第 23 回科学技術委員会技術セミナー5

謎解き臨床化学検査　わかりにくい言葉・あやふやな事を明確に
ー 精度管理、その他 ー

感度、特異度、ROC、陽性・陰性的中率
個体内内変動と個体間変動～検査データのモニタリング～

山本　慶和 （天理医療大学　臨床検査学科）

1．はじめに

臨床検査の目的の一つに“測定値から特定の疾患・病態の有無を判別すること”がある。このために目的とする疾患群と非疾患群（例えば疾患群が肝がんに対して非疾患群は肝硬変という場合、また疾患群が糖尿病に対して非疾患群は健常者のように、非疾患群の選択は目的によって異なることに注意したい）の判別指標として感度、特異度、ROC 曲線による臨床的有用性の評価が行われる。こうした情報に基づいてどの測定法（試薬）を選ぶのか常に意識しておく必要がある。また、検査室でこうした検討を行って施設における臨床的有用性の情報をまとめ、臨床側に提供できるとよい。

最初に、ここでは基礎的なカットオフ値と陽性、陰性の 2×2 分割表の作成、感度・特異度の計算、カットオフ値の設定、陽性的中率および検査後確率等について例を交えて紹介する（第 2 章）。

次に、個体内変動・個体間変動は外部精度管理調査の許容誤差限界の指標にされ、検査室の精度管理（品質管理）の一つとなっている。ここでは個体内生理的変動を考慮したモニタリング時の変化量の有意な差を評価する方法について述べる（第 3 章）。これらを通して、検査臨床的有用性の把握、これらを発展的に用いて臨床への情報発信・支援につながることを期待したい。

2．診断特性指標（2×2 分割表）、感度・特異度 [1,2,3)]

１）感度・特異度

検査では疾患群と非疾患群を完全に判別する理想的な検査はない、通常の検査は図 1 に示すように疾患群と非疾患群を判別できない領域が存在する。そこで感度・特異度を用いて検査の診断特性を評価することになる。感度・特異度を求めるには、まず検査項目の目的とする疾患・病態を選択する。

例えば急性冠症候群であれば早期診断を目的としてトロポニン、CKMB、ヒト心臓由来脂肪酸結合蛋白（H-FABP）などを用い、発症からの時間ごとの検査結果の陽性・陰性を調べ感度・特異度、ROC 曲線を使って診断特性の評価を行う。

臨床検査の診断的有用性を評価する際の診断特性指標（2×2 分割表）を表 1 に示す。

表の縦軸に疾患群および非疾患群、行には疾患群、非疾患群における検査結果の陽性・陰性（カットオフ値以上：陽性、未満：陰性）を配置する。

●疾患群で検査結果が陽性を真陽性 **a**、陰性を偽陰性 **c**

●非疾患群で検査結果が陰性を真陰性 **d**、および陽性を偽陽性 **b** とする。

これらのデータより以下の指標を得ることができる。

◆**感度 sensitivity** は疾患群 **a+c** で検査結果が陽性 **a** となる確率　感度＝$\frac{a}{a+c}$

◆**特異度 specificity** は非疾患群 **b+d** で検査結果が陰性 **d** となる確率

特異度＝$\frac{d}{b+d}$

◆**陽性的中率（陽性予測値）** Positive Predictive Value は検査で陽性 **a+b** と判定された中で疾患群（真陽性）の割合　陽性的中率＝$\frac{a}{a+b}$

◆**陰性的中率（陰性予測値）** Negative Predictive Value は検査で陰性 **c+d** と判定された中の非疾患群（真陰性）の割合　陰性的中率＝$\frac{d}{c+d}$

