

簡略化された外部精度評価法案を作成した。1) 異なる 2 つの濃度のサンプルを測定し、その結果がどちらも、 $\text{mean} \pm 3\text{SD}$ 内に入ったものに認定書を与える。(本稿では、「2 サンプル法」と呼ぶものとする)。2) 異なる 3 つの濃度のサンプルを測定し、 $\text{mean} \pm 3\text{SD}$ 内に入れば 3 点、 $\text{mean} \pm 9\text{SD}$ 内に入れば 1 点を与える。点数を発表、または 7 点以上を目安として認定書を与える。(本稿では、「3 サンプル法」と呼ぶものとする)。全衛連の平成 12 年度の報告書から、労働衛生検査クロスチェック 8 項目、すなわち、血中鉛、尿中デルタアミノレブリン酸、尿中馬尿酸、尿中メチル馬尿酸、尿中マンデル酸、尿中総三塩化物、尿中三塩化酢酸、尿中 2,5-ヘキサシジオンの自機関測定結果のうち、それぞれ共通 3 濃度の測定値を用いて、先に示した簡便法の (1) 2 サンプル法と (2) 3 サンプル法によって、それぞれの検査機関の評価を行った。この評価と現行の評価方法による評価とを比較することによって、それぞれの簡便法がどの程度実用可能であるかについて検討した。Mean と SD の計算法は、一般的に用いられている方法によった。すなわち、まず、全施設の測定結果から mean^* と SD^* を計算する。次に $\text{mean}^* \pm 2\text{SD}^*$ の範囲にある測定値のみを用

いて、再度、mean と SD を計算しなおす。これらの値をもって判定に用いる mean と SD とするというものである。なお、3 サンプル法では、評価は 1 つであるが、2 サンプル法では、共通 3 濃度のうちどの 2 サンプルを使うかによって 3 通りの評価が可能であり、これら 3 通りすべてについて評価を行った。

4. 職域健康診断における胸部放射線診断のあり方とその精度管理および系列データの利用法に関する調査研究 (分担研究者 曾根脩輔)

初年度の平成 14 年度に引き続いて、CR システムに関する研究として、システムの最近の性能の調査、各種 CR システムの特徴、機能の比較などを行った。現在国内で市販のシステムの画像処理パラメータのターミノロジーの整理、包括的理解と整合性を実現するために必要なデータ収集や解析を行った。CR 法の特徴である画像処理パラメータの多様化の実体を明らかにした。また CR の最近の動向を紹介するために、CR の基礎的事項を示し、CR システム毎に異なった名称で供給されているシステムの理解、他のシステムのものとの比較をたすけるための解説書の作成を開始した。胸部ファンとームを用いた CR 撮影実験を行い、画質評価を試みた。肺癌の CT 検診に

については、現在これが国際的に注目され、その有効性の調査が行われ始めている。我が国における普及が先行している。国際的あるいは国内的な利用状況を学会や研究会における発表などを通じて調査した。具体的なデータとして長野県下で平成 8～10 年度に実施された CT 検診の結果を参照して考察した。さらに、全国労働衛生団体連合会エックス線写真専門委員会で行われた胸部 X 線写真の画質審査の結果を経年推移の面から整理・検討した。

5. 循環器疾患の効率的なスクリーニングと一次予防のための職域健診のあり方に関する基礎的検討（分担研究者 久代登志男）

研究 1 では、平成 13 年の都内 2 健診機関の健診受診者（健診機関① 6,831 例、および健診機関② 37,233）のうち、治療中の疾患が無く、心電図所見と問診結果が得られた 15070 例を対象とした。心電図所見は①N（正常範囲内）②Af（心房細動）、③Long QT（QT 延長）、④LVH（左室肥大）、⑤Q（異常 Q 波）、⑥ST（ST 低下）、⑦T（T 波平低、陰性 T 波）を選択し、問診票は①胸痛、②動悸、③失神に関する所見の有無を調べた。心血管系疾患危険因子として AGE（年齢）、SBP（収縮期血圧）、DBP（拡張期血圧）、TC（総

コレステロール）、TG（中性脂肪）、HDL-C（HDL コレステロール）、LDL-C（LDL コレステロール）、FBS（空腹時血糖）、HbA1c、UA（尿酸）を検討項目とした。①胸痛、②動悸、③失神の有無により心血管系疾患危険因子に違いがあるかを検討した。さらに各問診所見の有無と各心電図所見の有無について χ^2 検定を行った。

研究 2 では、一般に職域健診、ドックなどで使われている問診票よりも効率的に循環器疾患スクリーニングが可能な問診票の作成を試み、従来からの問診票との疾患検出率を比較検討する。また、冠動脈造影検査、心臓電気生理学的検査、Head up tilt 試験などの循環器疾患精密検査を施行した症例に対し、同問診票を適応し、validation study を行った。

6. 時系列健康診断データの有効利用と精度管理（分担研究者 吉田勝美）

健診の連続受診者における高血圧、高脂血症発症の予見方法に関する評価を目的として、Matched Pair を基に、発症直前の血圧及び総コレステロール値を調整して、それ以前の値を考慮した際の発症予見性について検討した。もう一つは前述とは異なる集団で、高血圧発症直前の血圧値を調整して、発症に至るまでの血

圧変動の評価の有用性について検討した。まず、某健診機関受診者のうち、7年連続の受診者で血圧、総コレステロール(TC)の測定値、治療歴のデータの揃った男性2,318名を対象とした。解析方法は1年目に血圧は140mmHg未満かつ高血圧の治療を受けていない者のうち、7年目の受診までに高血圧(収縮期血圧(SBP)が140mmHg以上または高血圧の治療を開始した者)を発症した者をケースとし、発症しなかった者をコントロールとした。同じくコレステロール値が1年目に220mg/dl未満かつ高脂血症で治療していない者のうち、高コレステロール血症(総コレステロール(TC)220mg/dl以上または高脂血症の治療を開始した者)を発症した者をケース、発症しなかった者をコントロールとした。これらのケースとコントロールによるmatched case-control studyを高血圧と高コレステロール血症でそれぞれ行った。もう一つの方法は、血圧値について、某製造業の健診受診者データを用い、Matched Pairを基本として、個人毎の推移値について、移動平均(MA;AV)、標準偏差(SD)、一次回帰直線の勾配(SLOPE)を算出し、比較作業を行った。対象項目はsBP(収縮期血圧)を用いた。血圧高値者(Hypertensive)とする定義は、sBP値が140mmHgを超え、

