

# 乳房X線画像における乳房構成解析 (Breast Density Assessment) ソフトウエアの有用性

Usefulness of breast structure analysis (Breast Density Assessment) software in mammograms

櫻井 典子1),甲斐 千遥1),長 和弘2),近藤 敏志3),児玉 直樹1),笠井 聡1)

1) 新潟医療福祉大学 医療技術学部 診療放射線学科 2) 新潟県立がんセンター新潟病院 3) 室蘭工業大学 情報電子工学系専攻

Key words: Mammograms, Breast content analysis, Observer Performance Study

#### [Abstract]

The incidence of breast cancer is the highest in Japan, and mammograms are essential for breast cancer screening for early detection. Breast structure analysis is classified into four levels based on the percentage of mammary gland content, with high density breasts having a higher risk of breast cancer. The bifurcation line between non-dense and dense is therefore very important. However, the classification is subjective, based on the physician's judgment, and the inter and intra observer variation tends to be large. Therefore, we evaluated the usefulness of the breast composition analysis software through an observer performance study conducted by radiological technologists. The results showed that the agreement rate with an expert mammographer increased and the inter-observer variability decreased significantly when the software was used. In particular, it was suggested that the software may improve the accuracy of diagnosis and reading efficiency by accurately classifying the bifurcation line.

#### 【要旨】

日本における乳がんの罹患率は1位であり、早期発見のための乳がん検診にマンモグラムは必須である。乳房構成の分類は乳腺含有率の割合から4段階に分類され、その中でも高濃度乳房は乳がんのリスクが高くなる。そのため非高濃度と高濃度の分岐ラインは非常に重要であるが、分類は医師の判断による主観的なもので、そのバラツキが問題となっている。そこで乳房構成解析ソフトウエアの有用性を、診療放射線技師による観察者実験により評価した。その結果、乳腺専門医との一致率は向上し、観察者間のバラツキも有意に減少した。特に非高濃度と高濃度の分岐ラインが的確に分類され、診断精度と読影効率の向上に寄与する可能性が示唆された。

## 1. はじめに

本邦では、乳がんは女性におけるがんの死亡率第4位であり、年々増加の一途をたどっている<sup>1)</sup>. さらに罹患率では第1位であり、40歳代の罹患率が高いことが特徴的である。若い世代のがんの進行は早いが治療効果も高いため、より早期の発見が求められる。早期に発見することができれば、治る可能性が高いがんともいわれており、マンモグラム(乳房 X 線画像)を用いた乳がん検診の重要性がうたわれている。

SAKURAI Noriko<sup>1)</sup>, KAI Chiharu<sup>1)</sup>, CHO Kazuhiro<sup>2)</sup>, KONDO Satoshi<sup>3)</sup>, KODAMA Naoki<sup>1)</sup>, KASAI Satoshi<sup>1)</sup>

- 1) Department of Radiological Technology, Faculty of Medical Technology, Niigata University of Health and Welfare
- 2) Niigata Cancer Center Hospital
- 3) Division of Information and Electronic Engineering, Muroran Institute of Technology

Received July 4, 2022; accepted February 2, 2023

乳房構成の分類は、米国放射線学会 (ACR: American College of Radiology)が作成した乳腺 画像診断報告とデータシステムを基に標準化したも ので、BI-RADS (Breast Imaging Reporting and Data System) アトラスにおいて4つの乳房構成に分 類されている2). 日本でもそれに従い、日本乳がん検 診精度管理中央機構(以下,精中機構)が「脂肪性」 「乳腺散在」「不均一高濃度」「極めて高濃度」の4段階 に乳房構成を分類することをガイドラインとして推奨 している<sup>3)</sup>. 判定は、乳房内の乳腺組織が存在してい たと考えられる領域を分母とし、分子は大胸筋と等濃 度以上の面積の総和とする. この割合を乳腺含有率と 定義する. その乳腺含有率が10%未満を「脂肪性」, 10%から50%未満を「乳腺散在」、50%から80%未 満を「不均一高濃度」,80%以上を「極めて高濃度」と している. さらに「不均一高濃度」および「極めて高 濃度」は乳腺実質内の乳腺含有率が50%以上であり、 まとめて「高濃度乳房」と定義されている. 乳房構成 分類は、病変が正常乳腺に隠されてしまう危険性の程 度を示す<sup>3)</sup>. 高濃度乳房は白い領域が多く腫瘍と重な るため、腫瘍の検出を困難とするマスキングリスクが 高い. また高濃度乳房においては、乳がんの発症リスクも高いことが報告されている<sup>4)</sup>. 日本人は、欧米人に比べ高濃度乳房の割合が高く、日本人女性においても高濃度乳房が発がんリスクを高める報告がある<sup>5,6)</sup>. この乳房構成分類は、医師の判断による主観的なものであり、経験にもよるため、読影者間のバラツキが問題となっている<sup>7)</sup>. 4段階の乳房構成分類の典型例では一致率が高いが、その分岐ラインの症例では一致率が低くなる. Redondoらは、特に高濃度の分岐ラインである「乳腺散在乳房」と「不均一高濃度乳房」の間で顕著であり、致命的な問題としている<sup>8)</sup>.

