

平成 15-17 年度 厚生労働科学研究費補助金

健康科学総合研究事業

国民健康・栄養調査における
各種指標の設定及び精度の向上に関する研究

総合研究報告書

2006 年3月31日

主任研究者 吉池 信男

(独立行政法人 国立健康・栄養研究所)

目次

総括研究報告書	国民健康・栄養調査における各種指標の設定及び精度の向上に関する研究 吉池信男	4~12
分担研究報告書	「標本抽出方法及びデータ解析手法の検討」	横山 徹爾
	13~38
	「運動・身体活動の評価方法の検討」	下光 輝一
	39~72
	「糖尿病及び肥満関連指標の検討」	田嶋 尚子
	73~86
	「国民、県民健康・栄養調査等での血液精度管理手法の検討」	中村 雅一
	87~121
	「循環器疾患関連指標及び精度管理方法の検討」	
	「飲酒・喫煙等の指標についての検討」	吉池 信男・析久保 修
	122~160
	「歯科領域における指標と評価手法の標準化」	安井 利一
	161~166
	「保健行動、生活習慣関連指標の検討」	坪野 吉孝
	167~183
	「ストレス等生活習慣関連指標の検討」	内山 真
	184~203

厚生労働科学研究費補助金（健康科学総合研究事業）

総括研究報告書（平成 15～17 年度）

国民健康・栄養調査における各種指標の設定及び精度の向上に関する研究

主任研究者 吉池 信男 独立行政法人国立健康・栄養研究所

研究要旨

目的

「健康日本 21」を推進するための健康・栄養モニタリングの基盤を強化させる。

方法

○現行の国民健康・栄養調査において、一定の誤差率で推計するために必要な標本サイズを検討した。○運動・身体活動評価方法について、国民健康・栄養調査方式、JALSPAQ 質問紙、10 日間の歩数計装着と装着状況記録用紙による比較を行った。○従来法の欠点を克服するために開発した二重カフ法を用いた半自動血圧計について、聴診法と比較しその精度と実用性を検討した。○平成 14 年度糖尿病実態調査データを用いて、糖尿病と診断されていない HbA1c 6.1%以下・食後 6 時間以上で採血した者を対象に、メタボリックシンドローム診断基準（腹囲の代替として BMI、空腹時血糖の代替として HbA1c を使用）を満たす集団を抽出する HbA1c 値を検討した。○糖尿病関連検査の標準化に関する委員会（日本糖尿病学会；2002 年）及び臨床検査精度管理調査（日本医師会；2004 年度）への参加施設の HbA1c 値の精度を検討した。○血液検査受託分析機関の過去の測定値データを解析した。○質問調査による口腔内認知度と実際の所見の一致度を評価するため、一般住民を対象に調査を行った。

結果・考察

○メタボリックシンドロームの有病率=15%と仮定すると、誤差率 10%で推計するためには 70 単位区必要であると考えられた。○国民健康・栄養調査方式の問診により「運動習慣あり」と判定された者では、JALSPAQ 質問紙による評価において、有意に運動頻度、時間が多く、運動によるエネルギー消費量が大きかった。両方法の相関係数は、運動頻度で $r=0.59$ 、運動時間で $r=0.83$ であった。○二重カフ法による血圧計を用い 87 例について精度検定を行ったところ、誤差 $\pm 5\text{mmHg}$ 以内に 88%あり、臨床や疫学に応用可能であると考えられた。○メタボリックシンドロームを判別する至適 HbA1c カットオフ値は 5.15%であった。○HbA1c 測定の精度管理状態に関して、共通 C.V. は 1.56-1.44%と極めて高いレベルであった。○一般の対象者では、歯の補綴の種類について高度の識別認識を有するが、歯数についてはほとんどが認識できなかった。

結論

国民健康・栄養調査の企画や解析方法の検討を行う際の技術的な課題を検討する上で必要な情報や実証データを提供した。この成果は、国レベルでの「健康日本 21」の評価に資するのみならず、「健康日本 21」地方計画の推進及び評価や、個別の研究として行われる疫学的調査等の基盤を強化させる上でも有用であると思われる。

【研究組織】

分担研究者

下光輝一(東京医科大学衛生学公衆衛生学)

田嶋尚子(東京慈恵会医科大学内科学)

内山 真(国立精神神経センター精神保健研究所)

横山徹爾(国立保健医療科学院技術評価部)

朽久保修(横浜市立大学大学院医学研究科

システム情報予防医学:平成 16, 17 年度)

中村雅一(大阪府立健康科学センター脂質基準分析室:平成 16, 17 年度)

安井利一(明海大学歯学部口腔衛生学講座:平成 16, 17 年度)

坪野吉孝(東北大学大学院医学系研究科社会医学講座:平成 15 年度)

A. 研究目的

厚生労働省は、平成 14 年度に「健康日本 21」の評価手法検討会を行い、健康増進法の下に新たな枠組みで行われる国民健康・栄養調査については、この検討会の下部に調査分科会が設けられ、平成 15 年 3 月に報告書が出された。それによれば、国民健康・栄養調査は、従来の国民栄養調査に加えて、運動、喫煙、飲酒等の生活習慣や生活習慣病に関する事項に拡充されることとなっており、特に新たな項目については、指標としての妥当性、適切な調査手法のあり方や精度管理の方法について十分な検討が必要であるとされている。

このようなことを踏まえ、「健康日本 21」を推進するための健康・栄養モニタリングの基盤を強化させることを目的として、本研究を開始した。

B. 研究方法

研究班の構成メンバーの内 4 名は、「健康日本 21 評価手法検討会調査分科会」委員であり、この委員会で検討された事項を踏まえ、それぞれの専門領域を中心に以下の検討を分担して行った。

1) 調査対象客数及び標本抽出方法に関する検討

第 3 回米国健康栄養調査(NHANES III)では、非常に複雑な標本抽出が行われ、その集計のための最適なデータ解析手法が幅広く検討されてきている。そこで、NHANES III の標本抽出方法とデータ解析手法