かつ140mmHgを超えた値以降の2個分を含めた計3個分のsBP平均値が140mmHgを超える者とする。また、高値開始前件数が5個以上、開始を含め開始後3個、計8個以上件数がある者を対象とする。正常者(Reference)の定義は、sBP値が140mmHgを超えない者で、問診情報から高血圧症の病歴がない者とした。Matched Pair対照群の抽出は、性別、年齢が一致することを前提とし、Hypertensiveと定義する1回前(発症直前)のsBP値のレベルがReferenceと一致(± 2 mmHg)するものとした。なお、何らかの介入が加わっている可能性のある人は除いているので、上記の要件が満たされる標本数は110組となった。

C. 研究結果

1. 職域健康診断における臨床検査のあり方と健康管理(分担研究者 徳永力雄)

調査報告値の正確度とその評価を調査した。1) 総コレステロールの各年度の報告値はほぼ正規分布し、CVは少数例を除いて2%以下で、2SDを超える報告件数は少なかった。極端値を示す数機関を除いて、ほとんど一致した値で実用上(保健予防と臨床判断の観点からみて)問題はないと考えられた。2) 中性脂肪のCVは3~5%、測定値

は正規分布に近いが基準値よりやや低濃度側に偏っていた。その差は約 2% で実用上の問題なく小さいが測定技術上の課題が示唆された。3) HDL コレステロールの最頻値は基準濃度近傍にあるが正規分布にほど遠く、2SD を超える測定値が多く分散していた。各種の測定原理・試薬・試薬メーカーが存在するほか、調査試料の調整方法にも課題があることが推測された。評価標準偏差が小さいことは、レファランス機関と参加機関の技術水準に差があることを示しており、特に約 10% の機関で差が大きい。精度管理の上で重要な努力目標であることがわかった。4) AST は基準値よりやや高濃度 (活性) 側によっていた。各年度とも低濃度では中央値近傍に集積し、中等度濃度 (異常値下限) では正規分布、高濃度では 2SD 範囲にはあるが両側に分散していたが、従前からの頻用検査項目であり、実用上はほぼ問題がない技術水準であると判断された。5) ALT は各年度とも正規分布であったが、平成 13 年度 (第 10 回) の結果は 14、15 年度の結果に比較して基準濃度との差が大きく、基準値用試料が報告値用試料のいずれかの測定時の活性度などに差異があったことが疑われた。しかし、14、15 年度の CV は 5% 以下で 2SD を超える機関

数も少なく、実用上は問題のない精度であると判断された。6) γ -GT は各年度とも低濃度 (活性) ではほぼ正規分布しているが、中等度から高濃度になると 1SD の範囲でプラトーを示した。しかし 2SD を超える値は少なく、基準活性との差異も 200 国際単位以上の高活性域で 3~5% (6~10 単位相当) であり、実用上問題のない技術水準にあると判断された。7) 血糖の報告値は各年度とも基準値あるいは最頻値近傍に集中しており、SD 及び CV も小さく、測定精度が極めて高いと判断した。8) ヘモグロビン A1c の報告値は、HPLC・不安定分画を除去する方法では年度・濃度に拘わらず CV が約 2% でほぼ正規分布しており、実用上問題がない水準にあった。他の方法でも報告機関が各 10、8 機関と少ないが、CV が 2% 前後で、HPLC 法と比較してほとんど差がなく、特別な課題はないと思われた。しかし、測定方法、機器、メーカー別の試薬、検量用標準品の開発等が今後変化することが予測されるので、精度管理上注目しておく必要があると考えられた。9) ヘモグロビンは、年度・濃度を問わず CV が 1% 前後で測定精度は極めて高く、技術的には全く問題がない水準であった。SD 準拠の評価方法では差別化の意味がないレベルにあり、評価法

の工夫が求められると考えられた。

10) 赤血球の報告値の CV は 2% 以下で精度上も実用上もほとんど問題がなかった。しかし、測定機器の機種間で測定値が異なるため、後述の白血球や血小板と共に、機種の差を克服した精度管理用の標準血液の開発が望まれた。11) 白血球の報告値は機種別にみるかぎり CV も 2% 前後でほとんどが 2 SD の範囲内にあり、7~8% の機関を除いて実用上問題はない水準にあった。12) 血小板の報告値も、白血球とほぼ同様のことが言えた。

臨床検査自施設実施機関の調査では、回収率は 60% で、単純集計と自由記載意見の集約のみであるが、検査機関の現状と意見の概要を把握するという目的は達せられた。自由記載意見を纏めると、精度管理が日常業務の中で実施されていない機関が多い、そのために日常業務に影響し負担に感じている、国内の精度管理団体が多く統合化を望んでいる、主な評価方法である標準偏差による評価は SD の狭小化によりいたずらに精細化しており不満が多い、臨床上あるいは保健予防活動の目的に照らした臨床検査の意義と評価の概念について共通の理解が得られていない、検査技術者の人材確保・高齢化対策・育成研修の充実が望まれている、設備の維持・

投資が困難である、検査事業の競争が激しく大手機関への統合・外注化等の経営戦略の波に曝されている、などの現状と課題が窺われた。現在進行中の国レベルでの精度管理体制の再編成の動きと関連して、より適切な精度管理のあり方について、有用な意見が得られた。

2. 職域健康診断及び地域医療における臨床検査のあり方と精度管理 (分担研究者 中 甫)

