

臨床検査室の認定を視野に置いたものであり、ここに含まれない実施母体を実施する精度管理調査は、ここで確立した標準化手法に強制されることはない。しかし、可能な限りその概念および手法を活用することにより調査結果の相互比較が容易になることはいうまでもない。

3) 外部精度評価における調査項目の現状とあり方

表2に臨床検査精度管理調査を実施している上記6組織が実施している調査項目を示した。ただし特定の実施母体のみで実施されている微生物検査、病理検査、細胞診検査は表から除いた。また調査項目名に続く()内は精度管理調査の標準化の項で述べた基幹検査項目第一類、第二類、第三類の項目であることを意味する。表から明らかなように、臨床検査実施施設の全てを対象としている日医調査、日臨技調査、日衛協調査をみると基幹検査項目第一類は日衛協調査でCKの調査を実施していないことおよび日医調査でアルブミンの調査を実施していないことを除き、これら3実施母体の全てが調査対象項目としている。基幹調査項目第一類について予防医学検査を対象としている全衛連調査、中央会調査、総合健医調査も含めてみるとブドウ糖、AST、ALT、 γ -GT、総コレステロール、中性脂肪、ヘモグロビンの7項目が表にあげた全調査で

調査対象項目として実施されていることが分かる。基幹検査項目第二類は、HDL コレステロール、HbA_{1c}が日臨技調査、総合健医調査で実施されていないこと、総蛋白が日医調査、全衛連調査で実施されていないこと、赤血球数が日臨技調査で実施されていないこと、白血球数、血小板数が中央会調査で実施されていないこと、ヘマトクリットが日臨技調査、中央会調査で実施されていないことを除き、多くの実施母体で調査対象項目としてあげられている。予防医学検査の中でとくに重要な尿蛋白、尿糖半定量検査は、日臨技調査と全衛連調査のみが調査対象項目としてあげている。また基幹検査項目第一類である尿素窒素、尿酸、クレアチニンは、表中全衛連調査のみが実施していないが、これは労働安全衛生法に基づく検査項目として指定されていないことによると考えられる。以上をまとめると、基幹検査項目第一類および第2類は、医療機関検査室および予防医学検査実施施設を含めた全ての臨床検査室に対する調査項目として実施することが望ましい調査項目といえることができる。

4) 外部精度評価における調査試料の現状とあり方

(1) 調査試料の種類

外部精度評価に用いる調査試料は、日常検査の対象となる新鮮な被検者の試料に限りなく近似した性質を有

する試料が望ましい。精度管理調査開始当初は内部精度管理調査用コントロール血清が調査試料として広く用いられた。それらのコントロール血清はとくに酵素、脂質などヒト血清とは異なる動物由来の成分が添加されているものが多く、施設内の測定値の精密さを管理する内部精度管理用試料としては十分使用できたが、ヒト血清と反応態度が異なり、測定結果が測定法により異なり日常検査の性能が直接反映されないことが多かった。そのことから測定値のかたよりの有無を主として評価する外部精度評価では不適切であることが指摘されてきた。その後酵素についてはヒト培養細胞由来酵素が添加されるようになりヒト血清に急速に近づいたが、一部の酵素でヒト血清と異なる性質を示した。1996年日本臨床化学会学術連絡委員会が“常用酵素標準物質の規格”を委員会報告として発表し⁹⁾、その中で酵素の起源としてヒト体液、ヒト組織、ヒト組織培養系、ヒト細胞培養系、およびヒト酵素をコードする遺伝子を導入した組替え体(recombinant)とするとした。この規格に基づいて1997年11月にヒト赤血球、ヒト組替え体を酵素の起源とした常用酵素標準物質(CERM)が完成しその頒布が開始された。その後CERMと同じ起源の酵素が調査用試料に添加されるようになり、わが国では酵素項目におけ

る調査用試料の問題がほぼ解消された。しかし、調査試料は調製試料であることから試料マトリックスが通常の被検者試料と異なり、一部の項目または測定法でマトリックス効果による測定値のかたよりが生じることが問題として残されている。とくにマトリックスの影響を受ける測定法群としてドライケミストリー・システムがあげられる。全衛連調査では調査対象項目が限定されていることから以上のような調製試料は使用せず、予めプールした検査済み高濃度(または高酵素活性)ヒト血清と検査済み正常ヒト血清を段階的に混和して8濃度の試料を作製して調査試料としている。

ヘモグロビン、血球計数を調査対象とする血液検査用試料は、試料の安定性の問題から大規模の外部精度評価プログラムではヒト全血を調査試料として用いることはできない。したがってヒト半固定細胞や動物由来の細胞を含む試料が用いられるが、計数装置の原理により機器間差が生じて日常検査の性能を直接反映できない問題がある。この問題を解消するために日衛協調査および地域レベルの精度管理調査のような比較的小規模な調査では、新鮮なヒト全血試料を短時間内に参加施設に送付する工夫をしてこれらの問題を解決している。

(2) 調査試料の形状

調査試料は同一試料を多数の施設に送付することと、遠距離施設も存在することから試料の安定性が重要なポイントとなる。安定性に関して通常取られる手法は調査試料を凍結乾燥品として供給することである。しかし、凍結乾燥によりリポ蛋白の変性が生じることから、とくに沈殿操作を行わず直接測定を行う HDL コレステロール測定試薬において測定値の試薬間差が生じて新鮮なヒト血清による性能を直接反映しないことがある。以上のことから、日医調査では総コレステロール、中性脂肪、HDL コレステロールなどの脂質測定用試料は液状試料（-20℃の凍結試料）を調査試料として送付している。全衛連調査では 1992 年の第 1 回調査から生化学検査項目について凍結血清を凍結状態で送付している。

（3）調査試料の本数

調査試料は一般に 2 濃度以上の複数試料が送付されることが多い。2 濃度の場合は正常濃度および異常濃度、3 濃度以上の場合は低、中、高濃度の組み合わせ試料が送付される。外部精度評価では日常測定シリーズ中での 1 回測定が原則であり同一試料の繰り返し測定は行わないことから、基本的には測定の精密さの直接評価はできない。しかし、一般に複数濃度の試料を送付することから、試料間の測定値比または散布図の作成により間接的に施設内測定値の精

