

健診データの信頼性の評価

とくに臨床生理検査について

岡本 登 (愛知県総合保健センター 成人病診断部)
岩塚 徹

はじめに

システム化された総合健診のデータは、すぐれた精度管理下で標準的方法で収集された品質のよい多項目情報である。特に個人の時系列データファイルは従来の臨床や疫学では得られない有用な情報を提供し、病態の推移や異常の早期発見、あるいは疾病の予防や健康増進の為に利用される。しかし生体情報は個体差が大きく、かつ性、年齢、民族、体型、環境その他の要因で変動し、たゞ1回の検査ではその断面像を観察するにすぎない。従って健診データの信頼性を高めるには、検査項目毎に変動要因を詳細に分析しその要因をとり除くよう努力すると共に、個体差および個人変動の範囲と限界を明らかにし、検査機器、検査技術の誤差と出来るだけ除くことである。以下主として総合健診における生理機能検査データの信頼性について、我々のデータを中心に述べる。

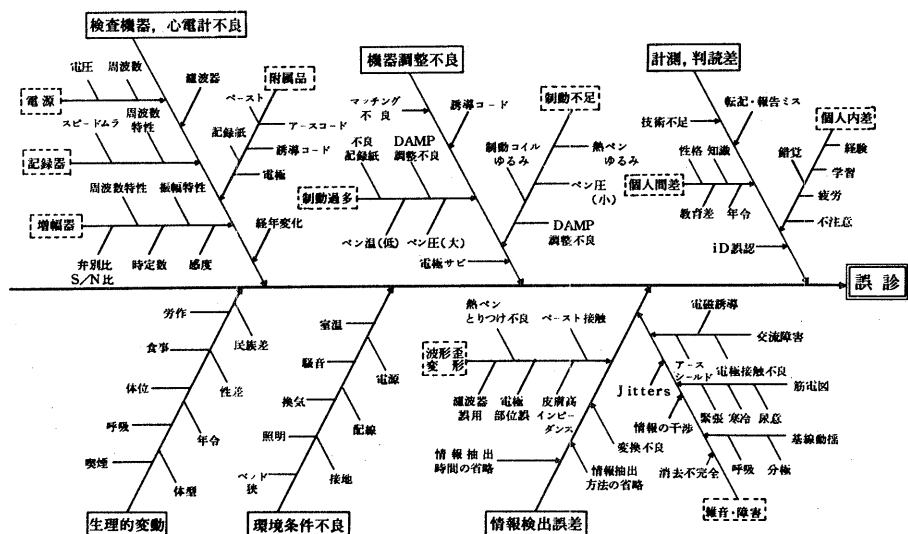
1 生体情報の特徴

生体情報は性、年齢、体型その他生理的、精神・心理的および環境などの要因で微妙に変動する。一方生体の組織、臓器は充分な予備力を備え、健康人は Homeostasis の範囲内で正常生活を営んでおり、臓器機能が一時的に範囲外に逸脱しても、自然の回復機能により健康が保持される。たとえば各部品には規格外の異常があっても機械系全体としての機能を果たし得れば故障とみなすことは現実的ではないと同様に、生体は Homeostasis 機構を備え臓器、組織の予備力、冗長性のある fail-safe 設計されたシステムである。従って組織、臓器の一部欠損や異常を直ちに疾病と定義づけることは問題である。

生理検査は直接生体を検査対象とし、ME 機器は直接受診者と接触あるいは結合する為、できるだけ Non-invasive (非侵襲的)、Non-destructive (非破壊的) で安全第一の検査である(図1)

生理検査の特性要因図 (心電図)

る必要がある。従って情報収集方法には限界があり化学検査や実験データに比べて不完全である。また生体情報には血圧、基礎代謝、肺機能検査など数値で表現される情報はむしろ

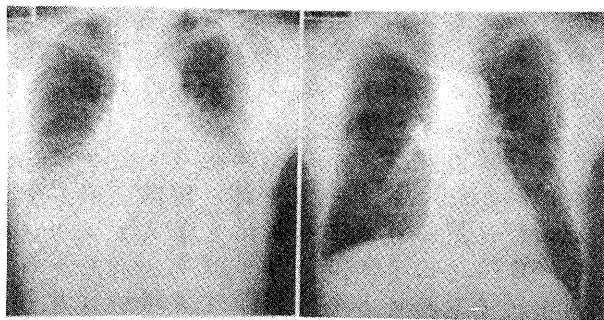


少なく、心電図、心音図、脳波、胃-胸部X線、眼底写真など図形データバンクとして表現される情報が多い。従って微妙な人間の感覚を計測機器とするいはゆる官能検査であり、Inter-observer および Intra-observer variation で表現されるバラツキやカテゴリーは避けられない。また生理検査の諸項目は正常から異常への移行がゆるやかで両者の境界は明らかでなく、grey zone と呼ばれる境界域が比較的広いことが判読上の個人差の要因となる。診断上の誤差は、生理変動、検査機器と設備、検査技術の誤差および判読上の誤差の総合であり、その特性要因図は心電図診断を例に挙げれば図1の如くなる。