1) 参加機関は自機関測定と外部(外注)機関の 2 群にほぼ折半され、10 年間連続して参加している機関は意外に少なく、自機関で 37%、外部機関で 24% しかなく、また 5 回以下しか参加していない機関は自機関で 35%、外部機関で 45% であり予想以上に多いことが判明した。2) 外部機関へ外注している機関で 10 回連続参加機関及び 6~9 回参加機関のうち同一外部機関に連続して依頼している機関は約半数であり、残りの多くは 1~2 回外注先を変更している。今回の検討ではその理由を明確にできなかったが、低い成績が得られたときに次回外部機関を変更している例が散見された。しかし 3 回以上変更している例が見られるように安易に外注先を転々と変更することには問題があることがわかった。3) 各項目の参加回数、自機関、外部機

関別成績から、 γ -GT が他の項目に比較して高得点の機関が少ないが、総コレステロール、中性脂肪、AST、血糖の4項目について全ての年度で80~100点の機関が全体で65%以上を占め、とくに外部機関では全ての項目において自機関を上回った。また10年間に60~79点2回までを加えると γ -GTが最も低く75%であったが、他の項目は80%以上であった。とくに59点以下が1回以上存在する機関は精度規格を設定しその中に入るよう努力する必要があることがわかった。4) 同一機関から変更することなく連続受注している外部機関の成績を経年的に見るといずれの項目も一部の機関において大きい変動を示している例が見られたが、これらの機関を除いて全体的に見ると総コレステロール、血糖の2項目は長期にわたって性能が安定していることが分かった。中性脂肪は経年的に90~100点の範囲内に収束してきていることが分かった。AST、 γ -GTの酵素2項目は経年的に90~100点の範囲内に著しく収束し、基準分析法の勧告と酵素標準物質の供給による効果が著しいことが分かった。

3. 職域健康診断における労働衛生検査のあり方と精度管理(分担研究者 栗原伸公)

2 サンプル法による結果を見ると、

血中鉛の場合、現行の方法による評価では低い点数のものが認定され、それより高いものが認定されないケースが見られるが、少なくともAランクのものはすべて認定され、Dランクのものは認定されていない。尿中デルタアミノレブリン酸、尿中三塩化物も血中鉛と同様の傾向を示したが、尿中馬尿酸、尿中マンデル酸では、現行の方法による評価がAのものも一部認定されないことになった。

また、尿中メチル馬尿酸、尿中三塩化酢酸では、組み合わせによってはAのものが認定されなかったり、Dのものが認定されてしまったりする現象が見られた。尿中2,5-ヘキサジオンでは、一部Dの施設が認定されたものの、非認定施設は、すべてどの認定施設よりも実際の評価が下であった。次に、3サンプル法による点数と現行の評価による点数を比較すると、すべてのケースで両者に強い相関が見られた。また7点以上を認定、6点以下を非認定とすると尿中三塩化酢酸での1施設、尿中2,5-ヘキサジオンでの1施設において、現行の方法ではDであったにもかかわらず、簡便法では認定されるということがある点を除けば、認定も概ね良好に行い得るものと考えられた。以上のことから、2サンプル法では、多くのケースで現在の

方法による評価とほぼ同様の結果を示したが、いくつかのケースでは、現行の方法による評価と大きく異なる結果が得られた。一方、3 サンプル法では、現在の方法による評価と比べ、ほぼ遜色のない結果が得られた。本研究におけるシミュレーションにより、3 つのサンプルのみを調べる簡便法でも、かなり高い精度、確度で評価が可能であることが示されたことから、基本的には、やはりこうした簡便法を採用し参加検査機関の1回あたりの費用と時間の負担を軽減した上で、調査の回数を増やすことが望ましいと考えられた。

4. 職域健康診断における胸部放射線診断のあり方とその精度管理および系列データの利用法に関する調査研究（分担研究者 曾根脩輔）

第1にCR(computed radiography)システムの最近の機能、ソフトウェア関係の調査、各種システムのコンセプトや機能、画質の調査、比較を行ってまとめた。現在国内で市販されているCRはソフト面で急速にその性能向上がはかられているが、その利点を使用者における利便性の向上につなげるための整理、まとめを行い、CRシステムの可能性をよく引き出すための準備とした。具体的なCRイメージを作成し、すなわちフアントームを撮影して、その画

質により検証した。第2にCRシステムの最近のハード、ソフト面の基礎的概説をこころみた。現在は各種システムの統一的理解が、横断的な解説書の不足から困難であるので、新たなCRシステム利用機関における担当者に役立つ簡潔なマニュアルの作成にとりかかった。今後、適宜付図を加えて視覚的な理解を促進する方向で完成を目指す。第3に肺癌やじん肺に対する胸部検診へのCT(computed tomography)の利用は国内的、国際的にその有効性の調査が国歌プロジェクトとして始まっており、具体的な普及活動、一般的な利用を行うものもでてきた。平成15年度にはその国際的調査研究への参加協力するための準備を行い、我が国では唯一の参加組織として安曇総合病院が加わり、国際的データ集積を開始した。肺がんのCT検診を普及させるために重要なことは、その処理能力や精度管理であり、そのために必要な読影者の養成、読影補助システムの研究開発を次年度以降行う。CAD(computer assisted diagnosis)の開発導入も重要とみておりその可能性をさぐる。第4に、従来から、全国労働衛生団体連合会エックス線写真専門委員会では、胸部検診に用いられている胸部X線写真の画質審査を行い、各検診機関の評価得点の経年変化も検討してきた。そのなか

で、X線写真の画質の向上に重要な事項として、X線撮影系についての最新の情報が撮影担当者に適切に提供されていないことや機関における画質に関する理解の不統一、あるいは画質向上に対する現場の熱意不足あるいはそのための動機付け不足が明らかになった。精度管理が行き届いていて評価の高い優良な機関を公表するなど、機関の精度管理に対するインセンティブを高めることが大切と思われた。

5. 循環器疾患の効率的なスクリーニングと一次予防のための職域健診のあり方に関する基礎的検討（分担研究者 久代 登志男）

研究1では、胸痛あり群は、AGE、BMI、SBP、DBP、TC、TG、LDL-C、FBS、HbA1c、UAが高く（ $p<0.01$ ）、HDL-Cは低かった（ $p<0.01$ ）。動悸あり群は、TC、HDL-Cが高く（ $p<0.01$ ）、BMI、FBS、HbA1c、UAは低かった（ $p<0.01$ ）。失神あり群は、AGE、BMI、SBP、DBP、TC、TG、LDL-C、UAが低かった（ $p<0.01$ ）。 χ^2 検定の結果は、Afあり群は胸痛あり、動悸ありと答えた例が多かった（ $p<0.01$ ）。Long QTあり群は胸痛ありが多かった（ $p<0.05$ ）。LVH、Qの有無と各問診項目の有無とは無関係であった。ST変化あり群とT波変化あり群は、

