

Quantra™ 2.2 ソフトウェア 開発コンセプトおよび臨床的性能

Ashwini Kshirsagar, Ph.D, Chief Scientist, Clinical Solutions, Research and Development, Hologic, Inc.

はじめに

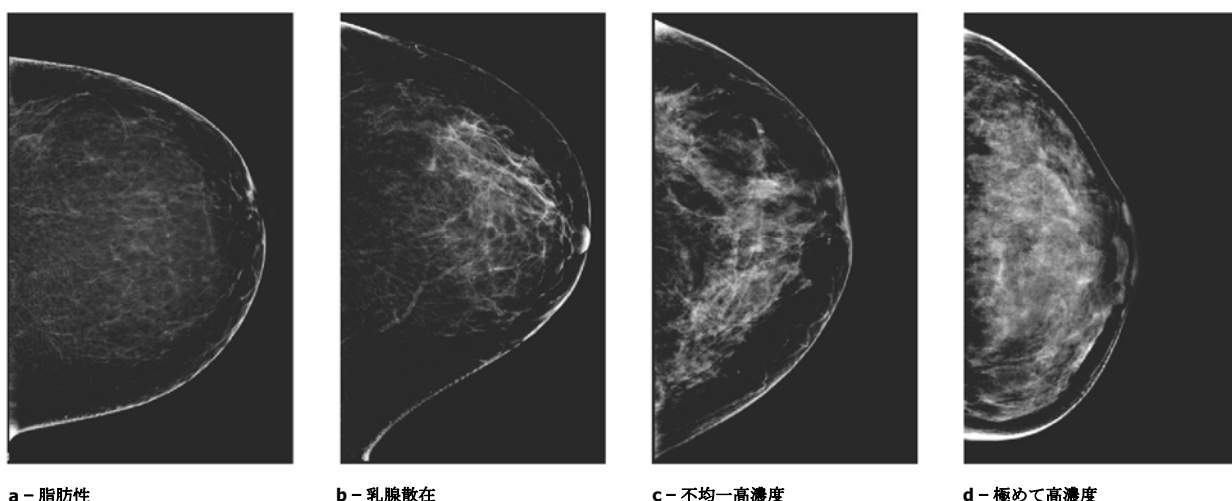
本稿の目的は、Hologic の乳腺密度評価製品 Quantra™ 2.2 ソフトウェアの最新バージョンを紹介することである。本稿では、乳腺密度の自動カテゴリ分類がなぜ有益であるのかを考察し、米国放射線学会(ACR)の BI-RADS Atlas 乳房構成カテゴリ分類版による説明内容の変更をレビューするとともに、最新バージョンの Quantra™ ソフトウェアと比較しながら、Quantra™ 2.2 ソフトウェアのアルゴリズム設計の背景にある論理的根拠について考察する。また、FDA 承認に向けて Quantra™ 2.2 ソフトウェアの臨床的性能を評価する試験の結果も報告する。

BI-RADS 乳房構成カテゴリ

米国放射線学会(ACR)は、読影医によるマンモグラフィ読影レポート（読影結果）の標準化を促進するために BI-RADS Atlas を発行している。画像の目視による乳房構成カテゴリの評価は、マンモグラフィの結果を伝える上で不可欠な要素であり、マンモグラフィの病変検出感度を伝える手段である。図 1 に、4 つのカテゴリに該当するマンモグラフィ画像を例示する。