の詳細を整理し、日本の国民健康・栄養調査への応用の可能性を検討した。HNAHES III に関する情報は、主に米国の National Center for Health Statistics のインターネットリソースおよび出版物から得た。さらに、現行の国民・健康栄養調査における推計値の誤差率について確認し、都道府県別や性・年齢階級別に一定の誤差率で推計するために必要な標本サイズについても検討を行った。都道府県の推計に必要な標本数の推定のためには、平成 14 年国民栄養調査データに基づく Monte Carlo simulation により以下の手順で推定した。(1)平成 14 年国民栄養調査データ 300 単位区の分布を、ある都道府県内での近似的な母分布とみなす。(2)これら 300 単位区から $K(=2\sim 100)$ 個の単位区を無作為に復元抽出して誤差率を計算するという作業を 1000 回ずつ行い、 $K=2\sim 100$ のそれぞれの場合ごとに誤差率の分布を作成する。(3)誤差率の分布の中央値が目標の誤差率となる単位区数を“5 割の県で所定の誤差率を達成するために必要な単位区数”とし、同様に分布の 80% 点に対応する単位区数を“8 割の県で所定の誤差率を達成するために必要な単位区数”とする。また、都道府県別の平均値の推定のために、経験ベイズ推定を応用することも試行した。

2) 身体活動・運動の評価方法の検討:

国民健康栄養調査における運動・身体活動評価方法について、従来の調査方法との整合性を保ちつつ、運動頻度、時間、強度について評価する方法を検討し、本研究班の案として提案した。平成 15 年度調査よりこの案が採用された。そして、この評価方法について以下の方法で妥当性を検討した。

対象は、東京近郊、佐賀県、鹿児島県在住の成人 219 名(男性 61 名、女性 158 名)である。調査は横断調査とし、対象者に以下の調査を依頼した。①国民健康栄養調査で行なわれている運動習慣に関する聞き取り調査と同様の方法で実施する調査(以下、NHNS 方式問診)、②公益信託日本動脈硬化予防研究基金身体活動質問紙(JALSPAQ 質問紙)、③10 日間の歩数計装着および装着状況記録用紙の記入(対象者のうち 181 名のみ)。

NHNS 方式問診は「健康上の理由で運動できない」「それ以外の理由で運動できない」「運動習慣あり」の3つの選択肢から一つを選ぶ形式となっているが、本研究では前2者を合わせて「運動習慣なし」と判定し、JALSPAQ 質問紙との比較検討を行った。歩数調査に関する検討は、歩数計の装着時間の分布を検討し、歩数計装着状況記録用紙により、装着時間の短い不適切なデータを除外することが可能かどうかを検討した。

3) 飲酒・喫煙、ストレス・休養等の生活習慣、保健行動の評価方法の検討:

3)-① 飲酒・喫煙指標に関する検討

平成14年まで施行されていた国民栄養調査においては、飲酒・喫煙習慣に関して「身体状況調査」の一部として調査員による聞き取り調査(旧調査法)が行われていた。平成15年から国民健康・栄養調査として調査方法が改められ、喫煙・飲酒習慣に関しては、質問紙(「生活習慣調査票」)による自記式調査(新調査法)となった。このような2つの方法による飲酒・喫煙習慣に関する評価指標に関して相互の比較性を検討した。具体的には、平成16年11月に実施された三重県民健康・栄養調査を受けた者のうち20歳以上の男女523名を対象として、身体状況調査の中で平成14年まで施行された国民栄養調査と同一内容及び手法による面接聞き取り調査(=「旧調査法」)を追加した。なお、旧・新両法によるデータ収集は別の調査員が独立に行った。

3)-② ストレス・休養指標に関する検討

平成9年2~3月に行われた健康づくりのための意識調査データを用いた。これは、住民基本台帳より層化無作為抽出した日本国内に居住する満20歳以上の男女4000名を対象とし、調査員による個別面接調査を行ったものである(有効回収数3030名、回収率は75.8%)。初年度は、心身の訴えの頻度および心身の訴えと生活習慣との関係を明らかにした。2年度目は、休養に関する指標として自覚的睡眠不足、睡眠時間の短縮、おそらくこれらの結果として生じる日中の眠気の3項目について、心理的・身体的愁訴との関連から検討した。3年度目は、これらの結果を総

合して、国民健康栄養・調査におけるストレス、こころの健康、休養に関する質問票を作成するとともに、その妥当性について検討した。

4) 循環器疾患、糖尿病等の疾病調査における評価方法の検討

4)-① 血圧・腹囲測定手法の標準化

血圧測定及び新たに導入された腹囲測定に関して、施設間及び経年データの標準化を目的として、測定者のトレーニングのための視覚的教材(DVD)を作成した。策定に際しては、ワーキンググループを組織し、各手技を視覚的に説明する教材の設計を行った。それに基づき、測定手順に関する実演及び器具などの画像収録、コトコフ音と水銀柱の動きの収録を行った。後者については、約30名のボランティアの協力を得て、聴診器に小型マイクロフォンを装着し録音を行った。収縮期及び拡張期血圧の判定については、5名の専門家(循環器、疫学)が独立して聞き取り・判定を行い、最もよく一致した値を「正解」とした。試作したDVDを国民健康・栄養調査を実施する計127自治体(政令市、中核市、特別区)に送付し、試行及び普及を図った。

4)-② 新たな血圧測定手法の検討

これまで水銀血圧計で行われていた血圧測定について、適切な自動血圧計を導入できないかどうかを検討するために、以下の検討を行った。従来のオシロメトリック法の欠点を克服するために、カフ中心に小カフを置いたダブル(二重)カフ法を用いた半自動血圧計(TERUMO ES-H55)をすでに開発している。この新しい機器について、実際の検診や臨床現場において、聴診法と比較しその精度と実用性を検討した。さらに、24時間血圧測定で白衣高血圧と判定された45例、正常血圧者84例、固定性高血圧者164例を対象として、自動血圧計に心電図計を組み合わせ、血圧測定時のRR間隔の変動パーワースペクトログラムより低周波成分(LF)と高周波成分(HF)の比を求め、心拍数、基底心拍数から基準化した基礎血圧を求める方法を検討した。

4)-③ 糖尿病関連指標の活用と標準化

日本糖尿病学会糖尿病関連検査の標準化に関す

る委員会(JDS委員会2002年)による全国775施設からのデータ、及び日本医師会による平成16年度臨床検査精度管理調査に参加した2,917施設から集計されたHbA1c値の精度について検討した。

