

UNIVERSITA' DEGLI STUDI DI PADOVA

DIPARTIMENTO di MEDICINA - DIMED

**CORSO DI LAUREA TRIENNALE IN "TECNICHE DI RADIOLOGIA MEDICA,
PER IMMAGINI E RADIOTERAPIA"**

Sede di Padova

Presidente: Prof. Roberto Stramare

Tesi di Laurea:

**USO DI UN SOFTWARE AUTOMATICO PER LA VALUTAZIONE
DELLA QUALITÀ TECNICA DELLE IMMAGINI MAMMOGRAFICHE
PER OTTIMIZZARE LE PERFORMANCE DEL SINGOLO
TECNICO DI RADIOLOGIA**

Relatore: Prof. Roberto Stramare

Correlatrici: Dott.ssa Gisella Gennaro
TSRM Dott.ssa Sara Del Genio

Laureanda: Letizia Povoło

Anno Accademico 2021-2022

INDICE

1. INTRODUZIONE	1
2. MATERIALI E METODI	3
2.1 Proiezioni mammografiche standard	3
2.2 Software automatico per la valutazione della qualità tecnica delle mammografie.....	6
2.2.1 Valutazione del posizionamento	6
2.2.2 Valutazione della compressione	9
2.4 TSRM della UOC Radiologia Senologica dell'Istituto Oncologico Veneto di Padova	11
2.5 Immagini oggetto di studio	11
2.6 Dati di baseline	12
2.7 Formazione e monitoraggio	15
2.8 Analisi dei dati	17
3. RISULTATI	18
3.1 Dati di baseline	18
3.2 Performance dei singoli TSRM dopo l'uso consapevole del software	21
3.2.1 TSRM 1	21
3.2.2 TSRM 2	22
3.2.3 TSRM 3	23
3.2.4 TSRM 4	25
3.2.5 TSRM 5	26
3.2.6 TSRM 6	27
4. DISCUSSIONE	31
5. CONCLUSIONI	34
6. BIBLIOGRAFIA E SITOGRAFIA	35

1. INTRODUZIONE

La diagnosi precoce del tumore alla mammella rappresenta l'arma più efficace nella prevenzione ed è essenziale per poter attuare interventi immediati e conservativi, in grado di ridurre le probabilità di progressione e di aumentare le probabilità di guarigione della malattia stessa.

In Italia, secondo le indicazioni del Ministero della Salute, il Servizio Sanitario Nazionale ha attivato dei programmi di screening mammografico: un servizio gratuito di prevenzione secondaria, garantito dai LEA (livelli essenziali di assistenza) e rivolto a tutte le donne residenti nel territorio italiano nella fascia d'età compresa tra i 50 anni e i 69 anni (con estensione della copertura ai 45-74 anni in alcune regioni). Lo scopo è quello di riconoscere il cancro nelle fasi più iniziali, spesso prima della comparsa dei sintomi, permettendo di ottenere sopravvivenze a cinque anni molto elevate. [1] Il programma, per riconoscere il cancro al seno anche nelle fasi meno avanzate, propone come esame primario la mammografia ad entrambi i seni (bilaterale) ogni due anni. Tuttavia, questo tipo di indagine molto importante presenta degli aspetti critici che non contraddistinguono altri esami radiologici. L'elevata manualità che richiede l'esecuzione della mammografia e le variazioni dell'habitus corporeo delle pazienti e della loro capacità di collaborazione sono aspetti che rendono l'indagine poco standardizzabile e che richiedono una maggiore diligenza al fine di ottenere uniformità e riproducibilità. [2]

In quest'ottica, il controllo della qualità tecnica dell'esame mammografico risulta essere fondamentale per permettere la corretta interpretazione delle immagini da parte del medico radiologo. Diversi studi, infatti, dimostrano che i problemi più comuni nella scarsa qualità delle immagini mammografiche sono lo scorretto posizionamento e la scorretta compressione. [3,4]

Per monitorare e determinare la correttezza delle mammografie sono essenziali i metodi di valutazione, che delineano delle linee guida nel tentativo di standardizzare il processo di analisi e giudizio. Questi metodi si dividono in due categorie: quelli basati sull'auto-valutazione (come il programma di accreditamento proposto dall'American College of Radiology, il sistema adottato

dalla Commissione Europea, il protocollo della Regione Emilia-Romagna RER e il PGMI applicato in UK), e i metodi automatici di valutazione.

Lo scopo di un monitoraggio continuo della qualità delle immagini è fondamentale per poter evitare e correggere errori, la cui presenza può comportare una serie di "disagi": infatti, può ridurre significativamente la sensibilità delle immagini sia di screening che diagnostiche rendendo per esempio necessario il richiamo della paziente per eventuali approfondimenti.

Questo elaborato di tesi ha come obiettivo principale quello di verificare se l'utilizzo da parte dei Tecnici Sanitari di Radiologia Medica (TSRM) di un software automatico di analisi delle mammografie acquisite, basato sul metodo PGMI, associato ad una formazione continua, può essere considerato utile nella pratica quotidiana per migliorare le performance dei singoli tecnici di radiologia nella tecnica di esecuzione degli esami mammografici.

2. MATERIALI E METODI

2.1 Proiezioni mammografiche standard

L'esame mammografico standard prevede l'acquisizione di due proiezioni per ogni mammella (Cranio Caudale — CC e Medio-Laterale Obliqua — MLO), in modo tale da ottenere una visione completa del seno. Per eseguire correttamente un esame mammografico è fondamentale tenere conto di due fattori: il posizionamento e la compressione, entrambi responsabili della qualità dell'immagine e valutabili da tecnici di radiologia esperti o da sistemi computerizzati, come il software oggetto di questo elaborato.

La proiezione ideale per lo studio del quadrante mediale è la proiezione cranio-caudale, indicata con la sigla CC. La paziente viene disposta davanti al mammografo con le braccia lungo i fianchi, le spalle rilassate e il volto ruotato nel lato opposto al lato in esame, facendo appoggiare l'orecchio sulla visiera dello stativo. Il TSRM procede regolando il piano d'appoggio fino all'altezza del solco sottomammario e successivamente con il palmo della mano solleva e distende il quadrante inferiore della mammella in esame e la posiziona al centro del detettore (evitando la formazione di pieghe). A questo punto ha inizio lentamente la compressione, fino a raggiungere la pressione adeguata. Prima di acquisire l'immagine mammografica si controlla e si risolve la presenza di eventuali pieghe cutanee e di sovrapposizioni che possono generare artefatti. (Figura 1) [5].

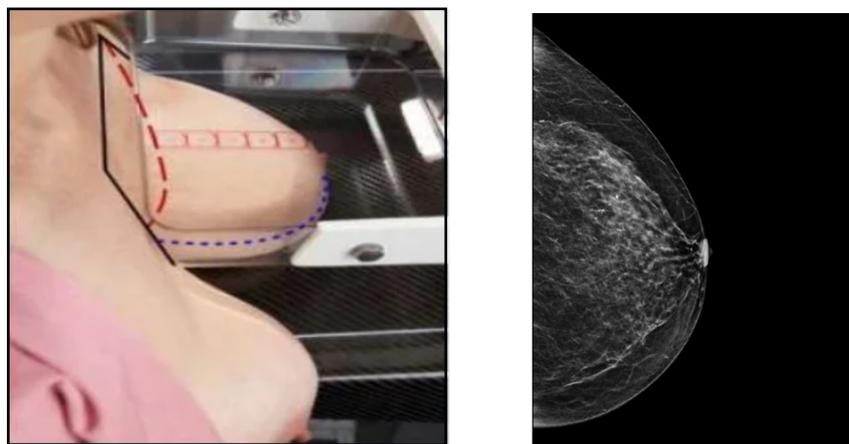


Figura 1: Posizionamento della proiezione Cranio-Caudale (CC). [da "Mammography technique, indication and contraindication", JSS Medical College]. Immagine CC presa dal PACS.

La proiezione più indicata per lo studio del quadrante supero-esterno è la proiezione medio laterale-obliqua (MLO), ottenuta ruotando lo stativo di 45° nel verso opposto al lato in esame. La paziente, in stazione eretta, viene disposta dritta di fronte al mammografo con il busto leggermente in avanti e inclinato verso il detettore. La spalla omolaterale viene intraruotata e fatta appoggiare sull'angolo del piano sensibile, mentre la mano viene sistemata sulla maniglia dello stativo. Il TSRM procede regolando l'altezza del mammografo e successivamente raccoglie e distende la mammella, posizionandola al centro del detettore. A questo punto, dopo aver fatto rilassare la spalla del lato in esame e invitato la paziente ad allontanare la mammella controlaterale, si procede lentamente con la compressione fino a raggiungere la forza adeguata. Prima di acquisire l'immagine mammografica, come per la proiezione CC, si controlla e si risolve la presenza di eventuali pieghe cutanee e di sovrapposizioni che possono generare artefatti. (Figura 2) [6].

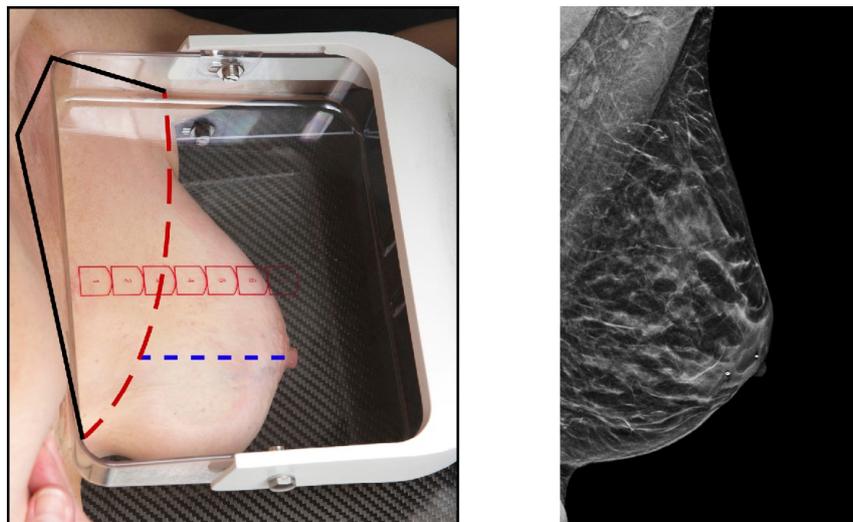


Figura 2: Posizionamento della proiezione Medio-Laterale Obliqua (MLO) [da "Mammography technique, indication and contraindication", JSS Medical College]. Immagine MLO presa dal PACS.