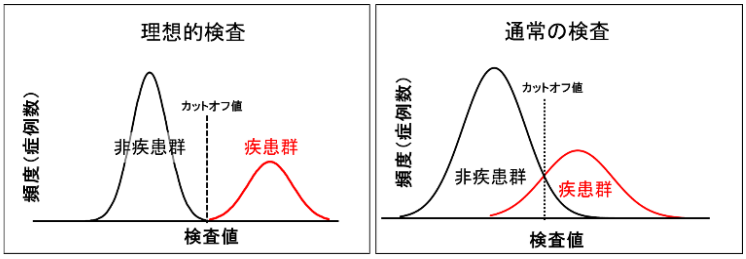


図1　理想的検査と通常の検査　感度･特異度

表1 臨床検査の診断的有用性の評価

表　診断特性指標（2×2分割表）

		疾患の有無		
		疾患群	非疾患群	合計
診断的検査結果	陽性	真陽性 a	偽陽性 b	a+b
	陰性	偽陰性 c	真陰性 d	c+d
	合計	a+c	b+d	

a：疾患群で　検査結果 陽性　真陽性
c：疾患群で　検査結果 陰性　偽陰性
b：非疾患群で 検査結果 陽性　偽陽性
d：非疾患群で 検査結果 陰性　真陰性

２）ROC 曲線（Receiver Operating Characteristic curve）とカットオフ値の関係

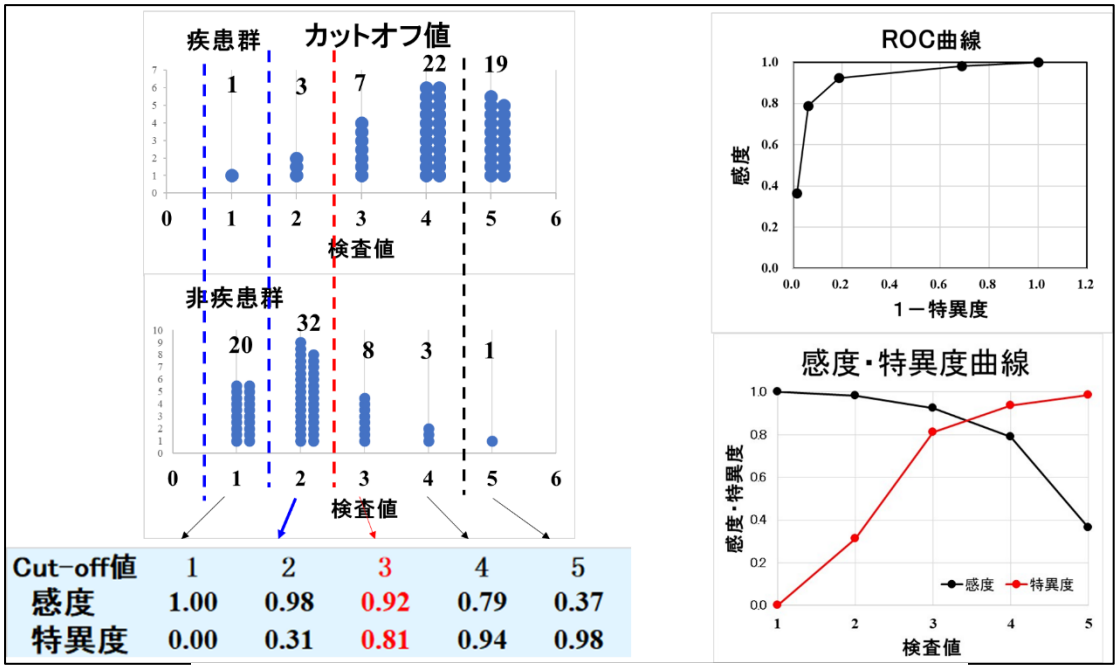


図 2　カットオフ値から感度・特異度、ROC 曲線

定量検査データではカットオフ値（陽性の判断基準）を変えると感度・特異度が変化する。感度・特異度がどのように変化するかをプロットしたものが ROC 曲線といわれ、検査の診断特性を示す。

＊ROC 曲線からカットオフ値の決め方

ROC が得られれば、そこから疾患群を合理的に判別するカットオフ値を設定することができる。ROC 曲線上の「図 2 の左上の隅の点から最小の距離となる交点」から決定する方法、あるいは感度・特異度曲線図 2 の右下図の「感度と特異度曲線の交点」から決定する方法があるが、後者の方が分かりやすい。

３）ROC 曲線と AUC

ROC 曲線は上述したように感度・特異度を視覚的に表し、検査の診断特性を示し、診断カットオフ値の設定と複数検査の優劣比較が可能である。AUC は ROC 曲線下面積（Ares under the curve）の大きさを表す。0.5～1 の範囲の値をとり、1 に近いほど判別特性能が優れている。

