

厚生労働科学研究費補助金
労働安全衛生総合研究事業

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

平成16年度 総括・分担研究報告書

主任研究者 櫻井治彦

平成17年3月

平成16年度目次

I. 総括研究報告

- 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 …………… 1
櫻井 治彦

II. 分担研究報告

1. 職域健康診断における臨床検査のあり方と精度管理 …………… 16
徳永 力雄
2. 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 …………… 31
栗原 伸公
3. 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 – 胸部 CT 画像読影の能率と精度
および学習効果の評価法に関する研究 – …………… 88
曾根 脩輔
4. 時系列健康診断データの有効利用と精度管理 …………… 91
吉田 勝美
5. 循環器疾患の効率的なスクリーニングと一次予防
のための職域健康診断のあり方に関する基礎的検討 …………… 103
久代 登志男

III. 個別研究

- 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 …… 110
中 甫
- 胸部 CT 画像読影の能率と精度および学習効果の評価法に関する …… 131
研究
曾根 脩輔

厚生労働科学研究費補助金（労働安全衛生総合研究事業）
総括研究報告書

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あり方として、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。自機関測定の場合には、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866機関の9.1%に相当した。しかしそのうち5年間で1項目のみが60点未満であった機関数は56機関で全体の6.5%で最も多い。一方、外部機関では過去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20機関で自機関より明らかに少ないが、これら外部機関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合計58機関となり、全体としては過去5年間の機関数総計765機関の7.6%であり自機関測定に比較して僅かに少ない結果となっている。しかし外部機関は健診機関の事業の一部として測定している自機関とは異なり、測定を専門的事業としているにもかかわらず7.6%も存在することは、受託機関にそれぞれ多数の

健診データが含まれていることを考慮すると自機関以上に重大な問題であると考えられた。60点未満の成績を示した項目についてみると自機関では5年間に最大3項目で、4項目以上は見られなかったのに対し、外部機関では1項目のみ；10機関、2項目：5機関、3項目：3機関、5項目：2機関が60点未満であり、5項目が2機関も存在していた。しかし、大きな影響を与えた外部機関3箇所を除き、その他の機関は複数項目であっても影響を与えている機関が1機関でありしかも同一機関であることは、問題がありながらも健診機関への影響は最小限に抑えられていると考えることができた。精度管理調査の成績が60点未満を示す理由として第一に測定誤差の存在が考えられるが、その内、比例系統誤差は最も起こりやすい誤差の一つで、今回の調査でも自機関、外部機関ともに最も多く見られる誤差であった。一定系統誤差はごくまれに見られる誤差であるが、今回の調査では、後に述べる特定の項目と外部機関においてそれを示唆するデータが存在した。その他、通常測定誤差とは異なるが、測定後のデータを転記する際の転記ミスや測定の際の試料の取り違えなども少なくなかった。また、測定法の性能に関連する誤差である直線性の欠如が示唆される問題も存在した。最後に調査試料の送付に関連

する問題も残っていた。測定誤差はこれら種々の原因により生じるが、各機関内における内部精度管理の普及と実践により測定の精密さすなわち偶然誤差はきわめて小さい範囲内に管理・維持できるようになり、測定値の施設間差は、確認ミスのような偶発的な誤差を除けば、測定値の正確さに関連するかたよりが誤差原因の主流を占めるようになっていく。調査結果からもかたよりの中でも比例系統誤差が大部分を占め、その原因が主として測定の標準となる検量用試料の正確さを含むキャリブレーション時に生じる誤差であることが判明している。これらのかたよりの原因となる比例系統誤差を縮小するには、正確さが明確な高位標準物質を日常一般法で繰り返し測定を行い、その平均値とこれら高位標準物質の標準値の一致性を確認するトレーサビリティの確認を実施し、もしかたよりが存在したときにはそれを校正する作業を行うことにより達成できる。しかし、現状では高位標準物質を用いてトレーサビリティの確認を実施している機関は著しく少なく、全く実施していない機関が70%も存在することは憂うべき状況であると考えられた。この状況は日本医師会が年1回実施している臨床検査精度管理調査の結果においても同様である。測定値の施設間差の原因の大部分が校正可能な比例系統誤差であ