密さの評価が可能である。さらに 3 濃度以上の試料を用いることにより測定値の試料濃度に対する直線性の評価が可能になる。しかし、各施設に同一試料が送付されることから、参加施設間での測定値のカンニングが生じる可能性があり外部精度評価プログラムの問題点の一つになっている。

全衛連調査では、（1）の調査試料の種類で述べた 8 試料を最低濃度または次に低濃度の試料のいずれかおよび最高濃度または次に高濃度の試料のいずれかが必ず送付されるという条件設定を行い他の濃度試料をコンピューターによりランダム化して合計 5 試料を参加施設に送付している。したがって測定値のカンニングができないプログラムになっており、さらに直線性の評価および誤差の分類評価が容易になるという長所がある。

（4）調査の頻度

大規模な精度管理調査では年 1 回実施されることが多い。その場合は実施時点での技術水準の提示および評価に限定され、生じている測定値の施設間差の厳密な原因究明は行いがたい。測定誤差の客観的な評価のためには少なくとも年 3～4 回は実施する必要がある。先述の外部精度評価プログラムの統合により、将来この条件を満たすことが可能になるであろう。とくに検査室の認定に関

連する技能試験による評価は継続的な性能の評価を欠かすことができない。

C. 測定誤差とその管理

測定法、測定試薬および測定機器を用いて測定された測定値には必ず誤差が含まれるが、許容できる最小限の誤差を長期間一定に維持することが測定値の信頼性を保証する上で重要である。測定誤差は測定値のばらつきに起因する偶然誤差と測定値のかたよりに起因する系統誤差に大別される。偶然誤差は突き止められない原因により生じる誤差で測定値の精密さを表し、各施設内で実施される管理図法を用いる内部精度管理法で容易に管理できる。系統誤差は何らかの原因で測定値にかたよりを与えることにより生じる誤差で測定値の正確さに関連するが、管理図を用いる内部精度管理法のみでは管理できない。測定値の正確さは、理論的には正確さの根拠が明確な基準分析法で測定した測定値または基準分析法を用いて標準値を求めた標準物質と比較することにより確認でき、また校正できる。一方、外部精度評価スキーム(EQAS)の結果は、参加各機関が施設内の測定値の精密さを十分管理している場合には、測定値のかたよりにより生じる系統誤差の客観的な評価法として有用性が高いことが知られている。しかし、測定

値のかたよりの原因となる因子は複数存在することから、関連する各因子に分類してその原因の究明と是正を行う必要がある。

1) 偶然誤差とその管理

偶然誤差は測定のばらつきとなって現れる誤差で、そのばらつきの小さい程度を精密さといい、ばらつきを標準偏差またはその指定した倍数で表した値を精密度と呼んでいる¹⁰⁾。1994年日臨技では定量検査の精密さ・正確さ評価法標準化ワーキンググループを組織して検討を開始し、1997年“定量検査の精密さ・正確さ評価法指針を公表した¹¹⁾。その後1999年にその改訂版を日本臨床検査標準協議会(JCCLS)から発行している¹²⁾。この評価指針によると精密さの評価法として、管理試料(コントロール血清)を用いる方法と患者試料を用いる方法をあげている。基本的には日常検査における同一試料の反復測定により得られた標準偏差を評価の物差しとしての個体内生理的変動の標準偏差と比較することにより評価する方法である。得られた評価が個体内生理的変動の1/2以下であれば精密度は許容できると判断し、その精密度を内部精度管理法としての管理図法(例えば $\bar{x}-R$ 管理図法)で管理する。偶然誤差の指標である精密度の評価方法の概略を下記に示した。

(1) 管理試料による評価¹²⁾

被検法が安定な状態にあるときに、管理試料を毎日2本以上、日常検査の中にランダムに挿入して測定し、これを20日以上繰り返す。その結果から分散分析法により日間精密度(SD_A)、日内精密度(SD_E)、総合精密度(SD_S)を計算し、総平均値が基準範囲上限以下のときは SD_S を個体内生理的変動の標準偏差($SD_w/2$)と比較し、上限以上のときは変動係数(CV_S)で評価するとしている。

(2) 患者試料による評価¹²⁾

被検法が安定な状態にあるときに、患者試料(試料数: n)に順番を付け、それらを試料として1回目の測定を実施し、2回目の測定を試料の順番をランダムに変更するかあるいは全く逆の順番にして同日内に測定する。それぞれの患者の2回測定値の平均値を横軸にとり、1回目と2回目の差を縦軸にとって差のプロット図を作成する。分散分析法により群内平方和(S_E)、誤差分散(V_E)、標準偏差(SD_E)を求め、この SD_E を $SD_w/2$ と比較して評価するとしている。

2) 系統誤差と測定体系

系統誤差は測定結果にかたよりを与える原因によって生じる誤差で、そのかたよりの小さい程度を正確さといい、推定したかたよりの限界の値で表した値を正確度と呼んでいる¹⁰⁾。具体的にはかたよりは測定値または推定量の分布の中心である平均

値と真の値(true value)との差で、濃度または平均値の真値に対する比率(%)で表現する。ここでいう真値とは測定量の正しい値と定義されるが、先述のごとく誤差を含まない測定値は存在しないことから、観念的な値であって実際には求められないことから、真値とみなし得る値が真値と見做して用いられる¹⁰⁾。真値とみなし得る値については化学分析における正確さを基本とした測定体系の考え方が参考となる。

(1) 化学分析における正確さを基本とした測定体系

1977年米国 NIST (National Institute of Standards and Technology, 当時は NBS) の Uriano および Gravatt は、化学分析における正確さを基本とした測定体系を発表した¹³⁾。その後米国 NCCLS ではこの考え方を臨床検査に導入し一連のガイドラインを発表している^{14~16)}。わが国でも日本臨床化学会を中心にこれらの測定体系の考え方を導入している¹⁷⁾。図1に分析化学における正確さを基本とした測定体系を臨床化学検査の測定体系として示した。以下に図中の各部分について解説する。