これらの背景より、乳房構成の分類を自動で行うソフトウエアが開発されている。方法としては2種類あり、2次元的に診断レベルの画像から乳腺濃度を推定する方法と、3次元的分析により乳腺密度から推定する方法が挙げられる<sup>9.10,11)</sup>.

本研究の目的は前者である、2次元的に濃度ベースで乳房構成を推定するソフトウエアを使用することで、日常的にマンモグラムを撮影する診療放射線技師と乳腺専門医の乳房構成分類の一致率が向上する可能性を検討した。

# 2. 方法

#### 2-1. 使用機器

本研究では、コニカミノルタ社製の乳房構成解析ソ フトウエア (Breast Density Assessment (Bda): 以下、ソフトウエア)を使用した. 本ソフトウエア は、マンモグラム上で乳腺組織がもともと存在してい た領域を抽出するプロセスと、乳腺が現存する領域を 抽出するプロセスから構成され、2次元的に濃度ベー スの閾値処理で乳腺含有率を算出している。現時点で の乳房構成の分類は、精中機構により視覚的に濃度で 判定することが推奨されており、本ソフトウエアは精 中機構に沿ったソフトウエアといえる12). そして判 定結果を5MPの高精細モニター (MS-S500BG0JA JVCKENWOOD) 上に表示した. ソフトウエアの表 示図をFig.1に示す. 大きな数字が乳房構成解析の結 果であり、「脂肪性乳房」「乳腺散在乳房」「不均一高濃 度乳房」「極めて高濃度乳房」の4段階を1~4と割り 当てている. さらに左右乳房のそれぞれの結果も右下 段の小さな数字で表記される. また左右乳房の乳腺含 有率の高さはバー表示されており、上方ほど高い乳腺 含有率を示す.

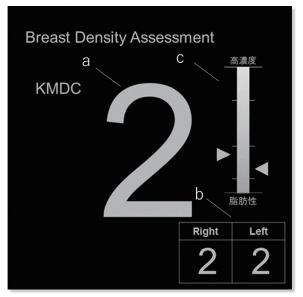


Fig.1 An example of software display

- a: Breast composition classification results are assigned as 1-4.
  - 1. Almost entirely fatty (fatty)
  - 2. Scattered areas of fibroglandular density (scattered)
  - 3. Heterogeneously dense, which may obscure small masses (heterogeneous dense)
  - 4. Extremely dense, which lowers the sensitivity of mammography (extremely dense)
- b: Displays the results of breast composition classification for each of the left and right breast.
- c: Display of the degree of density for the left and right breast as a bar.

#### 2-2. 観察者

マンモグラフィーの撮影経験が3年以上あり,総撮影枚数が10,000枚以上の診療放射線技師11人(男2人,女9人)とした。平均経験年数は12.8 ± 4.7年であり、平均撮影枚数は26,118枚であった。観察者実験を行うに当たり、新潟医療福祉大学の倫理審査委員会で承認を得た(18733-210903).