自動化された乳腺密度カテゴリ評価はなぜ有益なのか

乳腺密度カテゴリの視覚的評価には、読影者自身¹および読影者間^{2,3}で大きなばらつきがあることが十分に裏付けられている。経験や訓練の程度が BI-RADS のカテゴリ分類結果に影響を及ぼすことも明らかになっている⁴。さらに、密度カテゴリの結果に基づき追加検診の決定が下されることで、意図しないバイアスが生じる可能性があることも否定できない。また、一部の州には、患者に乳腺密度や、マンモグラフィの感度が低下する可能性、癌のリスクが高い可能性を伝えることを求める法律があり、これも密度カテゴリの結果に意図しないバイアスを引き起こす役割を果たすことがある。このため、自動化された方法は、乳腺密度カテゴリの評価、報告に対して、バイアスを一切引き起こすことのない再現可能な選択肢となりうる。Hologic Quantra™ ソフトウェアはこの自動化を実現し、既存のマンモグラフィワークフローに組み込むことができる。



a - 脂肪性

b - 乳腺散在

c - 不均一高濃度

d - 極めて高濃度

図 1. ACR の BI-RADS Atlas に記載されている乳房構成カテゴリ

ACR の BI-RADS 密度カテゴリー分類ガイドラインの修正

ACR は、BI-RADS Atlas 第 5 版の一部として、乳房構成カテゴリー報告のガイドラインの修正を発表した⁵。修正はあったものの、カテゴリー分類の全体的な報告目的に変わりはなかった。すなわち、乳房内の乳腺組織含有量を目視により推定し、4 つのカテゴリーに分類することである。BI-RADS 第 4 版では、乳腺の高濃度組織の割合（%）を 25% 毎の各カテゴリーに分類した。BI-RADS Atlas 第 5 版には、この割合の範囲は記載されていない。図 2 に、ACR が提案する第 4 版と第 5 版の変更点を示す。

BI-RADS のカテゴリー	説明
a	乳房のほぼ全体が脂肪である。
b	乳腺が散在している。
c	乳腺組織が不均一に高濃度を呈しているため、小さな腫瘍を発見しにくい。
d	乳腺組織が極めて高い濃度を呈している。このためマンモグラフィの感度が低くなる可能性がある。

この修正は、乳腺密度について、マンモグラフィにおいて軟組織病変の描出が高濃度組織で隠されてしまうこと（マスキング作用）を説明した本文の記述を際立たせるために行われた。BI-RADS 委員会によれば、主観的に推定した乳腺密度と不明瞭な病変による感度の変化との関係は、乳腺密度の割合の比較的小さな作用よりも臨床的に重要である。同委員会はまた、体積に基づく乳腺密度データと妥当性が確認された各割合のカットポイントの推定が再現可能であることを裏付ける十分なエビデンスが文献によって発表されるまで、BI-RADS 密度カテゴリー分類に乳腺密度の割合の範囲を含めないとしている。

高濃度組織パターンとそのマスキング作用の重要性

マンモグラフィの感度および乳がん発症の内在リスクが乳腺密度と関係していることを示す十分なエビデンスがある⁶⁻⁷。しかし、リスクおよびマスキングが密度に左右されることは、単に乳腺組織量の場合よりもはるかに複雑である。高濃度組織の割合が同じ乳房であっても、高濃度組織にぴったりと重なった領域がわずかでもあると、均等に分布した高濃度組織よりも不明瞭になりやすくなる。図 3 に、「ウォーリーをさがせ！」に似た図でこの概念を説明する。目的は、群集の中から柄シャツを着ている人物を見つけることである。

図 2. BI-RADS Atlas 第 5 版で第 4 版から修正された乳房構成カテゴリー。BI-RADS 第 5 版には高濃度組織の割合の範囲は記載されていない。



a. 大きな空間に分布している群集の方が、柄シャツを着ている人物を見つけやすい。

図 3 「ウォーリーをさがせ！」に似た図。対象を探す作業では対象物の分布およびパターンによる作用の方が密度自体よりも重要な役割を果たすことを示している。

(a) と (b) の人数は同じであるが、群集が一部に集中している図 (b) の方が、広い空間に散らばっている図 (a) よりも、柄シャツを着ている人物を見つけるのが難しい。

乳腺組織のパターンおよび構造が、マンモグラフィによる癌リスクの予想に同程度に重要な役割を果たすことを示したエビデンスが相次いで報告されている⁸⁻¹⁰。この事実は ACR の BI- RADS Atlas で提案された修正に反映されている。図 4 に、新しい BI-RADS ガイドラインのもとで、乳房構成カテゴリー分類がいかに関与を受けるのかを図示する。