① 基本的測定単位: 1960年国際度量衡総会において単位の統一化を図るために国際単位系(SI 単位)を決議し、国際標準化機構(ISO), 国際純正応用化学連合

(IUPAC)などの協力を得て普及の努力が続けられている。1977年 WHO は第 30 回世界保健会議で医学領域への SI 単位系導入を推奨した。1994 年日本臨床化学会学術連絡委員会では遅ればせながら SI 単位換算表を作成し¹⁸⁾、臨床検査関連学会に提示して普及の努力を始めた。SI 単位は長さ、質量、時間、電流、熱力学的温度、光度、物質の量の 7 種類の基本的単位を設定している。この中で分析化学に直接関連する単位は質量と物質の量であるが、物質の質量を天秤で直接測定できる場合は最も誤差の少ない結果が得られる。しかし臨床化学検査における微量の成分分析は直接測定はできないことから何らかの分析法により測定せざるを得ない。SI 単位で直接測定したときのかたよりは約 10⁻⁸%と考えられている¹⁹⁾。

- ② (絶対) 基準法: 測定原理的に誤差が最低の水準に抑えられる測定法で、これは理論的にも実験的にも証明されている方法と定義される¹⁷⁾。この方法は同位体希釈・質量分析法(ID-MS)に代表される測定法で²⁰⁾、特別な技術、費用、測定装置などを必要とすることから認定を受けた基準測定検査室(reference measurement laboratory)において実施できる

測定法である。この測定法で測定したときのかたよりは約 0.1~1%と考えられている¹⁹⁾。この方法は後に述べる实用基準法の開発および評価、一次標準物質の値付けに用いられる。

- ③ 一次標準物質: 標準物質の特性値が、(絶対) 基準法によって決定されたものと定義される¹⁷⁾。一次標準物質は实用基準法の開発と評価、正確度の高い二次標準物質の作製、厳密な精度保証の目的で使用される。
- ④ 实用基準法: 測定体系上(絶対) 基準法に次ぐ測定法であり、十分に研究され、必要条件ならびに手順が明確に記述され、目的とする用途に相応した正確さと精密さをもった値が得られる測定法と定義される¹⁷⁾。国際学会または国内学会において勧告される基準分析法がこの方法に相当する。この方法は特別な装置を必要とせず、熟練した技術を有すれば実施できる測定法である。この測定法で測定したときのかたよりは約 1~3%であると考えられている¹⁹⁾。この方法は後に述べる日常一般法の開発および評価、日常に使用する検量用試料の作製、内部精度管理、外部精度評価における測定の正確度の評価に使用される。
- ⑤ 二次標準物質: 標準物質の測定値

が、実用基準法によって決定されたものと定義される¹⁷⁾。二次標準物質は実用基準法と同様に日常一般法の開発および評価、日常に使用する検量用試料の作製、内部精度管理、外部精度管理における測定の正確度の評価に使用される。

- ⑥ 日常一般法: 多数の測定試料に応じることができ、日常的に用いる精密さを有する測定法と定義される¹⁷⁾。日常検査室で一般的にルーチン法として採用されている方法がこれに当たる。これらの測定法で測定したときのかたよりは5~10%であると考えられている¹⁹⁾。
- ⑦ 血清ベース検量用試料: 日常一般法の検量用試料としては目的成分の純品を秤量して精製水などの溶媒に溶解することにより作製する溶媒ベースの検量用試料とヒト血清（あるいはアルブミン）をベースとして必要に応じて目的成分を添加した血清ベースの検量用試料がある。血清ベースの検量用試料は上記基準分析法または標準物質による値付けの必要がある。現在日常一般法では自動分析装置を利用して分析を行うことが一般的であるが、自動分析装置では被検試料と検量用試料を同一試料採取ノズルで採取することから、溶媒ベースの検

量用試料では比重および粘性が異なることから、採取量に差が生じて系統誤差の原因となる。したがって血清ベース検量用試料が主として用いられるが、一方測定により値付けを行うことから値付けにおける測定誤差がかたよりの原因となる可能性がある。

- ⑧ トレーサビリティ: 測定体系の上でより高い正確さに下位から次々と合わせるができることと定義される¹⁷⁾。例えば日常一般法による測定値の正確さを実用基準法を用いて確認したり、一次あるいは二次標準物質を日常一般法を用いて繰り返し測定することにより確認することをトレーサビリティの確認と呼ぶ。
 - ⑨ 伝達性: 測定体系の上で高位の正確さを順次下位のものに合わせていくことと定義される¹⁷⁾。例えば一次あるいは二次標準物質の標準値を血清ベース検量用試料に伝達するときなどに用いられる。
 - ⑩ 校正: 標準物質あるいは基準分析法による測定値を用いてかたよりを補正することと定義される¹⁷⁾。例えばトレーサビリティの確認により標準物質の標準値と日常一般法の間にかたよりがみられたとき、そのかたよりを補正することを校正するという。
- (2) 酵素活性検査の測定体系

酵素活性の測定は、基質の種類および濃度、緩衝液の種類および濃度、pH、測定温度などの条件を至適化した一定条件における生物学的活性を測定していることから（絶対）基準法は存在しない。図2に日本臨床化学会による測定体系を示した。日本臨床化学会では、現在AST, ALT, CK, ALP, LD, γ -GTの6種類のヒト血清中酵素活性測定の方法を公表しているが、それらの酵素についてこの測定体系が適用される。以下に図中の各部分について解説する。

- ① JSCC 勧告法：この方法は臨床化学検査の測定体系（図1）における実用基準法に相当する。この方法は、pH、基質濃度はでき得る限り最適とし、条件を明記する、温度も必ず記載し、でき得れば30°Cが望ましいという1964年の国際生化学連合(IUB)の勧告に従って測定条件の設定を行い、測定温度は30°Cとした。この測定法を日本臨床化学会(JSCC)勧告法と呼んでいる。
- ② JSCC 常用基準法：先のJSCC 勧告法は測定温度を30°Cとして設定されているが、世界的に臨床検査領域における日常検査の酵素活性の測定は37°Cで実施されていることから、JSCC 勧告法の条件を測定温度のみ37°Cとする方法をJSCC 常用基準法と呼び、この方法で得られた活性値を日常

