

視覚評価の基礎

—視覚評価 ROC解析—

秋田大学医学部附属病院 中央放射線部 ○高橋 俊吾(Takahashi Shungo)

【はじめに】

デジタルX線撮影をはじめとした多くの放射線医学分野における画像に対して、診断を担保しながら適切な被ばく線量での画質提供は常に放射線技師が向き合っていかなければならない課題である。課題に対して線質・線量といった異なる条件から取得された放射線画像に物理評価や視覚評価といった手法により画質の評価・被ばく線量の最適化が行われている。今回は視覚評価のうちROC観察者実験・解析について理解を深めるとともに、実際に検証・解析を行った。

【視覚評価概論】

視覚評価には病変の検出能、良性・悪性の判断や疾患の状態の判断をする識別能、それら2つを組み合わせた総合評価、病変の見えやすさの比較評価といった方法がある。検出能評価にはROC (1症例1判定-位置情報無)、LROC (1症例1判定-位置情報有)、FROC (1症例多判定-位置情報有) が用いられる。識別能評価にはROCのみ用いられる。総合評価にはROC・FROCを組み合わせる場合が多くみられる。

視覚評価が目的とする「診断の正確さ」を、医用画像に病気が存在する・存在しないという判定で考えた場合、診断によって得られる結果は下表Table 1のように示される。

Table 1. 診断の正確さ

		反応(診断の結果)	
		異常	正常
刺激 (診断の対象)	異常	真陽性	偽陰性
	正常	偽陽性	真陰性

真陽性 (True Positive: TP) 偽陰性 (False Negative: FN)
偽陽性 (False Positive: FP) 真陰性 (True Negative: TN)

これらを用いて、診断の正確さの指標である

$$\text{感度}[\%] = 100 \times \text{TP} / (\text{TP} + \text{FN})$$

$$\text{特異度}[\%] = 100 \times \text{TN} / (\text{TN} + \text{FP})$$

が表される。感度 (sensitivity) は表記が存在している医用画像を正しく「病気あり」と診断した確率、特異度 (specificity) は病気が存在しない医用画像を正しく「病気なし」と診断した確率を意味する。

すべての対象に対して「病気あり」と判定を行うと感度は100%・特異度は0%となり、すべての対象に「病気なし」の判定を行うと感度は0%・特異度は100%となる。感度・特異度を「診断の正確さ」の指標として用いる場合には必ず、感度と特異度の両方を示す必要がある。

さらに感度・特異度に加えて「病気あり」・「病気なし」の判定がどれほど正しいかという観点から

$$\text{陽性的中率}[\%] = 100 \times \text{TP} / (\text{TP} + \text{FP})$$

$$\text{陰性的中率}[\%] = 100 \times \text{TN} / (\text{TN} + \text{FN})$$

が表される。陽性的中率 (PPV : positive predictive value) は判定が「病気あり」であった場合に、それが正しい判定である確率を示し、陰性的中率 (NPV : negative predictive value) は判定が「病気なし」であった場合に、それが正しい判定である確率を示す。

ROC解析は観察者実験から得られた評定データ、コンピュータから算出したROC曲線を用いて、曲線下の面積 (AUC : the area under a receiver operating characteristic curve) にて評価を行う。ROC曲線の縦軸は感度/100と同等の意味を持つ真陽性率 (TPF : true positive fraction)、横軸は(100-特異度)/100と同等の意味を持つ偽陽性率 (FPF : false positive fraction) で表される。座標が左上 (TPF:1 FPF:0) に近づくほど診断能が高いと評価される (Fig.1)。

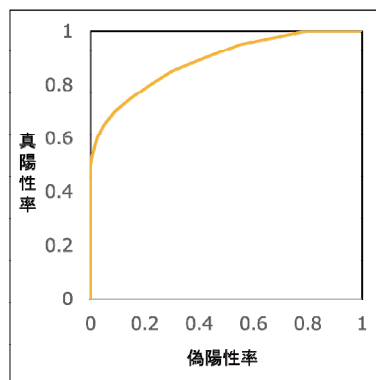


Fig.1 ROC曲線

ROC曲線描出についてはいくつかあるが、両正規分布モデルを用いた方法がある。描出観察者実験から得られた評定データにおいて「病気あり・

病気なし」それぞれの試料画像に対する観察者の反応はそれぞれ正規分布になるという仮定に基づいて理論が構築されている。横軸は試料画像に病気が存在するかどうかの判断基準、縦軸は実際に「病気あり」・「病気無し」である試料画像の頻度を表している。両正規分布モデルでは観察者の「病気あり・病気なし」の試料画像のそれぞれに対する反応（診断の結果）は判断基準に対して正規分布になると仮定している。