ることは、一方では測定値の信頼性の保証は容易な状態にあることを意味している。現在、定量検査が行われている全ての検査項目について高位標準物質の開発が実現しているわけではないが、比較的件数の多い日常的な検査項目については既にこれらの高位標準物質が存在するので、これらの項目については各機関において少なくとも年1、2回の頻度でトレーサビリティの確認と校正を実施すべきであると考えられた（櫻井、中）。

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との面接により意見を聴取した結果は主観的、定性的なものであり、特に専属産業医の面接者が限られていることによるバイアスは十分考慮しなければならない。しかし、全国的かつ国際的企業の産業医であるのでその視点は将来を洞察するにふさわしく、示唆に富む討議であった。健診の有効性や有用性の評価指標として、死亡率や寿命がどの程度妥当であるか、他の指標、例えば有所見に対する保健・生活指導など早期対応による就労期間と退職後のQOLの向上も指標になり得ないか、など評価軸に関する提案と疫学研究の必要性が指摘された。労働安全衛生の他のインフラとして、作業環境管理、作業管理、労働安全衛生マネジメントシステム、健康づくり・1次予防、などがある

が、これらは作業環境管理を除いて未整備の状態を実効性が担保されていないという指摘もあった(徳永)。

全衛連の現行の労働衛生検査項目に関する外部精度評価については、その回数、評価方法、評価項目数、検体数について、現行の方法を支持する機関が非常に多数を占めたが、これは、現行の方法により参加機関において一定の満足が得られていることを意味しており、やはり全衛連において長年検討され改良されてきたものを超える方法を見出すのは相当難しいことがわかった。しかしながら、今回提案された簡略法については、賛成、反対の両意見が各 10 数%ずつあり、多様な意見が見られたが、外部精度評価の回数を増やし、それぞれについては簡略化も導入するという方向性については、概ね 4 割の機関で支持された。今後この試案をもとに、改良された方法を、より明確な形で提示することができれば、現行法に加えることにより改善が可能と考えられた。情報交換については、多くの機関で懸念されており、早期に解決すべき問題であると考えられた。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、数%から 1 割程度の機関では、そうでない可能性があることが推察された。一方、外部精度評価において、若干の問題を含む機関も数%から 1 割程度存在

する。今回のアンケート調査では、内部精度管理と外部精度評価で問題となる機関が一致しているかどうかについては解析できなかったが、両者が共通している可能性はあるので、外部精度評価の結果が悪い機関については、内部精度管理の状況を詳しく調べる意義もあると考えられた(栗原)。

胸部 CT 検診では大量の受診者の CT 画像読影が対象となり、異常所見の有無判定に基づく要精検または精検不要の判定が読影医の主な仕事である。連続確信度で回答する習慣はない。CT 検診像の読影診断に不慣れな若手医師が、日ごろ見慣れていなかった画質の CT 画像を読影するにあたり、異常所見の拾い上げと鑑別診断の 2 段階の過程が要求される。異常所見の拾い上げについてはほぼ 100%の感度が要求され、質的診断については高い感度と、ほぼ 100%の特異度が望ましい。このような目標に向かって如何に効率的に学習成果をあげるかが重要であるため、その評価法を検討した。今回の検討では、さらに、CT 画像をモニター診断時の読影所要時間も測定して作業能率や診断確信度の関係もみたが、さらなる検討のうえの結論づけが必要と思われた(曾根)。