地域の糖代謝異常に関連する健康指標としてのHbA1cの有用性を検証するために、舟形町住民(約5,600人)を対象として、OGTT2時間値200mg/dlに対応する至適HbA1c値をROC曲線で検討し、年齢階級別・集団別のHbA1cの平均値と糖尿病およびIGTの有病率の相関係数を用いて、OGTTによる有病率が分っている山形県、福岡県、長崎県のデータとの比較を行った。以上からHbA1cによって有病率を推定することの妥当性を検討した。

メタボリックシンドロームの指標としてのHbA1cの検討をするために、平成14年度糖尿病実態調査データを用いて、糖尿病と診断されていない、HbA1c6.1%以下、食後6時間以上で採血の3項目を満たす1796名を対象に、メタボリックシンドローム診断基準(但し腹囲の代替としてBMI25以上、空腹時血糖110mg/dl以上の代替としてHbA1c5.6%以上を使用)を満たす集団を抽出するHbA1c値の推定を試みた。

4)―④ 血液生化学検査の標準化

国民健康・栄養調査等の血液生化学検査は、例年、㈱エスアールエル(SRL)が受託分析してきた。そこで、1999～2005年の7年間のSRLにおける測定値を分析対象とした。また、対象項目は、総コレステロール、HDLコレステロール、中性脂肪、尿素窒素、尿酸、クレアチニン、AST(GOT)、ALT(GPT)、 γ -GT(γ -GTP)、ブドウ糖の計10項目とした。外部精度管理法として、日本医師会により全国規模で実施される臨床検査精度管理調査と、米国の政府機関(CDC)と国際的な脂質標準化ネットワーク(CRMLN)が共同実施する脂質標準化プログラムの二つをモデルとして選択し、それを土台にして経年的なモニタリングシステムを構築し、その妥当性を多面的に検討することとした。

正確度は反復切断補正法による参加施設全体の平均値とSRLの測定値の両者から複数試料の平均%バイアスとして計算することを試みた。精密度は、

国民健康・栄養調査の検体がSRLで測定された時期を含む内部精度管理血清の成績を基に、1日当たり1個の測定値($n=1$)を20日間無作為に抽出して、変動係数を求めた。総合誤差は、脂質標準化で適用される計算式(総合誤差=%バイアスの絶対値+1.96x20日間の変動係数)を用いて単年度毎に計算した。

5) 歯科領域における指標及び調査手法の検討

質問調査による口腔内認知度と実際の所見の一致度を評価することを主な目的として、502名(男性197名、女性305名;平均年齢は56.0歳)を対象に調査を実施した。自記式調査票に関しては、口腔内の歯の状況、歯肉の状況、補綴の状況、咀嚼の状況等につき、口腔検診実施前に郵送法により送付を行い、検診当日に持参してもらう方法で調査を実施した。さらに、次年度には、同様の自記式調査票に加えて手鏡や歯のチャートを使用した調査を、480名(男性184名、女性296名)に対して実施した。

(倫理面への配慮) フィールド調査を行う場合には、研究者と市町村等の保健行政担当者との間で、調査の実施、データの取り扱い、被対象者に対する結果返し等に関して十分な協議を行うとともに、疫学研究に関する倫理指針(平成14年文部科学省・厚生労働省告示第2号)に則り、研究機関における倫理委員会の承認を得た。また、国民栄養調査及び糖尿病実態調査データの使用に際しては、厚生労働省健康局総務課生活習慣病対策室に目的外使用の申請を行い、許可を得た後に、データファイルの取り扱いに十分な注意を払いながら、解析を行った。

C. 結果および考察

1) 調査対象客数及び標本抽出方法に関する検討

NHANES IIIの標本抽出方法とデータ解析手法の詳細を整理したところ、日本の国民健康・栄養調査においても推定精度を上げるために、(1)抽出確率に応じた重み付けの計算、(2)応答バイアスの調整、(3)標本の性年齢等の構成を母集団の構成に調整するという手順で集計を行うことが望まれると考えられた。現行の日本の国民・健康栄養調査では、全国レ

ベルでは男女ともに標準誤差率 1%以下の栄養素が多い。性・年齢階級別ではやや誤差率が大きくなるが、主栄養素では標準誤差率 2%以下となっている性年齢階級が多い。しかし、都道府県別に推計を行うには、ほとんどの県で現行よりもかなり多い単位区の調査が必要である。例えば、脂肪エネルギー比の平均値を 80%の確率で誤差率 5%で推定するために必要な単位区の数 は 8 である。割合の推定ではさらに多くの単位区が必要であり、例えばメタボリックシンドロームの有病率=15%と仮定すると、それを 80%の確率で誤差率 10%(標準誤差=1.5%)で推計するためには 70 単位区必要であると考えられた。標本数の小さな県では偶然変動によって栄養素等の標本平均が極端に高め、あるいは低めになることが多く、経験ベイズ推定ではこれが全体の平均値により近い値となった。

表 1

都道府県別に、80%の確率で所定の誤差率を達成するために必要な単位区数

変数	仮定した保有率	性別	必要な単位区数	
			誤差率 10%	誤差率 5%
脂肪エネルギー比率	連続変数	男女計	<5	8
		男	<5	10
		女	<5	8
野菜摂取量	"	男女	7	20
		男	7	25
		女	8	30
日常生活における歩数	"	男女	<5	16
		男	8	25
		女	<5	18
運動習慣のある者(成人)	30%	男女	35	>100
	32%	男	70	>100
	29%	女	45	>100
喫煙率	25%	男女	40	>100
	44%	男	35	>100
	11%	女	>100	>100
睡眠による休養が不足している者	26%	男女	30	90
	27%	男	40	>100
	25%	女	40	>100
大量飲酒者	-	男女	-	-
	8%	男	>100	>100
	-	女	-	-
肥満者の率(成人の内臓脂肪型肥満)	28%	男女	30	95
	31%	男	35	>100
	25%	女	65	>100
糖尿病有病者・予備群の率	34%	男女	30	>100
	32%	男	50	>100
	34%	女	40	>100
高血圧症有病者・予備群の率	59%	男女	14	45
	70%	男	14	45
	52%	女	25	85
メタボリックシンドローム有病率	15%	男女	70	>100
	17%	男	>100	>100
	13%	女	>100	>100