#### 2-3. 評価画像

使用した臨床画像は、2020年1月から2020年8月までに医療法人社団恒聖会 大塚ブレストケアクリニックにおいて撮影されたマンモグラフィー画像、キヤノンメディカルシステムズ社製の装置で匿名加工し収集した。日本人の乳房構成分類の割合は「脂肪性」は5%、「乳腺散在」は58%、「不均一高濃度」は35%、「極めて高濃度」は2%との報告がある<sup>13)</sup>. その割合に沿うように選択した150症例の画像を2種のデータセットとして使用した。データセットAは「脂肪性」が5症例、「乳腺散在」が56症例、「不均一高濃度」が36症例、「極めて高濃度」が3症例の100症例とし、

データセットBは「脂肪性」が3症例,「乳腺散在」が27症例,「不均一高濃度」が18症例,「極めて高濃度」が2症例の50症例とした.

#### 2-4. 乳房構成分類のtruthの作成

乳房構成分類は読影者間にバラツキがあるため、評 価画像の正解は、精中機構においても読影・画像評価 に関わり、マンモグラフィーに習熟した乳腺専門医1 人に回答を求めた. 回答方法は通常の読影と同様に主 観的に評価し、画像表示後に口頭で回答したものを記 録した. 実験者側で画像の表示と記録を行った. 評価 内容は乳腺含有率(%)と乳房構成分類(1脂肪性, 2乳腺散在, 3不均一高濃度, 4極めて高濃度) とし た. それによりデータセットAおよびBの150症例の 全ての画像の正解を決定した. 以下,この正解をtruth とする. またtruthとソフトウエアの一致率は「脂肪 性」9%(2/23症例),「乳腺散在」79%(142/180症 例),「不均一高濃度」73% (102/140症例),「極めて 高濃度」47%(8/17症例)であった. また乳腺散在 と不均一高濃度の分岐ラインで乳房構成を2分類(1, 2非高濃度乳房と3,4高濃度乳房)とした非高濃度と 高濃度では、71%(144/203症例)と70%(110/157 症例)であり、全体では71%(254/360症例)の一致 率であった.

#### 2-5. 評価実験とトレーニング

画像診断に対してシステムの有用性を評価する場合に、独立観察者実験(independent test)と連続観察者実験(sequential test)がある。本研究では、独立観察者実験と連続観察者実験を両方実施し有用性を評価した。評価としては、乳腺含有率と乳房構成分類のtruthと同様に主観的評価とし、観察者に回答を求めた。

独立観察者実験は、システムを使用しない場合と使用した場合の観察者実験を2回行うこととなる。この場合の注意点としては、学習効果のバイアスを避けるため2回の実験の間隔や順番を考慮する必要がある。連続観察者実験は、システムを使用せずに読影を行い、直後にシステムを使用し再度読影を行う。2つの実験の共通した注意点としては、判断基準が変化しないように実験前に読影のトレーニングを行う必要がある<sup>14</sup>。今回は、実験前にソフトウエアの表示方法や特性を説明し、実験に使用していない計30症例(「脂肪性」4症例、「乳腺散在」11症例、「不均一高濃度」11症例、「極めて高濃度」4症例)を用いたトレーニング

を行った. ここでは,正しく4分類が行われている18 症例と専門医のtruthと一致しない12症例を含め,観察者に乳房構成解析結果の性能とソフトウエアの表示に慣れてもらうための十分なトレーニングとした.

#### 2-5-1. 独立観察者実験

独立観察者実験として実験1,3の実験を行った.実験1ではソフトウエアの結果は表示せず、観察者に評価の回答を求めた.実験3ではソフトウエアの結果を画像と同時に提示し、評価の回答を求めた. どちらもデータセットAの画像を提示し評価を行ったが、実験3ではデータセットAの順番は変動した. また実験1,3の間には実験2を行い、さらに1時間程度の時間を空けることで学習効果の影響がないよう考慮した. また実験1,3のそれぞれの評価に要した時間を計測した.

#### 2-5-2. 連続観察者実験

連続観察者実験として実験2を行った。実験2ではデータセットBの画像を提示し評価の回答を求めた。ソフトウエアの結果は表示せず回答を求め、次いで評価直後にビューワー上にソフトウエアの結果を提示し、再度評価の回答を求めた。実験1~3の実験概要図をFig.2に示す。

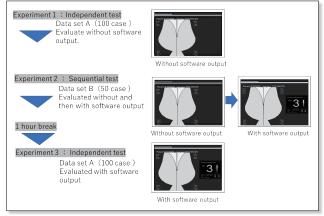


Fig.2 Schematic diagram of experiments 1-3.