検査法の基準とした。したがって、これらの酵素について日常検査法の正確さを確認（トレーサビリティの確認）するためにはこの方法と比較することになる。

- ③ JCCLS 認証常用酵素標準物質 (CERM)：1996年日本臨床化学会学術連絡委員会は、常用酵素標準物質の規格を委員会報告として発表し⁹⁾、1997年その規格に基づいて作製した酵素標準物質についてJSCC 常用基準法により値付けを行いJCCLS 認証委員会により認証された“JCCLS 認証酵素標準物質”を日本臨床検査薬協会を通して供給を開始した。現在第3ロットが供給されている。
- ④ 日常検査法(JCCLS 標準化対応法)：JSCC 実用基準法で測定した値が、CERM や検量用ERMなどを利用することにより伝達されている場合にその測定法をJSCC 標準化対応法と呼ぶ。このJSCC 標準化対応法は基本的には用手法であるが、JSCC 常用基準法に準拠した試薬を使用し自動分析装置で実測Kファクターを求め、温度を37°Cに管理して測定されていれば適用される。またCERMを用いてトレーサビリティが確認されている測定法および企業から供給されるCERMの標準値が伝達されている検量

用ERMを用いて検量を行った測定法も全てJSCC標準化対応法となる。すなわちJSCC実用基準法およびCERMの値を伝達することにより測定値の標準化を図ったと考えてよい。

- ⑤ 検量用ERM：これはJSCC標準化対応法の検量目的のため、企業がCERMの標準値を伝達した検量用酵素標準物質をいう。したがってこのERMの信頼性の責任はそれを提供した企業側に存在する。

3) 測定の正確さの評価

臨床化学検査における測定の正確さの基準は、先の臨床化学検査の測定体系における(絶対)基準法、実用基準法のような基準分析法により得られる測定値および一次標準物質、二次標準物質の標準値である。そのような基準分析法による測定値および標準物質の標準値を基準とした測定の正確さ評価法の概略について以下に示した。

- (1) 1種類の濃度の血清標準物質による評価^{1,2)}

正確さの明確な血清標準物質を添付されている認証書の指示に従って調製し、被検法が安定な状態にあるときに10回以上繰り返し測定する。その結果から平均値、標準偏差を求め、平均値の95%信頼区間を計算する。血清標準物質の標準値(X_r)、平均値(X_B)からかたより(B)を計算す

る($B = X_B - X_r$)。計算された平均値の95%信頼区間に X_r が含まれていれば、平均値の付近の濃度領域における有意なかたよりはないとみなす。信頼区間から外れた場合は $100 \times |B| / X_r$ が5%以下(N_a, K は2%以下)かどうかで正確さを判断する。

- (2) 3種類以上の濃度の血清標準物質による評価^{1,2)}

3種類以上の濃度の血清標準物質の使用が可能なときは、血清標準物質の標準値に対する被検法測定値の関係を直線回帰式($Y = a + bx$)で表し、被検法の正確さを一定系統誤差($a=0$ からのかたより)と比例系統誤差($b=1$ からのかたより)の大きさを評価できる。それぞれの血清標準物質を添付されている認証書の指示に従い調製し、被検法が安定な状態にあるときに各標準血清を5回以上繰り返し測定する。その結果から直線回帰式を計算し、切片 a および勾配 b のかたよりの有無の有意検定を行う。切片、勾配の両方に統計的にかたよりがないとみられたときは、一定系統誤差、比例系統誤差を認めないと判断する。もし一方にかたよりが認められたときは、評価対象となる濃度〔医学的意思決定濃度(x_0)〕における予測値($Y_0 = a + bx_0$)のかたより B の x_0 に対する割合($|B| / x_0 = |Y_0 - x_0| / x_0$)を求め、この値が5%以下(N_a, K は2%以下)

かどうかで正確さを判断する。

(3) 患者試料による比較対照法との比較実験による評価^{1,2)}

多数の患者試料の比較対照法による測定値に対する被検法による関係を直線関係式($Y=a+bx$)で表し、被検法の正確さを一定系統誤差($a=0$ からのかたより)と比例系統誤差($b=1$ からのかたより)の大きさを評価できる。比較対照法と被検法が安定状態のときに、用意した 50 例以上の患者試料をそれぞれ両法で測定する。それぞれの測定結果について散布図と差のプロット図を作成し、直線性とはずれ値の有無を検討し、はずれ値があれば統計的手法により除外する。比較対照法にも誤差が存在することを仮定した直線関係式を計算する(詳細は省略)。一定系統誤差と比例系統誤差の有無についてブートストラップ法を利用して検討し(詳細は省略)、信頼区間の中に $a=0$, $b=1$ が入っていれば一定系統誤差、比例系統誤差がないと判断する。もし一方にかたよりが認められたときは、評価対象となる濃度〔医学的意思決定濃度(x_0)〕における予測値($Y_0=a+bx_0$)のかたより B の x_0 に対する割合($|B|/x_0 = |Y_0-x_0|/x_0$)を求め、この値が 5%以下(N_a, K は 2%以下)かどうかで正確さを判断する。

D. 精度管理調査における測定値の

評価

種々の試薬および測定機器を用いて測定された測定値には必ず誤差が含まれるが、許容できる最小限の誤差を長期間一定に維持することが測定値の信頼性を保証する上で重要である。測定誤差は先述のごとく偶然誤差と系統誤差に大別される。偶然誤差は各施設内で実施される管理図法を用いる内部精度管理法で容易に管理できる。系統誤差は測定値の正確さに関連するが、管理図を用いる内部精度管理法のみでは管理できない。測定値の正確さは、理論的には正確さの根拠が明確な基準分析法で測定した測定値または基準分析法を用いて標準値を求めた標準物質と比較することにより確認でき、また校正が可能となる。一方、同一試料を多数の検査室で測定することにより測定値の信頼性を評価する精度管理調査の結果は、参加各機関が施設内の測定値の精密さを十分管理している場合には、測定値のかたよりにより生じる系統誤差の客観的な評価法として有用性が高いことが知られている。しかし、測定値のかたよりの原因となる因子は複数存在することから、関連する各因子に分類してその原因の究明と是正を行う必要がある。