対象年齢30~75歳

2) 身体活動・運動の評価方法の検討:

NHNS 方式問診により「運動習慣あり」と判定された者は、「運動習慣なし」と判定された者と比較して、JALSPAQ 質問紙による評価において、有意に運動頻度、時間が多く、運動によるエネルギー消費量が大きかった。NHNS 方式問診と JALSPAQ 質問紙で評価した運動習慣の相関係数は運動頻度で $r=0.59$ ($p<0.001$)、運動時間で $r=0.83$ ($p<0.001$)であった。運動強度の評価では NHNS 方式問診による低強度は JALSPAQ 質問紙による 3.9METs に相当した。同様に、中強度は 4.7METs、高強度は 6.6METs に相当した。

歩数計の装着に関する検討では、装着時間が 12 時間以上のデータの割合は 94%、10 時間以上は 96%、8 時間以上は 97%であった。装着時間と装着記録への回答を比較すると、装着時間 4 時間未満までは全員が「装着しなかった」と回答した。装着時間が 4 時間以上になると「装着した」と回答するものが出現し、装着時間が長くなるに従って「装着した」と回答するものの割合が増えた。逆に、「装着しなかった」と回答した者の割合は 8 時間以上 10 時間未満では 31%、10 時間以上 12 時間未満では 8%、12 時間以上 14 時間未満では 3%だった。

3) 飲酒・喫煙・ストレス・休養等の生活習慣、保健行動の評価方法の検討:

3)-① 飲酒・喫煙指標に関する検討

飲酒習慣に関しては、旧調査法では新調査法よりも「飲酒習慣あり」に分類される者が明らかに多かった。旧・新両法を比較したところ、旧調査法で「飲酒習慣あり」に分類された者の約半数(46.6%)が、新調査法において「飲酒習慣なし」と異なる回答をしていることが分かった(κ 係数 0.55)。また飲酒量に関しても両法では飲酒者の飲酒量が一致しないことも明らかになった。喫煙習慣に関しては、両法でほぼ同じ割合となり、さらにクロス表により比較したところ、高い一致度を示した(κ 係数 0.83)。喫煙者の喫煙本数も両法でほぼ同様の割合となった。以上のことから、飲酒に関しては、飲酒習慣及び飲酒量は一致度が低く、両法の切り替えを挟んで経年的に比較するには問題があると考え

られた。喫煙に関しては、経年的比較性は良好であると考えられた。

3)-② ストレス・休養指標に関する検討

身体的訴えに関する質問のうち、4項目以上持つ人を身体的不調者とし、心理的訴えに関する項目のうち、2項目以上持つ人を心理的不調者とした。身体的不調者は男性では9.2%、女性では14.5%、心理的不調者は男性では13.4%、女性では17.4%と有意な性差を認めた。身体的不調者と心理的不調者について、生活習慣に関する項目を独立変数とし、多変量ロジスティック解析を行った。心理的不調者と有意に関連のあった生活習慣に関する要因は、男女共にストレス、生活満足感の欠如、男性のみでストレス対処困難、女性のみで不規則な食事であった。身体的不調者と有意に関連のあった生活習慣に関する項目は、男女共にストレス、睡眠による休養の不十分さ、男性のみで生活満足感の欠如、女性のみで運動不足であった。身体的不調及び心理的不調ともにストレスの関連が最も大きかった。

自覚的睡眠不足、睡眠時間の短縮、日中の眠気に共通した身体的の訴えでは、“疲れやすい”は日中の過剰な眠気、睡眠時間不足、主観的睡眠不足において共通に見られる訴えであった。心理的な訴えでは、“イライラする”と“気持ちのゆとりがない”は日中の過剰な眠気、主観的睡眠不足、睡眠時間不足において共通に見られる訴えであった。これら共通して

いた心身の訴えは背景にある日中の眠気、睡眠不足の共通要因となりうるものと考えられた。

こうした解析に基づいて、平成12年度に行われた保健福祉動向調査(心身の健康)におけるストレス、こころの健康、休養に関連した質問項目を参考に国民健康・栄養調査におけるストレス、こころの健康、休養に関する質問票を作成した。さらに近年大きな問題となっている自殺に関する意識調査項目を含め、今後行われることが予想される調査の結果を自殺予防対策に使用しうるようにした。

4) 循環器疾患、糖尿病等の疾病調査における評価方法の検討

4) — ① 血圧・腹囲測定手法の標準化

作成した視覚的教材の構成を図1に示す。試作したDVDを、国民健康・栄養調査を実施する計127自治体(政令市、中核市、特別区)に送付し、その評価を行ったところ、「大変有用」または「有用」との回答は、腹囲の測定手順の説明98%、血圧測定手順83%、血圧(コトコフ音の説明、聞き取り問題)75%であった。しかし、保健所の現場においては、DVDの再生装置が無いために、メディア形態としてはビデオテープを希望する施設も少なくなかった。

また、コトコフ音の聞き取り練習及びテストについては、収録した血管音のS/N比が十分で無いケースもあったとともに、コトコフ第4音～第5音が再生装置の性能(特に低音の再生)によって聞き取り結果の

**国民健康・栄養調査
身体状況調査手技のトレーニング**
(Video instructions for measurement procedures of blood pressure and abdominal circumference for the National Health and Nutrition Survey, Japan)
2004年9月1日(試用版)

厚生労働科学研究健康科学総合研究事業
国民健康・栄養調査における各種指標の設定及び精度の向上に関する研究班
【主任研究者】
吉池 徳男 独立行政法人国立健康・栄養研究所健康・栄養調査研究部
【ワーキンググループ・研究協力者】
齊本 伸雄 浜松医科大学衛生学教室
柏原 徳志 関西学院大学保健学研究所
甲田 道子 中京女子大学健康科学部栄養科学科
松久保 修 横浜市立大学大学院医学研究科情報システム予防医学
横山 徹郎 国立保健医療科学院技術研経部
【作成及び出演協力】
国立健康・栄養研究所健康・栄養調査研究部
藤井 純子 吉田 真起子 斉藤 京子

