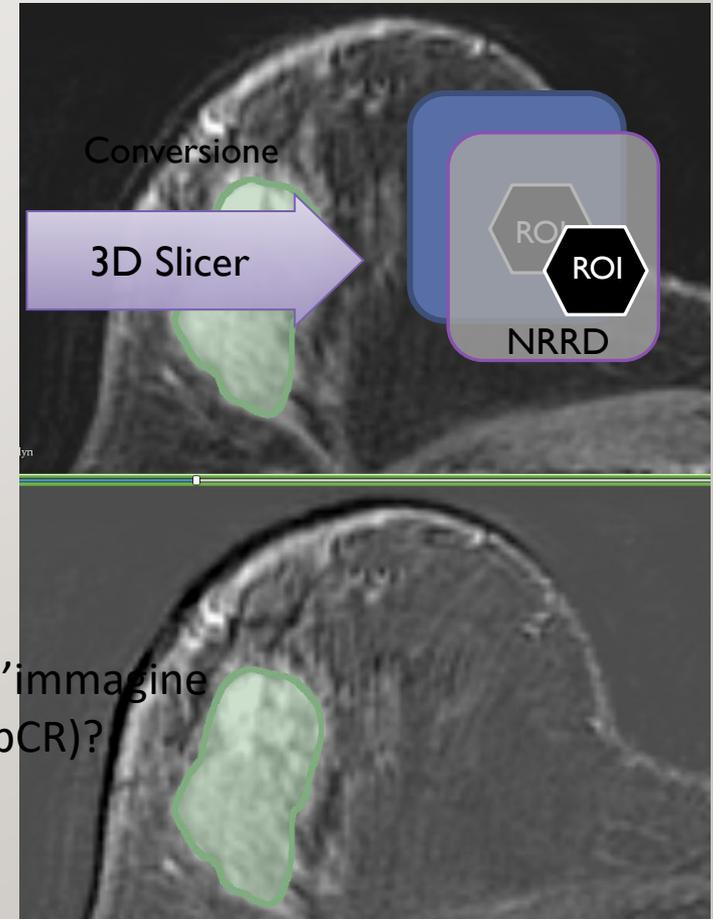
Conversione

3D Slicer



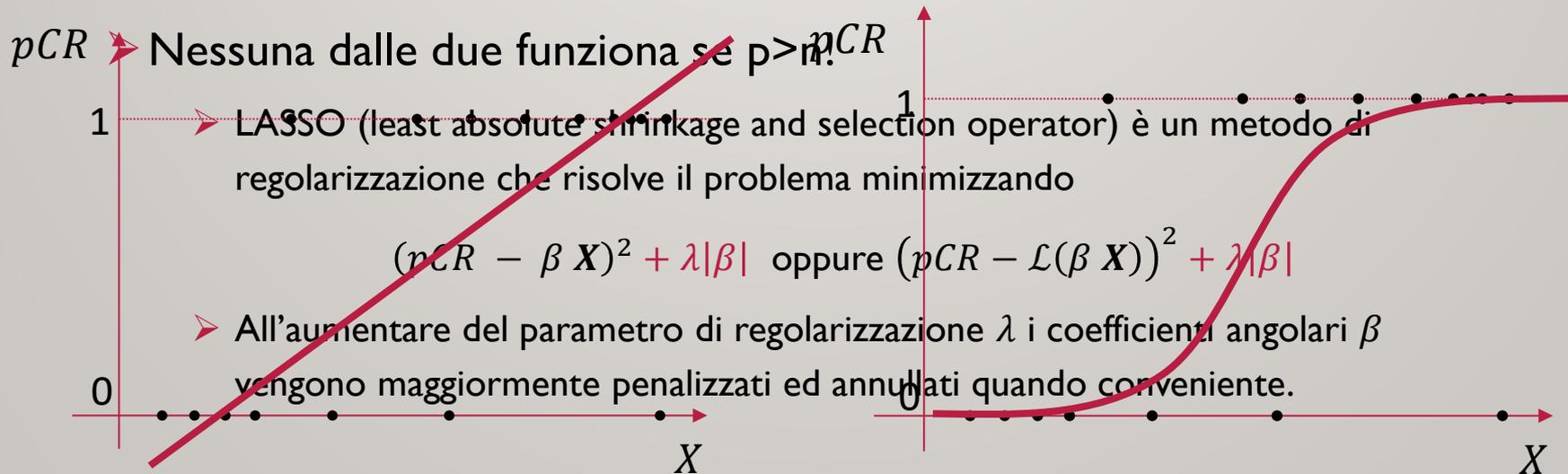
Fat Suppressed
DCE MRI @ 3T
(eTHRIVE)

- Ogni feature rappresenta una proprietà dell'immagine
- Quale di queste è correlata con l'outcome (pCR)?
- 210 features (variabili) > 60 pazienti
- Problema large p – small n
- Come modellizzo la correlazione?