- Experiment 1 (Independent test): Evaluation without software output as an independent observer experiment (using data set A).
- Experiment 2 (Sequential test): Evaluation without software output as a continuous observer experiment, with results immediately after (data set B).
- Experiment 3 (Independent test): Evaluation with software output as an independent observer experiment (using data set A with random order).

#### 2-6. 検討方法

以下,5つの検討を行った.1)独立観察者実験(実験1,3)と連続観察者実験(実験2)における,乳腺含有率(%)のtruthと観察者の回答との差を比較するため,乳腺含有率(%)の観察者の回答とtruthとの差分を,t検定により有意差検定を行った.2)観察者間のバラツキを評価するために,症例ごとの観察者の乳腺含有率(%)の回答の標準偏差を,t検定により有意差検定を行った.3)各乳房構成分類において,truthと観察者の回答の一致率を算出した.4)最も重要となる乳腺散在と不均一高濃度の分岐ラインで乳房構成を2分類(1,2非高濃度乳房と3,4高濃度乳房)とし、そのラインを越えた変更の症例数を検討した.5)独立観察者実験(実験1,3)においては,評価時間を計測しその変化を検討した.1)3)4)の検討は,観察者11人のデータをまとめて行った.

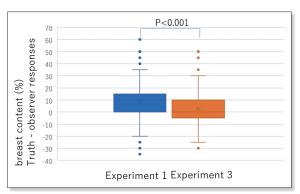


Fig.3 Difference in breast content (%) between truth and observer's response in independent test (Experiments 1 and 3)

Observer's response were closer to the truth in experiment 3 than in experiment 1, with statistical significance. (p < 0.001)

Table 1 Agreement with truth in observer's response of independent test (Experiments 1 and 3)

xperiment	1	Agree	ment rate	: 66.7%	Experimen	nt 3	Agree	ment rate	. 17.170
Observer Truth	1	2	3	4	Observer Truth	1	2	3	
1	47	8			1	27	28		
2	58	528	30		2	1	567	48	
3		237	148	11	3		137	246	1
4		3	19	11	4		1	24	

The number of agreed cases was higher in experiment 3 than in experiment 1, the agreement rate was improved by approximately 11.0%.

# 3. 結果

#### 3-1. 独立観察者実験

独立観察者実験(実験1,3)における乳腺含有率の truthと観察者の回答との差(Fig.3)および観察者の 回答の標準偏差 (Fig.4)において、統計的有意差が認 められた(p<0.001)。また各乳房構成分類のtruth との一致率は66.7%(実験1)から77.1%(実験3)と なり、およそ11.0%の向上が見られた(Table 1).

一方,乳房構成分類の乳腺散在と不均一高濃度の分岐ラインを越えた変更の症例数の結果をTable 2に示す。実験3において非高濃度から高濃度への変更が,146症例のうちtruthと一致したものは110症例であり、割合としては75.3%(110/146症例)であった。逆に高濃度から非高濃度への変更が、26症例のうちtruthと一致したものは12症例であり、割合としては

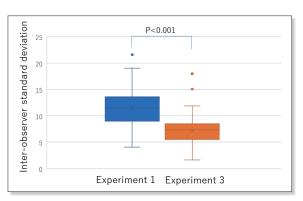


Fig.4 Inter-observer variation in breast content (%) in independent test (Experiments 1 and 3)

Inter-observer standard deviation was reduced in experiment 3 compared to experiment 1, with statistical significance. (p < 0.001)

Table 2 Evaluation between observer and truth in independent test (Experiments 1 and 3) (Total number of cases in agreement between observer and truth/Total number of cases evaluated by observer)

Experiment 3 Experiment 1	non-dense breast (1or2)	dense breast (3or4)
non-dense breast	582/735	110/146
(1or2)	(79.2%)	(75.3%)
dense breast	12/26	144/193
(3or4)	(46.2%)	(74.6%)

The agreement with truth in judging dense breast cases improved in 75.3% (110/146), confirming that the use of software improved the agreement between the specialist and the observer.

46.2% (12/26症例) であった.

評価所要時間の変化の結果を Table 3 に示す. 11人中10人において時間は短縮されており, 100例の分類に要した平均時間は300秒 (5分)の短縮 (1症例当たり3秒の短縮) につながった.