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。既に、我々は健診データの時系列的な解析により、ホメオスタティック（恒常的）な項目は時系列的な解析に適しているという結果を得ている。一方、上記の項目の中で TG 及び UA はホメオスタティックな項目とは異なり、通常はデータの変動が大きく時系列的な解析には適さない項目であると考えていた。しかし、それらの動態は病的で高値のレベルを含めた全体のデータ解析の結果であり、今回のようにデータが基準範囲に近いレベルを推移している場合、それらの項目に関しても、ホメオスターシス機能が維持され、ホメオスタティックな項目と同様に経時的評価を適用することが可能となったと考えられた。また、本来、FBS はホメオスタティックな項目であると考えられるが、ROC 曲線の年次の順序性が不規則であることから、経時的評価に適さないという結果が得られた。これは、今回使用した健診データの精度の信頼性に問題があったのではないかと思われた。次に、既に得ている知見をもとに、時系列的なデータ解析に適している項目の sBP, dBP, FBS, TC を健康変動 battery とし、データの変動性が高く、ホメオスタティックな項目ではない UA, GPT（肝炎な

し）、TG, γ -GTP, GOT を病態変動 battery としてグループ化した。健康変動 battery に属する項目は、時系列的なデータ解析により、異常値を予見できる可能性が高い。一方、病態変動 battery の項目に関しては、異常値を予見することは難しく、発症前の介入時期を特定することが難しいと考えられた。判定結果の言語的な表現については、従来の一時点のデータから判定する方法では、悪化しても改善しても同じ判定分類で説明されることが多いのに対し、時系列的なデータ解析結果を利用すれば、「データの状態を正確に評価できる」、「今後のアクションを的確に示せる」という 2 つの要素が得られる可能性が指摘された（吉田）。

定期健康診断等の結果、脳・心臓疾患を発症する危険性が高いと判断された労働者に対して、脳血管及び心臓の状態を把握し、二次健康診断及び脳・心臓疾患の発症の予防を図るための労働者災害補償保険による二次健康診断等給付制度があるが、一次健康診断結果が給付要件を満たさない場合は検査が行われず、虚血性心疾患の診断には運動負荷心電図検査が有用であるが、職域健診の場で全例に運動負荷心電図検査を施行することは困難であり、かつ効率が悪い。そこで、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虚血性心疾患が疑われ、冠動脈

造影検査を行った症例に施行し、検査結果と問診表の回答とを照らし合わせ validation study を行い循環器疾患の効率的なスクリーニングとための職域健診のあり方について基礎的検討を行った。症例数の制限から今回は虚血性心疾患に関する検討にとどまった。偽陰性率が高かったが、問診票による各質問の回答に対す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%であった。今後は、失神、動悸に関する問診回答と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との対比を含め、症例数を増やした検討が必要であり、さらに問診票の精度の向上（感度を保ち、偽陰性率を低くする）のために問診結果のスコア化を試みるなどの課題が残ると考えられた（久代）。

E. 結論

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、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、自機関で総機関数の 9.1%、外部機関で 7.6%であったが、外部機関の 7.6% は複数の健診機関に影響を与えている例があるので、その影響は重大であることがわかった。最も頻度の高い比例系統誤差は、容易に校正可能な誤差であることから、年 1、2 回はトレーサビリティの確認と校正を行うべきであり、以上を実践することにより臨床検査の信頼性向上が果たせるとともに個人の検査データについて信頼性の高い経年的管理が可能になると考えられた（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 機関の大半は、ほぼ適正な人的資源を有し、かつ臨床検査の精度（正確さ）の水準も高く、相当高度な予防医学活動を展開している機関であると考えられた。46 都道府県で健診実績が最も多い各 1 機関の検査項目別精度管理の成績は、全衛連調査の全機関の傾向とほぼ同じ水準にあり、自機関測定機関は外部依託機関に比して 70 点未満の低評価の割合がやや高率であった。健康診断及び臨床外部検査精度管理のあり方について、産業医及び臨床検査の研究者と討議して健康診断の目的・役割、法規則の課題、企業責任と産業医責任、労働衛生活動の変容と課題、