図 4 は、BI-RADS Atlas 第 4 版の記載に従い、乳腺密度が 2 番目に少ない 25~50% に相当する 2 つの乳房の図である。このガイダンスでは両方とも「b」に分類された。しかし、上の画像では密度分布が均一であるため、小さな病変のマスキング率は低くなる。それに対して下の画像では右乳房の外側上部に集中している乳腺組織は、小さな病変が不明瞭になる可能性がある十分な密度であるため、BI-RADS Atlas 第 5 版では「c」に分類される。

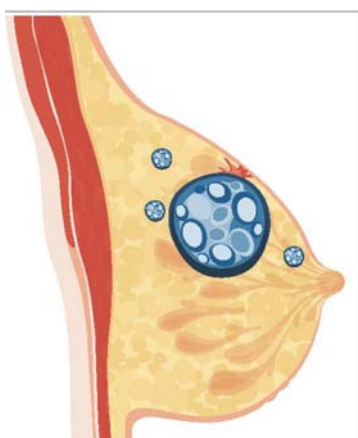
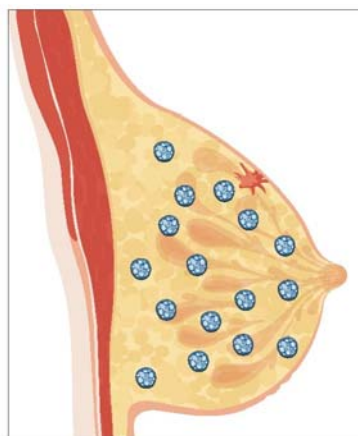


図 4. BI-RADS Atlas 第 4 版の記載に従い、乳腺密度が 2 番目に少ない 25~50% に相当する 2 つの乳房を示した図。



b. 同じ群集でも限られた空間に集中している場合の方が、柄シャツを着ている人物を発見するのは難しい。

Quantra™ 2.2 ソフトウェアの設計目的

前バージョンの Quantra™ ソフトウェアは、高濃度組織の割合を考慮した BI-RADS Atlas 第 4 版に基づいて設計されていた。この方法ではまず、マンモグラフィ画像取得パラメータを用いた脂肪組織および高濃度組織に対する X 線減衰モデルを使用して、乳腺体積密度の割合 (%) を推定した。次に、妥当性が確認された推定体積密度の閾値のみを用いて、乳房を 4 つのカテゴリに分類した。このソフトウェアでは、このような閾値や各体積密度の数値、各面積密度、高濃度組織の体積測定に基づき算出した他のパラメータも表示された。この方法では乳房体積に占める高濃度組織の量を推定することができるが、実質組織分布パターンによって決まる局所的な密集状態により生じるマスキング作用を特に考慮していないのは明らかである。

Hologic® は、Quantra™ ソフトウェアの向上に常に努めている。当社では、本製品の初期バージョンを使用している顧客層から得たフィードバックを解析し、乳房組織構成を推定する画像処理アルゴリズムの強化を図っている。BI-RADS Atlas 第 5 版に記載されている乳房構成カテゴリの修正を受けて、Hologic は、乳腺体積密度単体よりも、乳腺組織の分布パターンと構造の理解に重点を置く方法に注力した。Quantra™ 2.2 で判定された乳房構成カテゴリでは、ACR による最新のガイダンスに従い、乳腺組織分布のマスキング作用が考慮される。その結果、現在販売されている Quantra™ 製品に含まれていた体積密度と関連パラメータは、Quantra™ 2.2 ソフトウェアの出力では表示されなくなる。Quantra™ 2.2 ソフトウェアの表示からの乳腺密度に関係する数値パラメータの削除は、BI-RADS 第 5 版に準拠した製品に対して FDA から提示された要件であった。