#### 3-2. 連続観察者実験

連続観察者実験(実験2)における乳腺含有率のtruthとの差では統計的有意差は得られず(p=0.063)(Fig.5), 観察者間のバラツキは減少した(p<0.001)(Fig.6) 乳房構成分類のtruthとの一致率の向上は見られなかった (Table 4).

一方,乳房構成分類の乳腺散在と不均一高濃度の分岐ラインを越えた変更の症例数の結果をTable 5に示す。ソフトウエアを用いた場合に非高濃度から高濃度への変更が20症例であり、そのうちtruthと一致した

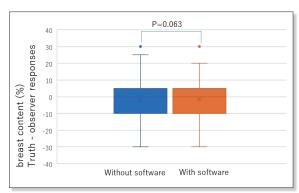


Fig.5 Difference in breast content (%) between truth and observer with and without software in the sequential test (Experiment 2)

There was no statistically significant difference. (p=0.063)

Table 4 Agreement with truth in breast contents of sequential test (Experiment 2)

Without so	ttware	Ag	reement ra	te: 76.2%	With so	ftware	Agre	eement rate	: 76.2%
Observer Truth	1	2	3	4	Observer Truth	1	2	3	4
1	23	10			1	20	13		
2		267	30		2		269	28	
3		74	120	4	3		67	122	9
4			13	9	1			14	8

The agreement rate was 76.2% with and without software, but the number of matched cases increased in categories 2 (scattered) and 3 (heterogeneous dense).

Table 3 Difference in evaluation time (seconds) in independent test (Experiments 1 and 3) per observer.

Time required (second)						
	Experiment 1	Experiment 3	shortening time			
Observer01	1,597	1,035	562			
Observer02	1,370	805	565			
Observer03	1,311	1,055	256			
Observer04	1,103	765	338			
Observer05	1,409	1,048	361			
Observer06	1,552	785	767			
Observer07	962	707	255			
Observer08	1,033	1,083	-50			
Observer09	560	483	77			
Observer10	675	610	65			
Observer11	1,260	1,158	102			
Average	1,167	867	300			

Evaluation time was reduced in 10 of 11 observers, with an average reduction of 300 seconds. (3 seconds/case)

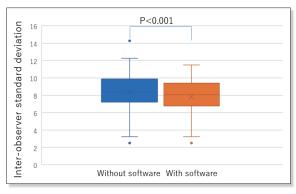


Fig.6 Inter-observer variation in breast content (%) with and without software in sequential test (Experiment 2)

Inter-observer variability was reduced with use of the breast composition analysis software, with statistical significance. (p < 0.001)

Table 5 Evaluation between observer and truth in sequential test (Experiment 2)
(Total number of cases in agreement between observer and truth/Total number of cases evaluated by observer)

With software Without software	non-dense breast (1or2)	dense breast (3or4)
non-dense breast	230/354	13/20
(1or2)	(65.0%)	(65.0%)
dense breast	9/15	117/161
(3or4)	(60.0%)	(72.7%)

The agreement with truth improved in 65.0% (13/20) and 60.0% (9/15) of both dense breast and non-dense breast cases, with use of the software.

ものは13症例であり、割合としては65.0%(13/20症例)であった. 逆に高濃度から非高濃度への変更が15症例であり、そのうちtruthと一致したものは9症例であり、割合としては60.0%(9/15症例)であった.

## 4. 考察

乳がんの早期発見・早期治療には検診が欠かせない. スクリーニング検査での乳がんの見落としは高濃度乳 房が大きな原因となっている. 乳房構成分類の問題は 主観的であり、バラツキが大きい点である $^{7}$ )。また4段 階の分岐ラインの症例の場合は迷うことが多いといわ れている.独立観察者実験(実験1,3)の結果より. ソフトウエアを使用した場合にtruthの結果との一致 率は66.7% (実験1) から77.1% (実験3) となり,お よそ11.0%の向上が見られ観察者間のバラツキも減 少した. この実験では、実験1でソフトウエアを見ず に評価を行い、実験3で画像とソフトウエアを同時に 提示して評価を行うものであった. 一方, 連続観察者 実験(実験2)の結果より、ソフトウエアを使用した 場合にtruthの結果と観察者間のバラツキは減少した が一致率は変化がなかった. この実験はソフトウエア を見ずに評価を行い、その直後にソフトウエアも提示 して再評価を行うものであった. 自分で評価結果を判 断した直後では、ソフトウエアの結果を見ても修正し にくいといった心理的な要因が作用したのではないか と考えられる. ソフトウエアの使用方法は限定されて いないが、これらの結果を踏まえると、ソフトウエア の表示は同時に提示する方が乳腺専門医による結果に 近づき、観察者間のバラツキも減少し良好な結果とな ることが示唆される.