A ciascuna proiezione sono associati dei criteri che determinano la correttezza del posizionamento. Per la proiezione CC i criteri presi in considerazione sono:

- *Nipple in profile*, ovvero il capezzolo visibile di profilo;
- *Nipple Midline*, ovvero il capezzolo centrale e perpendicolare al seno, senza esagerazioni mediali o laterali di oltre 5 gradi;

- *PNL Met*, ovvero la lunghezza della linea tracciata posteriormente e perpendicolarmente dal capezzolo verso il muscolo pettorale;
- *No cutoff*, ovvero l'inclusione di tutto il tessuto mammario senza "tagliare pezzi".

Per la proiezione MLO i criteri presi in considerazione sono:

- *Nipple in profile*, ovvero il capezzolo visibile di profilo;
- *IMF visible*, ovvero il solco sotto-mammario visualizzato e svolto;
- *PEC to PNL Met*, ovvero il muscolo pettorale rappresentato fino alla linea posteriore e perpendicolare al capezzolo (PNL);
- *Adequate PEC*, ovvero l'adeguata quantità di muscolo pettorale visibile, misurata tenendo conto dell'angolo, della lunghezza e della larghezza;
- *No PEC skin folds*, ovvero l'assenza di pieghe cutanee sopra il muscolo pettorale;
- *PEC Shape*, ovvero la forma concava del muscolo pettorale;
- *No cutoff*, ovvero l'inclusione di tutto il tessuto mammario.

Oltre al corretto posizionamento, un altro fattore indispensabile per la qualità degli esami mammografici è la compressione della mammella. Se applicata in modo adeguato, infatti, garantisce una serie di vantaggi che permettono di ottimizzare gli esami e di rispettare la correttezza tecnica. La compressione contribuisce a mantenere la mammella ferma nella posizione voluta, evitando artefatti da movimento e sfocature; distende le strutture anatomiche riducendo il mascheramento dovuto alla sovrapposizione delle stesse; diminuisce le radiazioni diffuse e aumenta il rapporto contrasto - rumore grazie alla riduzione dello spessore; e contribuisce a limitare la dose e ad aumentare il tasso di rilevamento del cancro. [7]

2.2 Software automatico per la valutazione della qualità tecnica delle mammografie

Lo strumento per la valutazione automatica della qualità tecnica delle mammografie, ovvero la qualità del posizionamento e della compressione, utilizzato in questo progetto di tesi sperimentale è il software “Volpara Analytics”, versione 3.2. Il pacchetto è costituito da quattro moduli:

- TrueDensity®: usa una combinazione di fisica dei raggi-X e di machine learning per generare una misura accurata ed oggettiva di densità volumetrica della mammella.
- TruePGMI®: valuta automaticamente la qualità del posizionamento di ogni immagine mammografica.
- TruePressure®: misura la pressione di compressione applicata alla mammella, usando la forza di compressione del gantry e l’area di contatto della mammella col dispositivo di compressione.
- TrueRadDose®: analizza la dose (di radiazione) data alla paziente, basandosi sulla densità della mammella.

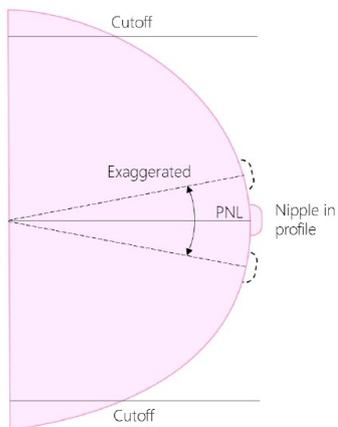
In questo lavoro di tesi sono stati utilizzati i dati prodotti dai moduli TruePGMI® e TruePressure®, resi disponibili tramite dashboard dinamiche ed interattive, progettate specificamente per i principali ruoli presenti in un’Unità Operativa di Radiologia Senologica, tra cui quella del TSRM.

2.2.1 Valutazione del posizionamento

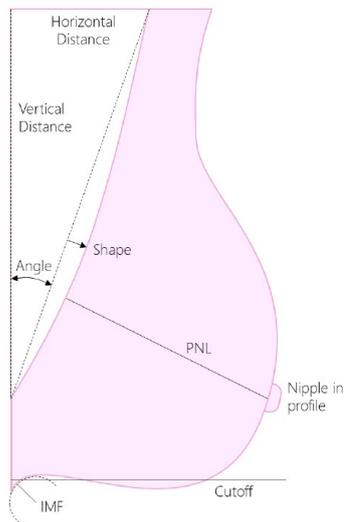
L'applicazione clinica Volpara Positioning (TruPGMI) valuta il posizionamento della mammella utilizzando come standard di riferimento il metodo di valutazione denominato “PGMI”, proposto dallo *United Kingdom Mammography Trainers Group*, con il supporto del *Royal College of Radiographers*, e basato sulla classificazione degli studi in 4 classi: Perfetto, Buono, Moderato e Inadeguato (PGMI). Classificare uno studio come *Perfect* (P) indica un’eccellente esecuzione e corrisponde ad un’alta qualità diagnostica; *Good* (G) indica una buona esecuzione e corrisponde ad una media-alta qualità diagnostica; l’attribuzione di *Moderate* (M) indica una qualità accettabile, ed infine *Inadequate* (I) indica una scarsa qualità per cui il Medico Radiologo potrebbe richiedere una ripetizione dell’immagine. L’analisi del posizionamento vPGMI dell’immagine è determinata

da algoritmi che localizzano specifici punti di riferimento nella mammella e, tramite delle misurazioni geometriche, definiscono se i criteri di correttezza sono stati rispettati. [8]

Non tutti i criteri di correttezza hanno lo stesso peso nel determinare la qualità dell'immagine: ci sono dei parametri meno impattanti, il cui mancato rispetto non determina gravi perdite di qualità e di informazione, e altri parametri il cui peso è invece maggiore. Nella figura successiva sono riportate per ogni parametro di ciascuna proiezione, le variazioni delle valutazioni PGMI a seconda del rispetto o meno degli stessi. (Figura 3)



P	G	M	I	Criteri di correttezza CC
V	V o F	V o F	V o F	Nipple in profile
V	V o F	V o F	V o F	Nipple Midline
V	V	V o F	V o F	PNL Met
V	V	V	F	No cutoff



P	G	M	I	Criteri di correttezza MLO
V	V o F	V o F	V o F	Nipple in profile
V	V o F	V o F	V o F	IMF Visible
V	V	F	V o F	Adequate PEC
V	V	V o F	V o F	PEC to PNL Met
V	V o F	V o F	V o F	PEC Shape
V	V o F	V o F	V o F	No PEC Skin folds
V	V	V	F	No cutoff

Figura 3: Valutazioni PGMI in base al rispetto o meno dei criteri di correttezza, per la proiezione CC e per la proiezione MLO, dove V significa "vero" (criterio verificato), F significa "falso" (criterio non-rispettato), V o F significa che il criterio può essere rispettato oppure no.

Per esempio, nella proiezione MLO, il parametro “PEC shape”, ovvero la forma del pettorale, è utile per la standardizzazione delle immagini, ma dal punto di vista diagnostico non è particolarmente rilevante ed è per questo che la sola violazione di questo criterio porta comunque a giudicare l’immagine “Good”. Lo stesso discorso non può essere fatto per quei criteri che hanno un peso maggiore, come il parametro “No cutoff”, il cui mancato rispetto potrebbe nascondere la presenza di una lesione e rendere la mammografia inefficace; per questo motivo la presenza di *cutoff* rende l’immagine “Inadequate” anche se tutti gli altri criteri risultano rispettati.

Nella proiezione CC i criteri più influenti sono “No cutoff” e “CC PNL”, mentre nella proiezione MLO sono “No cutoff”, “PEC to PNL Met” e “Adequate Pectoral”.

Ciascun TSRM, nella sua area riservata ha la possibilità di: vedere la propria percentuale di punteggi perfetti, buoni, moderati, inadeguati in un periodo di tempo selezionato; identificare quali parametri di posizionamento sono meno rispettati per cercare di correggerli, potendo visualizzare anche le singole immagini acquisite (Figura 4,5); accedere a materiali di formazione che affrontano in modo specifico i problemi di posizionamento.

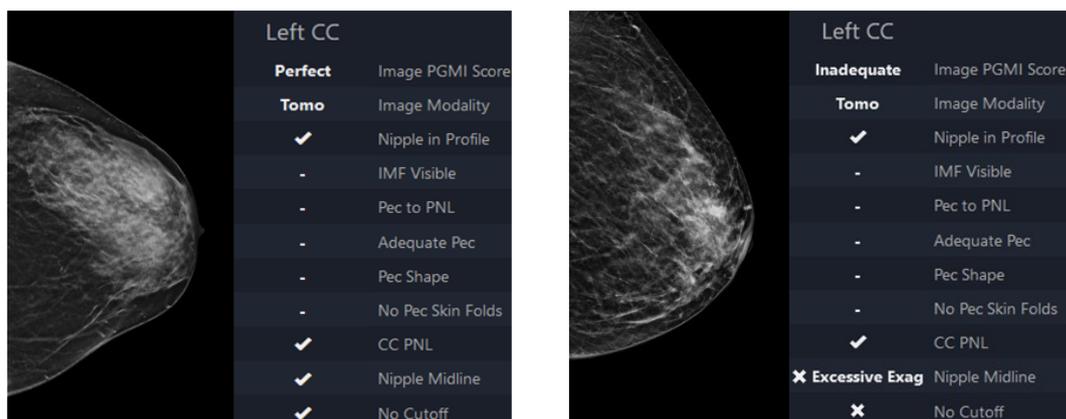


Figura 4: Confronto tra immagini CC valutate da Volpara Positioning, disponibili nel pannello “Technologist Performance Review” dell’area riservata di ogni TSRM. A sinistra quella giudicata “Perfetta” dal momento che rispetta tutti i criteri di correttezza; a destra quella giudicata “Inadeguata”, poiché il tessuto mammario è stato tagliato e il capezzolo non è centrale.