1. 測定値のかたよりの原因となる因子

表 3 に測定値のかたよりの原因と

なる主な因子をあげた。これらの各因子が最終的に測定値の正確さに影響を与える。

1) 測定法

同一生体成分の定量を目的とした種々の測定原理の測定法が数多く存在する。項目により国際または国内学会から勧告されている基準分析法が既に存在するが、これらの基準分析法は試薬の作製および測定に関して用手法が基本であり、自動化測定装置を用いる検査が中心となっている日常検査法として基準分析法を直接適用することはできない。また基準分析法が勧告されている成分はごく一部にすぎないので、実際の日常検査は測定原理が異なる複数の測定法が利用されている。臨床化学検査では、反応原理に化学反応および酵素を試薬とした酵素的測定法が主流を占めているので、測定法間差は抗原抗体反応を原理とする免疫学的測定法に比較して一般に小さいが、反応特異性の差異による測定法間差が存在することがある。したがって、精度管理調査では参加施設の報告値を現在使用されている主な測定法に分類後評価している。

2) 試薬

同一反応原理の測定法であっても原料となる化学試薬（あるいは酵素試薬）の原料および試薬キットとしての組立の差異により測定値の試薬間差が生じることがある。したがっ

て、精度管理調査では測定値の試薬間差の有無をチェックするため、参加施設の報告値を試薬キットメーカー別に分類して評価の参考としている。もし特定の試薬キットにおいて有意の差が生じている場合は、その結果の全体群への影響を避けるため独立評価される。

3) 測定装置

生体成分の定量検査において全自動分析（測定）装置または検出系に測定装置が使用される。全自動分析（測定）装置には、いずれの試薬でも適用できるオープン試薬方式と装置製造元が適用試薬を限定している固定試薬方式がある。オープン試薬方式の装置は試薬と装置がそれぞれ独立してかたよりを生じる原因となり、固定試薬方式の装置は、試薬を含む分析（測定）システムとしてのかたよりの原因となる。精度管理調査では、全ての項目で使用している測定装置の調査を実施しているが、臨床化学検査の分野では固定試薬方式のドライケミストリーシステム以外の測定装置は評価上の参考としてのみ利用している。

4) 検量用試料（キャリブレーター）

臨床化学検査では、一般に濃度既知の標準物質あるいは検量用試料（キャリブレーター）を被検試料と同時に測定し、その吸光度から被検試料の濃度が計算される。これらの標準物質には先述のごとく溶媒ペー

スの標準液と血清中の成分を基準分析法により測定することにより値付けを行った血清ベースのキャリブレーターがある。現在市販されている多くの試薬キットでメーカー指定の標準液または血清キャリブレーターが添付（あるいは別売り）されているが、とくにメーカー指定の血清ベースキャリブレーターは、メーカーが値付けを行うことから、その信頼性の責任は試薬メーカーに帰属する。

5) トレーサビリティ確認の有無

各施設で測定されている成分の測定値の正確さを確認するには、より上位の基準分析法と日常一般法との比較実験の実施または基準分析法により値付けされた標準物質を日常一般法で繰り返し測定することにより達成できる。したがって、トレーサビリティの確認を実施している施設群の測定値が、全測定法群の中で最も正確な測定値を反映していると考えられることができる。日医調査では、2000年(平成12年度)から測定体系が確立している成分についてトレーサビリティの確認の有無について調査を行っている。調査成分により多少異なるが、トレーサビリティの確認を行っている施設は参加施設全体の十数%に過ぎないのが現状である。

2. 同一測定法群の分類と付帯調査

同一測定項目の測定値は本来誤差許容範囲内で一致しなければならな

いが、測定値にかたよりを与える因子は数多く存在することから、報告された測定値には許容誤差範囲を超える施設間差が存在する。その原因の明確化と是正を図るため、測定値の評価に先だって大分類として全参加施設の測定値を測定法の測定原理、試薬などが共通する同一測定法群(peer group)に分類して評価される。項目によっては測定装置によりかたよりが生じる可能性がある場合は同一測定装置を使用する施設を同一測定法群として分類される。さらに小分類として同一測定法群内でかたよりが生じる可能性がある検量用試料(キャリブレーター)の種類、トレーサビリティ確認の有無などの付帯調査が行われる。

3. 一次統計処理

1) 極端値の棄却

同一測定法群に分類されたデータには、極端な誤差、記入ミスなどによる極端値が存在する。このような極端値は統計計算された統計量に影響を与えることから極端値を棄却する必要がある。一般に用いられる手法は、同一測定法群の全データから平均値(M)、標準偏差(SD)を計算し、 $\pm 3SD$ を越えるデータを極端値として棄却する。場合によっては1回棄却後のデータから再度 $\pm 3SD$ の棄却をすることもある。

2) 統計計算

同一測定法群ごとに極端値棄却後

に残ったデータから平均値(M)、標準偏差(SD)、変動係数(CV)を計算し、そのデータを一次集計として以後の評価の基本データとする。

4. 精密さの評価

測定の精密さの評価と管理は、本来施設内における内部精度管理により達成すべきものであって、精度管理調査の報告値から精密さの直接評価はできない。しかし、多くの精度管理調査において濃度の異なる2濃度以上の複数の調査試料が配布されることから、施設内での試料間の測定値の比を計算することにより施設内における精密さの間接的評価が可能になる。

5. 正確さの評価

精度管理調査における測定値の真値(true value)とみなし得る目標値(target value)としては、歴史的に極端値棄却後の同一測定法群の平均値が用いられてきた。しかし、同一測定法群の平均値を真の値とみなし得るかについてはかねてから論議的になっていた。臨床化学検査および酵素活性検査の測定体系と正確さの概念が明確になった現在、基準分析法および標準物質が存在する項目については調査試料をそれらの基準分析法または標準物質を基準として調査試料を測定することにより、試料ごとの目標値を設定すべきであるとの議論が高まっている。1992年全衛連臨床検査専門員会は、問題点改善