製作:
NTV映像センター メディア制作部

※血圧測定におけるコトコフ音の聞き取り(聴聴・聴取問題)については、ノートPC内蔵の
小口径スピーカーでは、低音の再生がほとんどできないので、できるだけ大口径のスピー
カーをご利用下さい。また、装置の低音の再生の能力により拡張系血圧の読み取りが
変わる場合がありますので、ご注意ください。

※このDVDは、厚生労働科学研究での評価テスト用に試作されたもので、それ以外の
用途で使用される場合には、事前に主任研究者(nobuyos@nh.go.jp)まで御連絡下さい。

※血圧測定聞き取り練習の自己評価のための自動計算エクセルシートは、下記のURLから
ダウンロードできます。
http://www.nh.go.jp/iken/oss/system_data.html

国民健康・栄養調査における各種指標の設定及び精度の向上に関する研究班

図1 血圧及び腹囲測定の標準化のための視覚的教材の構成内容

異なる場合があった。この点については、今後特に改良を加える必要があると思われる。

4)―②新たな血圧測定手法の検討

ダブル(二重)カフ法を用いた半自動血圧計(TERUMO ES-H55)は、アルゴリズムが明確で、AAMI(SP-10)の基準を満たす血圧計であり、小型で耐久性や保守が簡単で、マンメーターも水銀圧力形より精度が高く、しかも自動判定ができ、客観的な血圧値が得られるので、検診などの現場に用いることができる血圧計と考えられる(表2)。なお、現場の医師の聴診法との比較で、AAMIプロトコールで行った成績よりも差が大きかったのは、医師の聴診能力の差がある可能性もある。むしろ血圧測定は聴診法よりも客観的精度の保証されている自動血圧計で行った方が良いと思われる。

健診などの随意血圧測定時には白衣現象が伴わない正常血圧を高血圧の誤診とされる可能性があるため、血圧測定時に心電図を記録することにより交感神経活性(LF/HF)を測定し、随意血圧を基底血圧に基準化する装置を試作した。本装置は真の基底血圧を高い精度で推定でき、白衣高血圧を88%の精度で判定できた。

4)―③ 糖尿病関連指標の活用と標準化

HbA1c 測定の精度管理状態に関しては、HbA1c 値の共通 C.V.は 1.56%-1.44%と極めて高いレベルであり、わが国では HbA1c 測定の標準化が維持されていた。

HbA1c によるカットオフポイントに関しては、HbA1c 値は OGTT による各判定区分におけるヒストグラムの重なりが大きいいため、その測定は OGTT の代用にはならないが、「糖尿病の可能性を否定できない」集団および「糖尿病が強く疑われる」集団を、糖尿病ではないと考えられる集団から抽出するための指標とすることは妥当性があると考えられた。そのカットオフ値は従来の値よりもやや低い 5.3~5.6%と推定された。

集団における指標の検討については、既知の糖尿病を含めた地域住民の年齢別 HbA1c の平均値は 4.90%~5.35%で、集団の HbA1c 値が 0.1%上昇すると有病率が 2.5%上昇する可能性が示唆された。これ

表2 水銀血圧計と新しい方法の特徴の比較

	水銀圧力計による聴診法	本測定法(二重カフ法)
価格(円)	6000~10000	25000(販売数により廉価となる)
装置の大きさ	6×35×10cm	17×8×5cm
重量	約1000g	120g
耐久性と保守	毎年点検要	5年間補償(単三電池1個で1000回以上測定可)
マンメーターの精度	±2mmHg	±1mmHg 聴診法との差0.2±3.9/ 0.1±4.6mmHg
客観性	劣る(聴診能力による)	優れる(自動判定可)
測定時間	1分以内	1分以内
備考	水銀圧力計の水銀による害有り	本装置の圧力計は水銀圧力計の替わりとしても用いられる

を平成14年度の糖尿病実態調査の都道府県別の平均HbA1cを用いて糖尿病の有病率(40歳以上)推測したところ 7.1~18.9%であったが、地域ブロック別にみると地域差はなかった。

メタボリックシンドロームの判定基準に関して検討を行い、メタボリックシンドロームを判別する際の至適HbA1cカットオフ値は 5.15%であった。本研究で用いたメタボリックシンドローム診断基準の構成因子のうち、HbA1c値によって比較的良好に識別しえたのは随時血糖値(食後6時間以上)とBMIであった。

4)―④ 血液生化学検査の標準化

解析対象とした6年間では、尿素窒素以外の9項目では、毎年の総合誤差は許容限界に収まり、経年変化を追跡できる可能性が示唆された。尿素窒素では、平成14年の総合誤差のみが5.1%と許容限界の4.5%を0.6%超えていた。このことから、平成14年分の成績評価では参考に留めるのが適切と思われる。総コレステロールとHDLコレステロールの2項目については、脂質の標準化を併用して運用したが、二つのシステムで得られた測定精度間には特に大きな矛盾は認められなかった。

以上の検討から、経年的モニタリングシステムとしては、わが国で広く活用されている①日本医師会による臨床検査精度管理調査と世界各国の特に研究領域で利用されている②CDC/CRMLNによる脂質標準化プログラムの二つの外部精度管理法を基に新たに構築することが望ましいと考えられた。ただし、国際的な比較可能性については、今後の更なる検討課題として残ると考えられる。

5) 歯科領域における指標及び調査手法の検討

自己申告のみ(手鏡無し)の自記式質問票による歯数の一致率はきわめて低く、信頼性も低いと推察された。すなわち、歯数の一致率は、総計で 43.0%という結果であった。一方、同様の質問紙に加え、手鏡を持たせて確認をさせることにより、一致率は向上し 49.7%となった。特に、男性では手鏡なしでの一致率が 25.8%であるのに対し、手鏡使用により 48.5%と著しく増加した。歯の状況から、一致率の内容について分析をすると、欠損のない者や義歯の場合には手鏡なしでも 50%を超える一致率を得たが、自分の口腔内に何本の歯が存在するかは国民にとって相当難しいことであることが推察された。また、ブリッジが装着されると、どれが自分の歯か判別できなる傾向が明確に示されていた。