9 DALLA REGRESSIONE LINEARE A LASSO

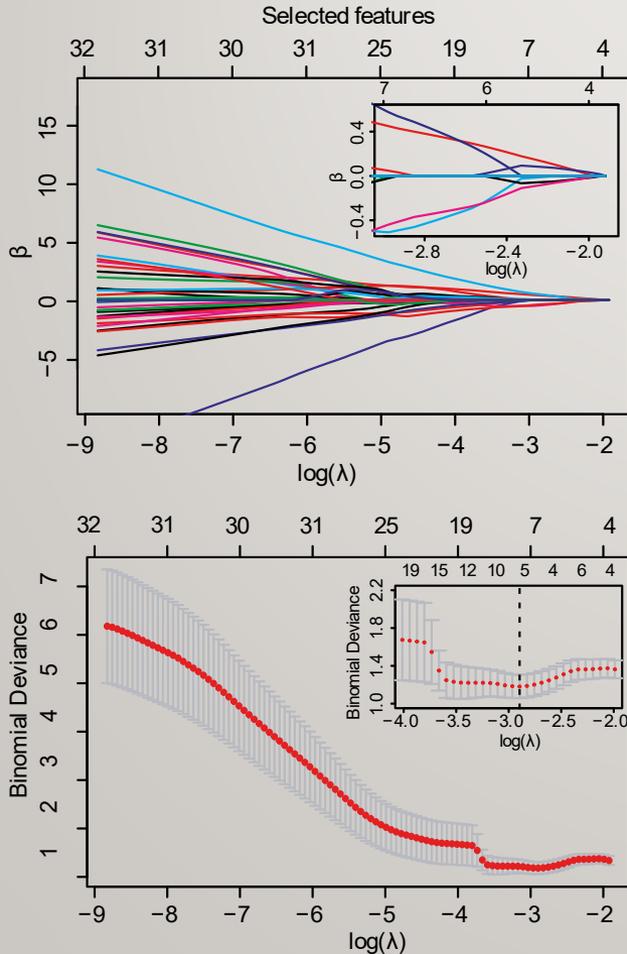
- Le features X devono essere messe in correlazione con la pCR
 - Regressione lineare? $pCR = \beta X$; minimizzo $(pCR - \beta X)^2$
 - Predizioni > 1 o < 0 non hanno senso
 - Regressione logistica? $pCR = \mathcal{L}(\beta X) = \frac{e^{\beta X}}{1+e^{\beta X}}$; minimizzo $(pCR - \mathcal{L}(\beta X))^2$
 - \mathcal{L} è la funzione logistica
 - Può rappresentare la probabilità di pCR



10

LA SELEZIONE DELLE FEATURES E LA LOOCV

LASSO + Logistico



Più λ è alto più lo scarto quadratico medio è elevato
Quale λ scelgo?



Utilizzo una cross-validazione per capirlo
LOOCV = Leave-one-out cross-validation



Costruisco un modello temporaneo su 59 pazienti e lo testa su uno ripetendo il processo 60 volte.

Il λ ottimale minimizza la devianza dei 60 modelli sui rispettivi pazienti di test

LE VARIABILI CLINICHE E ISTOLOGICHE

Alle 5 variabili di radiomica sono state integrate le informazioni cliniche e istologiche

	Responders	Non-resp.	All	p-value
Numero	20	40	60	---
Età (y)	49.2 (±11.6)	52.8 (±12.2)	51.6 (±12.0)	0.273
ADC (x10⁻⁶ mm²/s)	842 (±270)	875 (±197)	864 (±222)	0.629
PET SUV Max	8.44 (±5.08)	6.79 (±5.51)	7.34 (±5.39)	0.257
Ki67 (%)	39.7 (±23.0)	20.0 (±11.3)	26.6 (±18.5)	0.0013
PgR (%)	12.3 (±23.9)	34.6 (±33.9)	27.2 (±37.5)	0.0046
ER (%)	34.0 (±40.5)	80.8 (±26.2)	65.2 (±38.4)	0.0001