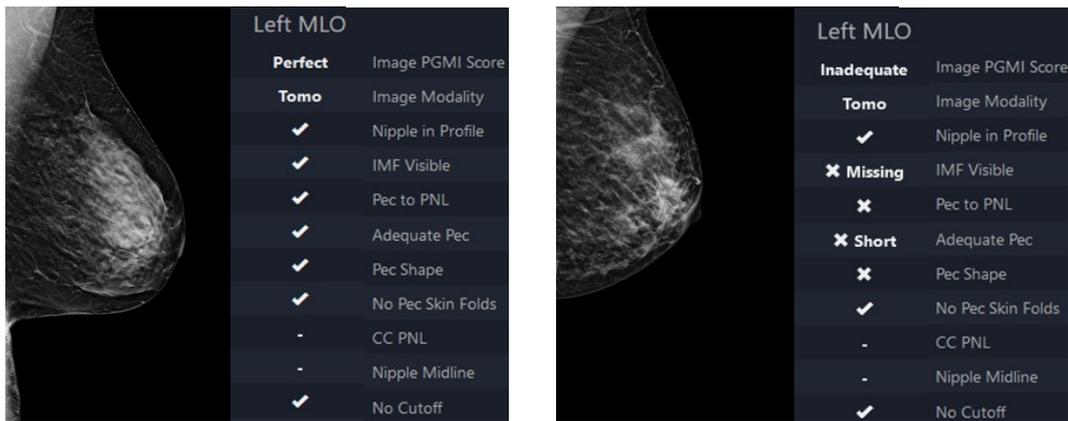


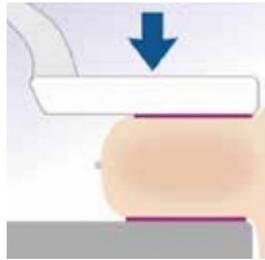
Figura 5: Confronto tra immagini MLO di sinistra valutate da Volpara Positioning, disponibili nel pannello “Technologist Performance Review” dell’area riservata di ogni TSRM. A sinistra quella giudicata “Perfetta” dal momento che rispetta tutti i criteri di correttezza; a destra quella giudicata “Inadeguata”, poiché non è stato rappresentato correttamente il muscolo pettorale e il solco inframammario non è visibile.

2.2.2 Valutazione della compressione

Le linee guida europee affermano che la forza (Newton) con cui si comprime la mammella deve essere “decisa ma tollerabile”, con una compressione massima applicata di 130-200 N [9]. Tuttavia, considerare unicamente la forza come criterio per valutare l’adeguatezza della compressione è un errore poiché non è correlabile al disagio e al dolore percepito e non tiene conto delle differenze nelle dimensioni delle mammelle (Figura 6).

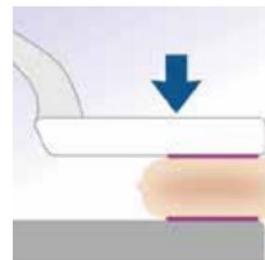
La grandezza fisica e l’unità di misura più corretta da tenere in considerazione per applicare una compressione adeguata al singolo seno è la pressione (Kilopascal), calcolata indirettamente come rapporto tra la forza applicata e l’area di contatto da comprimere.

$$Pressione = \frac{Forza (N)}{Area (m^2)}$$



Seno grande - area di contatto: 180 cm^2

- Forza: 180N
- Pressione: 10 kPa
- Minor dolore



Seno piccolo - area di contatto: 110 cm^2

- Forza: 180N
- Pressione: 16 kPa
- Maggior dolore

Figura 6: Esempio di come cambia la pressione (e conseguentemente il dolore percepito) al variare della dimensione del seno, a parità di forza di compressione applicata [da "Optimal Breast Compression in Mammography Volpara"].

Un recente studio olandese ha dimostrato che il tasso di rilevamento del cancro raggiunge un picco a una pressione di circa 10kPa e al di sotto o al di sopra di tale livello, le prestazioni cliniche diminuiscono [10]. A tale proposito, si ritiene che il range di pressione adeguato per ottenere una compressione ottimale rientri tra i 7 kPa e i 15 kPa.

Durante l'esecuzione dell'esame, il TSRM non conosce la pressione che sta esercitando alla mammella. Volpara TruPressure, disponibile come componente di Volpara Analytics, è uno strumento in grado di determinare l'area di contatto della mammella e quindi di calcolare retrospettivamente la pressione applicata al seno, a seconda della forza utilizzata. I risultati vengono resi disponibili nella piattaforma di Volpara Analytics e classificati in 3 possibili categorie: *target*, se la pressione applicata è compresa tra 7kPa e 15kPa; *low*, se la pressione applicata è minore di 7kPa; *high*, se la pressione applicata è maggiore di 15kPa.

Ogni tecnico può accedere alla sua area personale e consultare i propri dati, in modo da poter verificare la pressione media applicata durante gli esami per capire se la forza impressa è adeguata e se viene adattata alla dimensione del seno. Questo feedback rende i TSRM più consapevoli e li aiuta ad ottimizzare la qualità degli esami successivi. [11]

2.4 TSRM della UOC Radiologia Senologica dell'Istituto Oncologico Veneto di Padova

Tutti i componenti del personale tecnico considerato in questo progetto si dedica esclusivamente all'attività senologica. Tra di loro, l'esperienza nell'esecuzione delle mammografie è variabile: tre di loro svolgono l'attività mammografica da oltre 20 anni (TSRM 1, TSRM 2, TSRM 5), uno di loro da gennaio del 2020 (TSRM 4), un altro da giugno 2021 (TSRM 3) e un altro ancora da marzo 2022 (TSRM 6).

2.5 Immagini oggetto di studio

Per questo progetto di tesi sperimentale l'analisi è stata limitata alle immagini acquisite sulle donne arruolate nel Progetto *Screening Giovani Donne* presso l'UOC Radiologia Senologica dell'Istituto Oncologico Veneto di Padova, utilizzando i mammografi Hologic Selenia Dimensions e Hologic Selenia 3Dimensions (i quali differiscono solamente per la risoluzione spaziale delle immagini di tomosintesi). La scelta di considerare questo tipo di popolazione e di escludere le pazienti oncologiche (seguite anch'esse dal centro), nasce dalla necessità di valutare la qualità tecnica su pazienti con anatomia normale, che quindi non fossero state precedentemente sottoposte a trattamenti chirurgici e radioterapici. Questi interventi, infatti, oltre ad aumentare la sensibilità al tocco e alla pressione, lasciano sulla mammella dei segni tipici, come cicatrici, ematomi e clip metalliche che possono compromettere il corretto posizionamento e la corretta compressione. [12] Di conseguenza, il software Volpara Analytics, valutando in modo automatico e oggettivo la qualità di ogni mammografia, potrebbe assegnare dei giudizi che, essendo basati sull'ipotesi di anatomia normale, sono tendenzialmente inferiori.

Per riuscire a verificare l'utilità del software per ottimizzare l'esecuzione delle immagini mammografiche da parte di ciascun TSRM, è stato necessario raccogliere i dati afferenti a due periodi diversi, intervallati da un momento di formazione in cui è stato spiegato ai TSRM come si utilizza il software. Per avere un confronto ragionevole, è stata scelta una quantità di immagini pre-formazione che fosse dello stesso ordine di grandezza delle immagini disponibili post-formazione. I dati di posizionamento e compressione (in totale 9672 immagini di

screening) relativi al primo periodo (Set - Dic 2021) sono stati considerati di “baseline”, ovvero descrittivi della “condizione di partenza” di ciascun tecnico. Il periodo post-formazione (Mar - Set 2022), nel corso del quale sono state prodotte 9937 immagini di screening, è stato utilizzato come dataset di controllo per verificare se l’uso del software ha permesso ai TSRM di migliorare le proprie performance di posizionamento e/o compressione.



Figura 7: Linea del tempo che mostra i due differenti periodi considerati per la raccolta dati, intervallati dal periodo di formazione.

2.6 Dati di baseline

Per baseline si intende il punto di riferimento e di partenza rispetto al quale calcolare gli scostamenti delle principali variabili; in questo caso consistono nei parametri di posizionamento e di compressione, i quali sono responsabili della qualità delle immagini.

Per mostrare la qualità delle immagini il software Volpara Analytics si serve di differenti metriche, più o meno di dettaglio. Tra questi, la rappresentazione più generale utilizzata è il “quality score”. Tale punteggio viene calcolato a partire dai valori di TruPGMI e TruPressure e viene confrontato con i dati globali degli utenti di Volpara Analytics (oltre 5200 TSRM e 67.5 milioni di immagini). La formula del quality score (QS) per le immagini del periodo di studio selezionato è la seguente:

$$QS = 4[(0.6 \times PG\%) + (0.4 \times T\%)]$$

dove, “PG” è la percentuale di immagini con un valore PGMI perfetto o buono, e “T” è la percentuale di immagini con un valore di compressione all’interno dell’intervallo target.

Con lo scopo di stabilire i singoli parametri, a cui ciascun TSRM avrebbe dovuto prestare maggior attenzione, è stato necessario servirsi anche di un indicatore più preciso, che per ogni criterio di posizionamento e di compressione indicasse la sua correttezza. A tal proposito, è stato preso in considerazione lo star rating, disponibile nel pannello “Compliance > Technologist performance review”. Questa valutazione assegna un numero di stelle, da 1 a 5, proporzionale alla percentuale di immagini che rispettano un preciso parametro di posizionamento o di compressione. Gli intervalli percentuali, specifici e diversi per ciascun criterio, sono stati costruiti tenendo in considerazione la distribuzione delle performance degli utenti globali di Volpara Analytics e a seconda del numero di stelle assegnato al criterio, viene associata una categoria.

- Focus (lowest 10%): raccomanda una maggiore attenzione per il parametro;
- Okay (sotto la mediana): segnala che c'è una possibilità di miglioramento;
- Global Median (pari o sopra la mediana): suggerisce di continuare così;
- Eccellente (top 25%): indica che la prestazione è di alta qualità;
- Top 10%: significa che il rispetto di quel parametro è ottimo.

Nella tabella seguente (Tabella I) vengono mostrati per entrambe le viste, gli intervalli percentuali che attribuiscono a ciascun parametro uno star rating e una classe precisa. Fondamentale è notare il fatto che “*Low compression*” e “*High compression*” sono complementari rispetto a “*Target compression*”.