胸痛あり、動悸あり例が多かった（ $p<0.01$ ）。心電図所見と失神有無とは関連が見られなかった。すなわち、問診票で胸痛ありとした例に従来の心血管系疾患危険因子が多く、一方、失神ありとした例には心血管系疾患危険因子はむしろ少ないことがわかった。胸痛ありとした例は、虚血性心疾患による症状を反映していた可能性があることから、健診時安静時心電図検査と問診票を併用することにより、職域健診受診者のリスク層別化と高リスク受診者の抽出が効率的に行なえる可能性が示唆された。なお問診票で捕らえられた失神および動悸は虚血性心疾患との関与が低いと考えられた。また、問診項目と心電図との関連からは、胸痛ありとした例にAfあり、Long QTあり、ST変化あり、T波変化ありが多く、動悸あり例にはAfあり、ST変化あり、T波変化ありが多かった。このことから、胸痛、動悸がある例では、安静時心電図に所見を有することが示唆された。失神例には安静時心電図検査は有用性が低いと考えられた。

研究2においては、職域健診受診者、総合健診受診者を対象とし、循環器疾患スクリーニングを目的とした問診票を新たに作成した。職域健診受診者を中心としたデータ収集が進行中であるが、現在までのところ

十分な解析結果が得られていない。
平成 16 年度内に集計、解析が終了
する見込みである。

6. 時系列健康診断データの有効利用と精度管理 (分担研究者 吉田勝美)

マッチングの成功率は観察開始後 1 年目は SBP、TC ともに約 80% の成功率であったが、観察年数が増えるほどマッチングの成功率は上昇した。ケースとコントロールの検査値の対応ある差の検定によると、SBP、TC とも発症前の 2、3、4 年目はケースとコントロールの検査値の差が大きく、5 年目、6 年目になると差が小さくなることが明らかとなった。発症前の検査値で発症を判別する感度と特異度から ROC (Receiver Operating Characteristics) 曲線と AUC (Area Under Curve) を求めたところ、AUC より、SBP、TC とも発症 2、3、4 年前の値が発症の判別能力が高いことが明らかとなった。コンディショナルロジスティック分析により、独立変数として、SBP、TC の発症前の各値と 3 回の検査値による平均値、標準偏差、3 回目と 1 回目の検査値の差を変数として分析したところ、SBP では、発症前 1 年から 4 年とも検査値の平均値が取り込まれ、TC では発症前 1~4 年では、発症 4 年前の時点での前後の検査値

の標準偏差が取り込まれた。発症前 1~3 年では、発症 1、2 年前の検査値が取り込まれた。発症前 1~2 年ではそれぞれの年の検査値、発症 2 年前の時点での平均値が取り込まれた。発症前 1 年では、平均値と傾きが取り込まれた。発症 1 年前の検査をマッチングさせたケースコントロール研究では、高血圧と高コレステロール血症の発症は発症前 2~4 年前の検査値、特に平均値が重要であることが明らかとなった。この結果から、高血圧と高コレステロール血症の発症は発症前 2~4 年前の検査値、特に平均値が重要であることが明らかとなった。したがって、健診連続受診者においては、単年度の検査値の高低による判定だけでなく、数回の検査値の変動によって将来の発症が評価されることより、健診の保健指導にも有用な手法となるものと期待された。

血圧高値発症前の Matched Pair により血圧変動を解析した研究では、Matching 前の各々の移動平均の Hypertensive group と Reference group のペア比較から、個人値 AV は、発症時期に近づくに従って Hypertensive の方が Reference より高値での出現頻度が高くなっていた。個人値 SD も、発症時期に近づくに従って Hypertensive の方が Reference に比してわずかに高い方

へシフトしていた。個人値 SLOPE ではペア間の差がみられなかった。次に、Hypertensive group の Matching 前 3 個の各 4 点における移動平均の時期による比較において、個人値 AV では、早期は低値に分布していたが時間経過とともに高値側へシフトし、発症に近づくに従って値が上昇していることが分かった。個人値 SD では、発症に近づくに従って僅かながら変動が大きい方にシフトしているが、明確なものではなかった。個人値 SLOPE では時期による差が殆どみられなかった。以上のことから、上述の健診連続受診者における高血圧、高脂血症発症リスクの評価と手法は異なるが、同じような結果が得られ、高血圧症については、約 2 年半から 4 年前という長期間の pre stage があることから、それを見出し、予防対策に結びつけることが有用と考えられた。

D. 結論

全衛連精度管理参加施設に対して行われた精度管理調査において、目標値との一致率を 1993 年から 2001 年までについて解析したところ、総コレステロール、 γ -GT および血糖の参加施設平均値は $100 \pm 1\%$ に範囲内でよく一致しており、トレーサビリティの確認が広く行われていることが示された。中性脂肪は当

初 96%程度であったが、2001 年には 98%程度まで向上した。AST は 99~102%の範囲内にあり今後も $100 \pm 2\%$ の範囲で推移することが期待できると考えられた。同様の調査が 2000 年に日本医師会臨床検査精度管理調査によって行われ、Na、K、Cl、尿酸では 99%以上の一致率であったが、Ca、Hb では若干の問題点が指摘された。総合的に見てこれら精度管理事業の成果により現在の我が国の多くの検査結果についてはかなり高い精度が確保されていることが示された。

さらに過去 11 年間実施された全衛連臨床検査精度管理調査から 10 年間の参加各機関の成績を経年的に調査検討しその特徴及び傾向について解析したところ、参加機関は自機関測定と外部（外注）機関の 2 群にほぼ折半され、同一外部機関に連続して依頼している機関は約半数であり、残りの多くは 1~2 回外注先を変更しており、低い成績が得られたときに次回外部機関を変更している例が散見された。安易に外注先を転々と変更することは一定の信頼性を保証するためには避けるべきことが示唆された。全体的に見ると総コレステロール、血糖の 2 項目は長期にわたって性能が安定していることが分かった。また中性脂肪、AST、 γ -GT についても経年的に高得点の