その一例として、Mark Hatherill（Arch Dis Child 1999;81:417–421）ら[4]の結果を示す。図 3 では小児（3 か月から 193 月年齢）の敗血症ショック（重症感染症）の鑑別におけるプロカルシトニン（PCT)、CRP および白血球

（WBC）の感度・特異度の ROC 曲線および AUC を示す。AUC 値はそれぞれ 0.96、0.83 および 0.51 となり、PCT が幼児や子供の重篤な細菌性疾患の鑑別に役立つ可能性があること述べている。

■Diagnostic markers of infection:
procalcitonin（PCT）
C reactive protein（CRP）and
leucocyte count（WCC）

■AUC値
PCT : 0.96
CRP : 0.83
WCC: 0.51

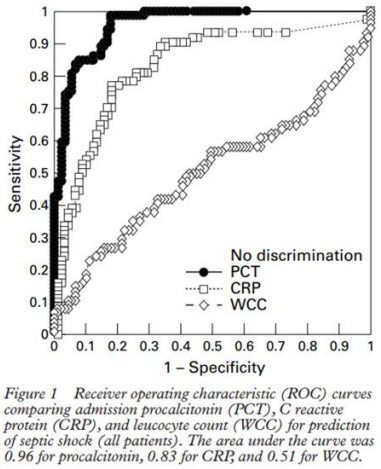


図3 敗血症性ショック（小児の重症感染症）に対するPCT、CRPおよび白血球数の感度と特異度を示す ROC 曲線

Figure 1 Receiver operating characteristic (ROC) curves comparing admission procalcitonin (PCT), C reactive protein (CRP), and leucocyte count (WCC) for prediction of septic shock (all patients). The area under the curve was 0.96 for procalcitonin, 0.83 for CRP, and 0.51 for WCC.

Mark Hatherill. Arch Dis Child 1999;81:417–421

４）尤度比（Likelihood ratio; LR）による確定診断・除外診断

感度・特異度の性質は、感度を上げれば特異度は低下するトレードオフの関係にある。感度と特異度を組み合わせた指標で評価するために、尤度比という指標が用いられる。

尤度は「ありそうな状態のこと；公算」を意味し、

「There is a strong likelihood of rain.　雨が降る**見込みは大い**にある」のように使う。

尤度比は目的の疾患に対する鑑別能力を示し、陽性尤度比 LR+ と陰性尤度比 LR- がある（表 2）。

◆検査が陽性の場合、陽性尤度比 $LR+ = \frac{感度}{(1-特異度)}$

◆検査が陰性の場合、陰性尤度比 $LR- = \frac{(1-感度)}{特異度}$

●陽性尤度比が 10 以上あれば、確定診断に有効な検査

●陰性尤度比が小さければ（0.1 以下）、除外診断に有効な検査といえる。

同様に

●感度の高い検査が陰性ならば、その疾患はほぼ否定（除外診断）できる

●特異度の高い検査が陽性ならば、その疾患にり患している（確定診断）といえる。

表 2　陽性尤度比、陰生尤度比

感度：0.90、特異度：0.95

	疾患	非疾患	陽性尤度比/陰性尤度比
陽性	90	5	LR+ ＝ 0.9/(1－0.95) ＝ 18
陰性	10	95	LR－＝(1－0.9) / 0.95 ＝ 0.105