Grado				
2	2	23	25	0.0006
3	18	17	35	

HER2				
Pos	11	35	46	0.008
Neg	9	5	14	

Margini				
Irregolari	16	19	35	0.025
Spiculati	4	21	25	

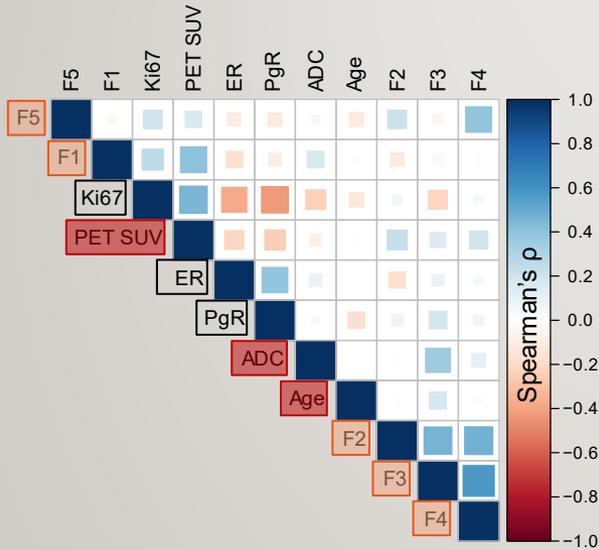
	Responders	Non-resp.	All	p-value
IntEnh				
Eterogeneo	14	33	47	0.37
Omogeneo	1	3	4	
Rim	5	4	9	

Curva				
I	0	3	3	0.66
II	3	6	9	
III	17	31	48	

Tipo				
MultiCentrico	6	18	24	0.45
MultiFocale	5	6	11	
Unifocale	9	16	25	

Forma				
Irregolare	8	27	35	0.018
Ovale	6	11	17	
Rotonda	6	2	8	

12 I GRUPPI DI COVARIATE



Correlazioni contenute



Tutte le variabili sono state considerate

5 gruppi:

GI = {Age, ADC, SUV_{max} , Grade, Shape, Margin, IntEnh, Curve, Type}

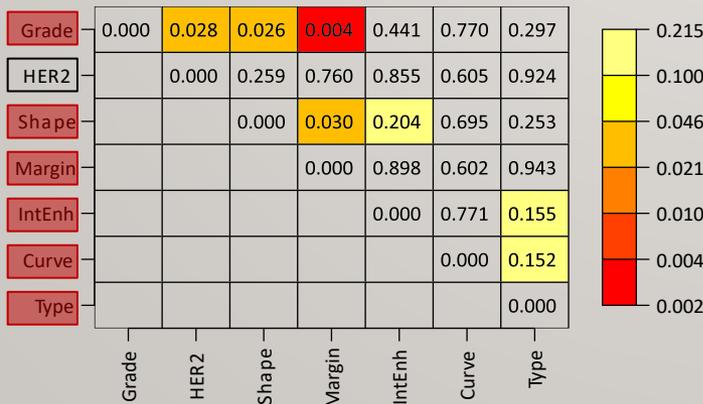
Rad = {F1, F2, F3, F4, F5}

Hist = {Ki67, ER, PgR, HER2}

NoRad = **GI** \cup **Hist**

All = **GI** \cup **Rad** \cup **Hist**

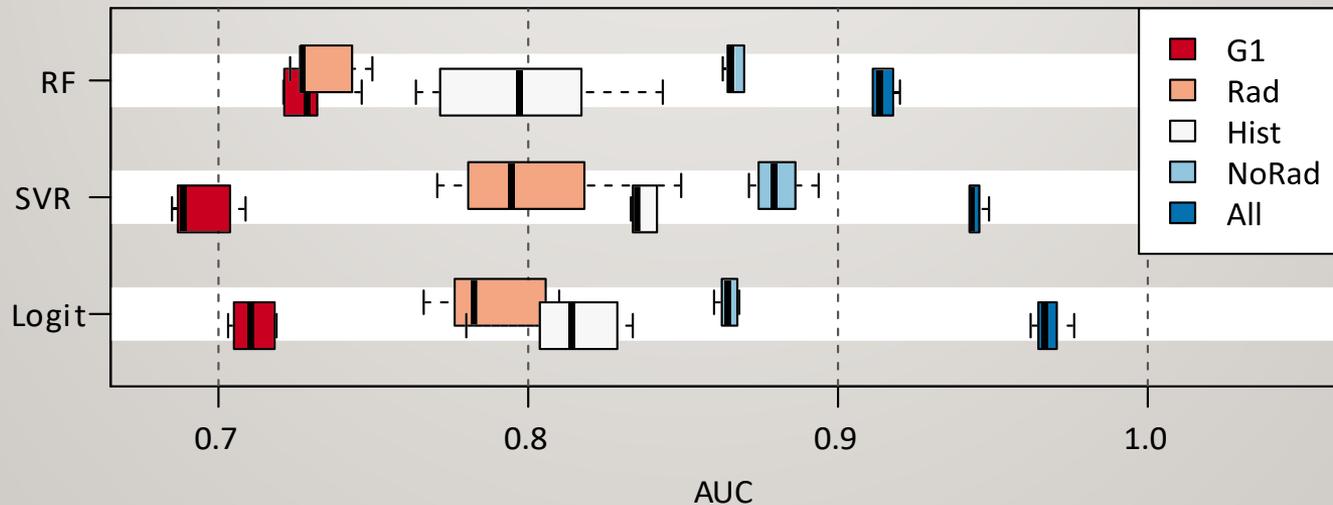
Fisher's p-values



13 IL CONFRONTO TRA TRE CLASSIFICATORI

3 diversi classificatori (Random Forest, Support Vector Machines e Regressione Logistica) con approccio esaustivo

Average LOO and 30-fold ROC AUC (best 6 models)



p-value medio, corretto per la molteplicità del test

p^{best6}	Logit	SVR	RF	$p^{best6} / p^{best6}_{All}$	Logit	SVR	RF
G1	1,33 E-1	4,65 E-1	9,64 E-2	G1	29983	165069	1651
Rad	1,01 E-3	6,32 E-4	1,81 E-2	Rad	228	224	310
Hist	1,74 E-4	2,59 E-5	5,18 E-4	Hist	39	9	9
NoRad	3,86 E-4	4,99 E-5	5,29 E-4	NoRad	87	18	9
All	4,44 E-6	2,82 E-6	5,84 E-5	All	1	1	1

14 LO STUDIO PROSPETTICO

- Per validare il modello è importante utilizzare un test indipendente
- Il modello finale scelto per lo studio prospettico è di tipo logistico regolarizzato, con i seguenti coefficienti angolari

Covariata	β_i
Intercetta	-1.26268420176
F1	3.09122287103
F2	0.58728830096
F3	-0.06482682640
F4	-0.00893927588
Ki67	0.02676369214
ER	-0.01917022894
Grado_2	-0.11673269623
Grado_3	0.11036353233
HER2_0	-0.68869467641
HER2_1	0.67981316688
Margini_I	0.19412601516
Margini_S	-0.18993705152

- Il modello è già stato implementato nel futuro software per la raccolta dati:

BIRRa

15 BIRRa



BIRRa (Birra Is Real Radiomics) è un software che nasce con un duplice intento:

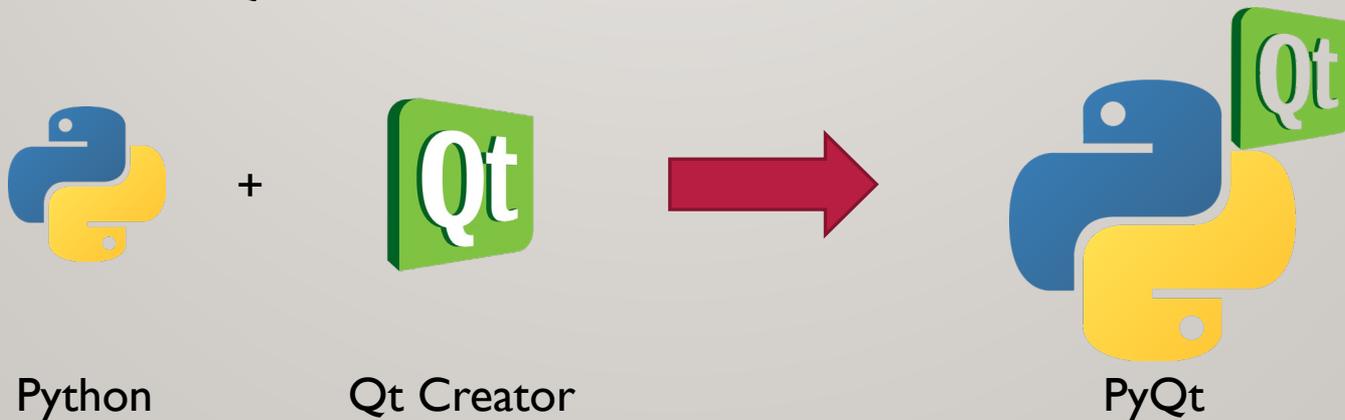
- Raccogliere in modo pratico ed ordinato i dati per lo studio di validazione prospettico, alternativo al canonica database excel
- Fornire un tool user-friendly con cui medici, anche se non esperti di machine learning, possano accedere ad un modello radiomico