範囲内に収束してきていることから、基準分析法の勧告と酵素標準物質の供給による効果が著しいことが明らかになった。

次に全衛連臨床検査精度管理調査に参加し特に自機関で臨床検査を実施している約 180 機関の平成 13 年度から 15 年度の最近 3 年間のデータを解析したところ、総コレステロール、中性脂肪、HDL コレステロール、AST、ALT、 γ -GT、血糖、ヘモグロビン A1c、ヘモグロビン、赤血球、白血球、血小板の 12 項目のうち総コレステロール、中性脂肪、AST、ALT、 γ -GT、血糖、ヘモグロビンの 7 項目の検査精度はほとんどの機関において実用上問題がなく適正な水準にあると判断された。とりわけヘモグロビン、血糖、総コレステロールは、ほとんど全ての機関が極めて高度な測定水準を維持していると考えられた。しかし、HDL コレステロール、ヘモグロビン A1c、赤血球・白血球・血小板の 5 項目については、精度管理調査方法、各機関の測定技術においてより向上が望まれると判断された。

職域健康診断における労働衛生検査については、我が国では平成元年より全衛連による精度管理が行われており、これらの結果から我が国の外部精度評価は、労働衛生検査の「品質管理」に優れた実績をあげ

ていると考えられた。ただし、併せて調査した諸外国のシステムと比較すると、我が国の外部精度評価は、同じ物質の検体サンプル濃度数が多く、また異常値を定める基準が厳格である一方、1 年あたりの調査回数が少なかった。したがって、調査方法の簡略化と実施回数の増加が、今後の検討課題になりうるものと考えられたために、具体的な簡便法案を作成し、過去に行われた外部精度評価の実際のデータを用いたシミュレーションを行いその実用性を検討した。その結果、異なる 3 つの濃度のサンプルを測定し、 $\text{mean} \pm 3\text{SD}$ 内に入れば 3 点、 $\text{mean} \pm 9\text{SD}$ 内に入れば 1 点を与え、9 点満点で点数を発表、または 7 点以上を目安として認定を与えるという方法を用いると、現行の方法による評価にほぼ矛盾のない結果が得られる可能性があることが示された。今後解決すべき問題点は残っているものの、こうした簡便法を用いることにより、我が国においても、外部精度評価 1 回あたりの費用を削減し、実施回数を増やすことを検討すべきであると結論された。

胸部放射線診断のあり方に関しては、CR システムの多様な機能の利用方法、精度管理を中心にした撮影方法、表示方法についての追加検討を行い、前年度の暫定的まとめの部

分的改定と至適パラメータの最終まとめを作成した。そして胸部ファントムの CR 写真による具体的な検討を追加した。

時系列データを有効に活用するために必要とされる精度管理のあり方について検討する前提として、時系列データの有効な利用方法そのものについて研究を行っているが、今年度は前年の結果を踏まえ、発症直前の血圧及び総コレステロール値を調整して、それ以前の値を考慮した際の発症予見性について検討した。その結果、高血圧と高コレステロール血症の発症は発症前 2~4 年前の検査値、特にそれらの平均値が予測指標として重要であることが明らかとなった。健診連続受診者においては、単年度の検査値の高低による判定だけでなく、数回の検査値の変動によって将来の発症が評価されることより、健診の保健指導にも有用な手法となるものと期待された。

循環器疾患のスクリーニングには、従来問診結果が有効に活用されていなかったが、これも含めた効率的なスクリーニング法を検討し、その際の情報の精度管理のあり方を検討することを目的とした研究では、問診票で胸痛ありとした例に従来の心血管系疾患危険因子が多く、虚血性心疾患による症状を反映していた可能性があると考えられた。健診時安静

時心電図検査と問診票を併用することにより、職域健診受診者のリスク層別化と高リスク受診者の抽出がより効率的に行なえる可能性が示唆された。

E. 健康危険情報

特になし。

F. 研究発表

特になし。

G. 知的財産権の出願・登録状況

特になし。

資料3

職域における健康診断と精度管理のあり方に関する研究

主任研究者 櫻井治彦 中央労働災害防止協会労働衛生調査分析センター

研究要旨

①職域健康診断に係る精度管理の現状を調べ問題点と今後の改善の方向を明らかにすること、②個人の時系列データを利用して健康診断の意義を高めること、③従来利用されることの少なかった問診情報を取り入れてスクリーニングの効率を高めることを目的とした。

現状と問題点については職域健康診断の大部分を実施している健康診断機関に対して全国労働衛生団体連合会が毎年行ってきた外部精度管理事業の成績を分析評価した。個人の時系列データの利用に関する研究では事業場従業員について長期間にわたり継続的に行われた健康診断結果データベースを解析した。問診情報を取り入れてスクリーニングの効率を高める方法を検討した研究では、スクリーニング、精密健診、治療を一体的に行っている医療機関の受診者を対象として、問診情報の有効活用について検討した。

全国労働衛生団体連合会（全衛連）臨床検査精度管理調査によると比例系統誤差の頻度が高かった。年1、2回のトレーサビリティの確認と校正を行うべきと考えられた。大規模検査機関の大半はほぼ適正な人的資源を有し臨床検査の精度の水準も高かった。労働衛生検査項目については、内容と評点を簡略化した外部精度評価を追加することを提案した。CT装置による肺がん検診については存在診断については学習効果が得られやすかったが、低X線曝射CT像による質的診断は困難であることがわかった。一般定期健診については健診データの経年的評価モデルにより、データの動態が的確に把握できるため、異常性の予見が可能となることがわかった。循環器疾患スクリーニングを目的とした問診票を用いた問診結果と冠動脈造影検査結果とを照合することにより職域健診受診者のリスク層別化と高リスク受診者の抽出が効率的に行える可能性が示唆された。