臨床検査の精度管理の課題等に関して有意義な意見と提言が得られた(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結果、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が、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。結論として、外部精度評価に関して次のような案が提言された。①現在の全衛連の外部精度評価を、情報交換がより起こりにくい形で年1回行う。②8項目、共通3サンプルについて、適、不適のみの判定の精度評価を1に示した年1回の調査の4ヶ月後および8ヶ月後に年2回、計3回行う。③1、2の精度評価は同一評価団体、あるいは連合体で、すべてワンセットとして行う。④評価については、A、Bの評価よりも、「適」をどれだけ続けてとるか重要視するように宣伝する。⑤「不適」となった機関には、通常の原因究明および指導に加えて、内部精度管理の方法、検体の採取・保

管方法についてもより徹底的に指導する(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結果では、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。CT画像読影による存在診断と質的診断の学習および研修の効果判定面での有用性が示唆された(曾根)。

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結果、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、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)の予見が可能となり、早期発見、早期介入が可能となることがわかった。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について個人に分かりやす

い言語的表現を提案した。これにより、今後、時系列的なデータ解析アプローチを発展させることにより、高精度な EBM に基づく保健指導の介入指標を開発できる可能性が示唆された（吉田）。

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 を行ったところ、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（久代）。

F. 健康危険情報

特になし。

G. 研究発表

1. 論文発表

Armato SG III, Altman MB, Wilkie J, Sone S, Li F, Doi K. Automated lung nodule classification following automated nodule detection on CT: A serial approach. Med Phys 2003; 30(6):1188-1197.

2. 学会発表

伊津野 孝, 小林 祐一, 上原 正道, 中根 弥枝, 座間 聡子, 松木 彩子, 柴山 順子, 佐々木 敏雄, 東 敏昭, 吉田 勝美, 杉田 稔

健康診断データの時系列的評価方法の検討 1、2

第 78 回産業衛生学会

H. 知的財産権の出願・登録状況

特になし。

厚生労働科学研究費補助金（労働安全衛生総合研究事業）
分担研究報告書

職域健康診断における臨床検査のあり方と精度管理

分担研究者 徳永力雄 関西医科大学常務理事・名誉教授

研究要旨

(1) 全衛連臨床検査精度管理調査参加 310 機関から健診実績が年間 20 万件以上の大規模健診機関 66 機関を選び、医師数、技師数、1996～2004 年度の総合評価点、検査項目別平均評価点、70 点未満の機関数等から精度管理の水準を分析した。66 機関の健診総数は 3,476 万件余、平均常勤医師数は 8、平均技師数は 44、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、等であった。66 機関の平均総合評価点は 9 年間とも 90 点以上で、90 点未満は 9 年間延べ 577 機関の 15.4%、2000 年以降の 5 年間では延べ 328 機関の 10.1%であった。80 点未満は 8 機関であった。検査項目別平均点は全項目とも 90 点以上で、例年成績が相対的に悪い HDL コレステロール、白血球、ヘモグロビン A1c も 90 点以上であった。以上より、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と推察された。

(2) 46 都道府県から健診実績が最も多い機関を 1 つ選んで、臨床検査項目ごとに 2002～2004 年度の項目別平均評価点が 70 点未満の機関数を調べた。精度管理成績が例年悪い HDL コレステロールは 3.6%、 γ -GT は 5.1%、白血球は 9.4%、血糖は 5.8%、ヘモグロビン A1c は 8.7%が 70 点未満の低い評価であったが全国の傾向と同水準であった。70 点未満の機関の割合は自機関測定の方が外部委託測定に比してやや高率であった。

(3) 健康診断等のあり方について産業医及び臨床検査の研究者と討論し以下の意見・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る・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い、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) 健康管理事業を年単位で実施する制度は有用である、6) 産業医の業務は企業の責任で契約等で定めるのがよい、包括的に産業医責任を問うのは負担が重い、7) 営利・効率主義が強調され労働衛生の地盤が沈下している、衛生管理者等の人材育成の土壌が劣化している、8) 第1次予防活動を強化する制度を導入すべき、9) 企業外健診機関の低コスト主義による品質低下防止策が必要である、10) 臨床検査外部精度管理は重要で一層充実すべきである、11) 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、12) サーベイ実施機構・団体の認証制度、第3者評価が必要である。13) 検査技師の生涯教育体制を充実すべきである。