Tabella 1: intervalli delle percentuali di correttezza che assegnano il relativo numero di stelle e la relativa categoria per la proiezione CC e la proiezione MLO.

Proiezione CC					
STAR RATING	★	★★	★★★	★★★★	★★★★★
CATEGORIA	Focus	Okay	Global median	Eccellente (top 25%)	Top 10%
Nipple in Profile	<67%	67-78%	79-83%	84-87%	>88%
PNL Met	<54%	54-67%	68-73%	74-80%	>80%
No cut off	<95%	95-97%	98-99%	-	100%
Nipple Midline	31%	31-45%	46-51%	52-55%	>55.5%
Target Compression	-	<58%	58-63%	64-68%	>69%
High Compression	>30%	13-30%	12-3%	-	<2%
Low Compression	>50%	27-50%	<26%	-	<10%
Proiezione MLO					
STAR RATING	★	★★	★★★	★★★★	★★★★★
CATEGORIA	Focus	Okay	Global median	Eccellente (top 25%)	Top 10%
Nipple in Profile	<69%	69-81%	82-85%	86-89%	>90%
IMF Visible	<23%	23-35%	36-44%	45-51%	>52%
PEC to PNL Met	<48%	48-64%	65-72%	73-78%	>79%
Adequate PEC	<86%	86-92%	93-94%	95-96%	>97%
No cut off	<96%	96-98%	99%	-	100%
PEC Shape	<63%	63-75%	76-81%	82-86%	>87%
No PEC skin folds	<90%	90-95%	96-97%	98-99%	>99%
Target Compression	-	<55%	55-67%	68-75%	>76%
High Compression	>4%	3-4%	<2%	-	<1%
Low Compression	>75%	46-75%	45-21%	-	<20%

I dati di baseline sono stati definiti per 5 TSRM su 6, dal momento che il TSRM 6 è entrato nel *team* a Marzo 2022. Si sono considerate le 3 metriche di valutazione generale (ovvero il Quality score, la percentuale di immagini con un valore PGMI perfetto o buono e la percentuale di immagini con compressione target), le quali sono state confrontate con la mediana globale (e per il Quality score anche con il valore limite); successivamente sono state raccolte le valutazioni in stelle assegnate dal software Volpara Analytics a ciascun parametro di posizionamento e di compressione.

Dai dati di baseline, sono state delineate le performance di partenza di ciascun TSRM e rintracciati quei parametri di compressione e di posizionamento con un giudizio di 1 o 2 stelle, afferenti quindi alle categorie “Focus” e “Okay”. Tale valutazione indica che la percentuale di immagini in cui il criterio viene rispettato è al di sotto della mediana globale, e pertanto richiede una maggiore attenzione e un maggiore impegno per aumentare la sua corretta presenza nelle mammografie.

2.7 Formazione e monitoraggio

Il mese di Febbraio 2022 è stato contraddistinto dalla fase di formazione, in cui ad ogni TSRM è stata presentata una scheda personale riassuntiva contenente i dati di baseline (con i criteri meno rispettati maggiormente marcati), e in cui sono state illustrate le principali finestre del software. Nella scheda sono state anche fornite delle linee guida per la correzione dei parametri più violati: per il posizionamento è stato riportato l’iter per poter accedere ai video di formazione forniti da Volpara Analytics; per la compressione è stato ricordato quanto sia importante tenere conto della pressione. (Figura 8)

Lo scopo di questa fase è stato quello di rendere consapevole ogni TSRM degli errori più frequentemente commessi e di formare ciascuno di essi all’utilizzo del software Volpara Analytics, in modo tale da poter essere consultato autonomamente almeno una volta a settimana.

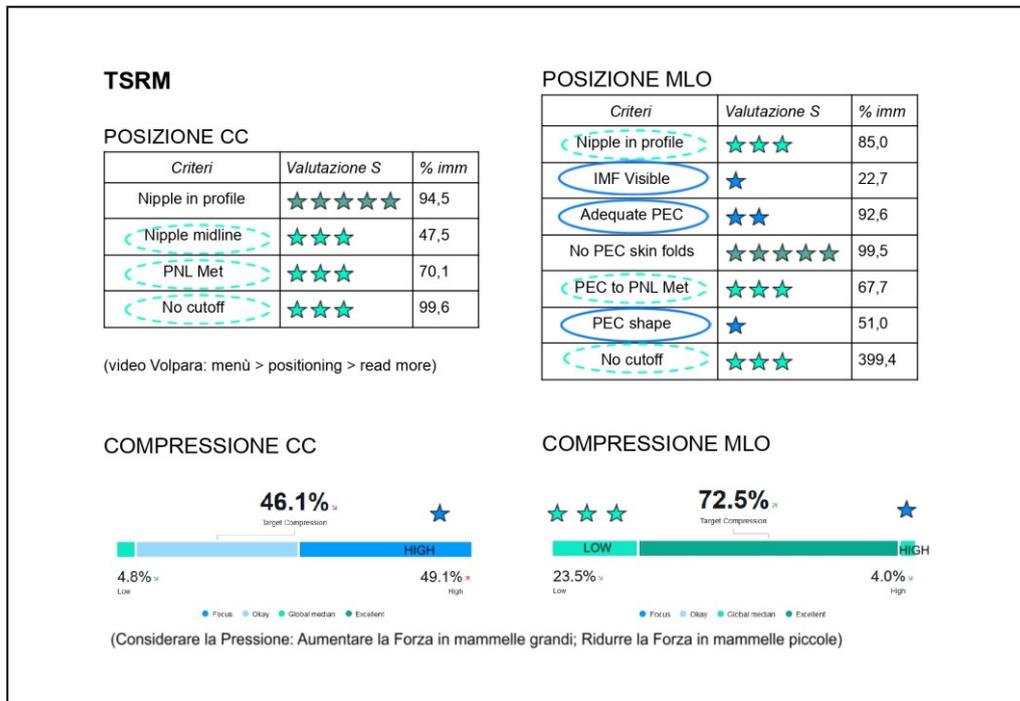


Figura 8: Esempio di una scheda di valutazione personale creata raccogliendo i dati di Volpara Analytics relativi al posizionamento e alla compressione delle immagini di screening nel periodo compreso tra Settembre 2021-Dicembre 2021.

Da Marzo a Settembre 2022, al termine di ogni mese e per ogni singolo TSRM, sono state raccolte le valutazioni di Volpara Analytics, relative ai parametri di posizionamento e di compressione, provenienti dalle immagini di screening. Questi valori sono stati messi a confronto sia con i dati analizzati nel periodo antecedente alla formazione (Settembre 2021 - Dicembre 2021), sia con i dati relativi al/ai mesi precedenti. Per ciascun operatore sono state quindi create delle tabelle riassuntive. Per il posizionamento sono stati presi in considerazione i criteri poco rispettati e già segnalati nel/nei mese/i precedente/i (o nel periodo pre-formazione), per cui, in caso di stabilità sono stati comunque riportati, in caso di variazione è stata riportata la nuova valutazione; inoltre, si sono mostrati anche i parametri che hanno subito un cambiamento significativo, sia in negativo che in positivo. Per la compressione invece, sono stati riportati direttamente i grafici forniti dal software Volpara Analytics. (Figura 9)

Ogni TSRM ha ricevuto la sua tabella personale al termine di ogni mese, in modo tale da poter essere al corrente dell'andamento delle sue performance e poter capire a quali parametri avrebbe dovuto prestare maggior attenzione.

TSRM					
POSIZIONE CC	SETT 2021-DIC 2021 SCREENING	MARZO 2022 SCREENING	APRILE 2022 SCREENING	MAGGIO 2022 SCREENING	GIUGNO 2022 SCREENING
%Good+Perfect	52.5%	73.6%	81.3%	71.6%	67.8%
Nipple Midline	★ 30.8%	★★★★ 51.3%	★★★★★ 59.3%	★★★★★ 52.0%	★★★★ 48.3%
POSIZIONE MLO	SETT 2021-DIC 2021 SCREENING	MARZO 2022 SCREENING	APRILE 2022 SCREENING	MAGGIO 2022 SCREENING	GIUGNO 2022 SCREENING
%Good+Perfect	31.8%	31.4%	51.7%	48.8%	41.1%
IMF Visible	★★ 26.5%	★★ 23.9%	★★ 34.5%	★★ 29.3%	★★ 27.7%
No cutoff	★★★★ 99.6%	★★★★★ 100%	★★★★★ 100%	★★★ 98.4%	★★★★ 99.1%
PEC Shape	★★ 68.8%	★ 47.8%	★ 60.3%	★ 61.8%	★★ 65.2%
COMPRESSIONE CC	SETT 2021-DIC 2021 SCREENING	MARZO 2022 SCREENING	APRILE 2022 SCREENING	MAGGIO 2022 SCREENING	GIUGNO 2022 SCREENING
	54.5% 	48.7% 	44.1% 	69.3% 	58.0%
COMPRESSIONE MLO	SETT 2021-DIC 2021 SCREENING	MARZO 2022 SCREENING	APRILE 2022 SCREENING	MAGGIO 2022 SCREENING	GIUGNO 2022 SCREENING
	88.6% 	79.7% 	89.7% 	83.9% 	87.6%

Figura 9: Esempio di scheda di valutazione personale consegnata al termine del mese di Giugno 2022.

2.8 Analisi dei dati

Nello stesso modo in cui sono stati definiti i dati di baseline, sono stati delineati anche i dati afferenti al periodo successivo alla formazione. I dataset sono stati organizzati all'interno di fogli elettronici al fine di poter effettuare un'analisi statistica descrittiva che valutasse l'esistenza di differenze tra le due raccolte di valori. Il confronto è stato svolto sia in termini di "star rating", sia facendo una comparazione tra percentuali, ponendo in particolare l'attenzione sui criteri che tra Settembre 2021 e Dicembre 2021 rientravano nelle categorie "focus" e "okay". Per ogni parametro sono state confrontate le differenze tra proporzioni, considerando significativo un "p-value" inferiore a 0.05.

Un'analisi diversa è stata eseguita per il TSRM 6, entrato a far parte dell'UOC Radiologia Senologica dell'Istituto Oncologico Veneto di Padova a Marzo 2022. In questo caso sono state considerate le percentuali cumulative mese per mese, per verificarne l'evoluzione temporale.