Quantra™ 2.2 ソフトウェア: 機械学習に基づく解析

機械学習に基づくアルゴリズムは、カテゴリが既知の観察結果が含まれたトレーニング用データセットをベースとした知識に基づき、新たな知見のカテゴリの特定に伴う分類の問題に適切に対処することが知られている。分類の問題とは、実質のパターンおよび構造を用いた乳腺密度のカテゴリ分類である。Hologic では、各種製品の改良へと向けた持続的な取り組みの一環として、マンモグラフィ検査による画像と放射線科医による報告書を日常的に収集している。これによって、何千枚ものマンモグラフィ画像と放射線科医が判定した BI-RADS 密度カテゴリから成るトレーニングセットは、BI-RADS ライクなカテゴリに自動分類させるための機会学習アルゴリズムのトレーニングデータセットとしてすぐに利用できた。Quantra™ 2.2 ソフトウェアは、Hologic が、人工知能と機械学習技術を活用して自動化された乳腺密度カテゴリ分類を実現するため、乳腺密度の BI-RADS カテゴリが既知の膨大な画像データベースを用いた努力の成果である。

Quantra™ 2.2 ソフトウェアは、Hologic の自動乳腺密度測定技術を強化したシリーズとして最新版となる。今回のアルゴリズムでは、マルチクラスサポートベクターマシン (SVM) ¹¹ をベースとして、乳房を乳腺実質組織のパターンおよび構造に基づき 4 つのカテゴリに分類する方法を用いている。検診時に放射線科医が判定した BI-RADS カテゴリがグランドトゥールースとして使用できた 6,000 件超のマンモグラフィ検査を用いて、4 クラスの SVM モデルにトレーニングを行った。トレーニングでは、絶対的濃淡値のみではなく、乳房内でのピクセル値の分布およびパターンを示す固有の統計的特性を重視した。このようなパターンおよび構造に基づく特性を用いてトレーニングを行ったモデルは、体積測定値のみではなく、組織の分布およびマスキング作用に基づいて乳房構成カテゴリを分類しようとする。この技術を用いた Quantra™ 2.2 ソフトウェアのアルゴリズムは、BI-RADS Atlas 第 5 版に示されている新しいガイドラインに沿って乳房構成カテゴリを自動的に推定する。

Quantra™ 2.2 ソフトウェアの臨床的評価

Hologic®はエビデンスの一部として、4 枚撮影する標準的なマンモグラフィ検査による検診 230 件から得た画像を用いて試験を実施した。画像はすべて、Hologicの Selenia® Dimensions®の「コンボ」モードを使用して得た。つまり、同じ圧迫を加えた同一の被検者から、トモシンセシス画像と従来の 2D 画像を得たということである。

グラウンドトゥルースの推定

読影者 5 名が、この 230 件を検討して BI-RADS 乳腺密度カテゴリを個別に判定した。各検診では、トモシンセシス画像と従来の 2D 画像を検討に使用した。読影者には、ACR が発行した BI-RADS Atlas 第 5 版のガイドラインに従うよう指示した。各検診のグラウンドトゥルースとなる乳腺密度カテゴリは、読影者 5 名が判定したカテゴリースコアの中央値を計算して求めた。

Quantra™による密度カテゴリの評価

前バージョンと同様に、Quantra™ 2.2 ソフトウェアでも、コンボモードの使用時に 2 種類の画像を解析に用いることができる場合、ユーザーは従来の 2D 画像を解析するか、トモシンセシス画像を解析するかを選択することができる。画像の一方のみ（従来の 2D またはトモシンセシス HD のいずれかのみ）を使用しているユーザーの場合、Quantra™ソフトウェアで、乳腺密度の結果の推定に利用可能な画像モードが解析される。このため、従来の 2D 画像とトモシンセシス画像のいずれも、読影者 5 名による判定を解析して得られたグラウンドトゥルースとしての BI-RADS カテゴリと個別に比較、解析した。トモシンセシス画像を解析する際、Quantra™ソフトウェアではトモシンセシスデータによる 0 度の位置のプロジェクション画像を使用する。