歯周病の状況についての一致性に関しては、健康であると認知した者のうち、検診において臨床的にも健康であると診断された者は 66.4%であった。一方、健康ではないと認知した者が健康でない一致率は 42.7%であった。

以上の調査研究の結果から、自記式調査による口腔内認知度と実際の所見の一致度については、一般の対象者では、歯の補綴の種類については高度の識別認識を有するが、歯数についてはほとんど認識できないことが判明した。歯周病の質問票についても、特異度が高いものの組み合わせによる質問が必要であることを示唆した。これらのことから、単に、質問紙によって口腔内の情報の的確性を向上させることは難しいと考えられ、手鏡など補助的な手段の応用、あるいは質問の組み合わせ等を検討していくことで、信頼性の向上が図ることができると考えられた。

D. 結論

本研究班では、国民健康・栄養調査の企画や解析方法の検討を行う際の技術的な課題を検討する上で必要な情報や実証データを提供している。具体的には、運動習慣や身体活動量、飲酒・喫煙、ストレス・休養等の生活習慣を質問紙調査等により適切に把握するための指標、循環器疾患・糖尿

病等の疾病状況を把握するために必要な血圧や血液指標等に関する精度管理方法、対象者の標本抽出方法やデータの解析方法等である。また、これらの研究成果については、国レベルでの「健康日本 21」の評価に資するのみならず、「健康日本 21」地方計画の推進及び評価や、個別の研究として行われる疫学的調査等の基盤を強化させる上でも有用であると考えられる。

E. 健康危険情報

この研究において健康危険情報に該当するものはなかった。

F. 研究発表

1. 論文発表

- 1) Zaman MM, Yoshiike N: Video-instruction for blood pressure measurement. *Regional Health Forum of WHO South-East Asia Region* ; 7(1); 63-65, 2003
- 2) Nakamura M, Yoshiike N: Current systems of national and regional nutrition surveys and future direction. *J Community Nutrition* 5(2); 59-64 , 2003
- 3) Yoshiike N, Lwin H: Epidemiological aspects of obesity and NASH/NAFLD in Japan. *Hepatol Res*, 2005
- 4) 吉池信男:わが国における糖尿病患者数と管理状況. *Modern Physician* ; 25(3); 245-250, 2005
- 5) 吉池信男, 藤井紘子, 渋谷克彦:糖尿病実態調査 *糖尿病学の進歩*—基礎から臨床まで 178-183, 2005, 金原出版
- 6) Tominaga M. Japanese standard reference material for JDS Lot 2 haemoglobin A1c. I: Comparison of Japan Diabetes Society-assigned values to those obtained by the Japanese and USA domestic standardization programmes and by the International Federation of Clinical Chemistry reference laboratories. *Ann Clin Biochem* ;42:41-46, 2005.
- 7) Nishimura R, Kanda A, Sano H, Matsudaira T, Miyashita Y, Morimoto A, Shirasawa T, Kawaguchi T, Tajima N. Glycated albumin is low in obese, non-diabetic children. *Diabetes Res Clin Pract.* 2005 (e-pub)
- 8) 田嶋尚子. 糖尿病ハイリスク群の効率的なスクリーニング. *Progress in Medicine*; 25:15-21, 2005
- 9) 内山真:日本人の睡眠の特徴-国際睡眠疫学調査の結果を踏まえて. *医学のあゆみ* 205:529-532, 2003.
- 10) Kaneita Y, Ohida T, Uchiyama M, Takemura S, Kawahara K, Yokoyama E, Miyake T, Harano S,

Suzuki K, Yagi Y, Kaneko A, Tsutsui T, Akashiba T: Excessive daytime sleepiness among Japanese General population. *J Epidemiol* 15: 1-8, 2005.

- 11) Kaneita Y, Ohida T, Uchiyama M, Takemura S, Kawahara K, Yokoyama E, Miyake T, Harano S, Suzuki K, Fujita T: The relationship between depression and sleep disturbances: a nationwide general population survey. *J Clin Psychiatry* (in press 2005).
- 12) 井上茂, 下光輝一: 身体活動・運動調査とその評価、エキスパートから学ぶ「健康教育・栄養相談・生活習慣改善指導」—生活習慣病の予防と管理—, ライフサイエンスセンター, 2003; 203-207

2. 学会発表

- 1) 横山徹爾, 由田克士, 吉池信男. 国民健康・栄養調査における都道府県別推計に必要なサンプルサイズ. 第 15 回日本疫学会学術総会. 2006.1. 名古屋
- 2) 吉池信男: 健康日本21の評価における国民健康・栄養調査の役割. 第 63 回日本公衆衛生学会総会, 2004.10.27, 松江
- 3) 斉藤京子, 横山徹爾, 野末みほ, 草間かおる, 印南京子, 吉池信男: 国民健康・栄養調査における飲酒・喫煙習慣の質問票の比較検討～三重県県民調査より 第 64 回日本公衆衛生学会, 2005.09.16, 札幌
- 4) 井上茂, 下光輝一, 吉武裕, 原田亜紀子, 小田切優子, 大谷由美子, 石井香織: 国民健康栄養調査方式の運動習慣評価の妥当性. 第 60 回日本体力医学会大会. 2005.9.23. 倉敷市
- 5) 中村雅一, 佐藤眞一, 嶋本喬: 国民健康・栄養調査における各種指標の設定及び精度の向上に関する研究—国民、県民健康・栄養調査等での血液精度管理手法の検討—. 第 64 回日本公衆衛生学会総会. 2005.9.16, 札幌市
- 6) 安井利一, 宮武光吉: 歯科領域における指標と質問紙評価手法の標準化. 第 64 回日本公衆衛生学会総会. 2005.9.15, 札幌市
- 7) Uchiyama M: Are there cultural differences in the alertness concept? 2nd International Sleep Disorders Forum. Paris, France, 2004.9.
- 8) 李嵐, 譚新, 尾崎章子, 渋谷佳代, 鈴木博之, 有竹清夏, 栗山健一, 田ヶ谷浩邦, 内山真: 日本の一般人口における自覚的な休養不足感と睡眠の問題. 日本睡眠学会第 30 回学術集会. 宇都宮. 2005.6.

