

放射線科学

画像評価の話

小寺 吉衛

診療に用いられるX線画像の善し悪しを決めるものは何か。大学で画像評価の講義を行っていることから、こういう質問をよく受ける。この答えが実に難しい。画像の良否を決定する因子が複雑に作用することも要因ではあるが、それよりもなにも私は診断に携わっている医師ではないので、本当のところどうい画像が診断に向いているのかよく分からない。「そんな人間がよく画像を評価しているな」とどこからかお叱りの声が聞こえてきそうであるが、わからないものはわからない。それどころか、こう言うては失礼であるが、普段診断を行っている医師の方々にもよく分からないのではないかと思う。身近の医師に聞いてみても、うーんと唸られることもしばしばである。そんな曖昧なものをよく評価などすると言われるかも知れないが、曖昧だからこそ面白い、と思っている。

名古屋大学医学部保健学科の放射線技術科学専攻では、2年生の後期で「医用画像工学」を履修する。内容は、光の定義から始まり、画像のできる過程をシステム工学的に解説する。要は、画像の劣化の因子（ボケや雑音）がどのように作用して真の像から劣化を含む画像になるかということ解析的に説明する訳である。数学的には実に美しく、説明する側も惚れ惚れするような過程である。この数学で説明のつく部分を抜き出して、画像のボケと雑音の解析方法を伝授するのが画像工学の初歩である。内容は至極明瞭であり、理詰めを考える学生からは、数学的な展開が面白いと好評である。ところが、片や感覚的な思考が中心の学生からは、さっぱり分からないと不評たらたらである。ここらあたりで、画像工学の講義に対する好き嫌いがはっきりしてくる。そもそも画像は目で見て判断するわけであり、そう言う意味では感覚的であることから、理詰めを考えるだけではダメである。ところが往々にして最初に数学的な事柄を中心に説明するものだから、画像工学は理屈っぽいと取られやすい。日々は反省。

これまでに「画像評価法」として考えられてきた方法は、人が画像を見て判断する方法（視覚評価、心理的評価、心理物理的評価、主観的評価などと呼ば

れている)と、画像の濃度などを光学的機器を用いて計測し、得られたデータを解析して評価する方法(物理的評価、客観的評価などと呼ばれている)に分けられる。初めに述べた、数学的な解析による方法とは、大半が物理的評価法のことであるが、人が画像を見て判断する方法の中にも、大変数学的な、そして見事な手法がある。その代表的なものに、受信者動作特性曲線(receiver operating characteristic curve、ROC 曲線)を用いたROC解析と呼ばれるものがある。名前を聞かれた方は多いであろうし、実際にこれを用いて評価を行ったという方もおられることと思う。この手法が評価法として高い評価を受けている理由は、主観的な評価にもかかわらず、得られた結果に客観性があるからである。方法をご存知ない方のために、簡単に説明すると、まず、評価したい画像システムを含む複数の画像システムを選ぶ。新しく出た増感紙フィルム系を調べたいということであれば、その増感紙フィルム系と従来から用いている増感紙フィルム系の2種類でよい。つぎに、このシステムで診断する場合の対象となる症例を想定し、その中で淡い陰影となるようなものを考える。ROC解析は見えるか見えないかという状態を必要とするので、みんな見えてしまうコントラストの強い陰影は不適である。胸部であれば小さな腫瘍陰影などがよいであろう。ROC解析では、このような腫瘍陰影を含んだ画像と含まない画像をそれぞれ50枚程度、各々のシステムについて集めなければならない。実際の臨床例が集められればそれに越したことはないが、同一患者を従来のシステムと新しいシステムで2度撮影するということは患者の被曝という観点から許されることではない。そこで、腫瘍陰影を模したもので撮影を行うことになる。たとえば、胸部の腫瘍陰影ということであれば直径5mm程度のプラスチックのビーズ玉などがよく用いられる。通常は、10cm×10cm程度の大きさの中に先ほどのビーズを1個置いて適度な濃度(実際に腫瘍陰影を観察するのと同じ程度の背景濃度)になるようにX線を照射して撮影する。このとき、ビーズの陰影が見えるか見えない程度になるようにアクリル板などを置いてコントラストを調整する。二つのシステムで、ともに50個のビーズがすべて見えてしまっても、あるいは全く見えなくても評価にはならない。どちらか一方ですべて見え、もう一方で全く見えないということであれば特別な解析を行うことなく、すべて見えるシステムの方がよいに決まっている。見えたり見えなかったりする状態、これが重要である。では、同じものを同じように撮影したにもかかわらず、なぜ見えたり見えなかったりするのか。そこに、画像のゆらぎが関係してくる。具体的に考えてみよう。X線を一様に照射したとしてもX線の分布が均一になることはない。このX線のゆらぎによる雑音を量子モトル(quantum