感度：0.70、特異度：0.80

	疾患	非疾患	
陽性	70	20	LR+ ＝ 0.70/(1－0.80)＝ 3.2
陰性	30	80	LR－＝(1－0.70) / 0.80 ＝ 0.375

５）検査後確率の推定　～ベイズの定理の利用～

検査後確率は

事後オッズ＝事前オッズ×尤度比（ベイズの定理）で求めることができる。

●事前オッズは検査前の確率のこと、また事前オッズは有病率ととらえることもある。

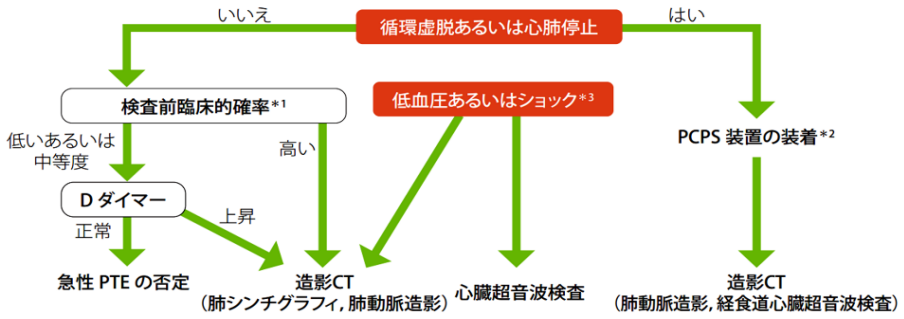
●事後オッズは検査後の確率のこと（検査の結果から目的疾患の確率を示す）

検査の前の目的疾患の事前確率を 40％とすると

事前オッズは＝40／（100-40）＝40／60＝0.67

事後オッズを検査後確率に直すときには　確率＝$\frac{\text{オッズ}}{(\text{オッズ}+1)}$　にて計算する。

ここで、Paul D Stein（Ann Intern Med. 2004 Apr 20;140(8):589-602.）[5]および肺血栓塞栓症および深部静脈血栓症の診断，治療，予防に関するガイドライン（2017 改訂版）[6]の急性肺血栓塞栓症（Pulmonary Thromboembolism、PTE)の除外のための D ダイマー検査について説明を加えた事後確率の計算例をみてみよう（図 4、表 3）。



PTEを疑った時点でヘパリンを投与する．DVTも同時に探索する．
*1：スクリーニング検査として胸部X線，心電図，動脈血ガス分析，経胸壁心臓超音波検査，血液生化学検査を行う．

図４　急性肺血栓塞栓症PTE の診断手順

肺血栓塞栓症および深部静脈血栓症の診断，治療，予防に関するガイドライン（2017改訂版）より

表３　急性肺血栓塞栓症 PTE 検査前確立の評価法

■改訂ジュネーブ・スコア	
•66歳以上	＋1
•PTEあるいはDVTの既往	＋1
•1ヵ月以内の手術，骨折	＋1
•活動性の癌	＋1
•一側の下肢痛	＋1
•下肢深部静脈の触診による痛みと	＋1
•片側性浮腫 心拍数 75～94/分	＋1
•95/分以上 ＋2 血痰	＋1

■臨床的確率	
合計スコア 0～1	低い
2～4	中等度
5以上	高い

肺血栓塞栓症および深部静脈血栓症の診断，治療，予防に関するガイドライン
（2017改訂版）より

- 病歴と診察所見（表 3 参照）より検査前確率を中等度 40％とすると、事前オッズは＝0.40/（1.00－0.40）＝0.67 となる。
- 急性肺血栓塞栓症の D ダイマー（ELISA 法）の感度 0.95、特異度 0.36 とする。
  検査結果が陰性の場合
  陰性尤度比＝（1－感度）／特異度＝（1－0.95）／0.36 ＝ 0.13 となる。
- この条件から事後確率を計算するために事後オッズより
  事後オッズ＝事前オッズ×陰性尤度比＝0.67×0.13＝0.087
- 事後オッズより検査後確率を求めると
  検査後確率＝0.087／（0.087+1）×100 ＝ 8.0％

D ダイマー検査結果が陰性となると検査後確率が 8.0％と相当低く、急性肺血栓塞栓症は除外された。つまり、検査前確率が高くないときに除外診断特性の高い（陰性尤度比の小さい）検査（ここでは D ダイマー）を実施し、陰性であれば目的の疾患を除外できる。

ただし、D ダイマーの実施は検査前確率が高い（臨床的に PTE である可能

性が高い）場合にはDダイマー検査の利用価値は下がるとされている。

６）まとめ

以上、感度・特異度等の診断特性指標、確定診断、除外診断の指標となる陽性・陰性尤度について説明し、最後に目的疾患の検査後の確率について述べた。

参考文献

1)山本慶和．感度、特異度、ROC曲線、陽性的中率、陰性的中率．謎解き臨床化学検査わかりにくい言葉・あやふやな事を明確に．医療検査と自動化．47: Suppl.1 144-146, 2022