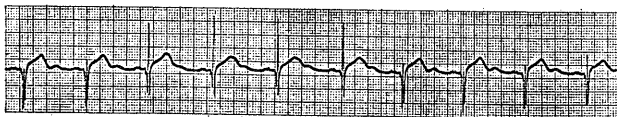
2 生理的変動要因と変動範囲

生体は安静時、活動時を問わず内部環境の恒常性を保つ為外部環境に適応しており、従って生体情報は動的である。すなわち性、年齢、体型など生物学的差異は勿論、食事、呼吸、体位、喫煙、労作、精神心理的ストレス、日変動、季節変動は明らかである。生体の短期あるいは長期の生物時計(バイオリズム)の機構を考慮し、できるだけ検査条件を明らかにする必要がある。(図2)

図2は胸部X線写真の呼吸性変化である。呼吸時に大きい心陰影は深呼吸時には正常範囲に縮小している。心陰影を経時的に追跡する場合、X線管球の位置、距離、体位、呼吸位など撮影条件を一定にすべきである。また麻酔による振動装置による撮影を考慮する。



その他心電図の呼吸性変動(図3)

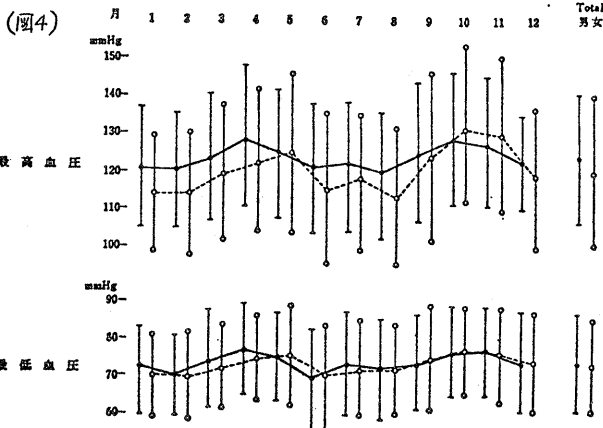


(図3) 心電図波形の呼吸性変動 (V₂誘導)

吸気ではR波がほとんど消失し、呼気相ではR波が増高する周期的変動を繰返す

や食事の影響も明らかで、特に便診時ブドウ糖負荷後の心電図では便常者にもST-T変化が出易く、10%近い偽陽性がみられ注意を要する。便診では耐糖能検査は必須項目であり、便診システム全体の流れにおいて心電図検査と糖負荷の時間関係を明確にする必要がある。また日中は勿論、夜間睡眠中も脈拍数の変動範囲は広く生理的変心拍数を決り難い例がある。血圧値についても同様に体位、呼吸、労作、気温による動揺のほか、日内変動、季節変動も明らかでストレス下では上昇し睡眠中や早朝は低い。また病院や診療所における血圧値が自己測定値より高く、医師が変わるごとに変動することはしばしば経験される(Pressure effect)。従って1回のみの血圧測定で正常、異常を判別することは不適当であり、また便診の血圧値がどのような条件下で測定されたか明記すべきであらう。

図4は便診における血圧値の月別平均値と分散を示す。最高・最低



血圧の季節変動(対象数: 男 15,445 女 3,354)

血圧は男女、いずれも春秋に高い山がみられるが、季節の変わり目で暖房がなく室温低下の時期に相当した変動である。

3 情報収集技術の信頼性と誤差

変動する生体情報と正しく収集するには、検査機器の調整と検、検査環境、検査条件を一定とし、正常・異常の判定に有用な情報の傾および量の選択、特徴パラメータの種類と抽出方法、情報の数量化、情報処理方式の客観化が必要である。また検査に対する受診者の不安をとり除き、協力が得られる様、受診者への接し方にも慎重さを要する。たとえば肝機能検査の如く、個人の最大能力を引き出す為には術者の技術と共に受診者の理解度、協力度のよし悪しが成績に影響することは明らかである。