mottle)と呼んでいる。X線画像の世界では、雑音の中でこの量子モトルの占める割合が非常に大きい。増感紙フィルム系の雑音としては、この量子モトルの他に、フィルム銀粒子のゆらぎ、増感紙の蛍光体の構造的なゆらぎなどがあるが、通常のシステムで量子モトルの占める割合は80~90%といわれている。CR(computed radiography)のようなデジタル系では、X線の線量を下げても画像を得ることができることが知られているが、実際には線量を下げると量子モトルがどんどん大きくなり、その雑音によって像が著しく劣化する。したがって、従来の増感紙フィルム系とほとんど変わらない線量で撮影されているのが現状である。それどころか、多い線量で撮影された場合には雑音が小さくなるため、線量が増えているのではないかという危惧もされている。いずれにしても、デジタル系では、新たな撮影条件の設定が必要である。さて、話をROC解析に戻そう。同じように撮影されても量子モトルの影響で画像がゆらぐことがわかった。これが試料を沢山作る理由である。少ない試料では、ゆらぎによる結果の偏りを補正することはできない。こうして作られた腫瘍陰影（これを信号と呼ぶことにする）のある画像とない画像をランダムに並べ、順次観察者に提示していく。観察者は、信号があると思えば、高い点数を与え、ないと思えば低い点数を与えることにする。提示する側は、どの画像に信号があり、どの画像に信号がないか知っていることから、信号のあるものかないものについてそれぞれ点数の順番に並べることができる。ROC解析の面白いところはここからで、点数の高い方から任意の点で二つにデータを分割する。例えば100点満点で80点以上を境に信号があったと考えることにする。この中で実際に信号のあるものが8枚、ないものが1枚であったとすると、正解は $8/50=0.16$ 、不正解は $1/50=0.02$ となる。前者を適中確率(true positive、TP)、後者を誤報確率(false positive、FP)と呼ぶ。システムとしては、前者が1に近く、後者が0に近いことが理想である。先ほどの境の点数を100点より大きくすると、TP、FPとも0になる。また、境の点数を0にして考えると、両者とも1になってしまう。この点数が、観察者の判断基準の境界を表す。ROC曲線とは、横軸に誤報確率FPをとり、縦軸に適中確率TPをとって、判断基準を100点から0点まで順番に下げたときにできる曲線のことである。理想的なシステムでは、誤報確率は0で適中確率が1であるから、正方形のグラフでは左下端から真直ぐ上に向かって左上のコーナーにいき、そのまま右にいく線になる。信号があるかないか全く分からないシステムでは、左下から右上に斜45度の直線（対角線）となる。通常はこの直線から上側に左上のコーナーに向かって凸の曲線となる。よく、一番悪いシステムは、左下端から右下の角を通り、右上に上がっていく

直線と思われることがあるが、これは全く逆に判断しているだけで信号ありとなしを確実に分離していることになり、システムとしては悪いものでないことが分かる。

さて、このようにして得られた ROC 曲線から何が分かるのであろうか。曲線が左上のコーナーに近いほど、信号ありとなしをよく分離できるシステムであることが分かる。先の例で、新しい増感紙フィルム系と従来の増感紙フィルム系で ROC 曲線を求めた結果、新しいシステムの方の曲線が上にあったとすれば、新しいシステムの方が信号を見つけやすいシステムであるということが示唆される訳である。では、新しいシステムは本当によいシステムなのか。ここで考えなければならないのは、この ROC 曲線は、胸部の腫瘍陰影を模した信号の結果であり、それ以外の何ものでもないということである。したがって、信号の大きさが変わったり、背景濃度が変わったら、結果も異なる可能性がある。X線写真を観察する診断行為の中で、腫瘍陰影の大きさが限定されていたり、特定の背景濃度のみを対象とすることはない。視覚評価では、常に用いた試料の条件に依存する結果しか得られない。反対に、物理的評価では、そのような条件をすべて排除することは可能である。しかし、目で見た状態を忠実に再現することは困難である。また、画質を表す種々の因子は互いに相反する方向に動いており、これらを総合的に評価する方法は今のところない。このようなジレンマの中で画像評価は行われている。まだまだ道は遠い。

X線に対する感度が高く、鮮鋭度も高く、雑音は低く、コントラストの高い、そしてラチチュードのひろい理想のシステムなどというものは存在しない。これはデジタル系でも同様である。このような限られた条件の中で、最適のシステムを考えるということは、穴の空いたバケツに水を入れているようなものかも知れない。しかし、X線が医療に使われて 100 年、その間、画像は格段によくなってきた。この後の 100 年でさらに画像をよくすることは非常に難しいかも知れないが、挑戦するに足るテーマであると考えている。また、近年のデジタル化の傾向は、画像に付加価値をつけるという意味では大きな可能性を秘めている。医用画像の評価とは、そういう可能性も一つの価値観として対象に含めるべきものと考えている。さらに道は遠い。さて、画像の善し悪しとは一体なんであらうか。

(名古屋大学医学部教授・保健学科放射線技術科学専攻)