を図るため専門委員会内に「参考値検討委員会」を設けた。同委員会では基準分析法と基準分析法を用いて標準値が付けられた標準血清が存在する項目について、標準血清を用いて調査試料の目標値に相当する参考値を求め、1993年度からその参考値を目標値として参加施設の評価を行ってきた²¹⁾。一方、日医精度管理調査検討委員会は、改善の試みの一つとして1996年からpeer group内における測定値の分布の歪みを是正するため、peer groupの3SD1回切断後のデータに反復切断補正法を適用して調整平均値(M_{adj})を求め、それを目標値として用いてきた。また目標値設定の必要性に関する論議が高まっていることから、1999年に検討委員会内に「目標値設定小委員会」を組織して、目標値設定に関する研究を行い、翌2000年に研究結果の報告書を発行した²²⁾。

1) 全衛連における参考値設定とその結果

(1) 対象項目

基準分析法および標準血清が存在する総コレステロール、中性脂肪、AST、ALT、 γ -GT、グルコースの6項目を対象とした。

(2) 基準測定検査室

参考値検討委員会委員が属する6施設とした。

(3) 精度管理調査試料

① 総コレステロール、中性脂肪測定

用試料：日常検査終了後の高コレステロール血清をプールし、別途プールした日常検査終了後の正常血清と段階的に混和することにより1～8濃度の試料を作製して除菌濾過後凍結試料とした。

② AST、ALT、 γ -GT、グルコース測定用試料：日常検査終了後の高AST、ALT血清をプールして製造直前にグルコースを添加し、別途プールした日常検査終了後の正常血清と段階的に混和することにより1～8段階の試料を作製して除菌濾過後凍結試料とした。

③ 配付試料：上記試料のうち低値試料1, 2のいずれかおよび高値試料7, 8のいずれかが必ず入る前提条件を設定し、残りをランダム化して合計5試料を凍結状態で各参加施設に送付した。

(4) 参考値測定方法

- ① 各基準測定検査室において各試料を1日5重測定、3日間繰り返し測定する。
- ② 同時に福祉・医療技術振興会(HECTEF)の標準血清およびJCCLS 認証常用酵素標準物質(CERM)を①と同様の手法で測定する。
- ③ 個々のデータについて標準値による校正を行う。
- ④ 全データ(15×6=90)から平均値±2SDを越えるデータを棄却後、

平均値を求めて参考値とする。

- ⑤ 参考値を目標値として調査参加各施設の結果を評価する。

(5) 結果

- ① 総コレステロール：図3に2001年度の脂質測定における参考値と参加施設の平均値の関係を示した。図中45°の対角線は平均値と参考値が完全に一致したときの位置を示しており、線で結ばれたプロットは各試料における参加施設の平均値の位置を示している。総コレステロール(全測定法)は左図から明らかなように平均値は参考値に比較し僅かに低値を示した。参考値に対する平均値の%(一致率)を求めると試料濃度に関係なく約99%であり、良好な一致を示した。
- ② 中性脂肪：図3の右図は中性脂肪(グリセロール消去法)の結果である。図から明らかなように総コレステロールに比較して参加施設の平均値がやや低く濃度に関係なく一致率は約98%で、約2%の負の比例系統的な差を示した。
- ③ AST：図4に2001年度の酵素活性測定における参考値と参加施設の平均値の関係を示した。左図のAST・JSCC標準化対応法の参加施設の平均値は参考値に対して僅かに高いがほぼ一致した。低活性試料では1単位ずれても%は大きく変動するので、高活

性部分の一致率をみると 102% 前後の一致率を示した。

- ④ γ -GT: 図 4 右図に γ -GT・JSCC 標準化対応法の結果を示した。図から明らかなように参考値と参加施設の平均値はほぼ一致した。参考値に対する平均値の一致率は 101%であった。
- ⑤ グルコース: 図 5 に 2001 年度のグルコース測定における参考値と参加施設の平均値の関係を示した。グルコース測定では日常検査法として種々の測定法が採用されているため、ここではヘキソキナーゼ・UV 法 (HK-UV)、グルコース酸化酵素・電極法 (GOD-電極)、グルコース脱水素酵素・UV 法 (GDH-UV)、グルコース酸化酵素比色法 (GOD-POD) について記号を変えて示した。図から明らかなように測定法により参加施設の平均値が異なるが、参考値に対する一致率は $100 + 1 \sim 2\%$ の範囲に入った。

(6) 全衛連検討結果の総括

全衛連は労働安全衛生法に基づく職場健診を実施している機関の全国組織で、全衛連が実施する 2001 年度精度管理調査参加施設は 325 施設であった。全衛連調査では過去 10 年間にわたって基準分析法を用いて値付けされた標準血清が存在する項目について、標準血清の標準値により校正された参考値を求めて参加施

設の結果の評価を行ってきた。酵素項目は、CERM 供給以前は実測 K 値を用いて参考値を求めてきたが、CERM 供給開始以来 CERM により校正された参考値で評価を行っている。ここでは 2001 年度の調査試料で求めた参考値と参加施設の平均値を比較した。総コレステロールは、測定法に分類せず全測定法を peer group としている。総コレステロールの参考値は調査開始当初から HECTEF から供給されている脂質測定用標準血清を用いて校正されている。参考値に対する参加施設平均値は当初から $100 \pm 1\%$ の範囲で極めてよく一致しており、参加施設および検量用試料を供給する試薬メーカーともに基準分析法または標準血清を用いることによるトレーサビリティの確認を広く行っていることを示唆している。中性脂肪は参加施設の 98% がグリセロール消去法を使用しているため、結果にはグリセロール消去法のみを示した。中性脂肪の参考値も脂質測定用標準血清を用いて校正されている。中性脂肪は総コレステロールに比較して参考値に対する参加施設の平均値の一致率がやや低く 98% 程度である。この傾向は過去の年度においても同じで、96% から経年的に一致率が上昇している。参考値の校正に用いた標準血清は、日本臨床化学会 (JSCC) による勧告法²³⁾を用いて標準値を求め、