生理検査の信頼度を高めるには、安静時の他、しばしばストレス負荷や不当な環境下で検査を行ない、生体の反応様式や程度から生体の部の Homeostasis の乱れを診断する賦活テストが用いられる。心電図における運動負荷や薬剤負荷、脳波検査における音、光などの物理刺激や薬剤負荷、寒冷昇圧試験、起立血圧試験などは通常用いられる検査である。また正常・異常の境界域に心電図や一過性の不整脈、夜間狭心症など、1回の検査で診断が困難な場合には、冗長性のある繰返し検査、長時間記録や夜間心電図検査などとさらに精密検査を要する。スクリーニングレベルの検査と精密検査では収集される情報に差があるのは当然であり、健診においては臨床検査の経済的、人的コストと効率とのバランスから検査項目とそのレベルが決定される。

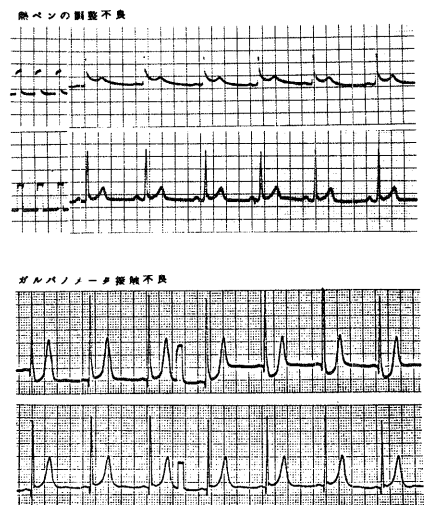
正しい情報収集には検査機器および附属器の調整と検は必須である。生理検査機器ではその諸特性、操作性のほか、特に電源および安全性の定期と検や保全および点検シートによる始業・終業と検が重要である。

図5は心電計調整不良による波形歪を示す。熱本の調整不良(最上段)とガルバニータ接触不良(3段目)例である。血圧計にはアナログ型とリバロッチ型があるが、JIS規格のある水銀血圧計の方が広く普及し、最近では自動血圧計も健診に用いられている。血圧計の水銀量のと検、校正と共にマンシートの幅、長さがWHOの報告を満足するカフを用いるべきである。また腕の太さに比し幅の狭い短かいマンシートを使用すれば血圧値は高く誤るし、繰返し血圧測定する場合カフの空気を完全に除かないと実際より低値を示すことになる。なお、機器の故障には故障と共に作動しなくなる陽故障と、故障のまま、不良データを作り続ける陰故障があるが、安全性の低下につながるおそれがある陰故障の発見には細心の注意を要する。

4 時系列データの評価

総合健診の最大のメリットは時系列データの蓄積により変化と早期に見出すことである。この為にはデータ活用の目的を明確にし、必要最適な特性値を選択あるいは変換し、検索照合用にデータプールする必要がある。たとえば血圧や心電図の正常・異常という分類は不十分で、血圧の実数値や心電図分類、診断名を保存し、

(図5) ST-T 波の波形歪



さらに求めれば AD 変換された波形計測値で保存すれば、診断基準の修正による誤差も僅かである。なお健診の保存データは他の臨床検査成績や予後調査結果およびその他新しい情報と結合し、修正や補足が容易に行われなければならない。単記計量値は最高の情報伝達効率を有するが、非連続的な分類値も 4-5 分類以上あれば単記計量値のほぼ 80% の伝達効率であることが証明されている。総合健診の心電図、胃部・胸部 X 線所見分類などは数十カテゴリーにコード分類されており、時系列データとして有用である。

一般に総合健診と二次検診の相関は予想外に低く、個人の時系列データの相関はさらに低い。生理的変動と生活態度の変化や外的要因による一過性誤差変動が主因と考えられる。集団の生理的変動幅に比し個人の変動幅は狭く、個人の正常値を評価するには非正常性の外因の影響を考慮し、最低 4-5 回のデータが必要とされる。過去 4 回以上の時系列データを有する総合健診受診者 5000 例中、胃底心電図陽性および疑陽性の異常者 730 例を追跡調査した。その後も胃底陽性あるいは疑陽性を示す例は男女共僅かに 20-30% で、陽性者の約 40% (男 43%, 女 36%) および疑陽性者の半数以上 (男 54%, 女 58%) は運動負荷陰性すなわち正常反応を示した。これは生理的変動というよりも生活環境の改善あるいは治療によるネードバックと考えられ、同時に高血圧、肺機能、肥満、高コレステロール血症等の risk factor が改善した。また上室・心室期外収縮や右脚ブロックなどの不整脈が可逆的に出現あるいは消滅する無自覚健康者が意外に多く、これらの心電図所見を直ちに異常とするのは問題である。