G. 知的財産権の出願・登録状況

この研究において、知的財産権に該当するものはなかった。

厚生労働科学研究費補助金（健康科学総合研究事業）
分担研究報告書（3年分の統括）

標本抽出方法及びデータ解析手法の検討

分担研究者 横山徹爾 国立保健医療科学院技術評価部 主任研究官

研究要旨

国民健康・栄養調査に関する以下の検討を行った。【1】標本抽出方法と応答バイアスを考慮しつつ、調査日における日本人1人あたりの食物摂取量等の平均的な値をより高い精度で偏りなく推計するために適切なデータ集計手法を、第3回米国健康栄養調査(NHANES III)における標本抽出・集計方法を参考にして検討した。【2】層化クラスター抽出による調査での点推定、区間推定の方法を整理し、平成14年国民栄養調査の集計に適用を試みた。また、層（都道府県）毎に一定の誤差率を達成するために必要な単位区数を、Monte-Carloシミュレーションにより明らかにした。都道府県別等の層別集計の場合に、人口規模に応じた標本数の違いを調整するために、経験ベイズ推定の適用を試みた。【3】国民健康・栄養調査の1日調査による栄養素摂取量のデータと、別途行われた複数日の繰り返し調査のデータを用いて、栄養素摂取量の習慣的摂取量の分布を推定する方法を検討し、平成14年国民栄養調査のデータでの分析を試みた

A. 研究目的

日本の国民健康・栄養調査は、全国の国勢調査区から同年の国民生活基礎調査のために層化クラスター抽出した調査区を分割して作成した‘単位区’から他の調査が行われた地区を除き、さらにそこから無作為抽出した300単位区内の全世帯および全世帯員を調査客体としている。集計の際には、基本的に性年齢階級等および総数の単純平均等を算出して公表している。^{1,2)}

本研究では、以下の3テーマについて検討した。

【1】国民健康・栄養調査のこのようにやや複雑な抽出標本において、また全世帯から協力が得られるとは限らない場合に、「調査日における日本人1人あたりの食物摂取量等の平均的な値をより偏りなく推計する」という目的のために適切なデータ解析手法と標本抽出方法を、第3回米国健康栄養調査(NHANES III)を参考に、検討した。

【2】現行の国民・健康栄養調査における推定

値の誤差率を確認し、都道府県別に一定の誤差率で推計するために必要な標本サイズについて検討を行った。また、経験ベイズ推定の適用を試みた。

【3】日本人の食事摂取基準（2005年版）を用いて集団の評価を行うために必要な、集団における習慣的な摂取量の分布型を推定する方法を検討した。

【1】データ解析方法と標本抽出方法の整理

B1. 方法

NHANES IIIでは、非常に複雑な標本抽出が行われ、その集計のための最適なデータ解析手法が幅広く検討されてきている。そこで、NHANES IIIの標本抽出方法とデータ解析手法の詳細を整理し、日本の国民健康・栄養調査への応用の可能性を検討した。NHANES IIIに関する情報は、主に米国のNational Center for Health Statisticsのインターネットリソース³⁾および出版物から得た。⁴⁻⁶⁾

C1. 結果

NHANES III の標本抽出方法

図1-1はNHANES IIIの標本抽出方法の概要を整理したものである。全体で4段階抽出となっており、対象は、生後2ヶ月以上の施設に入所していない米国人である。60歳以上の高齢者、生後2ヶ月～5歳の小児、黒人、Mexican-Americanは、それぞれのグループ内で十分な精度での推定が可能ないように、過剰抽出が行われる。

1. 第1段階

郡(county)または小さい郡を併合したものを第1次抽出単位(Primary Sampling Units: PSUs)と呼ぶ。まず、全米で2,812個のPSUsを定義し、それらを47に層化する。47層のうち13層は個々が一つずつの大きな郡そのものであり、必ず調査対象とする。これら13層は、一部を論理的および実務的理由により分割して計21の調査地区(stands)とする。残りの34層からは、人口サイズや人種構成等から定めた重みに比例した確率(確率比例抽出法)で1層あたり2つのPSUを抽出する。従って、全体で $21+2\times 34=89$ 個の調査地区が選ばれる。

NHANES IIIはPhase I(1988-91年)とII(1991-94年)の2期に分けて行われる。89個の調査地区を無作為に2分割し、44調査地区をPhase I、45調査地区をPhase IIとする。この作業により、Phase IとIIは独立なサンプルとみなすことができるので、各Phaseおよび全体として偏りのない標本として集計可能である。

2. 第2段階

国勢調査区またはその集合体または国勢統計区をarea segmentsと呼ぶ抽出単位とする。Phase Iは1980年の国勢調査、Phase IIは1990年の国勢調査に基づいている。各調査地区を人種構成等により層化し、1調査地区あたり約24のarea segmentsを層化無作為抽出する。従っ

て全米で2,138のarea segmentsが抽出される(約 24×89)。

3. 第3段階

第2段階で選ばれたarea segmentsの全世帯リストを作成する。各層でみた場合に89個の調査地区を通して抽出確率がほぼ等しくなるように抽出率を決め、世帯単位で抽出を行う。ここまでで全米で93,653世帯が抽出される。

4. 第4段階

性年齢人種により層化して、世帯員を層化無作為抽出する。全米で39,695人(19,528世帯)が抽出される。抽出された世帯員のことをSampled Persons(SPs)と呼ぶ。

各段階での抽出数は表1-1の通りである。

表1-1. NHANES IIIにおける標本抽出の各段階における抽出単位の数

第1段階	
全米の第1次抽出単位(PSUs)の数	2,812
層(strata)の数	47
抽出されたPSUsの数	81
抽出された調査地区(stands)の数	89
第2段階	
抽出されたarea segmentsの数	2,144
第3段階	
抽出された世帯の数	93,653
第4段階	
抽出された世帯員(SPs)の人数	39,695
SPsがいる世帯の数	19,528
実施結果	
インタビューを受けたSPsの人数	33,994
移動検査場(MEC)で検査を受けた人数	30,818
自宅で簡易検査を受けた人数	493