Requisiti:

- Essere un programma stand-alone portatile
- Raccogliere le informazioni cliniche
- Convertire i volumi DICOM in NiftI
- Consentire il contornamento della lesione
- Estrarre correttamente le features radiomiche dalla ROI
- Applicare il modello creato per ottenere la probabilità di pCR
- Salvare i volumi e i dati clinici per futura referenza

16 Qt CREATOR + PYTHON

- L'ideale è sfruttare librerie già disponibili in Python (pyradiomics, simpleITK, dcm2niix, etc)
- Il modello precedentemente realizzato in R può comunque essere importato o definito analiticamente
- Per lo sviluppo della GUI (Graphic User Interface) si è scelto Qt Creator



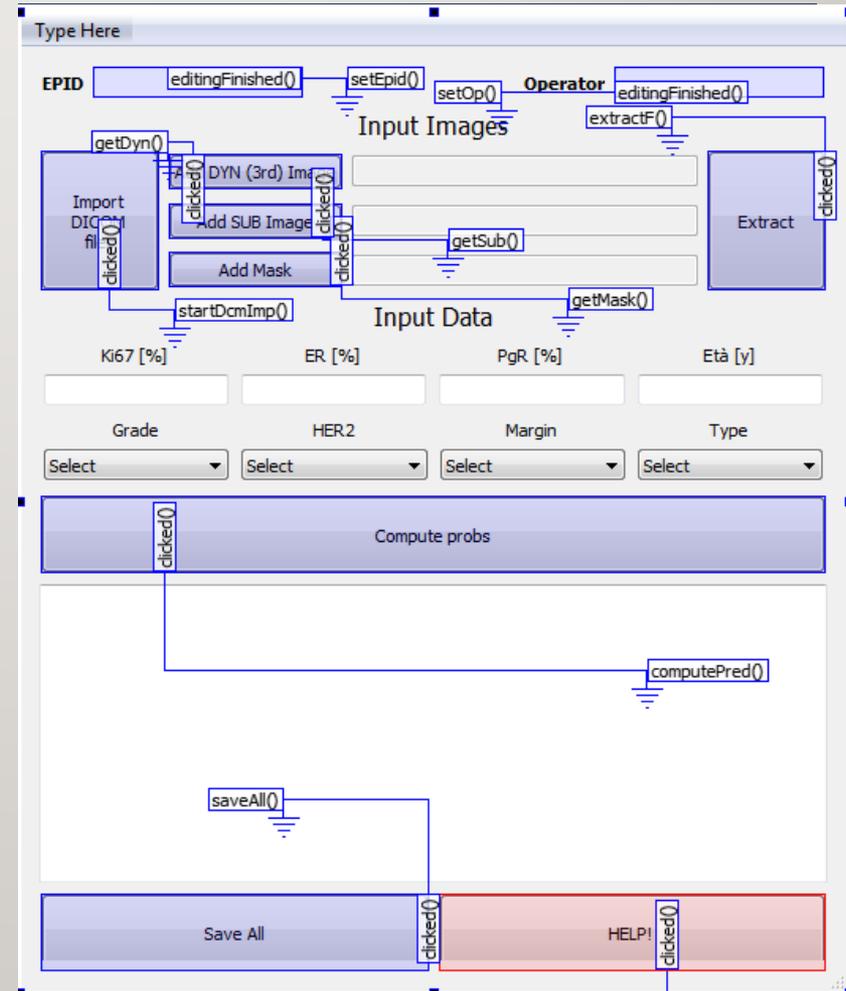
17 LA FINESTRA PRINCIPALE

Sistema a Signals e Slots:

- Un Signal (evento) è lanciato per ogni interazione con la GUI
- Lo Slot è la funzione che viene eseguita in seguito all'emissione di un Signal

Questa finestra consente la selezione delle immagini MRI, della maschera (contorno della lesione) e delle informazioni cliniche.