結論として、臨床検査精度管理の問題として比例系統誤差の頻度が高くトレーサビリティの確認と校正を行うべきことなどいくつかの問題が指摘された。簡略化した外部精度評価の追加が提案された。CT装置によるがん検診では存在

診断について学習効果が得られやすく質的診断は困難であることがわかった。一般景気健診データの経年的評価モデルによりデータの動態が把握でき異常化の予見が可能となることがわかった。問診と冠動脈造影検査を併用し受診者のリスク層別化と高リスク受診者の抽出が効率的に行える可能性が示唆された。

分担研究者

- 徳永 力雄（関西医科大学名誉教授）
栗原 伸公（神戸女子大学家政学部公衆衛生学助教授）
曾根 脩輔（長野県厚生農業協同組合連合会安曇総合病院院長）
吉田 勝美（聖マリアンナ医科大学予防医学教授）
久代登志男（日本大学医学部付属駿河台病院循環器科助教授）

研究協力者

- 中 甫（日本福祉総合医療研究所所長）

A. 研究目的

職域における健康診断は、労働環境の有害要因による健康異常の早期検出、有害要因曝露量の推定、作業関連疾患の早期発見、労働環境に必ずしも関連を持たない健康障害の早期発見、健康度の把握等、さまざまな目的で実施されてきた。今後は、事業者による自主的な労働衛生管理

の重要性がますます高まり、また職域健康管理を労働者個人の生涯にわたる健康管理の一部として位置付ける必要があることなどから、職域における健康診断の有効性及びその基礎としての精度を確保することの重要性が認識されてきている。

本研究では、従来の健康診断のあり方を一歩進め、より効果的に労働者の健康維持、増進に資するためには、健康診断の測定および判定における精度の向上がきわめて重要であるとの認識に立ち、各種健康診断における精度管理の現状を把握し、今後のあり方を検討することを第一の目的とした。その際、基本的に健康者を中心とする集団である職域の労働者に対する健康診断においては、地域あるいは臨床での一般人に対する健康診断より高い精度管理が必要であることを念頭において検討を行った。

また今後、職域、地域、などで得られる健康診断情報を一元化し個人毎の時系列データを有効に活用することが徐々に進められる機運にあるが、職域における健康診断に限っても、個人の時系列データの活用はき

わめて重要である。そこで本研究では、時系列データを有効に活用するために必要とされる精度管理のあり方について研究することを第二の目的とする。

更に、従来利用されることの少なかった問診情報を取り入れてスクリーニングの効率を高めることを第三の目的とした。

B. 研究方法

櫻井と中は、全国労働衛生団体連合会（全衛連）臨床検査精度管理調査の1999年度から2003年度までの5年間に参加した全機関を対象として全衛連参考値検討委員会において真値とみなし得る目標値を設定している総コレステロール、中性脂肪、AST、ALT、 γ -GT、血糖の6項目について、60点未満である項目が1回でも存在する機関のデータを抽出し、機関ごとに5年間の成績を入力し、その結果を自機関と外部機関別に想定できる誤差原因別に分類してそれぞれの件数を調査するとともに、誤差を生じる原因及びその誤差を是正する方法について検討した。

徳永は、全衛連臨床検査精度管理調査に参加した健診機関310機関のうち、健診実施人数が年間延べ20万人以上の66機関の1996年度から2004年度の9間の年度別精度管理調査総合評価点及び検査項目別評価点、2003年度延健診実施数、職員数、常勤医師数、非常勤医師数、検査技師

数を集計し、また各都道府県から健診実施人数が最も多い1機関ずつを選んで、2002～2004年度の検査項目別評価点、総合評価点、70点未満の検査項目を算出して、これら大規模健診機関における問題点を検討した。また健康診断等のあり方について、大企業の専属産業医等4名及び臨床検査精度管理研究者6名と健康診断の目的・意義、法規則、企業責任と産業医責任、検査方法・項目・実施間隔、等の課題を討議して意見を集約した。

栗原は、昨年度までに諸外国との比較の中から、わが国の労働検査項目に対する外部精度評価の在り方として、回数を増やすこと、また評価の内容（方法および評価点）を簡略化することを含めて検討することを提言し、さらにシミュレーションによってその具体的試案を作成してきたが、本年度はそれら提言や試案の実現可能性を調べるために、実際の現場で外部精度評価を受けている機関、すなわち、全衛連による労働衛生検査の外部精度評価に参加している全検査機関（全衛連会員138、非会員186、計324機関）の精度管理担当者に対して郵送法によるアンケート調査を行った。アンケートの内容は、A 外部精度評価について（回数・評価・チェックするサンプルの数に関する意見、評価に対する受け止め方、我々の提言・試案に対する意見、機関どうしの情報交換につい

て)、B 内部精度管理について(頻度、方法、マニュアルの有無、自己評価など)、C 検体保存方法について D 精度管理に対する意見、提言、であった。

曾根は、肺がん検診で近年重視されてきている CT 検診の精度管理、特にこれを担当する医師の養成、そのレベル向上に向けた検討、方法論の早急な確立が望まれていることを踏まえ、医師の診断能力及び学習効果の評価法を検討する目的で検討を行った。早期肺がん症例や非がん、正常例の CT 像を混在させて編集した CD-ROM をテスト用材料として、医師の診断能力を、正診率と誤診率で測定し、ROC 解析を用いて評価した。CD-ROM には肺がん、非肺がん、正常の CT 像、延べ 270 例分が収録されていた。新人医師が勤務に入る直前、一定期間の修練を経たあとなどに読影実験を行った。画像は CRT 装置画面(19 インチ)上に表示した。読影医は、その画像を読影、異常所見の有無、その診断などで回答した。読影実験に用いた CT 像は 1996 年度から 1998 年度の 3 年間に長野県の 29 自治体の一般住民からの希望者を対象にして行った検診から得たものであった。検診は同意書を得て行われ、本研究は JA 長野厚生連安曇総合病院の倫理審査委員会の承認を得て実施された。