研究協力者

川合陽子 慶応大学医学部附属
病院中央臨床検査部
講師

久保野勝男 株式会社エスアール
エル医科学分析セン
ターセンター長

高橋正宣 練馬区医師会医療健
診センター顧問

中 甫 日本福祉総合医療研
究所所長

西尾宏信 関西医科大学衛生学
助手

福田豊志 有限会社シードウィ
ン取締役

森 雄一 財団法人神奈川県予
防医学協会集団検診
センター副所長

A. 研究目的

健康情報の生涯一貫した一元的管
理の視点から職域健康診断の臨床検

査精度管理において全国の大規模健
康診断・検査機関等（健診機関）の
現状と課題を最新の資料に基づいて
分析する。また、産業医等からみた
今日の職域健康診断と健康管理の課
題を分析して、今後の健康管理、健
康診断、臨床検査精度管理のあり方
について考察し提言する。

14、15年度の精度管理調査資料の
分析により、健康診断実施機関の臨
床検査精度管理と臨床検査値の精度
は、検査項目、試料濃度、機関所在
地のそれぞれの観点において、約
10%の機関を除いて実用上問題とな
る課題は少ないことが判明した。一
方、国民のライフステージを通じた
一貫した健康管理に役立つ健康情報
の一元管理と効果的な活用はこれか
らの施策として必要である。職域の
健康診断で得られる臨床検査デー
ターは30年以上にわたって蓄積され
るために、そのデータの信頼性と

活用は、国民の健康保持の観点から重要な役割をもっている。

個人の臨床検査データは、医療機関受療中を除くとおおむね就学前、就学期、就労期、老齢期のそれぞれに所定の法律に基づく健康診断等によって得られる。臨床検査そのものは、本研究が対象としている就学期と就労期の集団健康診断実施機関によってその大部分が実施されている。そこで、生涯を通じた個人別健康情報の一括集積と利用の観点から、何時でも、どこでも、どの項目でも、信頼できる臨床検査データが得られるか否かについて推定する目的で、大規模健診機関の年次別精度管理成績を検討した。また、職域健康診断のあり方と課題を産業医等の面接調査により検討した。

B. 研究方法

全国労働衛生団体連合会（全衛連）が実施している臨床検査精度管理調査に参加した健診機関の精度管理成績及び 2003 年次同連合会基礎調査資料より 310 機関を選び、その中で健診実施人数が年間延べ 20 万人以上の 66 機関の 9 年間（1996～2004 年度）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 年度延検診実施数、職員数、常勤医師数、非常勤医師数、検査技師数を集計した。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、項目別精度管理調査成績が低い機関数を算出した。対象年度は 2002 年度（第 11 回調査）～2004 年度（第 13 回調査）で、検査項目別評価点、総合評価点、70 点未満検査項目、60 点未満検査項目を算出した。

健康診断と健康管理のあり方については、大企業の専属産業医 1 名、大企業基盤型総合労働衛生機関の産業医 1 名、中小企業の嘱託産業医 1 名と分担研究代表者が個別に面会して討論した。討論は、健康診断の目的・意義・役割、法規則の課題、企業責任と産業医責任、健診の検査方法・項目・実施間隔等の裁量性、企業活動の変容と国際化による課題、産業保健スタッフ間の課題、等の論点にそって行った。臨床検査精度管理のあり方については、7 名の関係者が会合して討議した。

C. 研究結果

a. 大規模健診機関の年次別精度管理成績の検討

(1) 年間健診実施数 20 万件以上の機関の特性

2003 年度実績で各種健診（健康診断、検診）の延べ実施件数が年間 20 万件以上の機関は、66 機関であった。その都道府県別機関数を<表 1>に示した。機関数が多かったのは、東京都 6 機関、神奈川県 5 機関、愛知県 4 機関、兵庫県 4 機関、福岡県 4 機関で、2～3 機関が 12 道府県、1