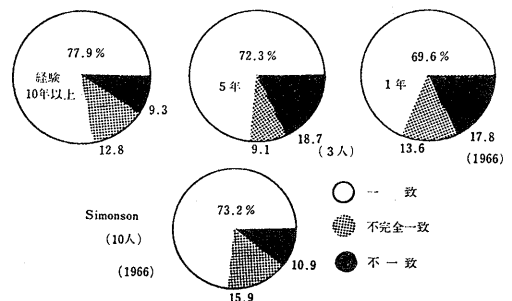
健診データの時系列分析では多くの項目で有意の経年変化が認められた。約 4000 例に亘る初診時と 1 年後の生理検査データの相関をみると、相関係数が 0.9 以上の高い相関は身長、体重など体型と肺活量にすぎず、最高血圧 0.62, 最低血圧 0.50, 眼圧 0.54, 心胸郭比 0.64, 1 秒率 0.75, 3 秒率 0.27 であった。眼底所見については Sheie および KW 分類いずれも相関係数 0.3 以下の低い相関であった。これは眼底写真判読上の誤差の影響も考えられた。

5 検査データ判読上の誤差

一般に正常・異常あるいは疾病の診断は、学習や経験を通じて得られた豊富な記憶パターンとの照合であり、最適の意思決定の論理である。この記憶パターンの差や性格の違いなどが判読差の要因となる。統計的に正常範囲と決める場合、対称的分布では偏差値が用いられるが、非対称的分布の運用データでは Percentile 分析が有用である。すなわち分布の上下 2.5% を除外した 95% 領域を正常範囲とする方法である。しかし分布上正常・異常にオーバーラップのある場合は異常の重症度や予後を考慮して最適判別点を用いる。Optimist は境界域を正常へ、Pessimist は異常へ包括して考える傾向がある。

一方総合健診では検査精度は Sensitivity (感度) と Specificity (尖鋭度) で表現される。判定基準を過大評価し甘くみれば感度はよくなるが第 1 種の誤り、すなわち虚報(疑陽性)が増え(真原病やコンヒューダ原病を作り、辛くあるいは過少評価すれば Specific であれば第 2 種の誤り、すなわち異常の

(図6) 心電図診断の恒常性



見逃し(偽陰性)が増加する。化学検査の如く真値あるいは基準値が明らかになる場合は、検査成績の客観的、統計的評価が可能であるが、画像情報やパターン認識に関する生理検査は数量化が困難で経験や勘に頼り易く判読差は大きくなる。

心電図判読の Inter および Intra-observer variation の誤差に関する報告は多く、いわゆる「心電図病」を作らなす特警告している。(図7) 眼底写真判読上の誤差

また集検における心電図異常率の判読する医師数と及比例するとの説がある。すなわち個人の判定では第1種の誤差(偽陽性)が高いことを示唆している。心電図の判読差に関する著者らの調査では、国際的な専門医間でも全体の約半に判読の不一致率を示し、また同一心電図の繰返し判読における診断の再現性(Repeat variation)も悪かった(図6)。国内の調査についてもほぼ同様で、125例の心電図を9人の専門医に判読させた場合、9人共判定が一致したのは28.4%にすぎず、9人中6人以上の診断が一致したのは72%であった。6ヶ月後の繰返し判読では約半の症例で、多少れ少なかれ診断名に変更がみられたが、経験の深い専門医ほど再現性がすぐれていた(図6)。

眼底写真の診断も判読差が大きい。撮影条件や技術の誤差と共に読み方の個人差が問題となる。

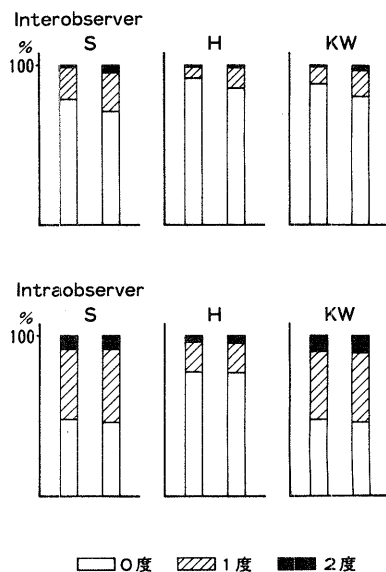
図7の上段は対象の属性に有意差を認めない「健常

群」に対し、2人の眼科専門医の変化性変化(S)、高血圧性変化(H)、KW分類の所見を比較した図である。下段は経験35年の同一医師による判定の恒常性を検討した図である。上段の2人の医師間には高血圧変化0度以外、すべて4%差($P \leq 0.01$)を認め、下段は上段2人と異なり経験の深い専門医であるが、肉眼的判定の恒常性に限界があることを示唆している。