両正規分布（Fig.2）において、ある判断基準での真陽性率（橙点線領域：\）・偽陽性率（青点線領域：/）からROC曲線上にプロットされる。（Fig.3）判断基準を変化させ（Fig.4）、それぞれの真陽性率・偽陽性率からROC曲線が描かれる（Fig.5）。

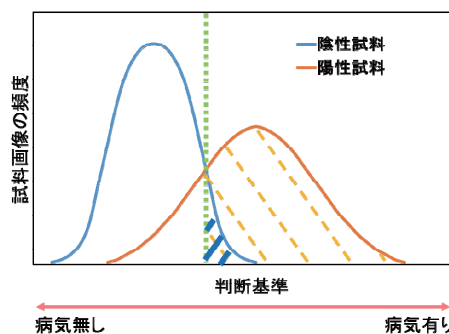


Fig.2 両正規分布と真陽性率・偽陽性率

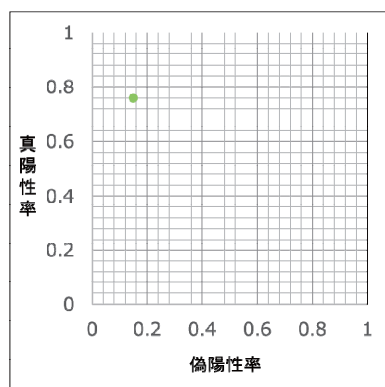


Fig.3 ROC曲線プロット

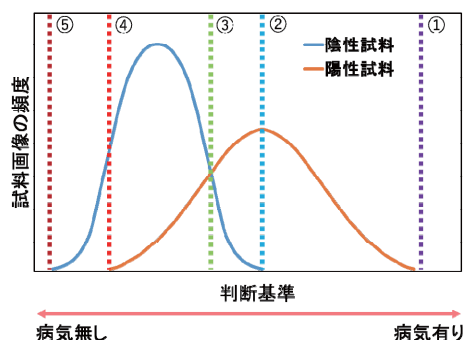


Fig.4 判断基準

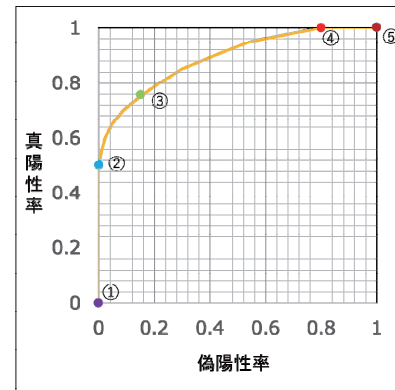


Fig.5 ROC曲線

2つの正規分布の距離が離れるほどROC曲線は左上角に近づき、分布の重なりが多いほど正の対角線に近づく。

ROC曲線から検査データを陽性・陰性に分ける数値をカットオフ値と呼ぶ。算出はROC曲線左上（真陽性率:1、偽陽性率:0）から最も近い点、ROC曲線の正の対角線から最も遠くなる点（Youden Index）の2種類がある。

放射線医学分野において、

- 1) モダリティ間の診断能の比較
 - 2) 検査法・手技間の診断能の比較
 - 3) 観察者の診断能の評価
 - 4) 新しいモダリティ・検査法・手法の評価
 - 5) 最適カットオフ値の推定
- に対してROC解析が用いられている。

【目的】

胸部ファントムを撮影して得られた胸部正面画像の模擬陰影（肺結節）の検出能評価から撮影線量の適正化を行うことを目的としてROC解析を行った。

【使用機器】

- ・X線撮影システム FUJIFILM DR BENE0-FX
- ・胸部ファントム 京都科学 胸部ファントム N-1
- ・読影モニター EIZO RadiForce MX215
- ・模擬陰影 300 mg/ml造影剤・粉寒天

【画像取得条件】

- ・管電圧: 110 kV SID : 180 cm
 - ・付加フィルタ : 0.1 mmCu
 - ・mAs : 0.5, 1.0, 1.6, 3.2 mAs (管電流200 mA)
- （上記、当院条件において標準体型患者にAEC使用下で撮影を行った場合、約1.6 mAs）