3. RISULTATI

3.1 Dati di baseline

I dati di baseline (pre-formazione) sono stati analizzati per i cinque TSRM che stavano già lavorando nel 2021.

Nella Figura 10 sono riportate le metriche generali di baseline (punteggio di qualità, percentuale di immagini P o G, e percentuale di immagini con compressione all'interno dell'intervallo target) per i cinque TSRM, sia per le proiezioni CC che per le MLO. La linea orizzontale rossa tratteggiata rappresenta la mediana globale ottenuta dalla distribuzione delle performance di oltre 5200 TSRM che utilizzano il software nel mondo. Il valore limite di 1.6 è disponibile solo per il punteggio di qualità calcolato da una combinazione di qualità del posizionamento e di compressione.

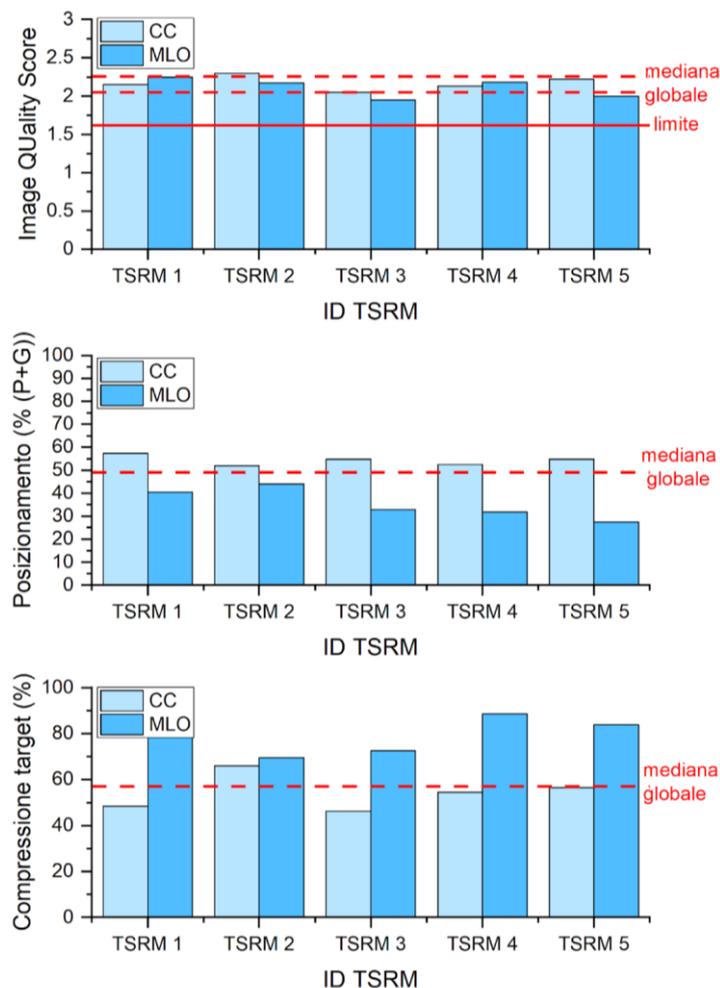


Figura 10: Punteggio di qualità delle immagini, percentuale di immagini P+G, percentuale di immagini con compressione target per ciascun TSRM e per ciascuna proiezione.

Il punteggio di qualità risulta superiore al limite di accettabilità per tutti i cinque TSRM e all'interno della mediana globale per la maggior parte di loro. Separando qualità del posizionamento e compressione si può notare che la percentuale di immagini classificate "Perfect" o "Good" (P+G) è molto maggiore per le proiezioni CC che per quelle MLO e per le proiezioni CC è superiore alla mediana globale, mentre è inferiore per le proiezioni MLO. Viceversa, per quattro dei cinque TSRM la proporzione di MLO ben compresse è molto maggiore delle CC; rispetto alla mediana globale, i TSRM della Radiologia Senologica IOV comprimono meglio le MLO e un po' peggio le CC.

Scendendo un po' più nel dettaglio, le tabelle II e III riportano le valutazioni del software in termini di star rating nel periodo di baseline per ciascun TSRM e criterio considerato, rispettivamente per le proiezioni CC e MLO.

Tabella II: Star rating per ciascun criterio e ciascun TSRM relativo al periodo di baseline per le proiezioni CC.

CC View: SEP 01, 2021 - DEC 31, 2021					
	TSRM 1	TSRM 2	TSRM 3	TSRM 4	TSRM 5
Nipple in Profile	★★★★★ 96.6%	★★★★★ 97.2%	★★★★★ 94.5%	★★★★★ 94.8%	★★★★★ 96.1%
CC PNL	★★★ 68.3%	★★★ 72.8%	★★★ 70.1%	★★★★★ 80.8%	★★★ 69.8%
No Cutoff	★★★ 98.9%	★★★★★ 100%	★★★ 99.6%	★★★ 99.2%	★★★ 99.6%
Nipple Midline	★★★★ 53.0%	★★ 36.7%	★★★ 47.5%	★ 30.8%	★★★ 46.2%
Target Compression	★★ 48.4%	★★★★ 66.0%	★★ 46.1%	★★ 54.5%	★★ 56.6%
Low Compression	★★★★★ 5.2%	★★★★★ 9.3%	★★★★★ 4.8%	★★★★★ 5.1%	★★★★★ 5.9%
High Compression	★ 46.4%	★★ 24.7%	★ 49.1%	★ 40.4%	★ 37.5%

Tutti i cinque TSRM partivano da una buona qualità del posizionamento della mammella nelle proiezioni CC. L'unico criterio non rispettato era "Nipple Midline" per il TSRM 2 e il TSRM 4. Inoltre, la percentuale di immagini con compressione target era inferiore alla mediana globale per quattro TSRM su

cinque e tutti i TSRM avevano una percentuale troppo alta di CC compresse eccessivamente.

Tabella II: Star rating per ciascun criterio e ciascun TSRM relativo al periodo di baseline per le proiezioni MLO.

MLO View: SEP 01, 2021 - DEC 31, 2021					
	TSRM 1	TSRM 2	TSRM 3	TSRM 4	TSRM 5
Nipple in Profile	★★★★★ 92.8%	★★★★★ 96.9%	★★★ 85.0%	★★★★★ 94.4%	★★★★★ 95.2%
IMF Visible	★★★ 39.9%	★★★ 40.6%	★ 22.7%	★★ 26.5%	★★ 23.8%
PEC to PNL Met	★★★★ 73.7%	★★★★★ 82.1%	★★★ 67.7%	★★★★ 74.0%	★★★ 65.4%
Adequate PEC	★★★ 95.3%	★★★★ 95.3%	★★ 92.6%	★★★ 94.0%	★★★ 94.7%
No Cutoff	★★ 98.9%	★★★★★ 100%	★★★ 99.4%	★★★ 99.6%	★★★ 99.5%
PEC Shape	★ 56.3%	★★ 63.6%	★ 51.0%	★★ 68.8%	★ 60.5%
No PEC Skin Folds	★★★ 96.9%	★★★ 97.7%	★★★★★ 99.5%	★★★★★ 99.3%	★★★★ 98.2%
Target Compression	★★★★★ 79.6%	★★★★ 69.6%	★★★★ 72.5%	★★★★★ 86.3%	★★★★★ 83.9%
Low Compression	★★★★★ 11.2%	★★★ 29.2%	★★★ 23.5%	★★★★★ 8.6%	★★★★★ 12.4%
High Compression	★ 9.2%	★★★ 1.2%	★★ 4.0%	★★ 2.8%	★★ 3.7%

Gli errori di posizionamento erano in numero maggiore nelle proiezioni MLO, con una percentuale di immagini considerate “corrette” (P+G) per tutti inferiori alla mediana globale. I criteri che richiedevano una maggior attenzione erano: “PEC Shape” per tutti i TSRM; “IMF Visible” per il TSRM 3, per il TSRM 4 e per il TSRM 5; “No cutoff” per il TSRM 1; “Adequate PEC” per il TSRM 4. La percentuale di immagini correttamente compressa è per tutti eccellente e superiore alla mediana globale, con una piccola proporzione di mammelle compresse eccessivamente per quattro TSRM su cinque.

3.2 Performance dei singoli TSRM dopo l'uso consapevole del software

Nelle sezioni successive sono riportate le tabelle di confronto per ciascun TSRM tra i dati di baseline e i dati di controllo, con un test sulle proporzioni. Nell'ultima colonna delle tabelle è riportato il p-value, che risulta colorato solo per le differenze significative; il colore verde è indicato per rappresentare un miglioramento rispetto alla baseline, il rosso per rappresentare un peggioramento. È stata riportata anche la classificazione in termini di star rating perché una variazione significativa (in un senso o nell'altro) può essere insufficiente per guadagnare/perdere una o più stelle.

3.2.1 TSRM 1

Il TSRM 1 aveva come obiettivi il miglioramento della compressione target per le proiezioni CC e il miglioramento dei criteri "No cutoff" e "PEC Shape" e la riduzione della compressione per le proiezioni MLO.

Nella Tabella IV è riportato il confronto tra il periodo di test e quello di baseline.

Tabella IV: Confronto tra periodo di test e periodo di baseline del TSRM 1 per le proiezioni CC e MLO.

TSRM 1						
Criteri CC	% PRE	° PRE	% POST	° POST	Differenza %	P
Nipple in profile	96.6	★★★★★	94.6	★★★★★	-2.00	0.0173
PNL Met	68.3	★★★	71.4	★★★	3.10	0.1011
No cutoff	98.9	★★★	99.7	★★★	0.80	0.0200
Nipple midline	53.0	★★★★	56.5	★★★★★	3.50	0.0878
Target compression	48.4	★★	59.8	★★★	11.40	< 0.0001
Low compression	5.2	★★★★★	7.5	★★★★★	2.30	0.0228
High compression	46.4	★	32.7	★	-13.70	< 0.0001
Criteri MLO	% PRE	° PRE	% POST	° POST	Differenza %	P
Nipple in profile	92.8	★★★★★	95.3	★★★★★	2.50	0.0105
IMF visible	39.9	★★★	39.7	★★★	-0.20	0.9212
PEC to PNL Met	73.7	★★★★	84.1	★★★★★	10.40	< 0.0001
Adequate PEC	93.5	★★★	97.2	★★★★★	3.70	< 0.0001
No cutoff	98.9	★★	99.6	★★★	0.70	0.0496
Pec shape	56.3	★	56.2	★	-0.10	0.9611
No pec skin folds	96.9	★★★	98.5	★★★★	1.60	0.0098

Target compression	79.6	★★★★★	73.1	★★★★	-6.50	0.0002
Low compression	11.2	★★★★★	24.2	★★★	13.00	< 0.0001
High compression	9.2	★	2.7	★★	-6.50	< 0.0001

La compressione target delle immagini CC è migliorata in modo significativo tanto da raggiungere la mediana globale, e in modo complementare la percentuale di mammografie eccessivamente compresse è stata ridotta. Per le proiezioni MLO, i criteri “No cutoff” e “High Compression”, hanno avuto un miglioramento significativo, mentre questo non è accaduto per il criterio “PEC Shape”.