吉田は、前年度に個人の経年的な健診結果の評価方法を検討し、評価

モデルを開発したが、本年度は、評価項目を生活習慣病に関わる 9 項目に広げ、項目毎に評価モデルの有用性の検討を行った。また、数式的表現である評価モデルを保健指導の現場で活用できるように個人に分かりやすい言語的表現への置き換えを試みた。収縮期血圧(sBP)、拡張期血圧(dBP)、総コレステロール(TC)、中性脂肪(TG)、GOT(AST)、GPT(ALT)、 γ -GTP、FBS、UA の 9 項目を対象として、matched case-control study を行い、項目毎の評価モデルの有用性を検討した。また、MA(移動平均)、SD(標準偏差)、Slope(傾き)の指標を用いた判定結果を言語的に表現した。

久代は、循環器疾患の効率的なスクリーニングと一次予防のための職域健診のあり方について検討しているが、昨年度までの成果により、健診時安静時心電図検査と問診票を併用することにより、職域健診受診者のリスク層別化と高リスク受診者の抽出が効率的に行える可能性が示された。そこで本年度は、職域健診受診者、総合健診受診者を対象とし、循環器疾患スクリーニングを目的として作成した問診票(①胸部症状の有無、②症状の持続時間、③症状の部位、④症状の広がりに関する設問)を冠動脈造影検査、心臓電気生理学的検査、Head up tilt 試験などの循環器疾患精密検査を施行した症例に行い、精密検査結果と問診票の回答

とを照合することにより循環器疾患スクリーニングにおける問診票の validation study を行った。冠動脈造影上①有意狭窄病変（狭窄度 75%以上）を認めたか、②造影時のアセチルコリン負荷試験により冠攣縮が誘発された場合を“IHD あり”（虚血性心疾患あり）とした。

C. 研究結果

全国労働衛生団体連合会（全衛連）臨床検査精度管理調査に過去 5 年間連続して参加した全機関を対象として全衛連参考値検討委員会において真値とみなし得る目標値を設定している総コレステロール、中性脂肪、AST, ALT, γ -GT、血糖の 6 項目について、60 点未満である項目が 1 回でも存在する機関のデータを抽出し、誤差を生じる原因について調査した結果では、自機関で 60 点未満の成績を 1 回でも示した機関は 79 機関で、5 年間の総機関数 866 機関の 9.1%であった。外部機関では、複数の機関から受託している例があり、60 点未満の成績が影響を与えた機関数から計算すると 58 機関存在し、5 年間の総機関数 765 機関の 7.6%であり自機関に比較すると僅かに少ない結果となった。しかし外部機関は健診機関の事業の一部として測定を行っている自機関とは異なり、測定を専門的事業として行っているにもかかわらず 7.6%存在することは重要な問題であると考えられた。精度

管理調査の成績が 60 点未満となった理由として、試料濃度に関係なく一定の比率で誤差が生じる比例系統誤差が多かった。この種の誤差は、高位標準物質を用いるトレーサビリティの確認により是正が可能であるが、約 70%の機関がこれを実施していないことが判明した（櫻井、中）。

全衛連臨床検査精度管理調査に参加した健診機関 310 機関のうち、健診実績が年間 20 万件以上の大規模検査機関 66 機関の年間健診総数は 3,476 万件余、平均常勤医師数は 8 人、平均技師数は 44 人、常勤医師 10 名以上を擁する機関 21、技師を 50 人以上擁する機関 19、技師 1 人あたり健診数が 2 万件以上の機関が 13 であった。大規模機関の大半はほぼ適正な人的資源を有しかつ臨床検査の精度の水準も高く高度な予防医学活動を展開している機関であることを示す結果であった。各都道府県で健診実績が最も多い機関と、全国平均とを比較したところ、白血球、ヘモグロビン A1c 以外では全国レベルよりやや良好な精度管理状態にあることがわかった。健康診断等のあり方について産業医及び臨床検査研究者より面接により意見を聴取し、健康診断と精度管理に付いて以下の意見・提言が得られた。1)健康診断の目的は 2 次予防から 1 次・3 次予防まで拡大している、2)健康診断は保健医療スタッフが従業員へ介入する手段として重要で、事後措置が労

働者の健康確保に大きな役割を果たしている、3)健康診断は労働衛生活動の重要なインフラであり日本固有の企業文化としてのニーズがある、4)業種・事業場など対象集団の特性に対応した柔軟な健康診断の法規定が望ましく、検査項目・対象・頻度は企業の裁量性に委ね、基本事項はガイドラインで示すべきである、5)企業外健診機関の低コスト主義による品質低下の防止策が必要である、6)臨床検査外部精度管理は重要で一層充実すべきである、7)精度管理試料のマトリックスへの配慮、試料の複数化・ランダム化、頻回のサーベイ実施、評価方法の改善が必要である、8)サーベイ実施機構・団体の認証制度、第三者評価が必要である、9)検査技師の生涯教育体制を充実すべきである(徳永)。

全衛連による労働衛生検査項目についての外部精度評価に参加している全検査機関(全衛連会員 138、非会員 186、計 324 機関)の精度管理担当者に対して、郵送法によるアンケート調査を行ったところ、現行年 1 回の評価回数、4 段階の評価指標、8 項目の評価項目、6 検体の検体数については、それぞれ現行の制度を支持する機関が圧倒的に多く見られた。しかし、アンケートにおいて、外部精度評価の回数増加と簡略化という提言を示し、意見を求めたところ、明確に反対とする意見も 15%ある一方で、全体の約 4 割からその基本

的な方針に一定の支持が得られた。ともに合格と考えられる A 評価と B 評価との機関側での受け止め方の違いを尋ねた結果、深刻に受け止める機関が多く、特に外部の見方が厳しいとのことであった。情報交換については懸念があり、簡略化は難しいとする意見も少なくなかった。だが、簡略化によって、評価についてはチェック程度のものであると考えるように外部の意識も含めて変えることができれば、情報交換はあえて必要なくなるのではないかと思われた。一方で現行の高度な評価も年に 1 度は残し、それに年に 2 度簡略化した方法を加えるのがよいと考えられた。内部精度管理については、8~9 割の機関で適切に行われていることがわかったが、1 割弱の機関では、そうでない可能性が推察された。検体の採取法は適切とはいえないところも多く見られた(栗原)。