結果¹²

試験集団の分布

前述したとおり、各検診のグラウンドトゥルースは、読影者 5 名がそれぞれ判定したカテゴリの中央値から求めた。以下の表に、ACR の BI-RADS 第 5 版を用いた読影者によるカテゴリ評価後の検診分布を示す。

密度カテゴリの コンセンサス (BI-RADS 第 5 版)	カテゴリあたりの件数
a	32
b	81
c	85
d	32
合計	230

結果の評価および分割表

BI-RADS 密度カテゴリなどの分類測定値を推定するための 2 種類の方法の間で結果が一致しているかどうかの評価は、一般的に分割表を用いて行う。分割表とは、2 種類の方法で測定したカテゴリの観察結果を記録する二次元マトリックスである。2 種類の方法がいかなる場合も完全に一致していれば、頻度数はすべて分割表の対角線上となり、対角線上でない位置にはすべてゼロが含まれる。しかし、乳腺密度の評価における読影者間の一致は一般的にわずかにすぎず²、セルの対角線上でない位置での頻度は 0 よりも大きい。

読影者間の変動

前述のとおり、放射線科医 5 名がマンモグラフィ検査 230 件の BI-RADS 密度カテゴリ評価を個別に実施し、グラウンドトゥルースを推定した。このデータから、230 件のそれぞれに対する読影者 2 名による 10 組の観察結果を用いて、読影者間の変動を理解することができる。同データから、放射線科医 2 名による BI-RADS カテゴリ評価が一致した検診は全体の 63～86%で、平均 76%であることがわかった。当社の社内試験から得られたこの一致率は、読影者間の一致について文献で報告された結果とほぼ同じである¹³。表 1 に、グラウンドトゥルース推定プロセスに参加した各放射線科医ペアが判定したカテゴリの累積分割表を示す。この結果は、読影者 2 名の間で観察結果が一致したのはわずか 76%であったことを示している。Quantra™ソフトウェアで推定された乳腺密度カテゴリと読影者が判定したグラウンドトゥルースのカテゴリを比較した結果を検討する際は、上記の結果に留意することが重要である。

	読影者 A				
読影者 B		a	b	c	d
	a	246	78	0	0
	b	94	673	84	0
	c	0	62	649	165
	d	0	0	63	186
			一致した観察結果		1754
			総数		2300
			全体的な一致率 (%)		76%

表 1. グラウンドトゥルース推定に参加した各放射線科医ペアが判定した BI-RADS カテゴリーの累積分割表。

放射線科医のグラウンドトゥルースと比較した Quantra™ の評価

Quantra™ソフトウェアの密度カテゴリーと放射線科医が BI-RADS 第 5 版を用いて判定したグラウンドトゥルースカテゴリーとの一致率も、Quantra™ソフトウェアで判定されたカテゴリーと各マンモグラフィ検査について推定された（前述の）グラウンドトゥルースカテゴリーを比較した二次元マトリックスである分割表として示すことができる。

表 2 に、2D 画像を用いて Quantra™ 2.2 ソフトウェアで判定されたカテゴリーと、BI-RADS 第 5 版を用いたグラウンドトゥルースを比較した分割表を示す。

		Quantra™ 2.2 ソフトウェア			
		従来の 2D 画像			
「グラウンド トゥルー ス」BIRADS 第 5 版		a	b	c	d
	a	21	11	0	0
	b	9	64	8	0
	c	0	1	61	23
	d	0	0	1	31
			観察結果が 一致した検診	177	
			総数	230	
			全体的な一致率 (%)	77%	

表 2. 従来の 2D 画像を用いた Quantra™ 2.2 のカテゴリーと、BI-RADS 第 5 版に基づく放射線科医によるグラウンドトゥルースカテゴリーとの比較。

表 3 に、トモシンセシスの 0 度の位置のプロジェクション画像を使用して Quantra™ 2.2 ソフトウェアで判定されたカテゴリーと、BI-RADS 第 5 版を用いたグラウンドトゥルースを比較した頻度分布を示す。