Ma vediamo come funziona...



18 PyInstaller

Il file .py non è pratico nell'utilizzo del medico:

- Necessita dell'installazione di Python
 - Non sempre facile nei computer aziendali
- Necessita dell'installazione delle librerie, creazione env, etc.

Soluzione: **PyInstaller**

- Crea un eseguibile (exe) includendo tutte le librerie e il motore Python
- Debugging diventa più complesso

```
pyinstaller --noconfirm --clean -w --onedir --paths ..\qtenv2\Lib\site-packages --icon "z_resources/birraVec.ico"  
--add-data "initW.ui;" --add-data "chat.ui;" --add-data "dcmImporter.ui;" --add-data "infoDcm2Niix.ui;"  
--add-data "mainwindow.ui;" --collect-all radiomics BIRRa.py
```

19 RISULTATI DELLA VALIDAZIONE

- Nuovo dataset di 66 nuovi pazienti (34 complete responders and 32 non-responders).*

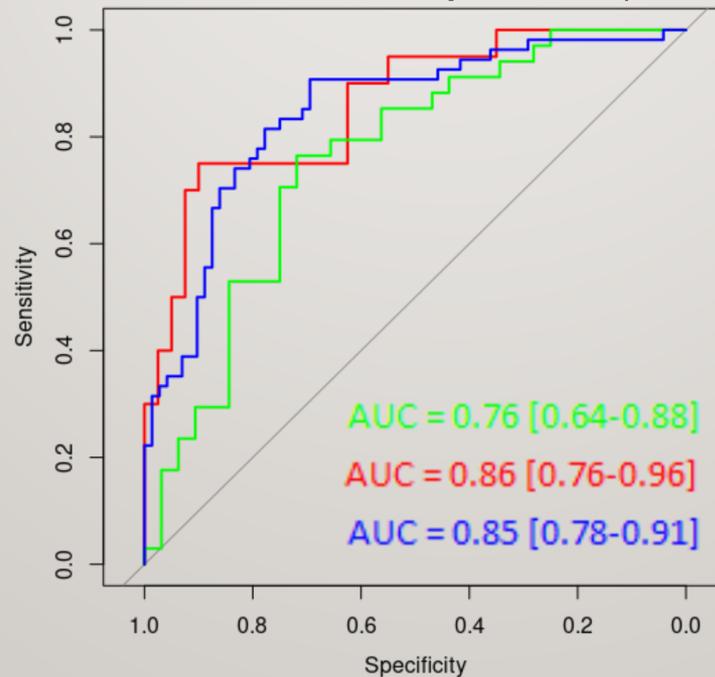


Figure 1. Model ROCs for the training set (red) and the validation set (green). The blue line is the roc curve for the model trained on the whole set and evaluated with a LOOCV. AUC values are shown together with their confidence intervals.

20 CONFRONTO CON REVISIONE SISTEMATICA

Motori di ricerca: Pubmed, Google Scholar, Microsoft Academic, Semantic Scholar, Base

Numero di articoli analizzati (fino al 02/09/2021)

PM	GS	MA	Base	SE	Tot
47	29	27	94	78	275

La meta-analisi degli studi che possono essere comparati con questo lavoro (trainig + validazione; utilizzo solo di informazioni pre NAC, etc.) forniscono un $AUC_{pre} = 74\% \pm 7\%$

Studio di validazione $AUC = 76\% [64 - 88]$

21 CONCLUSIONI

- La NAC è una terapia impiegata nel trattamento del tumore alla mammella, tuttavia la risposta del tumore talvolta è limitata.
- Il modello sviluppato (retrospettivamente) **unisce informazioni di radiomica MRI** (imaging quantitativo) **con informazioni cliniche** per predire prima della somministrazione se il paziente avrà una risposta completa alla NAC.
- Le prestazioni attese per il modello logistico sono elevate e la validazione prospettica ha dato anch'essa buone prestazioni.
- Per lo studio prospettico è stato sviluppato un software (BIRRa) che consente, oltre alla raccolta dati, di familiarizzare i medici con i concetti di radiomica e machine learning.



GRAZIE PER L'ATTENZIONE

Radiomica per la predizione della risposta alla chemioterapia nel tumore della mammella

Relatore: Prof. Marcello Lunardon

Correlatore: Dr. Carlo Cavedon

Giulio Benetti