2)日本臨床検査医学会ガイドライン作成委員会編.臨床検査のガイドライン JSLM2015.日本臨床検査医学会．2015

3)市原清志．多重ロジスティック分析による診断的有用性の評価方法.臨床検査．49：1341-35, 2005

4)Mark Hatherill, S N Tibby, K Sykes, C Turner, et al. Diagnostic markers of infection: comparison of procalcitonin with C reactive protein and leucocyte count. Arch Dis Child. 81:417–421, 1999

5)Paul D Stein, R D Hull, K C Patel, R E Olsen, et al, D-dimer for the exclusion of acute venous thrombosis and pulmonary embolism: a systematic review. Ann Intern Med. 20;140(8):589-602, 2004

6)肺血栓塞栓症および深部静脈血栓症の診断，治療，予防に関するガイドライン（2017改訂版）．合同研究班参加学会.2018　発行

3．個体内生理的変動を考慮したモニタリング時の結果の評価 [1)]

臨床検査データの信頼性を保証するには，信頼性評価の基準となる測定値の許容誤差限界を設定する必要がある。その許容誤差限界は生理的個体内・個体間変動に基づいて設定される。

継続的にデータをモニタリングする場合には、治療の効果判定、合併症の観察、病態の急激な変化などを観察するために、時系列的にデータを観察する。例えば 2 回の測定値間に変化が意味のある変化なのか、生理的変動を含めた分析の総変動を考慮すると意味のある変化とは言えない、といったことが生じているはずである。

さまざまな学会では診療・診断ガイドラインが作成され、検査データを基準とするものが多くある、その中で AKI（急性腎障害）診療ガイドライン [2)] では急性腎障害の診断基準では「CRE が 48 時間内に+Δ0.3 mg/dL の変化」としている。

動脈硬化性疾患・予防ガイドライン [3)] では「管理目標値は到達努力目標値であるが、一次予防では LDL-C 20～30%の低下により冠動脈疾患の発症が 30%低下することも示されていることより，20～30%の低下を目標となりうる」としている。

ここでは急性腎障害の診断の 48 時間内 Δ0.3 mg/dL の変化について触れてみる。

その前に、精密さおよびかたよりの許容誤差限界に触れておく。

許容誤差限界には①施設内の精密さの許容誤差限界 CV_A および
②真度・正確さの許容誤差限界 B_A の 2 つがある。

それぞれの許容誤差限界は

精密さ(施設内精度)：$CV_A(\%) < \frac{1}{2}CV_I$

真度（かたより）：B_A（%）$< \frac{1}{4} \times (CV_G^2 + CV_I^2)^{1/2}$

とされている。個体内生理的変動 CV_I 、 個体間生理的変動 CV_G

●$CV_A(\%)$の許容誤差限界を$\frac{1}{2}CV_I$ とする根拠 [5)] は、

総誤差＝$(CV_A^2 + CV_I^2)^{1/2}$ ・・(1) の式に

例えば（$CV_A = \frac{1}{2} CV_I$）の時、(1) 式に CV_A を$\frac{1}{2} CV_I$ に置き換えると

総誤差＝$(\frac{1}{2} CV_I^2 + CV_I^2)^{1/2} = (\frac{5}{4} CV_I^2)^{1/2} = 1.18\ CV_I$ となる。

$CV_A = \frac{1}{2} CV_I$ の時、個体内生理的変動が分析変動により 11.8%大きくなることを意味し、

$CV_A = CV_I$ とすると、個体内生理的変動が分析変動により 41.4%大きくなる。

このことより精密さ(施設内精度) $CV_A(\%)$の許容誤差限界を$\frac{1}{2}CV_I$ としている。

この結果を踏まえ Fraser[5] は CV_A と CV_I の関係が検査結果の変動に及ぼす程度を以下のごとくまとめている。

・CV_A が 3/4 CV_I のときは検査結果の変動に 25%の変動が加わる
・CV_A が 1/2 CV_I のときは検査結果の変動に 12%の変動が加わる
・CV_A が 1/4 CV_I のときは検査結果の変動に 3%の変動が加わる

1）モニタリング時の結果の評価の原理[5]

測定値には分析前変動、分析時変動 CV_A、および個体内生理的変動 CV_I などの変動が含まれる。これらの変動は正規分布すると考えられるので、SD あるいは CV%で表すことができる。ここでは分析前変動は軽減できることから無視できる程度であるとして扱わないこととする。