		Quantra™ 2.2 ソフトウェアによる トモシンセシス画像			
「グラウン ドトゥルー ス」 BI-RADS 第 5 版		a	b	c	d
	a	20	12	0	0
	b	11	61	9	0
	c	0	7	53	25
	d	0	0	1	31
			観察結果が 一致した検診		165
			総数		230
			全体的な一致率 (%)		72%

表 3. トモシンセシス画像を用いた Quantra™ 2.2 のカテゴリーと、BI-RADS 第 5 版に基づく放射線科医によるグラウンドトゥルースのカテゴリーとの比較。

表 1 および表 2 にまとめた結果に基づくと、Quantra™ 2.2 ソフトウェアと放射線科医が判定したグラウンドトゥルースが一致する全体的な割合は、典型的な読影者間の変動の範囲内にあるといえる。トモシンセシス画像の全体的な一致率は D 画像よりも低いが、この変動は、Quantra™ソフトウェアのトモシンセシス画像にシステム上の欠点があるわけではなく、単にソースデータの差によるものである。読影者間の一致率に 63～86%と幅があることを考慮すれば、この試験で従来の 2D 画像とトモシンセシス画像を用いて得られた結果の一致は、乳腺密度カテゴリーの推定における Quantra™が臨床データに対して有効に機能することを裏付けるものである。

さらに、a と b のカテゴリーを「脂肪性」、c と d のカテゴリーを「高濃度」にグループ分けすることは、検診現場での分類管理という観点から重要である。このため、分析を行い、同じデータを用いるが、カテゴリーを「脂肪性」と「高濃度」のみとして、以下の分割表をさらに作成した。

表 4 に、「脂肪性」と「高濃度」という評価を用いた、従来の 2D 画像を用いた Quantra™ 2.2 ソフトウェアとグラウンドトゥルースとの比較を示す。

	Quantra™ 2.2 ソフトウェア 従来の 2D 画像		
「グラウンド トゥルー ス」 BI-RADS Ed5		脂肪性	高濃度
	脂肪性	105	8
	高濃度	1	116
	観察結果が一致した検 診		221
	総数		230
	全体的な一致率 (%)		96%

表 4. 従来の 2D 画像を用いた Quantra™ 2.2 のカテゴリと、放射線科医が判定した「脂肪性と高濃度」のグラウンドトゥルースとの比較。

表 5 に、「脂肪性」と「高濃度」という評価を用いた、トモシンセシス画像を用いた Quantra™ 2.2 ソフトウェアとグラウンドトゥルースのカテゴリとの比較を示す。

Quantra™ソフトウェアによる脂肪性乳房か高濃度乳房かの分類方法は、放射線科医のグラウンドトゥルースと比較して、全体的な一致率が 90%をかなり上回っていることがわかった。

	Quantra™ 2.2 ソフトウェア トモシンセシス画像		
「グラウンド トゥルース」 BI-RADS Ed5		脂肪性	高濃度
	脂肪性	104	9
	高濃度	7	110
	観察結果が 一致した検診		214
	総数		230
	全体的な一致率 (%)		93%

表 5. トモシンセシス画像を用いた Quantra™ 2.2 のカテゴリと、放射線科医が判定した「脂肪性と高濃度」のグラウンドトゥルースのカテゴリとの比較。

結論

Hologic®は、臨床的ワークフローに適切に組み込むことができる自動化された乳房構成カテゴリ分類を実現するため、Quantra™ソフトウェアの継続的な改良に注力している。Quantra™ 2.2 ソフトウェアには、ACR の BI-RADS 委員会が提案した修正事項が反映されている。Quantra™ 2.2 の手法には、数千件のマンモグラフィ検査を用いてトレーニングを行った最先端の機械学習に基づくアルゴリズムが用いられている。Quantra™ソフトウェアは、Hologic の社内試験で得られた読影者の評価によるグラウンドトゥルースとの全体的な一致率が高いことが示された。Quantra™とグラウンドトゥルースとの全体的な一致率は、読影者間の一致率とほぼ同じであったことから、Quantra™を自動化された乳房構成カテゴリ分類法として臨床現場で使用できることが示された。さらに、被検者を「脂肪性と高濃度」に分類する際、Quantra™と放射線科医による評価との一致率が 90%を超えることも示された。