胸部X線写真上「心拡大」に判読誤差がみられる。主観的に「心拡大」と判定された1028例の心胸郭比(CTR)は41~76%間に広く分布し、正常上界とされる50%以下の例が23%に認められた。また正常心陰影群に心胸郭比50%以上の例が11%あり、両群の分布に重なりを認めた(図8)。なお心胸郭比に対しては、血圧(高血圧)よりも体型(肥満)の方が影響が大きいことは意外であった(図9)。

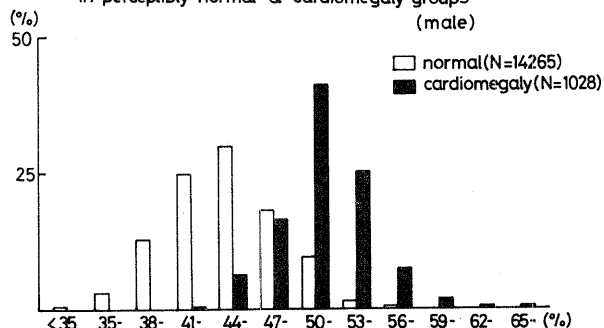
胸部(肺野)X線像や胃部X線像の判読についても、「異常なし」、「要精検」の分類には複数の専門医間に有意差(図8)を認めなかったが、病型分類や異常の細分類については判読者の専攻領域や経験の違いから有意差が認められた。

主として聴覚と計測器とする心音・心雑音の正常・異常の判定は一層誤差が大きい。2人の専門医が合意して基準とした心音・心雑音の中、約半数近くに不一致を認め、1ヶ月後再聴診時の再現性も



□ 0度 ▨ 1度 ■ 2度

% distribution of cardiothoracic ratio(CTR) in perceptibly normal & cardiomegaly groups (male)



78% にとどまった。長時間の聴診は(図9) 聴覚の疲労による誤差を生じ易く、従って総合健診における心音図自動解析の意義が認められる。

血圧測定は簡単に1か所血圧値は循環器情報として極めて重要である。

用手的血圧測定に関するWHOの勧告では、マンシェットの幅は腕の直径より20% 広く、長さ1/4半巻くに必要の長さとし、幅12cm以上、長さ22cm以上の基準をすすめている。腕の太さに

比し幅の狭い、短いマンシェットを使用すれば血圧値が20~30 mmHg 高くなり誤差を生ずる。また室温、体位、食事、喫煙などの影響を除いた標準条件下のCasual血圧測定を規定し、水銀中の目盛に従って末尾の数字は2mm単位の偶数値で記載することがすすめられている。

1か所聴診法で末尾数字を偶数読みすると、測定者個人の好みによる数字のカタマリを生じ末尾零(0)の頻度が高率となる。表1は総合健診19144例の血圧値を末尾の数字別にまとめた分布表である。末尾数字は収縮期、拡張期血圧いずれも零(0)が40%以上で圧倒的に高頻度にみられ、次いで(8)であった。性別、曜日別には有意差を認めず、すべての計測者にはほぼ共通の好み誤差(Zero preference)と考えられた。誤差は教育により改善をみたが、なお不平等な分布は避けられなかった。

健康成人男子拡張期血圧のヒストグラムを見ると、末尾数字が原則として偶数であること、およびZero preference による誤差の為、不規則、不平等な分布となる。血圧値の振幅を適当に

選べば分布の平滑化はみられるが、データの信頼性には問題がある(図10)。

以上の如く血圧測定は計測器、計測法および目盛の読みに対する誤差を除く為、自動血圧計を導入し検討した。

200例を対象として、収縮期、拡張期血圧につき、デジタル型自動血圧計と従来の用時計測法による血圧値と比較した(図11)。両者間には高度の相関があり、標準誤差は約8 mmHg、