総変動 CV_T は以下のように計算できる。 1回測定の時の変動（95.5%の確率の時）

$CV_T = \pm 1.96 \times (CV_A^2 + CV_I^2)^{1/2}$

●最初の検査の変動と 2 回目の検査の変動のそれぞれの上下限が比較する。

注：測定値はそれぞれに図 1 のように変動をもつとすると、この図 1 では 1.5 mg/dL,1.8 mg/dL の 2 試料がそれぞれ SD 0.05 mg/dL および SD 0.10mg/dL の条件で 20 回の測定値の変動を示した。左側 2 つの分布は全く重なりがなく異なる試料（スモークボックス）の可能性がたかい。
右側（白ボックス）の SD 1.00mg/dL の 2 つのデータは重なっており、異なる試料であると判定することができない。

この右側の試料をそれぞれ 1 回測定し、その値が 1.5 mg/dL, 1.8 mg/dL であった時、われわれはあたかも異なる試料であると判定するだろう。
しかし、このような変動（SD 1.00mg/dL）を考慮すると、やはり両者は異なる試料であるとは言い切れないことがわかる。

このようなことをわれわれは測定値の変化が有意な差があるものか判定できるように準備しておく必要があり、こうした情報を臨床側に発信できるようにしたい。

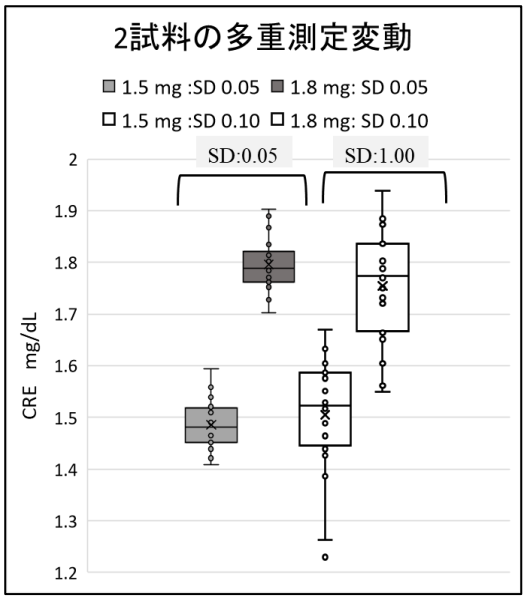


図 1 SD の違いによる 2 試料の分布幅

２）急性腎障害例によるモニタリング時の検査結果の評価

表１の急性腎障害の定義（AKI 急性腎障害診療ガイドライン 2016）[2] の「ΔsCr≧0.3 mg/dL48 時間以内」の条件を満たす血清クレアチニン CRE の室内再現性を加味して推定してみる。

表1　急性腎障害AKIN基準

定義	1. ΔsCr≧0.3 mg/dL（48時間以内）	
	2. sCrの基礎値から1.5倍以上（48時間以内）	
	3. 尿量0.5 mL/Kg/時以下が6時間以上持続	
	sCr基準	尿量基準
ステージ1	ΔsCr≧0.3 mg/dL or sCr 1.5～2.0倍上昇	0.5 mL/Kg/時未満 6時間以上

sCr: 血清クレアチニン
ΔsCr: 今回と前回のクレアチニンの差
AKIN:　Acute Kidney Injury Network
注: AKIN基準にはステージ3まで掲載されている。
・AKI（急性腎障害）診療ガイドライン　2016より改変引用

■　仮に１病日の CRE が 1.5mg/dL, 3 病日（48 時間以内）が 1.8 mg/dL であった，あるいは 3 病日 2.25 mg/dL（1.5 倍）であったとする。
この測定値の差 Δ0.3 mg/dL は有意な差があるといえるか検討してみよう。
●CRE の個体内変動 [4] を CV_I 5.4%，室内再現性を CV_I の 1/2 の CV_A 2.7%とした場合、1）の原理に基づくと