参考文献：

1. Misclassification of Breast Imaging Reporting and Data System (BI-RADS) mammographic density and implications for breast density reporting legislation. Gard, CC, et al. 2015, *The Breast Journal*, Vol. 21.5, pp. 481-489.
2. Inter- and intraradiologist variability in the BI-RADS assessment and breast density categories for screening mammograms. Redondo, A, et al. Nov 2012, *Br J Radiol*, Vol. 85(1019), pp. 1465-1470.
3. Categorizing breast mammographic density: intra- and interobserver reproducibility of BI-RADS density categories. Ciatto, S, et al. 4, Aug 2005, *Breast*, Vol. 14, pp. 269-275.
4. A dedicated BI-RADS training programme: Effect on the interobserver variation among screening radiologist. Timmers, JM, et al. 2012, *Radiology*, Vol. 81, pp. 2184-2188.
5. American College of Radiology (米国放射線学会). *BI-RADS Atlas*. 2013. Edition 5.
6. Digital mammography screening: sensitivity of the programme dependent on breast density. Weigel, S, et al. 2017, *Eur Radiol*, Vol. 27(7), pp. 2744-2751.
7. Mammographic density, breast cancer risk and risk prediction. Vachon, CM, et al. 2007, *Breast Cancer Research*, Vol. 9, p. 217.
8. A novel and fully automated mammographic texture analysis for risk prediction: results from two case-control studies. Wang, C, et al. 2017, *Breast Cancer Research*, Vol. 19, p. 114.
9. Mammographic density and structural features can individually and jointly contribute to breast cancer risk assessment in mammography screening: a case-control study. Winkel, RR, et al. 2016, *BMC Cancer*, Vol. 16, p. 414.
10. Mammographic texture resemblance generalizes as an independent risk factor for breast cancer. Nielsen M, Vachon CM, Scott CG, Chernoff K, Karemore G, Karssemeijer N, Lillholm M, Karsdal MA. 2014, *Breast Cancer Research*, Vol. 16(2), p. R37.
11. Text Categorization with Support Vector Machines: Learning with Many Relevant Features. T., Joachims. 1998, *European Conference on Machine Learning ECML-98*, vol 1398, Vol. 1398, pp. 137-142.
12. Hologic. Quantra 2.2 FDA clearance 510K K163623 社内資料 2017.
13. Comparing Visually Assessed BI-RADS Breast Density and Automated Volumetric Breast Density Software: A Cross- Sectional Study in a Breast Cancer Screening Setting. Van Der Waal, D, et al. 9, Sept 2015, *PLOS One*, Vol. 10.

販売名: デジタル式乳房X線撮影装置 Selenia Dimensions

認証番号: 222ABBZX00177000

文書管理番号: PMR 095

Hologic, Inc.

250 Campus Drive
Marlborough, MA 01752 USA
電話: +1-800.442.9892
営業: +1.781.999.7453
Fax: +1.781.280.0668
www.hologic.com

WP-00123 Rev 002 (05/18) ©2018 Hologic, Inc. Hologic, Quantra, Selenia, Dimensions は米国をはじめとする各国に所在する Hologic, Inc. とその子会社の商標または登録商標です。その他のすべての商標、登録商標および製品名の所有権は、各所有者に帰属します。本書に記載の情報は医療専門家を対象としており、購入勧誘やプロモーションが禁止されている場所でそのような活動を行うことを意図したものではありません。Hologic 社の資料はウェブサイトや eBroadcast、トレードショーなどを通して配布されているため、当該資料の場所を常に管理することはできません。各国で販売されている製品についての具体的な情報は、最寄りの Hologic 代理店または womenshealth@hologic.com にお問い合わせください。

HOLOGIC