問題となっている非高濃度と高濃度分岐ラインのバラツキであるが、独立観察者実験(実験1,3)と連続観察者実験(実験2)の乳房構成分類の結果(Table 1, Table 4)を詳細に見ると、乳房構成分類の乳腺散在と不均一高濃度においては、ソフトウエアを使用した方がtruthと一致した症例数の増加が見られた。独立観察者実験(実験1,3)では合計137症例の増加が見られ、連続観察者実験(実験2)では4症例の増加が見られた。この乳房構成分類の乳腺散在と不均一高濃度は、非高濃度と高濃度分岐ラインをまたぐ重要な部分であり、そこでの一致した症例数の増加は、ソフトウエアを使用した場合に分岐ラインが正しく分類できたと考えられる。またさらに非高濃度と高濃度で乳房構成分類を2分した場合に、ソフトウエアを使用した前

後でその分岐ラインを越えた変更をした症例では、変 更後にtruthと一致したかどうかも重要である。独立 観察者実験(実験1,3)では、非高濃度から高濃度へ の変更によって一致率が75.3%(110/146症例)であ り、高濃度から非高濃度への変更後の一致率は46.2% (12/26症例)であった。連続観察者実験(実験2)では、非高濃度から高濃度への変更後の一致率は65.0% (13/20症例)であり、高濃度から非高濃度への変更 後の一致率は60.0%(9/15症例)であった。重要なことは高濃度の検出率を上げることである。そのため 非高濃度から高濃度への変更によって一致率が両実験 で65%を超えたことは、観察者による高濃度の分類が ソフトウエアを使用した場合に的確になされたといえ る。このことは非常に重要であり、ハイリスクの症例 を逃さないことにつながることが示唆された.

本研究に使用したソフトウエアの乳房構成分類は. 2次元のソフトウエアであるため日本のガイドライン に沿った分類である. 特に検診においては定量的評価 が求められる. 現在, 検診においては, 乳房構成を通 知し乳がんのリスクについて理解を進める動きが米国 ではすでに始まっている. 日本では高濃度乳房の対応 が確立されておらず、一律に通知することは時期尚早 であるとされている3). 乳房構成に関する判定基準の 曖昧さも要因の一つとされているが、マンモグラムの 読影は2人の医師による二重読影が行われており、乳 房構成解析ソフトウエアの利用としては、一次読影と 二次読影の結果が割れた場合の参考とできる。先述し たように、乳房構成解析ソフトウエアを使用して非高 濃度と高濃度の分類が的確になることで、高濃度乳房 の通知を一層後押しできるものと示唆される。また時 系列的にも同一被験者の過去画像の結果と見比べ、乳 房変化の推移を定量的に評価でき、長期的な観点から も有益であると考える. さらに検診のマンモグラムは 膨大な画像の読影量となるが、今回の結果で評価所要 時間の短縮につながり、医師の負担の軽減ができるこ とも示唆された.

本研究の限界として、独立観察者実験では、実験1の症例を忘れるためのウォッシュアウト期間の長さが懸念される。これに関しては独立観察者実験(実験1、3)の間に連続観察者実験(実験2)を入れ、さらに1時間程度の時間を設けた。さらに学習効果のバイアスが避けられていたか確認のためのアンケートを行った。観察者全員より、実験1と3の症例が同一とは気付かなかった、もしくは特徴的な何症例かは印象に残るため同一症例と気付いたが、乳房構成分類の評価上では影

響がなかったと回答があった。この結果により、学習効果のバイアスは十分に避けられていたと考えられるが、作業の慣れによる効率化が時間の短縮につながった可能性も考えなければならない。また2種類のデータセットの症例は、日本人の乳房構成分類の割合に合わせてはあるが、やはり偏りが大きく今後の検討が必要であると考える。

