

— 実践 —
明日から使える視覚評価

熊本大学医学部保健学科
白石順二

Contents

- What? 視覚評価とは何か？
 - 物理評価 vs. 視覚評価
 - 様々な視覚評価法
- Why? なぜ、視覚評価が必要か？
 - 臨床的有効性評価のための6レベル
 - ROC解析と他の視覚評価法との違い
- How? どのように視覚評価を行うか？
 - 視覚評価に必要な基本的な統計学
 - 視覚評価に影響を与える様々な因子

視覚評価とは何か？

視知覚を駆使して画像に含まれる信号を検出し、その程度を定量的に表現することで、画像の持つ能力を評価すること。

視知覚 (Visual Perception)

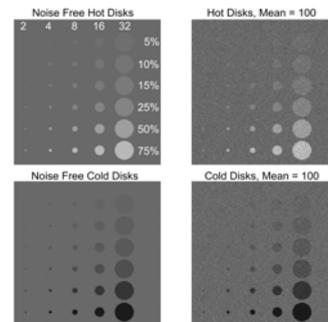
- 何が見えるかを知っている。
- 何が見えないかを知っている。
- 見え方の違い(優劣)を正しく表現できる。
- 再現性がある。

「見える」と「認識できる」の間には大きな違いがある。

様々な視覚評価法

1. ROC解析
(Receiver Operating Characteristic Analysis)
2. 2AFC, 一対比較法
(Two Alternative Forced Choice)
3. C-D ダイアグラム, バーガーズ・ファントム法
(Contrast-Detail Diagram, Burger's Phantom)
4. ハウレット チャート法
(Howlett Chart)

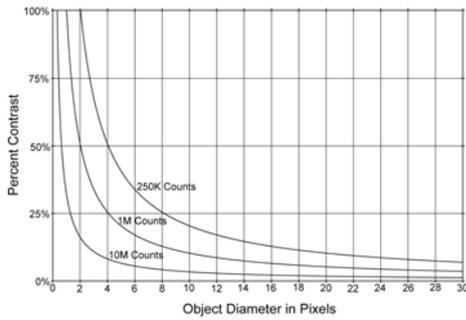
Sample images obtained with a contrast detail diagram



Rzeszotarski M S Radiographics 1999;19:765-782

RadioGraphics

Contrast detail diagram



Rzeszotarski M S Radiographics 1999;19:765-782

RadioGraphics

CT検査における評価項目

性能評価項目

- ノイズ特性
- コントラスト分解能
- 空間分解能
- スライス厚
- アーチファクト

画像評価項目

- フィルタ処理
- 再構成画像の画質
- 被曝線量の最適化
- 造影効果判定
- 保管時の圧縮処理

物理特性の評価

診断能の評価

物理特性評価のための視覚評価

解像特性, コントラスト特性, ノイズ特性の画質の3要素のすべて, または, いずれかの組み合わせの評価を目的とする.

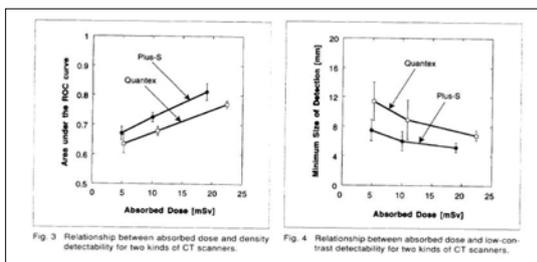
【例】

C-D ダイアグラム → 3要素すべて

視覚評価を用いない ROC解析による物理特性評価

The screenshot shows a research paper titled "ROC解析を応用したCT装置の密度分解能の測定" (Measurement of Density Resolution of CT Devices Using ROC Analysis). The authors listed are 白石順二, 上田容子, 宇部宮あかね, 花岡信子, 飯沼大作, 津田和良. The paper discusses the application of ROC analysis to evaluate the density resolution of CT devices, comparing it to traditional visual evaluation methods.

視覚評価を用いない ROC解析による物理特性評価



診断能評価のための視覚評価

診断能の評価の種類

- 検出の優劣 → C-D, ROC解析
- 鑑別の優劣 → ROC解析
- 見えやすさの優劣 → 2AFC
- 所要時間の長短 → Any
- 費用の多い少ない → ROC解析

なぜ、診断能の評価が必要か？

物理特性の評価と診断能の評価は、医療行為の効果を判定するという意味において、違ったレベルにある。

臨床的有効性評価のための6レベル (Fryback & Thornbury, *Med Decis Making*, 1991)