CT 装置を用いる肺がん検診での医師の診断能力の評価法を検討し、それを基に学習効果を調査した。早期肺がん症例、非がん症例、正常例の CT 像を 270 例混在させて編集した CD-ROM をテスト用材料として学習効果をみると、肺がん、肺がん以外の病変、正常の CT 像に対する正答率は、研修開始時にそれぞれ、45.8%、24.3%、95.7%であったが、研修後には、75.6%、60%、91.3%と上昇した。存在診断については学習効果が得られやすかったが、質的診

断については学習効果に個人差があり、低X線曝射 CT 像による質的診断は困難であることがわかった（曾根）。

Matched Pair を用いて生活習慣病に関わる収縮期血圧(sBP)、拡張期血圧(dBP)、総コレステロール(TC)、中性脂肪(TG)、GOT(AST)、GPT(ALT)、 γ -GTP、FBS、UAの9項目を対象として、検査項目毎に発症直前の値をマッチングさせ、それ以前の値の特徴量を見て発症予見性について検討した。某製造業の健診データを使用し、6年連続の受診者のうち男性4,756名を対象とした。解析方法は、1年目の収縮期血圧値(sBP)が140mmHg未満かつ高血圧の治療を受けていない者のうち、6年間に受診時に高血圧を発症した者(sBP値が140mmHg以上であった者)をケースとし、一度も発症しなかった者をコントロールとした。同様に他の項目についてもケースとコントロールを定義し、これらのケースとコントロールによる matched case-control study を行い、その結果から得られた項目毎の評価モデルの有用性について検討を行った。また、MA(移動平均)、SD(標準偏差)、Slope(傾き)の指標を用いた判定結果を言語的に表現した。項目毎に発症前の検査値で発症を判別する感度と特異度から ROC (Receiver Operating Characteristics) 曲線と AUC(Area Under Curve)を求め、有

効性(識別性)の評価を行った。(1)収縮期血圧値(sBP)では発症2年前>3年前>4年前>5年前>1年前の順で判別能力が高かった。同様に、(2)拡張期血圧値(dBP)では、発症2年前>4年前>3年前>5年前>1年前、(3)総コレステロール値(TC)では、発症2年前>3年前>4年前>5年前>1年前、(4)中性脂肪値(TG)では、発症2年前>3年前>4年前>5年前>1年前、(5)GOT値では、発症2年前>5年前>3年前>4年前>1年前、(6)GPT値では、発症5年前>4年前>2年前>3年前>1年前、(7) γ -GTP値では、発症4年前>5年前>2年前>3年前>1年前、(8)FBS値では、発症2年前>4年前>3年前>5年前>1年前、(9)UA値では、発症2年前>3年前>4年前>5年前>1年前の順で判別能力が高かった。次に、項目毎にケース及びコントロール群の値の変動を調査し、コンディショナルロジスティック分析による評価を行った。コンディショナルロジスティック分析では独立変数として、発症前の各検査値と3回の検査値による平均値、標準偏差、3回目と1回目の検査値の差を変数として分析した。これらの結果から、Matched Pairによる経時的評価モデルの適用に関しては、ROC曲線から判断すると、sBP、dBP、TCの有効性が高く、AUCの値とケース群とコントロール群の値の変動から判断すると、sBP、dBP、TC、TG、

FBS, UAの有効性が高いことがわかった。以上により、健診データの経年的評価ができるモデルにより、データの動態が的確に把握できるため、悪化傾向に向かうようなグループに対しては、異常性（悪化）の予見が可能となった。また、数式的表現である評価モデルを保健指導の現場で活用できるように、個人に分かりやすい言語的表現を提案した（吉田）。

職域健診受診者、総合健診受診者を対象とし、循環器疾患スクリーニングを目的として作成した問診票を用いた問診を冠動脈造影検査、心臓電気生理学的検査、Head up tilt試験などの循環器疾患精密検査を施行した症例に行い、精密検査結果と問診票の回答とを照合することにより循環器疾患スクリーニングにおける問診票の validation study を行った。臨床的に虚血性心疾患を疑われ、問診票の回答が得られ、冠動脈造影検査を施行した 53 例について解析した。①胸部症状のうち、“胸痛あり”については、感度 73.3%、特異度 12.5%、陽性反応的中度 82.5%であった。②症状の持続時間については、感度 46.9%、特異度 50%、陽性反応的中度 78.9%であった。さらに“～30 秒”と“30 秒～”とに分けて検討したところ、“IHD あり”のうち“30 秒～”と回答した場合は、感度 68.8%、特異度 37.5%、陽性反応的中度 81.5%であった。③症状の部位に関する設問の回答からは感度 0～

35.6%、特異度 62.5～100%、陽性反応的中度 75～100%が得られた。さらに感度、特異度が比較的良好であった部位をまとめることにより、感度 57.8%、特異度 50%、陽性反応的中度 86.7%が得られた。胸部症状の広がりに関する設問は、回答を組み合わせて解析した。指先で示せる範囲と胸全体指先で示せる範囲とをまとめて解析すると感度 60%、特異度 62.5%、陽性反応的中度 87.5%であり、こぶし位～手のひら位と胸全体とを合わせて解析した結果は、感度 77.1%、特異度 35.7%、陽性反応的中度 84.4%であった。今回の結果から、問診票による虚血性心疾患検出の感度は 60～80%、特異度は 12.5～100%、正確度は 37.5～62.8%、陽性反応的中度は 81.5～86.7%であり循環器疾患スクリーニングとして使えることが示唆された（久代）。

D. 考察

全衛連が過去 12 年間に年 1 回実施してきた臨床検査精度管理調査のうち、近々 5 年間の報告書に見られる 60 点未満の成績を示した項目及び機関について詳細な調査を実施したが、その理由は、12 年間の長期間では、測定法、試薬、分析装置等がその間に変更されている例数が多く含まれる可能性があるためである。自機関測定の場合では、過去 5 年間に 60 点未満の成績が 1 項目以上存在した機関数は合計 79 機関であり、全体としては過去 5 年間の機関数総計