1.5 mg/dL では $1.96\times(CV_A^2+CV_I^2)^{1/2}=11.8\%$　となる。
1.5 mg/dL の 11.8%の範囲は　1.50 ± 0.18 mg/dL　1.32～1.68 mg/dL
1.8 mg/dL の 11.8%の範囲は　1.80 ± 0.21 mg/dL　1.59～2.01 mg/dL
2.25 mg/dL の 11.8%の範囲は 2.25 ± 0.30 mg/dL　1.98～2.55 mg/dL

◆1.5 mg/dL から 1.8 mg/dL への変化は、CV_A 2.7%では 95%の確率では両者の上限(1.68 mg/dL)と下限（1.59 mg/dL）は重なり有意な差があるといえず、急性腎障害の存在が否定される。
1.5mg/dL の基礎値から 1.5 倍の 2.25 mg/dL の場合(下限 1.98 mg/dL)には 95%の確率で有意な上昇と判定することができる。

表2　各試料の総分析誤差CVt95%の上下限値

	mg/dL		1.5			1.8		2.25
CV_A	%	5.4	2.7	1.35	5.4	2.7	1.35	2.7
SD	mg/dL	0.08	0.04	0.02	0.10	0.05	0.02	0.06
CV_I	%	5.4	5.4	5.4	5.4	5.4	5.4	5.4
CVt95%	%	15.0	11.8	10.9	15.0	11.8	10.9	11.8
値-CVt95% *	mg/dL	1.28	1.32	1.34	1.53	1.59	1.60	1.98
値+CVt95%*	mg/dL	1.72	1.68	1.66	2.07	2.01	2.00	2.52

*：1.5，1.8mg/dLのそれぞれの試料における95%確率で変動する上限、下限を示す

上記表 2 において CV_A を CV_I（5.4%）の 1.0CV_I、0.5 CV_I および 0.25 CV_I で計算したものである。0.25CV_I（CVA2.7%）は最も高い精度の変動条件であるが、1.5 mg/dL の上限 1.66 mg/dL と 1.8 mg/dL の下限 1.60 mg/dL でわずかに重なるところがあり 95%の確率では有意な差とみなすことができない。

ここで、採用した個体内変動[4]はメタアナリスから求めたものである。様々な期間のものが含まれているものと思われる。

急性腎障害では 48 時間内の判定のため、かなり短期間であるので個体内変動 CV_I をそれの 1/2(2.7%)にして、再計算したものが表 3 である。この条件とすると両者に重なりはなく、95%確率で有意な差を認める判定となった。急性腎障害を疑うことができる。

CV_A2.7%は 1.5mg/dL で SD 0.04mg/dL に相当し、1.0mg/dL では 0.027mg/dL となり、かなり高い精度が求められていることに注意が必要である。

表3　各試料の総分析誤差CVt95%の上下限値

	mg/dL	1.5	1.8	2.25
CV_A	%	2.7	2.7	2.7
SD	mg/dL	0.04	0.05	0.06
CV_I*	%	2.7	2.7	2.7
CVt95%	%	7.5	7.5	7.5
値-CVt95% *	mg/dL	1.39	1.67	2.08
値+CVt95%*	mg/dL	1.61	1.93	2.42

CV_I*を仮に2.7%とした場合の上限、下限を示す

３）ガイドラインと測定値の変動

今述べてきたように、ガイドラインの指標が検査室の精度管理のレベルにとって非常に厳しい条件となる場合がある。臨床のガイドラインでは、もしかすると検査室の精度管理の状況を十分に把握されずに設定されている可能性がある。これに応えられる精度を持った測定法の開発が待たれる。また、このような場合には、不確かさによる評価、特定の観察期間の個体内生理的変動による

評価などが必要となるのではないかと考える。

参考文献

1)山本慶和.個体内変動、個体間変動. 医療検査と自動化. 47. Suppl.1：115-118, 2022.

2)急性腎障害診療ガイドライン作成委員会.AKI（急性腎障害）診療ガイドライン 2016　日腎会誌. 59（4）：419-533, 2017.

3)動脈硬化性疾患・予防ガイドライン 2017 版.　日内会誌　107：73-8, 2018.

4)生理的変動に基づいた臨床化学検査 36 項目における測定の許容誤差限界.　臨床化学.　35：144-153, 2006.

5)Callum G. Fraser.　中　甫訳.検査データの生理的変動－原理から実践へ-医歯薬出版 2004.