また高濃度乳房であれば、超音波検査やMRI検査を勧めることが重要であるといわれている <sup>15,16)</sup>. 特に超音波検査は非侵襲で簡便であり、超音波検査スクリーニングは良性と悪性の乳房病変を区別するための補足スクリーニングツールとして重要である <sup>17)</sup>. 本研究は診療放射線技師の評価で実験を行った. 今後、乳房構成解析ソフトウエアの結果が操作卓のコンソール上で見られるようになれば、さらに有意義な利用法につながると期待される. 診療放射線技師が読影の補助を行う際にソフトウエアを用いた乳房構成分類を行っておくと、医師の判断を早めることができると考えられ、読影補助時に乳房構成解析ソフトウエアが使えると有用性が増すと考える. そうすることで、診療放射線技師と乳房構成解析ソフトウエアで医師の業務の負担軽減につなげられる可能性が示唆された.

## 5. まとめ

本研究では、ソフトウエアの有用性を評価するために、11人の診療放射線技師による独立観察者実験と連続観察者実験を行った。乳腺組織の割合と乳房構成分類を評価したが、特に独立観察者実験において truth との一致率が向上し、観察者間のバラツキが減少した。 さらに評価時間の短縮にもつながっていた。乳房構成の分類で最も重要である非高濃度と高濃度の分岐ラインでは、ソフトウエアを使用することで分類が的確に行えた。乳房構成解析ソフトウエアの使用は、診断精度と読影効率の向上に寄与することが示唆された。

### 謝辞

本研究を行うに当たり、乳房構成の分類に関してご 指導いただきました名古屋医療センターの森田孝子先 生、ならびに評価に際し多大なるご協力を頂きました 新潟県労働衛生医学協会の山崎薫先生に深く感謝申し 上げます。またデータ取得に関し尽力いただきました コニカミノルタ株式会社の二村仁さまに深謝致しま す。

# 利益相反

本研究はコニカミノルタ株式会社との共同研究により実施した.

#### 表の説明

- Table 1 独立観察者実験 (実験1,3) における各乳房構成分 類におけるtruthと観察者の回答の一致率
  - 実験1よりも3の方が乳房構成分類が一致した症例数は多く一致率は実験1では66.7%で、実験3では77.1%でありおよそ11.0%の向上が見られた。
- Table 2 独立観察者実験(実験1,3)における観察者とtruth との評価(観察者とtruthとの一致症例総数/観察者 の評価症例総数)
  - 高濃度症例の判定においてtruthとの一致率が向上した症例は75.3% (110/146) あり、ソフトウエアの使用により、専門医と観察者の一致率の向上が確認された
- Table 3 独立観察者実験 (実験1,3) における観察者ごとの 評価時間の変化
  - 実験1の評価時間と実験3の評価時間の差分を取った結果、11人中10人の評価時間は短縮されており、平均で300秒の短縮となった。1症例当たり平均3秒の短縮となった。
- Table 4 連続観察者実験(実験2)のソフトウエアを見る前後におけるtruthと観察者の乳房構成分類との一致率 一致率はソフトウエアを見る前後で変化がなく 76.2%であったが、乳房構成分類の乳腺散在と不
- 均一高濃度において一致した症例数は増加した。 Table 5 連続観察者実験(実験2)における観察者とtruthとの評価(観察者とtruthとの一致症例総数/観察者の
  - 評価症例総数) 高濃度症例も非高濃度症例もtruthとの一致率が 向上した症例は65.0% (13/20) と60.0% (9/15) であり、ソフトウエアの使用により、専門医と観察 者の一致率の向上が確認された。

#### 図の説明

- Fig.1 ソフトウエアの乳房構成分類結果の表示イメージ
  - a: 乳房構成分類結果「脂肪性」「乳腺散在」「不均 一高濃度」「極めて高濃度」を1~4に割り当てて いる
  - b:左右乳房それぞれの乳房構成分類結果を表示
  - c: 左右乳房それぞれの乳腺含有率をバーで表示
- Fig.2 実験1~3の概要図
  - 実験1 (独立観察者実験): ソフトウエアの結果を表示なしで評価 (データセットA)
  - 実験2(連続観察者実験):ソフトウエアの結果を 表示なしで評価,直後に結果を表示し再度評価 (データセットB)
  - 実験3 (独立観察者実験): ソフトウエアの結果を表示ありで評価 (データセットA 順変動)
- Fig.3 独立観察者実験 (実験1,3) における乳腺含有率 (%) のtruthと観察者の回答との差
  - 実験1よりも実験3において観察者の回答はtruthに 近づき、統計的有意差が認められた(p<0.001).
- Fig.4 独立観察者実験 (実験1,3) における乳腺含有率 (%) の観察者の回答の標準偏差
  - 実験1よりも実験3において観察者間のバラツキは減少し、統計的有意差が認められた(p<0.001).