同時に米国 NIST (National Institute of Standards & Technology) から供給されている SRM 1951a²⁴⁾ を測定することにより正確さの確認が行われている。このように中性脂肪の一致率が低い原因の一つとしてメーカーが供給している検量用試料の値付けに JSCC 勧告法または脂質測定用標準血清の使用が徹底されてこなかったことがあげられる。その理由に JSCC 勧告法による測定には高度の技術を要すること²³⁾と、SRM1951a におけるグリセロール消去法の標準値の取り扱い²⁴⁾ に関して理解が乏しかったことが考えられる。しかし、経年的に一致率が上昇していることから、今後さらに一致率の向上が期待される。AST は参加施設の 98% が JSCC 標準化対応法を採用している。全衛連では 1999 年度調査から参考値の校正に CERM を用いてきた。2001 年度の参考値に対する参加施設の平均値の一致率は 102% 前後であるが、過去の年度の結果は 99%~102% の範囲にあり、今後も 100±2% の範囲内で推移することが期待される。γ-GT は参加施設の 89% が JSCC 標準化対応法を採用している。γ-GT の参考値は AST と同様に 1999 年度調査から CERM を用いて校正されてきた。2001 年度の参考値に対する参加施設の平均値の一致率は 101% であるが、過去の年度の結果は 100%

前後で極めて良好であり、今後も 100±1% の範囲内で推移することが期待される。グルコースは 4 種類の測定法に分散している。グルコースの参考値は HECTEF から供給されているグルコース測定用標準血清を用いて校正されている。2001 年度の結果からも明らかなように参考値に対する各測定法の参加施設の平均値はそれぞれ異なり 100~102% の範囲内に入っている。グルコースについては参考値による評価を 2001 年度から開始したので過去の十分なデータが存在しないので今後の推移を観察したい。

2) 日医における目標値設定とその結果

(1) 対象項目

基準分析法および標準血清が存在する Na、K、Cl、Ca、尿酸(UA)、ヘモグロビン(Hb)の 6 項目を対象とした。表 4 に対象とした項目および目標値設定に用いた基準分析法を示した。

(2) リファレンス検査室の選択

基準分析法を使用して測定可能な検査室を項目ごとに 5 施設以上 (ただし Ca は 4 施設) を臨床検査関連学会、協力施設、団体等に推薦を依頼した。

(3) 書類審査および熟達度試験の実施

推薦された検査室を書類審査の上、熟達度試験を実施した。その概要を

表5の上段に示した。熟達度試験に使用した試料はHb測定用試料を委員会主導で作製した以外は、HECTEFから供給されている2濃度の標準血清を使用した。測定は各試料について表5に示した多重測定と反復測定を実施した。推薦された全ての検査室が書類審査および熟達度試験に合格した。

(4) 目標値の測定法

- ① 各リファレンス検査室において2000年度の調査試料について表5の下段の条件で測定した。
- ② 項目、各試料についてリファレンス検査室の総平均、施設間CV、施設内CVを分散分析法(ANOVA)により計算した。
- ③ 精度管理調査参加施設の総平均、方法間CV、方法内CV、共通CV(後述)をANOVAにより計算した。リファレンス検査室の結果と調査参加施設の結果の比較を行った。

(5) 結果

表6に2000年度に実施した日医臨床検査精度管理調査における調査試料のリファレンス検査室による目標値設定の結果と同年度の参加施設結果を示した。リファレンス検査室の結果は、項目別、試料別に施設間CV、施設内CV、総平均を示した。参加施設の結果は、方法間CV、方法内CV、共通CV(後述)、総平均を示した。またリファレンス検査室

の総平均に対する参加施設の平均値の一致率(%)を表の右端に示した。

- ① Na、K、Cl：表から明らかなようにNa、K、Clでは、基準分析法による施設間CVが参加施設の方法間CVがやや大きく、施設内CVは参加施設の方法内CVより小さい結果が得られた。共通CV(common-CV)は方法内CVより僅かに小さい。総平均はNaは100%、K、Clは99%の一致率が得られた。
- ② Ca：表から明らかなようにリファレンス検査室の施設間CVは、試料1(S-1)では参加施設の方法間CVより僅かに小さいが、試料2,3(S-2,3)では大きく、施設内CVは参加施設の方法内CVより著しく小さい結果が得られた。総平均は、試料により異なり97~100%の一致率が得られた。
- ③ UA：表から明らかなようにリファレンス検査室の施設間CVは、全ての試料で参加施設の方法間CVより小さく、施設内CVも参加施設の方法内CVより著しく小さい結果が得られた。総平均は全試料で100%の一致率を示した。
- ④ Hb：表から明らかなようにリファレンス検査室の施設間CVは参加施設の方法間CVより僅かに大きく、施設内CVは参加施設の方法内CVより小さい結果が

得られた。総平均はいずれの試料も参加施設の総平均がリファレンス検査室の総平均より約 0.2g/dl 低値を示し一致率は 98% 程度であった。

(6) 日医検討結果の総括

日医における検討は、リファレンス検査室により基準分析法を用いて 2000 年度の調査試料について目標値の設定を試みた。表 6 では、最終的にリファレンス検査室の総平均と参加施設の総平均を比較した。総平均は Ca の一部の試料およびヘモグロビンを除き比較的良好な一致が見られたが、さらに詳細に検討した結果を踏まえて以下に述べる。Na、K、Cl はいずれも 99% 以上の一致を見ているが、中でも ISE 群の平均値はほぼ完全に一致した。この項目については、HECTEF から長年にわたりイオン電極用一次標準血清および常用標準血清が供給されており、機器メーカーは一次標準血清を用いて自社の測定機器の校正を行っており、ユーザーも一次標準血清または二次（常用）標準血清を用いてトレーサビリティの確認を日常的に行っていることによる成果と考えられる。したがって、測定機器が一次標準物質で校正され、二次標準物質でトレーサビリティが確認された peer group の調整平均値を目標値として採用できる可能性を示唆している。Ca は、低、中濃度試料では一致率