Tra i criteri, con un numero di stelle superiore a 1 o 2 nella baseline, ad essere stati soggetti a una crescita significativa dello star rating sono stati “Nipple Midline” per le CC; “PEC to PNL Met”, “Adequate PEC” e “No PEC Skinolds” per le MLO. Viceversa, ad aver subito un calo è stata la compressione target delle MLO, con un aumento complementare delle immagini ottenute con una scarsa compressione.

3.2.2 TSRM 2

Il TSRM 2 aveva come obiettivi il miglioramento del parametro “Nipple Midline” e la riduzione delle immagini ottenute con compressione eccessiva per le proiezioni CC e il miglioramento del criterio “PEC Shape” per le proiezioni MLO.

Nella Tabella V è riportato il confronto tra il periodo di test e quello di baseline.

Tabella V: Confronto tra periodo di test e periodo di baseline del TSRM 2 per le proiezioni CC e MLO.

TSRM 2						
Criteria CC	% PRE	° PRE	% POST	° POST	Differenza %	P-value
Nipple in profile	97.2	★★★★★	96.0	★★★★★	-1.20	0.1383
PNL Met	72.8	★★★	80.5	★★★★★	7.70	0.0001
No cutoff	100	★★★★★	99.3	★★★	-0.70	0.0051
Nipple midline	36.7	★★	47.0	★★★	10.30	< 0.0001
Target compression	66.0	★★★★	71.2	★★★★★	5.20	0.0133
Low compression	9.3	★★★★★	6.7	★★★★★	-2.60	0.0354
High compression	24.7	★★	22.1	★★	-2.60	0.1748

Criteria MLO	% PRE	° PRE	% POST	° POST	Differenza %	P-value
Nipple in profile	96.9	★★★★★	96.7	★★★★★	-0.20	0.8015
IMF visible	40.6	★★★	48.4	★★★★	7.80	0.0005
PEC to PNL Met	82.1	★★★★★	87.6	★★★★★	5.50	0.0008
Adequate PEC	95.3	★★★★	98.4	★★★★★	3.10	0.0001
No cutoff	100.0	★★★★★	99.9	★★★	-0.10	0.2910
Pec shape	63.6	★★	57.7	★	-5.90	0.0075
No pec skin folds	97.7	★★★	96.3	★★★	-1.40	0.0660
Target compression	69.6	★★★★	87.6	★★★★★	18.00	< 0.0001
Low compression	29.2	★★★	11.9	★★★★★	-17.30	< 0.0001
High compression	1.2	★★★	0.5	★★★★★	-0.70	0.1004

Si può notare che per le proiezioni CC il parametro “Nipple Midline” ha subito un miglioramento significativo; mentre la percentuale di immagini troppo compresse non ha visto variazioni importanti. Per la proiezione MLO invece il criterio “PEC Shape” ha subito un peggioramento significativo. Tra gli altri criteri, non rientranti nelle categorie “focus” e “okay” nel periodo pre-formazione, ad avere un miglioramento importante dello star rating sono stati “PNL Met” per le proiezioni CC e “IMF Visible” e “Adequate PEC” per le proiezioni MLO. Anche nella compressione in entrambe le proiezioni il TSRM 2 ha migliorato il proprio star rating. Invece, ad aver subito un calo del numero di stelle in entrambe le proiezioni è stato il parametro “No cutoff”, significativo nelle proiezioni CC e non significativo nelle proiezioni MLO.

3.2.3 TSRM 3

Il TSRM 3 aveva come obiettivi la riduzione delle proiezioni CC e MLO eccessivamente compresse e il miglioramento dei criteri “IMF Visible”, “Adequate PEC” e “PEC Shape” nelle proiezioni MLO.

Nella Tabella VI è riportato il confronto tra il periodo di test e quello di baseline.

Tabella VI: Confronto tra periodo di test e periodo di baseline del TSRM 3 per le proiezioni CC e MLO.

TSRM 3						
Criteria CC	% PRE	° PRE	% POST	° POST	Differenza %	P-value
Nipple in profile	94.5	★★★★★	91.0	★★★★★	-3.50	0.0028
PNL Met	70.1	★★★	73.4	★★★	3.30	0.2567
No cutoff	99.6	★★★	98.8	★★★	-0.80	0.1040
Nipple midline	47.5	★★★	53.2	★★★★	5.70	0.0115
Target compression	46.1	★★	61.2	★★★	15.10	< 0.0001
Low compression	4.8	★★★★★	16.4	★★★	11.60	< 0.0001
High compression	49.1	★	22.4	★★	-26.70	< 0.0001
Criteria MLO	% PRE	° PRE	% POST	° POST	Differenza %	P-value
Nipple in profile	85.0	★★★	92.4	★★★★★	7.40	< 0.0001
IMF visible	22.7	★	15.6	★	-7.10	0.0001
PEC to PNL Met	67.7	★★★	79.5	★★★★★	11.80	< 0.0001
Adequate PEC	92.6	★★	95.5	★★★★	2.90	0.0073
No cutoff	99.4	★★★	99.7	★★★	0.30	0.3254
Pec shape	51.0	★	57.3	★	6.30	0.0056
No pec skin folds	99.5	★★★★★	99.2	★★★★★	-0.30	0.4138
Target compression	72.5	★★★★	56.5	★★★	-16.00	< 0.0001
Low compression	23.5	★★★	42.9	★★★	19.40	< 0.0001
High compression	4.0	★★	0.6	★★★★★	-3.40	< 0.0001

Il TSRM 3 è riuscito a migliorare significativamente le sue performance di compressione delle proiezioni CC e anche l'importante criterio "Adequate PEC" nelle proiezioni MLO. Il criterio "PEC Shape" ha visto un leggero miglioramento significativo, mentre la percentuale di immagini MLO col criterio "IMF Visible" rispettato si è addirittura ridotta. Inoltre, come richiesto, le immagini MLO eccessivamente compresse sono calate, ma sono aumentate quelle scarsamente compresse: ciò ha comportato una riduzione dello star rating della compressione target.

Tra gli altri criteri, con un numero di stelle superiore a 1 o 2, "Nipple Midline" per le proiezioni CC è stato soggetto a una variazione positiva dello star rating, come "Nipple in profile" e "PEC to PNL Met" per le proiezioni MLO.

3.2.4 TSRM 4

Il TSRM 4 aveva come obiettivi il miglioramento del criterio “Nipple Midline” nelle proiezioni CC e quello dei criteri “IMF Visible” e “PEC shape” nelle proiezioni MLO, nonché la riduzione del numero di esami per cui le mammelle risultano eccessivamente compresse, sia in CC che in MLO.

Nella Tabella VII è riportato il confronto tra il periodo di test e quello di baseline.

Tabella VII: Confronto tra periodo di test e periodo di baseline del TSRM 4 per le proiezioni CC e MLO.

TSRM 4						
Criteria CC	% PRE	° PRE	% POST	° POST	Differenza %	P-value
Nipple in profile	94.8	★★★★★	93.2	★★★★★	-1.60	0.1924
PNL Met	80.8	★★★★★	81.3	★★★★★	0.50	0.8053
No cutoff	99.2	★★★	99.9	★★★	0.70	0.0449
Nipple midline	30.8	★	48.4	★★★	17.60	< 0.0001
Target compression	54.5	★★	56.1	★★	1.60	0.5341
Low compression	5.1	★★★★★	7.1	★★★★★	2.00	0.1059
High compression	40.4	★	36.8	★	-3.60	0.1531
Criteria MLO	% PRE	° PRE	% POST	° POST	Differenza %	P-value
Nipple in profile	94.4	★★★★★	94.5	★★★★★	0.10	0.9336
IMF visible	26.5	★★	31.2	★★	4.70	0.0477
PEC to PNL Met	74.0	★★★★	84.6	★★★★★	10.60	< 0.0001
Adequate PEC	94.0	★★★	97.3	★★★★★	3.30	0.0022
No cutoff	99.6	★★★	98.7	★★	-0.90	0.0594
Pec shape	68.8	★★	59.1	★	-9.70	< 0.0001
No pec skin folds	99.3	★★★★★	98.6	★★★★	-0.70	0.1881
Target compression	88.6	★★★★★	86.7	★★★★★	-1.90	0.2703
Low compression	8.6	★★★★★	12.6	★★★★★	4.00	0.0130
High compression	2.8	★★	0.7	★★★★★	-2.10	0.0025

Il criterio “Nipple Midline” delle proiezioni CC ha subito un netto miglioramento, passando dal 30.8% al 48.4% fino a raggiungere la mediana globale, mentre la compressione non ha subito variazioni importanti. Il criterio “IMF visible” nelle proiezioni MLO è migliorato (+4.7%) con una differenza statisticamente significativa; viceversa, si è significativamente ridotta la percentuale di MLO per

cui il criterio “PEC Shape” è stato rispettato. È migliorato il parametro di compressione delle MLO con una riduzione significativa delle mammelle troppo compresse.

Tra gli altri criteri con un numero di stelle superiore a 1 o 2, ad essere stati soggetti ad un aumento dello star rating tra Marzo 2022 e Settembre 2022 sono stati: “PEC to PNL Met” e “Adequate PEC”; invece i criteri “No cutoff” e “No PEC skinfolds” hanno visto un calo dello star rating, ma, essendo il p-value superiore al 5%, la differenza non è statisticamente significativa.

3.2.5 TSRM 5

Il TSRM 5 aveva come obiettivi l’aumento della percentuale di immagini ottenute nell’intervallo target di compressione nelle proiezioni CC e il miglioramento dei criteri “IMF Visible” e “PEC Shape”, nonché la riduzione della compressione nelle proiezioni MLO.