自動血圧計(Y)と用時計測(X)の相関式は次式の如くであった。

① 収縮期血圧(最大血圧) $Y = 1.00X + 3.97 (\pm 8.13)$

② 拡張期血圧(最小血圧) $Y = 0.91X + 6.94 (\pm 7.90)$

Mean & S.D. of CTR in relation to body built & blood pressure

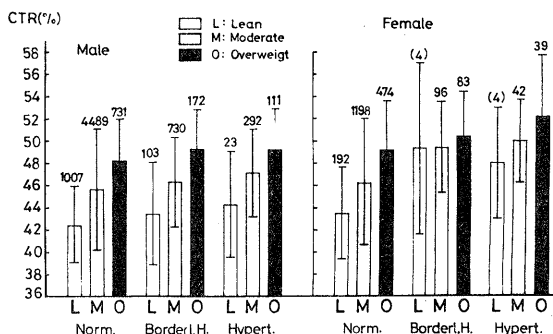
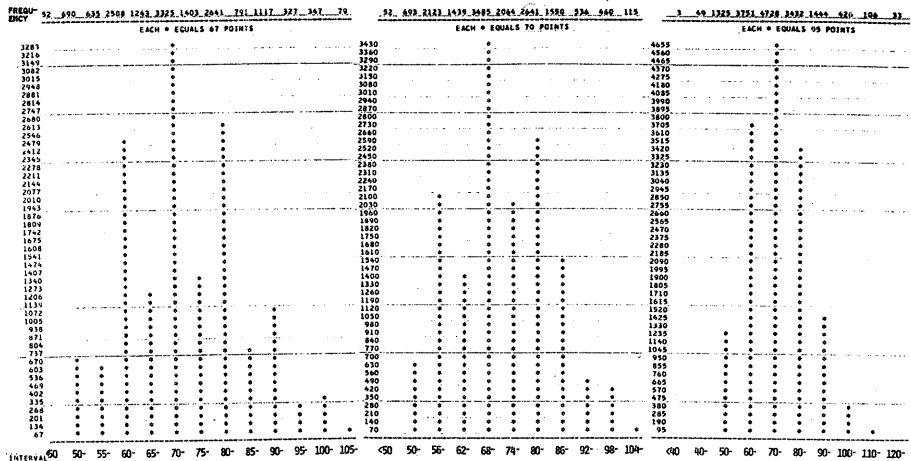


表1 血圧測定における末尾数字の分布

	末 尾 数 字 (男)										末 尾 数 字 (女)											
収縮期血圧	0	1	2	3	4	5	6	7	8	9	計	0	1	2	3	4	5	6	7	8	9	計
N	6401	2	2167	2	2216	14	2039	5	2447	4	15297	1610	0	569	0	572	1	506	0	589	1	3847
%	41.8	0.0	14.2	0.0	14.5	0.0	13.3	0.0	16.0	0.0	100.0	41.9	0.0	14.8	0.0	14.9	0.0	13.2	0.0	15.3	0.0	100.0
月	1270 43.2%	1	417	0	402	0	386	1	458	0	2935	321 43.6%	0	113	0	105	0	86	0	112	0	737
火	1288 41.2	1	445	0	449	0	426	1	515	0	3125	311 40.4	0	119	0	116	0	101	0	123	0	770
水	1362 43.8	0	419	0	436	1	422	1	472	0	3113	303 40.2	0	116	0	116	0	107	0	111	0	753
木	1295 39.8	0	474	2	516	5	430	1	530	1	3254	332 41.7	0	111	0	120	0	112	0	122	0	797
金	1186 41.3	0	412	0	413	8	375	1	472	3	2870	343 43.4	0	110	0	115	1	100	0	121	0	790
拡張期血圧	0	1	2	3	4	5	6	7	8	9	計	0	1	2	3	4	5	6	7	8	9	計
N	6741	1	1972	4	2050	15	1939	4	2570	1	15297	1695	0	509	0	519	3	506	0	614	1	3847
%	44.1	0.0	12.9	0.0	13.4	0.1	12.7	0.0	16.8	0.0	100.0	44.1	0.0	13.2	0.0	13.5	0.1	13.2	0.0	16.0	0.0	100.0
月	1243 42.4%	1	422	0	331	1	386	2	499	0	2935	322 43.7%	0	96	0	95	0	97	0	127	0	737
火	1396 64.7	0	425	2	385	1	376	1	539	0	3125	353 45.0	0	98	0	101	1	89	0	128	0	770
水	1406 45.2	0	346	0	460	3	390	0	508	0	3113	334 44.4	0	95	0	107	1	100	0	115	1	753
木	1413 43.4	0	419	2	438	1	431	1	549	0	3254	340 42.7	0	99	0	124	0	111	0	123	0	797
金	1283 44.7	0	360	0	386	9	356	0	475	1	2870	346 43.8	0	121	0	92	1	109	0	121	0	790

自動 (図10)

計測により
血圧値末尾
数字の Zero
preference は
消失し、期
待値に近い
分布を示し
(図12)。拡張
期血圧の
ヒストグラムも
不規則性は
なくなった。