が比較的低く、高濃度試料では良好な一致を見ている。しかし測定法別に結果を詳細に検討すると目標値に一致する測定法は見られなかった。現在 Ca の実用基準法は原子吸光分析となっているが、現在日常検査室で原子吸光法を実施できる施設が少なくなっていることから、早急に専門学会等による基準分析法の開発が必要である。当面は各製造会社において原子吸光法による社内標準の設定が望ましい。UA は、リファレンス検査室の総平均と参加施設の総平均は他の項目に比較して高い一致率が得られているが、とくにウリカーゼ・ペルオキシダーゼ法、ウリカーゼ・カタラーゼ法の参加施設の平均値は極めてよく一致した。UA については、HPLC 法およびその方法で値付けされた標準物質使用群の平均値は目標値として採用できる可能性を示唆している。Hb は、リファレンス検査室の総平均に対して参加施設の総平均が約 0.2g/dl 低値を示したが、可能性として各製造会社における社内標準の設定方法の際の全血試料のピペット操作の違いが原因となっている可能性があり、原因の解明が必要である。

3) 目標値設定による評価の可能性
外部精度管理調査における目標値設定による評価の可能性について、全衛連参考値検討委員会の活動成果および日医目標値設定小委員会の研

究結果を通して、主として臨床化学検査を中心に検証を行い以下の結論を得た。

(1) 基準分析法および基準分析法により値付けされた標準血清が存在する項目で求められた目標値は参加施設の平均値と高い一致率を示した。

(2) 項目により目標値と平均値が一致する測定法と一致しない測定法が存在した。これらの項目は、日常測定法、検量用物質、トレーサビリティの確認等、検討すべき課題が今後に残されている。

(3) 基準分析法または標準物質で校正された検査室が増加することにより、目標値と参加施設の平均値がさらに高まることが期待できる。

(4) 目標値を求めるには、リファレンス検査室の性能の維持と多大な労力を必要とし、将来 National External Quality Assessment Scheme(NEQAS)により少なくとも年3回実施されるようになると、各回ごと、各項目ごと、各試料ごとに目標値を設定することは極めて困難と考えられる。

最終結論として、基準分析法、基準分析法用いて値付けされた標準物質で校正可能な項目については、基準分析法または標準物質を用いてトレーサビリティの確認が行われている参加施設群の平均値を目標値として採用しても差し支えないと考えられる。

6. 評価規準幅

一般的に同一測定法群の目標値±2SD(95%信頼範囲)が現行の技術水準(state-of-the-art)を表現すると考えられていることから、目標値からの測定値のかたよりの評価規準幅については、歴史的に標準偏差を基本とした考え方が主流を占めてきた。しかし、測定体系の確立の結果として各施設における正確さの評価と校正が行われるようになり、測定値の施設間差が近年著しく縮小している。そのような現状を踏まえて考えると、統計的結果を基本とした評価範囲は、施設間差の大小に関係なく、常に一定比率で許容できないデータが発生することになり、同一測定法群の標準偏差が大きい測定法群では評価が甘くなり、施設間差の小さい項目では必要以上に許容限界が厳しくなることが指摘されてきた。以上のことから、項目ごとに医学的有用性に基づく誤差許容限界の設定の必要性が論議されてきた。日本医師会臨床検査精度管理検討委員会では、この点についてかねてから論議してきたが、測定値の医学的有用性についてはそのデータを利用する臨床医(または産業医)により目的と判断基準が異なることから画一的に決められないという結論に達した。1996年日本医師会は現状の欠点を補う評価範囲を検討し、共通CV法による評価を開始した^{2,5)}。この評価法の導入により

問題点がかなり修正されたが、評価規準幅（共通 CV）が日常検査業務における報告単位幅より小さく必要以上に厳しくなっている項目が存在することが問題になり、2001年補正共通 CV の概念が立案され²⁶⁾、2002年度から全ての項目に適用することになった。

1) 共通 CV による評価法²⁵⁾

(1) 同一測定法群の極端値棄却後の平均値(M)、標準偏差(SD)、変動係数(CV)を計算する。

(2) CV の小さい測定法群から順番に並べる。

(3) 80%の施設が含まれるところまでの上位の測定法群に限定する。

(4) (3)の方法を利用している施設のデータを対象に総平均(GM)を計算する。また一元配置分散分析法を適用して方法内変動を計算して、それを標準偏差の形で表してSD₀とする。このSD₀からCVの計算式を用いて共通CV(CV₀)を計算する。

(5) 平均値(M)が一群のかたよったデータに影響されるのを防ぐため、極端値棄却後のデータに反復切断補正法を適用して調整平均値(Madj)を求め、これに共通CVを適用する。

(6) 参加施設数 $n \geq 10$ のとき、調整平均値(Madj)に共通CV(CV₀)を適用して評価規準用のSD(SD*)を求める。

Madj ± 1SD* 以内に入っているもの
A 評価

Madj ± 1SD* を超え ± 2SD* 以内に入っているもの B 評価

Madj ± 2SD* を超えて ± 3SD* 以内に入っているもの C 評価

Madj ± 3SD* を超えるもの

D 評価

ただし $n < 10$ のときは「その他」に記入したときは評価対象外とする。

2) 補正共通 CV 法²⁶⁾

補正共通 CV 法は、以下の計算を除きその他の計算および取り扱いは、共通 CV 法と同じである。

共通 CV(CV₀)を報告単位幅(CV)で補正した補正共通 CV(CV₀')を計算する。

[報告単位 CV(CV_u) =

$$\frac{\text{報告単位幅(CV)}}{M} \times 100]$$

$$CV_0' = \sqrt{[CV_0' + (CV_u/2)^2]}$$

報告単位幅は、報告書に記入する最小桁の単位幅に相当する

ただし、尿素窒素は単位幅が 0.1 のところ、補正計算では 1 桁繰り上げて 1.0 とする。同様にクレアチニンは単位幅 0.01 を 0.1 として補正。

E. 結語

臨床検査における定量検査の測定値には必ず誤差が含まれる。この誤差は測定の精密さに関連する偶然誤差と正確さに関連する系統誤差に大別されるが、この両者を誤差許容範