【方法】

造影剤を寒天溶液で400倍に希釈し、固めたも

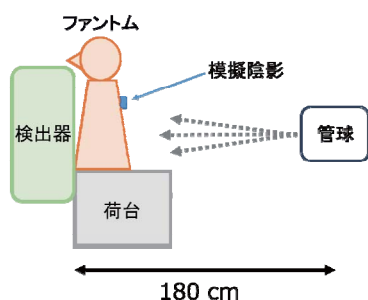


Fig.6 画像取得配置

Table 2. 評価・観察条件

評価条件		観察条件	
試料数	陽性20例・陰性20例	観察時間	制限なし
観察者	11人	観察距離	制限なし
観察回数	4回(4線量)/1評価者	階調調整	調整不可
評価順序	0.5,1.0,1.6,3.2 mAs の順	拡大縮小	調整不可
陰影位置	各線量で同位置 (縦隔・横隔膜付近を含む)	モニタ付近照度	約250lx
観察時間	1時間以内/1回		
検証間隔	1週間		

のを径6・8・12 mm、厚み4・6・8・10・12 mmとなるように型を取り、模擬陰影とした。径12 mm・厚み12 mmおよび径6 mm・厚み4 mmの模擬陰影は事前教育でのみ使用した。画像取得の際には模擬陰影を胸部ファントム背面に貼り付け、ファントムを動かす必要がないような配置とした (Fig.6)。

今回の測定における評価条件をTable 2に示す。評価は撮影線量が低い陰影の画像が後の評価への影響が小さくなると想定されたため、示した順序で行った。

事前教育には 1) 模擬陰影をPMMA上に配置して取得した画像 2) 模擬陰影全種を胸部ファントム同一位置に配置して取得した画像 3) 径12 mm・厚み12 mmおよび径6 mm・厚み4 mmの模擬陰影を胸部ファントム上数か所に配置して取得した画像 4) 模擬陰影を設置しない胸部ファントムを撮影した画像を用いた。

評価は連続確信度法にて実施し、得られた結果から「R」および「statcel」を用いてAUC値算出、統計学的検定を行った。

【結果】

11名の評価者による線量4条件のAUC値を示す (Table 3)。線量を上げるに従い、AUC値が向上した (検出能上昇) 結果を示したのは3名であり、他の評価者においてはAUC値が撮影線量に依存しない結果を認めた。また、Fig.7にプール法により算出された平均ROC曲線を示す。各線量に対するAUC値は、0.5 mAs - 0.825 1.0 mAs - 0.782 1.6 mAs - 0.864 3.2 mAs - 0.973 となり1.0 mAsが最も低い結果となった。フリードマン検定・ボンフェローニ法による有意差検定・多重比較の結果、0.5 mAs - 3.2 mAsにおいて5%水準で有意差を認めた。

Table 3. 各評価者のAUC

評価者	0.5 mAs	1.0mAs	1.6mAs	3.2mAs
A	0.734	0.782	0.830	0.791
B	0.788	0.720	0.798	0.855
C	0.795	0.554	0.746	0.758
D	0.670	0.530	0.781	0.722
E	0.698	0.780	0.908	0.968
F	0.804	0.799	0.764	0.899
G	0.631	0.675	0.737	0.705
H	0.614	0.618	0.769	0.800
I	0.546	0.615	0.662	0.736
J	0.566	0.620	0.646	0.611
K	0.750	0.816	0.710	0.769

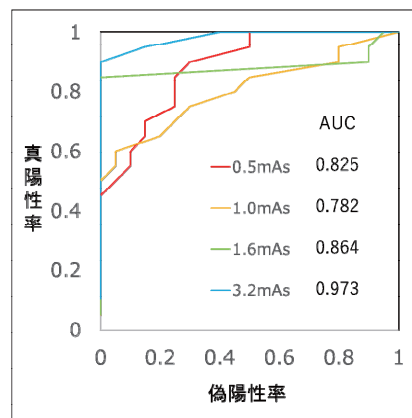


Fig.7 平均ROC曲線

【考察】

AUCの結果が線量の大小と必ずしも一致しないことについて、業務終了後に1時間程度の評価をしており評価時の疲労度の違いから来る検出力に影響があったことが考えられる。評価を1回30分程度に分割すること、毎回の評価基準が統一されるような事前教育 (径12 mm・厚み12 mmおよび径6 mm・厚み4mmの模擬陰影を含めた本番と同様形式での評価) が必要であったと考えられる。

平均ROC曲線において0.5 mAs - 3.2 mAs間に有意差を認めたことから、線量を1.0 mAs程度まで

低減可能であると推測される。また、1.0 mAsのAUCが最も低くなったことについて、評価に用いた画像でCNRを算出したが、明確に数値が下がりうる傾向は得られず、視覚評価の基準のばらつきによる影響が大きいと思われる。

【まとめ】

ROC解析について解説、実際に検証を行った。

ROC解析を行ったことで適切な難易度の試料を作成すること、事前教育を充実させ、判断基準を明確に定められるような工夫を取ることが特に重要な要素であることを実感した。

【参考文献】

- 1) 白石順二 標準 医用画像の視覚評価法 オーム社