また性、

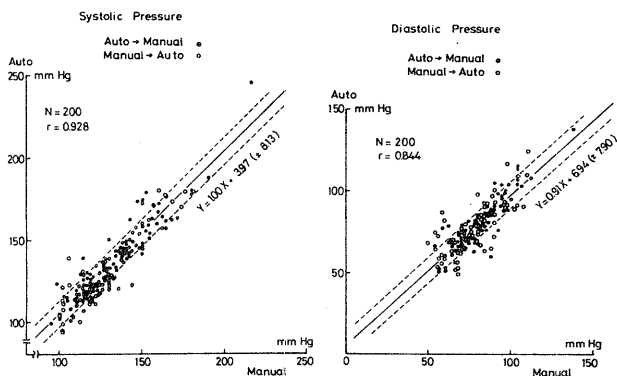
年齢別の血圧分布の特徴と健康
群における異常頻度を分析する
為に、正規確率紙上の累積度数
分布図は有用である。

(図11)

以上、生体検査の判読誤差は、
主として Perception に係るパター
ン認識、すなわち官能検査の信頼
限界を示唆している。判読誤差
の大きい検査は総合健診や臨床
に用いるのは好ましくなく、
データの数量化、客観化と判読
アルゴリズムの文書化および

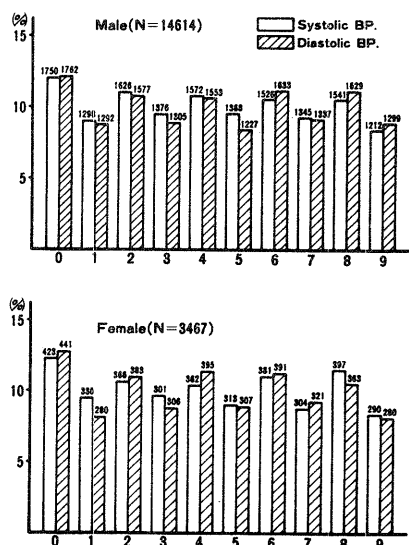
教育により計量診断を図る必要
がある。なお、緩数の専門医の
合意により確定した各種疾患群のデータフル
より基準パターンを作成し、教育普及を図る
ことが判読誤差を低下させる近道といえる。

近年自動化機器およびコンピュータシステムの
導入により、心電図、心音図、肺機能など、
生体検査の特徴抽出から診断に至る全過程を
自動処理する技術が実用化普及し、計測およ
び判読誤差は減少した。今後さらにパターン
認識手法の進歩と共に官能検査の数値分析、
数量化分析などハードウェア、ソフトウェア技術の
発展、4M (Man, Machine, Material, Method)
の品質管理と標準化、種々の品質管理手法の
運用情報処理への導入等により、健診データ
の誤差は減少し、信頼性が一層高まるものと
期待される。



(図12)

自動血圧測定における末尾数字の分布



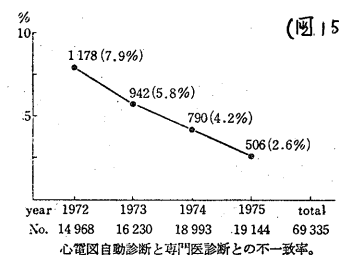
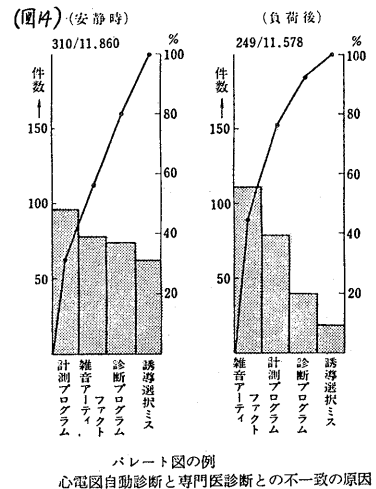
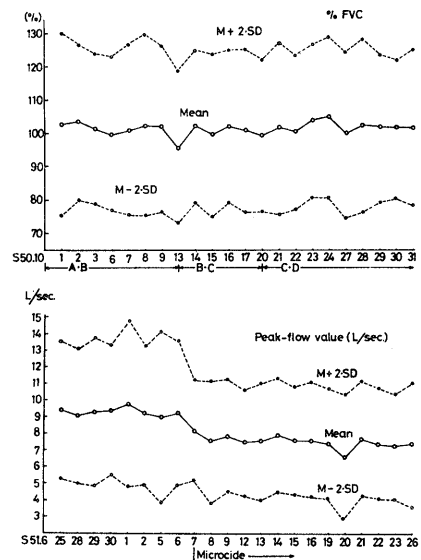
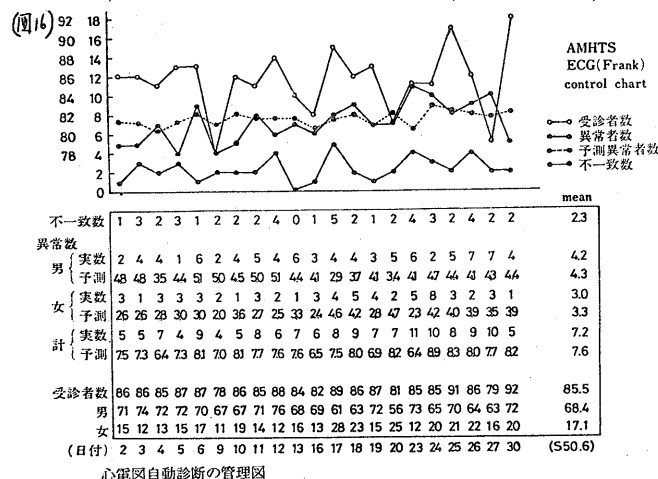
6 生理検査の統計的精度管理