Nella Tabella VIII è riportato il confronto tra il periodo di test e quello di baseline.

Tabella VIII: Confronto tra periodo di test e periodo di baseline del TSRM 5 per le proiezioni CC e MLO.

TSRM 5						
Criteria CC	% PRE	° PRE	% POST	° POST	Differenza %	P-value
Nipple in profile	96.1	★★★★★	96.9	★★★★★	0.80	0.4445
PNL Met	69.8	★★★	75.3	★★★★	5.50	0.0239
No cutoff	99.6	★★★	99.7	★★★	0.10	1.0000
Nipple midline	46.2	★★★	52.3	★★★★	6.10	0.0943
Target compression	56.6	★★	61.9	★★★	5.30	0.0969
Low compression	5.9	★★★★★	13.9	★★★	8.00	< 0.0001
High compression	37.5	★	24.2	★★	-13.30	< 0.0001
Criteria MLO	% PRE	° PRE	% POST	° POST	Differenza %	P-value
Nipple in profile	95.2	★★★★★	94.3	★★★★★	-0.90	0.5181
IMF visible	23.8	★★	24.2	★★	0.40	0.8670
PEC to PNL Met	65.4	★★★	82.2	★★★★★	16.80	< 0.0001
Adequate PEC	94.7	★★★	97.4	★★★★★	2.70	0.0226
No cutoff	99.5	★★★	100	★★★★★	0.50	0.1082
Pec shape	60.5	★	52.1	★	-8.40	0.0095
No pec skinfolds	98.2	★★★★	97.1	★★★	-1.10	0.1845

Target compression	83.9	★★★★★	76.3	★★★★★	-7.60	0.0039
Low compression	12.4	★★★★★	22.3	★★★	9.90	< 0.0001
High compression	3.7	★★	1.4	★★★	-2.30	0.0187

La compressione target delle immagini CC è migliorata grazie alla riduzione significativa dell'eccessiva compressione. Per quanto riguarda le proiezioni MLO, il criterio "IMF visible" non ha subito variazioni importanti, mentre il criterio "PEC Shape" ha subito un calo significativo. In compenso c'è stata una riduzione importante delle immagini troppo compresse.

Tra gli altri parametri, non rientranti nelle categorie "focus" e "okay", i parametri "PNL Met" e "Nipple midline" nelle proiezioni CC sono stati soggetti ad un aumento dello star rating, come pure "PEC to PNL Met" e "Adequate PEC" nelle proiezioni MLO. Le variazioni dei criteri "No cutoff" e "No PEC Skinfolde" nelle MLO non sono state significative.

3.2.6 TSRM 6

Come precedentemente detto, il TSRM 6, neo-laureato, è entrato a far parte dell'UOC Radiologia Senologica dell'Istituto Oncologico Veneto di Padova a Marzo 2022 e ha iniziato la sua attività lavorativa disponendo fin da subito del software Volpara Analytics. I dati cumulativi dei parametri, raccolti mese per mese, sono stati rapportati con i valori della mediana globale.

I grafici nelle figure 11 e 12 mostrano l'andamento delle performance per il posizionamento delle proiezioni CC e MLO.

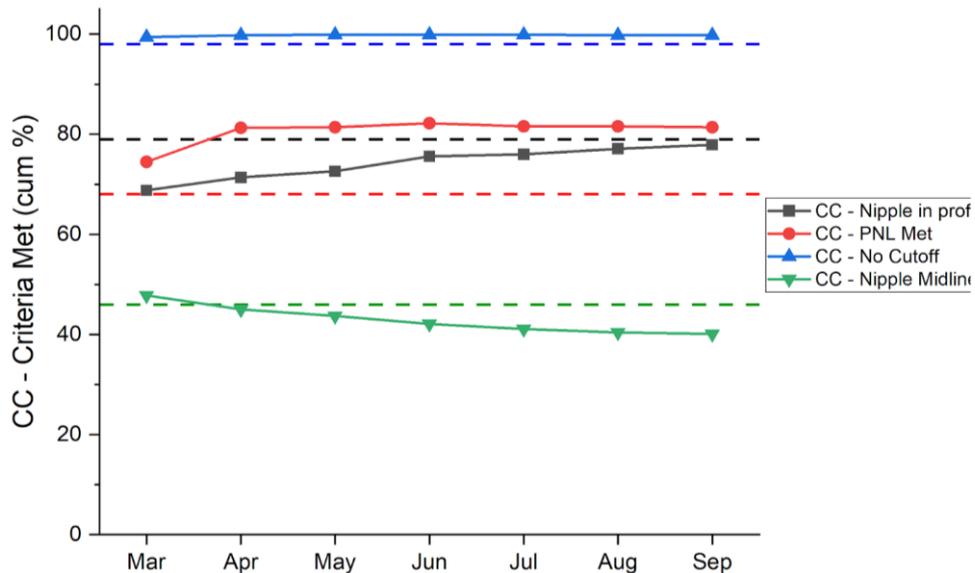


Figura 11: Percentuale cumulativa di immagini ottenute dal TSRM 6 rispettando i criteri di posizionamento in CC.

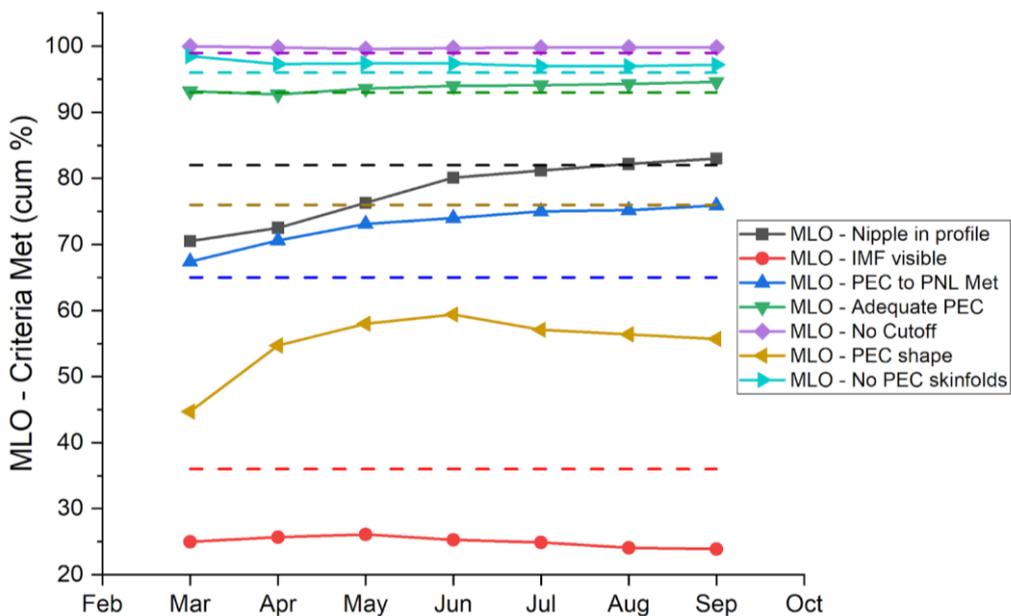


Figura 12: Percentuale cumulativa di immagini ottenute dal TSRM 6 rispettando i criteri di posizionamento in MLO.

Il TSRM 6, che ha potuto disporre del software fin da subito, ha ottenuto risultati molto buoni, almeno per i criteri di posizionamento più importanti, in tempi veramente brevi, raggiungendo la mediana globale nel giro di pochi mesi per

tutti i criteri ad esclusione di “Pec shape” e “IMF visible” nelle MLO e un peggioramento del criterio “Nipple midline” nelle CC.

Nelle Figure 13 e 14 sono riportate le percentuali cumulative relative alle performance di compressione del TSRM 6 rispettivamente per le proiezioni CC e MLO.

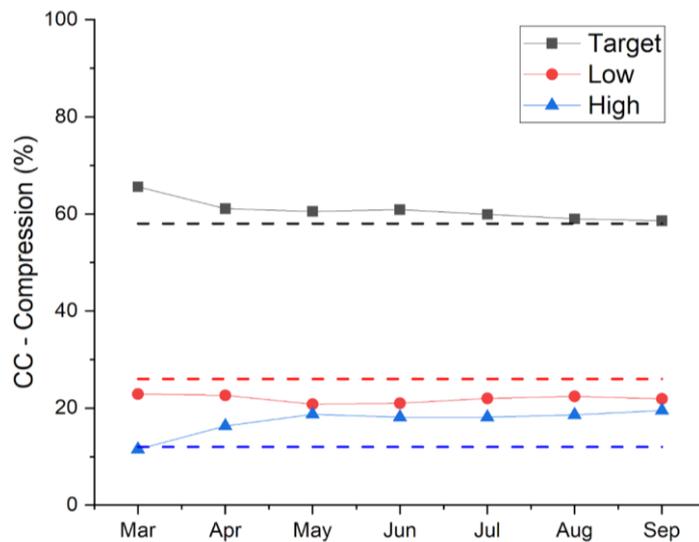


Figura 13: Percentuale cumulativa di immagini CC ottenute dal TSRM 6 con compressione dentro l'intervallo target, con compressione bassa e con compressione eccessiva.

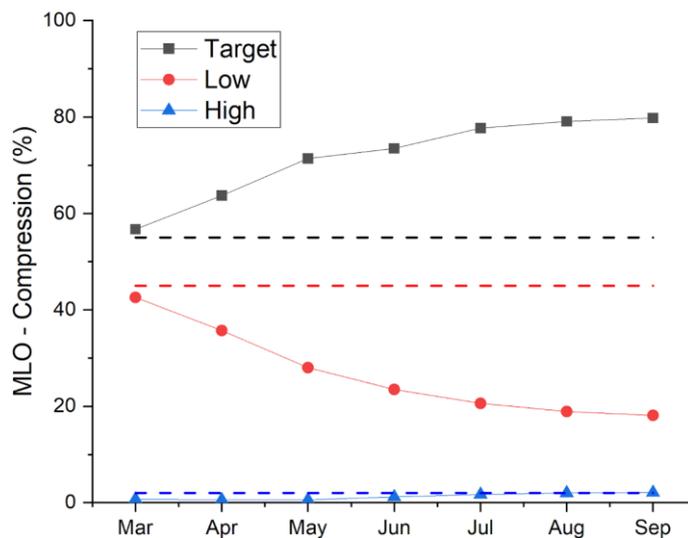


Figura 14: Percentuale cumulativa di immagini MLO ottenute dal TSRM 6 con compressione dentro l'intervallo target, con compressione bassa e con compressione eccessiva.