- 1) Technical quality: 技術的品質の評価
MTF, NPS, H&D curve, etc.
- 2) Diagnostic accuracy: 診断の正確さの評価
Agreement between diagnoses and "truth"
- 3) Diagnostic-thinking efficacy: 診療における効果
Impact of Dx test on physician's thinking about each patient
- 4) Therapeutic efficacy: 治療における効果
Impact of Dx test on patient management
- 5) Patient-outcome efficacy: QOLの効果
Impact of Dx test on patients' health
- 6) Societal efficacy: 社会的効果
Impact of Dx test on society as a whole

ROC解析は他の視覚評価法と比べて何が違うのか？

ROC解析以外の方法では、診断の正確さの評価を十分効果的にまたは有意に定量化することが不可能

感度と特異度

$$\text{感度} = \frac{\text{正しく病変と判定された症例の数}}{\text{実際に病変が存在した症例の数}}$$

$$\text{特異度} = \frac{\text{正しく正常と判定された症例の数}}{\text{実際に正常であった症例の数}}$$

異常例 << 正常例 → 感度 < 特異度
異常例 >> 正常例 → 感度 > 特異度

どのように視覚評価を行うか？

「人が見て評価した」というだけでは、視覚評価とは言えない。統計的検定があって初めて視覚評価と言える。

視覚評価に必要な統計学

- 観察者間の平均と変動
- 分類尺度, 順序尺度, 間隔尺度
- パラメトリック検定とノンパラメトリック検定
- 実験に用いる試料の読影の難易度の分布
- 平均値の間の統計的有意差検定
- 分布の間の統計的有意差検定

観察者間の平均と変動

- 観察者内変動と観察者間変動
 - 観察者内変動: 同じことを同じ観察者が繰り返し行った場合の観察者個人の変動
 - 観察者間変動: 同じことを別の観察者が行った場合の観察者グループ内の変動
- 変動の大きさは、観察者の経験量と実験に用いる試料の読影の難しさに関係がある。
 - 経験が少ない → 変動大きい
 - すごく易しい, すごく難しい → 変動少ない

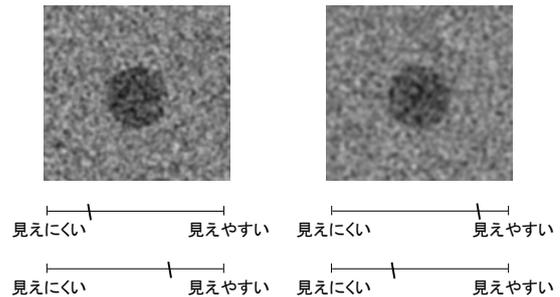
分類尺度, 順序尺度, 間隔尺度

- 分類尺度
 - 個体のある定性的な特性によって分類する尺度. 個体の大小関係は定義されない.
 - <例> 疾患A,B,C... プロトコールA,B,C
- 順序尺度
 - 個体のある量的な特性の順序関係で分類する尺度. 個体間の順序関係のみが定義され, その間隔は定義されない.
 - <例> 見えやすさの順位
- 間隔尺度
 - 個体のある量的な特性の順序および間隔で分類する尺度.
 - <例> 温度, 長さ, 時間のような物理的特性の測定値

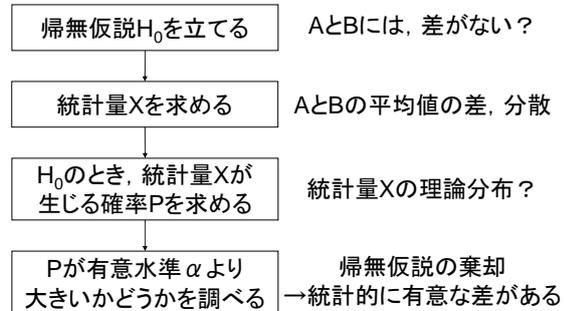
パラメトリック検定とノンパラメトリック検定

- パラメトリック検定 (Parametric Test)
- 母集団 (Population) の特性を利用
 - * 利用可能なのは間隔尺度のみ
 - 両側t検定, 分散分析 (ANOVA)
- ノンパラメトリック検定 (Non-Parametric Test)
- 母集団の特性は考慮しない
 - * 分類尺度, 順序尺度, 間隔尺度のいずれにも利用可能
 - Wilcoxon検定, Mann-Whitney検定

評定方法と尺度



統計的検定とは?



「有意差がない=同等である」は, 大きな間違いである!

まとめ

- 視覚評価は, 単に目で見て評価するということではない.
- 臨床的有用性を評価するためには, 診断能を評価しなければいけない.
- 視覚評価には統計解析が不可欠である.
- 実験方法によって統計解析の手法は異なる.
- 視覚評価の統計解析に伴う, 様々な変動要因を理解した上で, 視覚評価を行う必要がある.