図13は肺機能検査における努力性肺活量(%FVC)とピークフロー値の管理図である。毎日の平均値と標準偏差が記され、いずれも管理限界内にあるか否かにより、検査システム信頼性を向う方法である。グラフが管理限界を外れるが、バラツキやカテゴリーの異常が認められる場合は、検査対象(図13)

の質的变化が検査の調整不良、術者の技術が未熟な為か、その原因を説明しなければならぬ。図の上段は%FVCの(X-R)管理図で、13日の平均値の低下は勤務交替による不慣れた術者の技術的誤差である。下段ピークフロー値の平均値の低下は呼吸計に接続した殺菌用Microcideによる気流抵抗の増加が原因であった。

使診における心電図検査は専門医の判定を基準とし、自動診断における不一致例はその原因別のパレート図すなわち重点指向型の分析図(図14)を作成し、頻度の高い誤差要因に対する対策をとら精度管理にフィードバックされる。たとえば心電図自動診断では、安静時心電図では別点計測プログラムが最も重要な問題であり、負荷心電図では雑音、アーチファクトの除去が最重要であることを示している。

図15は昭和47年より50年までの心電図自動診断と専門医との不一致率をまとめたグラフで、不一致例のパレート分析結果が精度管理にフィードバックされ、不一致率が毎年減少したことを示す。また蓄積された心電図から、層別異常率と予備データとして心電図異常数の予測式を求め、毎日の予測数と実数との一致率から心電図自動診断システムの管理状態のより悪しが推定出来た(図16)。



まとめ

生理検査データの信頼性を高めるには、生体情報の入力から出力、報告までの各過程に介入するハードウェア、ソフトウェアの雑音や誤差の要因を詳細に分析し、品質管理手法により解決策を検査過程にフィードバックすることが必要である。また健診データの信頼性は品質管理手法の他、長期の予後調査から客観的に評価される。正常胃部X線所見から胃潰瘍や胃がんの発症、長期検診直後の肺結核や肺がんの発症、運動負荷正常心電図からの急性心臓死や正常血圧者の脳血管障害など、決してごく稀な現象ではない。疾病の要因がMultifactorialである以上、限られた検査から疾病を予測したり事故を防止するには限界がある。生命の本質にせまる分子レベルの酵素工学や免疫学、自律神経や内分泌検査が不十分な総合健診の検査項目に問題があるのか、検査の網の目が荒いのか、情報の収集や処理に問題があるのか、総合健診あるいは現代医学の限界であるのか、検査成績以外に受診者から得られる既知あるいは未知の追加情報が必要なのか、今後の医療に残された大きな問題と云えよう。

参考文献

- 1) Simonson, E.: The concept and definition of normality. *Ann. N. Y. Acad. Sci.* 134: 541, 1966.
- 2) Simonson, E.: Differentiation between normal and abnormal in electrocardiography, St. Louis, C.V. Mosby Co. (1961) 日比野進校訳, 中川喬市, 岡島光治訳著: 心電図, 正常と異常の鑑別, 医学書院 (1967)
- 3) 水野 康, 岩塚 徹, 他: 心電図判読の個人差, 最新医学, 23: 2147, 1968.
- 4) 岡本 登: 臨床生理検査の精度管理の現状と問題点, 日健診誌, 2: 21, 1974.
- 5) Spodick, D. H.: On experts and expertise: The effect of variability in observer performance, *Am. J. Cardiol.* 36: 592, 1975.
- 6) 岩塚 徹: 総合健診システム (AMHTS), 日内会誌, 65: 221, 1976.
- 7) 岡本 登: 心電図の正しいとり方, 日本衛生検査技師会, 近代出版, 1976
- 8) 岡本 登: 臨床検査の誤差論—生体検査の精度管理と問題点, 医用電子と生体工学 14: 473, 1976.