Nelle proiezioni CC il TSRM 6 ha mantenuto la compressione target sempre sopra la mediana, con una percentuale di mammelle poco compresse inferiore alla mediana e una percentuale di mammelle troppo compresse superiore. Nelle proiezioni oblique invece il miglioramento di performance di compressione è nettamente visibile, con un progressivo aumento delle mammelle all'interno dell'intervallo target di compressione ben al di sopra della mediana globale e una progressiva riduzione delle mammelle poco compresse, ben al di sotto della mediana globale.

Nella Tabella IX sono riportati i risultati riassuntivi dei cinque TSRM che hanno cercato di utilizzare il software per migliorare aspetti specifici della compressione e del posizionamento.

Tabella IX: Obiettivi di compressione o di posizionamento assegnati ai singoli TSRM raggiunti (✓) o non raggiunti (✗) per le proiezioni CC ed MLO.

CC View					
Criteri di correttezza	TSRM 1	TSRM 2	TSRM 3	TSRM 4	TSRM 5
Nipple Midline		✓		✓	
Target compression	✓		✓	✗	✓
High compression	✓	✗	✓	✗	✓
MLO View					
IMF Visible			✗	✓	✗
Adequate PEC			✓		
No cutoff	✓				
PEC Shape	✗	✗	✓	✗	✗
High compression	✓		✓	✓	✓

Il TSRM 1 ha raggiunto 4 dei suoi 5 obiettivi, il TSRM 2 ne ha raggiunto 1 su 3, il TSRM 3 è migliorato in 5 su 6 aspetti, il TSRM 4 in 3 su 6, e il TSRM 5 in 3 su 5.

4. DISCUSSIONE

Questo studio si poneva come obiettivo principale quello di verificare se l'utilizzo di un software automatico per la valutazione della qualità tecnica delle immagini mammografiche potesse essere considerato utile per migliorare le performance del singolo tecnico di radiologia nell'esecuzione delle mammografie.

I risultati emersi dall'analisi condotta dimostrano che la consultazione di tale strumento da parte dei TSRM per i 6 mesi successivi all'evento formativo, nel corso del quale è stato lanciato il progetto e a ciascun tecnico sono stati consegnati gli obiettivi di miglioramento, ha realmente contribuito in modo significativo alla risoluzione di molti punti critici relativi al posizionamento e alla compressione.

Per quanto concerne i criteri di posizionamento, gli obiettivi segnalati per la proiezione CC sono stati raggiunti da tutti i TSRM coinvolti; mentre, per la proiezione MLO, i criteri più importanti come "NO cutoff" e "Adequate PEC" sono stati ampiamente migliorati, ma rimangono ancora oggetto di attenzione i criteri "IMF Visibile" e "PEC Shape". È importante però ricordare che la visibilità del solco nella proiezione obliqua è un fattore che non dipende solo dalla correttezza del posizionamento da parte del TSRM ma anche dalla conformazione anatomica di ciascuna donna; in effetti la mediana globale è fissata al 36%, valore che indica che ci si aspetta che il solco possa essere visibile in una percentuale non molto elevata di immagini mammografiche. Il secondo criterio "PEC Shape" invece ha un peso inferiore sulla qualità dell'immagine, in quanto la forma del muscolo pettorale non è rappresentativa della copertura volumetrica della proiezione; pertanto, il mancato rispetto di questo criterio è da considerarsi di minor importanza.

Le performance di compressione sono migliorate significativamente in entrambe le proiezioni. Grazie al feedback del software, tre TSRM su quattro hanno raggiunto una compressione target pari ai livelli della mediana globale nelle proiezioni CC, e quattro TSRM su quattro hanno ridotto l'eccessiva compressione segnalata nelle proiezioni MLO.

In media, tre criteri segnalati su cinque (61,3%) hanno visto un miglioramento in termini di percentuali e di star rating.

Il software di valutazione automatica ha sia aiutato a migliorare le performance dei TSRM con più esperienza, ma ha anche contribuito a formare nella tecnica e a guidare nell'esecuzione il TSRM che partiva da "zero". I risultati infatti ottenuti da quest'ultimo sono stati molto buoni: con il passare dei mesi le percentuali cumulative di molti criteri hanno raggiunto e/o superato la mediana globale.

Insieme all'ottimizzazione tecnica dei criteri target sia CC che MLO, spesso e volentieri si è verificato anche un perfezionamento degli altri criteri, che avevano già un valore pari o superiore alla mediana globale. Una possibile spiegazione di questo risultato sta nel fatto che migliorare dei criteri in cui già si riesce bene può risultare più semplice rispetto che migliorare dei criteri, i quali, il più delle volte, non sono correttamente rappresentati. Infatti, nel primo caso è questione di prestare maggiore attenzione a qualcosa che già si sa fare; nel secondo caso è questione di cambiare una tecnica già consolidata ma non totalmente corretta. Osservando i risultati, anche se molto meno frequentemente rispetto i miglioramenti, si possono notare dei peggioramenti. Questi sono dovuti principalmente al fatto che le manovre e i movimenti per cercare di rispettare maggiormente i criteri obiettivo del miglioramento vadano a discapito di altri criteri.

In generale, cinque TSRM su cinque hanno ottenuto in totale più miglioramenti che peggioramenti e le variazioni maggiori dei criteri, sia positive che negative, sono state conseguite da quei TSRM con una minore anzianità all'interno della radiologia senologica (TSRM 3, TSRM 4). Questo a conferma del fatto che la loro tecnica essendo più recente e meno consolidata è più facilmente soggetta a variazioni e richiede del tempo per essere rinforzata e stabilizzata.

Un'altra considerazione da fare è che alcuni criteri, come "no cutoff", per raggiungere l'eccellenza hanno un intervallo molto ristretto e molto ravvicinato al valore che individua la categoria della mediana globale. Questo spiega il motivo per cui, nonostante una variazione in negativo non significativa della percentuale, la correttezza del criterio viene comunque declassata, passando per esempio da un numero di cinque stelle a un numero di tre stelle.

In questo progetto di tesi è stata verificata l'utilità di un software di valutazione automatica, ma un'alternativa molto diffusa per monitorare e migliorare le performance è l'utilizzo dei metodi di autovalutazione, come il protocollo dell'Emilia-Romagna. [13] Tali metodi richiedono ai singoli TSRM di esaminare un numero limitato di mammografie: questo li rende degli strumenti molto utili dal punto di vista didattico ma dal punto di vista statistico sono poco significativi e non propriamente adatti a fare analisi visti i grandi numeri che caratterizzano la normale pratica. Al contrario, un metodo di valutazione automatica da una parte gestisce in modo oggettivo una quantità di immagini elevatissima, il che rende lo strumento statisticamente potentissimo, ma dall'altra coinvolge in maniera minore i TSRM nel capire e nel correggere gli errori più frequenti. I due diversi metodi di valutazione, i cui giudizi sono concordanti secondo lo studio condotto da Waade GG et al. [14], non sono quindi dei metodi che si escludono a vicenda ma dei metodi che possono essere integrati tra loro per raggiungere un obiettivo comune, ovvero quello di migliorare l'esecuzione tecnica delle immagini mammografiche.

5. CONCLUSIONI

L'utilizzo costante di un metodo di valutazione automatica, senza un'eccessiva focalizzazione sul caso particolare, ma sfruttando l'elevata quantità di mammografie processate, può essere un valido alleato per ciascun TSRM, anche e soprattutto per chi si trova alle prime armi. Infatti, come dimostrano i risultati di questo studio, il software ha contribuito ad una ottimizzazione delle performance dei singoli TSRM nell'esecuzione delle immagini mammografiche, fornendo a ciascuno le conoscenze e le basi per capire gli errori e gli strumenti per poter risolvere la maggior parte di essi.

6. BIBLIOGRAFIA E SITOGRAFIA

1. Piano sanitario nazionale (Psn) 1998-2000
2. Bassett LW, Hirbawi IA, DeBruhl N, Hayes MK. Mammographic positioning: evaluation from the view box. *Radiology* 1993; 188:803–806
3. Taplin, Stephen H et al. Screening mammography: clinical image quality and the risk of interval breast cancer. *AJR. American journal of roentgenology* vol. 178,4 (2002): 797-803.
4. Bassett LW, Farria DM, Bansal S, et al. Reasons for failure of a mammography unit at clinical image review in the American College of Radiology Mammography Accreditation Program. *Radiology* vol. 215,3 (2000): 698-702.
5. Available:
<http://www.aitasit.org/wp49/2021/03/27/proiezione-cranio-caudale-cc/>
6. Available:
<http://www.aitasit.org/wp49/2021/03/27/proiezione-medio-laterale-obliqua-mlo/>
7. D. Kopans, *Breast Imaging*, 3rd ed., Lippincott Williams & Wilkins, London, 2007 (ISBN: 0-7817-4768-6).
8. Available:
VolparaPositioning_Data_Sheet
9. Perry, N., et al., European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition—summary document. *Ann Oncol*, 2008. 19(4): p. 614–22.
10. Holland, K., et al., Performance of breast cancer screening depends on mammographic compression, in *Breast Imaging: 13th International Workshop, IWDM 2016, Malmö, Sweden, June 19–22, 2016, Proceedings*, A. Tingberg, K. Lång, and P. Timberg, Editors. 2016, Springer International Publishing: Cham. p. 183–189.
11. Available:
Optimal Breast Compression in Mammography Volpara
12. Available:
<https://www.infoseno.org/il-tumore-del-seno/trattamenti-ed-effetti-collaterali/chirurgia-del-seno/effetti-collaterali-e-consigli-pratici/>
13. Protocollo di valutazione tecnica dell'esame mammografico, programma di screening mammografico della regione Emilia-Romagna.

14. Waade GG, Danielsen AS, Holen ÅS, et al. Assessment of breast positioning criteria in mammographic screening: Agreement between artificial intelligence software and radiographers. *Journal of Medical Screening*. 2021;28(4):448-455.