- Fig.5 連続観察者実験(実験2)におけるソフトウエアを見る前後での乳腺含有率(%)のtruthと観察者の回答との差
  - ソフトウエアを見る前と後では観察者の結果に変化はなく、統計的有意差は認められなかった (p=0.063).
- Fig.6 連続観察者実験(実験2)における乳腺含有率(%) の観察者間のバラツキ
  - ソフトウエアを見る前と後では観察者間のバラツキは減少し、統計的有意差が認められた (p<0.001).

## 参考文献

- 1) 厚生労働省:2020年人口動態統計の概況
- American College of Radiology: ACR BI-RADS®-Atlas der Mammadiagnostik: Richtlinien zu Befundung, Handlungsempfehlungen und Monitoring. Springer Heidelberg/Berlin, Germany, 2016.
- 3) 特定非営利活動法人日本乳がん検診精度管理中央機構 https://www.qabcs.or.jp/news/20200206.html
- Norman F Boyd, et al.: Mammographic Density and the Risk and Detection of Breast Cancer. n engl j med, 356, 227-236, 2007.
- Nagata, et al.: Mammographic density and the risk of breast cancer in Japanese women. Br J Cancer, 92, 2102-2106, 2005.
- Kotsuma, et al.: Quantitative assessment of mammographic density and breast cancer risk for Japanese women. Breast, 17, 27-35, 2008.
- 7) 東野英利子, 他:マンモグラフィの乳房構成と乳癌の検出 感度―複数の読影者による検討―. 日乳癌検診学会誌, 24 (1), 113-122, 2015.
- 8) Redondo, et al.: Inter- and intraradiologist variability in the BI-RADS assessment and breast density categories for screening mammograms. The British

- Journal of Radiology, 85, 1465-1470, 2012.
- Yoshida, et al.: Optimal Breast Density Characterization Using a Three-Dimensional Automated Breast Densitometry System. Current Oncology, 28, 5384-5394, 2021.
- 10) Gweon H M, et al.: Radiologist Assessment of Breast Density by BI-RADS Categories Versus Fully Automated Volumetric Assessment. AJR Am. J. Roentgenol, 201, 692-697, 2013.
- 11) Youk J H, et al.: Automated volumetric breast density measurements in the era of the BI-RADS Fifth Edition: A Comparison with Visual Assessment. AJR Am. J. Roentgenol, 206, 1056-1062, 2016.
- 12) 南 寛威, 他: デジタルマンモグラフィー画像を用いた乳房 構成解析ソフトウエア (Bda) の紹介. KONICA MINOLTA TECHNOLOGY REPORT, VOL.19, 78-82, 2022.
- 13) 厚生労働行政推進調査事業費補助金厚生労働科学特別研究事業「乳がん検診における乳房の構成(高濃度乳房を含む)の適切な情報提供に資する研究」班:高濃度乳房について、2017.
  - https://www.mhlw.go.jp/file/05-Shingikai-10901000-Kenkoukyoku-Soumuka/0000208392.pdf
- 14) Shigehiko Katsuragawa: ROC analysis and Evaluation of CAD system. Medical Imaging Technology, Vol.20, No.1, 29-35, 2002.
- 15) Zoey Zy Ang, et al: Classification of normal screening mammograms is strongly influenced by perceived mammographic breast density. J Med Imaging Radiat Oncol, 61(4), 461-469, 2017.
- 16) Nataliia Moshina, et al.: Positive predictive values by mammographic density and screening mode in the Norwegian Breast Cancer Screening Program. Eur J Radiol, 85(1), 248-254, 2016.
- 17) Wei in Li, et al.: Mammography breast density: an effective supplemental modality for the precise grading of ultrasound BI-RADS 4 categories. Gland Surgery, 10(6), 2010-2018, 2021.