NHANES III の集計方法

NHANES IIIや日本の国民健康・栄養調査のような標本調査の究極的な目的は、「対象集団全体の調査(悉皆調査)が行われたならば得られるであろう統計量を推定すること」と言ってよいだろう。この目的に近づくために、NHANES IIIでは標本のデータは単純に平均値等を計算するのではなく、以下のような重み付けの工夫が行われる。

- ① 抽出確率が異なる年齢・性別・人種グループの重みを調整する。
- ② 回答者 (respondents) と非回答者 (non-respondents) が異なる特徴を持つことによる偏り (応答バイアス) を減らす。
- ③ 標本データを母集団 (調査対象集団全体) の次元に合わせる。
- ④ 標本抽出フレーム (sampling frame) の不完全さ (世帯リスト作成段階で一部が脱落したり住所が定まっていない者が含まれないなど) を可能な限り補正する。
- ⑤ 推定手順において分散を減らすために、高い精度で分かっている補助的な情報を用いる。

重み付けは上記に対応して以下のように3段階で行われる。

1. 抽出確率に応じた重み付けの計算 (①に対応)

抽出された世帯員 SP の一人一人について、抽出される確率の逆数を基本の重みとする。抽出される確率は、その SP の年齢・性別・人種 (高齢者、小児、Mexican-American、黒人は高確率で抽出)、密度層 (Mexican-American の割合が高い area segments は高確率で抽出)、属する PSU (1つの PSU あたりの標本数はほぼ同じになるようにするため属する PSU によって抽出確率は異なる) によって決まる。

この重み付けによって、特定の性年齢層や人種等を過剰抽出することに由来する全体の推計値に与える偏りが減る。

2. 応答バイアスの調整 (②に対応)

応答バイアスを減らす最善の手段は高い回答率 (response rate) を得ることである。そのためには、NHANES III では、奨励金の支払いと依頼の繰り返しがとても効果的であったという。しかし、最善の努力によっても調査に非回答はつきものであるから、応答バイアスを減らす (完

全になくすことは不可能) ために統計学的な調整をすることもまた必要である。NHANES III は、自宅でのインタビューと MEC での調査とがあり、非回答も各段階でおこる。

(1) 重み付け回答率の計算

抽出された世帯員 SP の一人一人は上記 1. で述べたように重みが異なるので、応答バイアスの評価・調整のために用いる回答率も重み付けしたものを計算しておく。

(2) 応答バイアスの評価

応答バイアスを量的に推定するためには、外的な確認のためのデータが必要であるが、標本調査では一般にそれが得られないので、最終的な応答バイアスの大きさを量的に厳密に評価することはできない。しかし、その可能性についての評価はある程度可能である。

応答バイアスの可能性を評価するために最もよく用いられるのは、回答者と非回答者で、年齢、人種、世帯人数等の既知の属性を比較することである。最も強く回答率に影響する要因を検出する方法の一つに、CHIAD (Chi-square Automatic Interaction Detection) があり、NHANES III でも採用している。

(3) 応答バイアスの調整法

欠損値は、ユニット (個人) 単位またはアイテム (調査項目) 単位に生ずる。欠損値から生じる問題に対処する方法には、大きく分けて、重み付け調整 (weighting adjustment) と補完 (imputation) がある。

重み付け調整とは、非回答者を全く無視して、回答者の重みを単純に大きくする方法である。人種、年齢、世帯サイズ (インタビューが行われた者では健康状態) の各グループについて別々に重み付け調整が試みられた。

補完は、欠損値を補完値で置き換える方法である。補完値の決め方には多くの方法があるが、NHANES III では3つの補完方法が検討された。(a) Hot-deck は、欠損値を有する者と性・年齢・人種等が一致する (欠損値がない) 者を無作為

に選び、後者の値を前者の欠損値の代わりに用いる方法である。(b) 回帰補完 (regression imputation) は、線形回帰モデルによって補完値を推定する方法である。予測値そのものを補完値にする場合と、予測値に残差を無作為に加えた値を補完値にする場合とがある。(c) 多重補完 (multiple imputations) は、ベイズ的手法によって回答者の平均値を応答バイアスで調整するものである。いずれの方法も、点推定値はほぼ同じであり簡単に優劣はつけがたいようである。また、他にも多くの補完法が考案されており、十分な検討が必要と思われる。

3. 標本の性年齢等の構成を母集団の構成に調整する (③④⑤に対応)

この操作は Poststratification と呼ばれる。Poststratification によって、米国全体および様々な副集団レベルでの推定値のばらつきを減らし、また、応答バイアスを減らすためにも役立つ (ただし (2) の応答バイアスの調整の後に poststratification を行う)。

NHANES III の Poststratification は 2 段階で行われる。第 1 段階は国勢調査区レベル、第 2 段階は性・年齢・人種グループレベルで行われ、いずれも調査中央時点 (Phase I は 1990 年 3 月、Phase II は 1993 年 3 月) の Current Population Survey (CPS) のデータが対照人口として用いられる。また、CPS のデータそのものも、poststratification によって国勢調査人口 (住所が定まっていない者等を含む) に調整される。従って、poststratification によって (施設に入所していない) 米国人全体の推定値が得られる。

D1. 考察

NHANES III の標本抽出方法は、米国の状況 (日本と比べて多人種・多民族で、精度の高い住民登録システムがない) を考慮してかなり複雑なものとなっており、その複雑な標本抽出方法を

考慮した集計方法が検討されてきている。これら集計方法のいくつかは、日本の国民健康・栄養調査において、「調査日における日本人全体の平均的な食物摂取量等の推定を行う」という目的においてその精度向上のために応用可能であると考えられる。以下はそのような集計方法の一案である (図 1-2 参照)。

1. 抽出確率に応じた重み付けの計算

日本の国民健康・栄養調査は、同年の国民生活基礎調査のために層化クラスター抽出した調査区を分割して作成した単位区から他の調査が行われた地区を除き、さらにそこから無作為抽出した 300 単位区内の世帯および世帯員を調査客体としている (他の調査を行う地区が無作為抽出されている限り、その残りから無作為抽出された国民健康・栄養調査の標本も無作為抽出とみなしてよいだろう)。そのため、都道府県等の層によって単位区の抽出確率は異なる。NHANES III の最終段階の無作為抽出の単位が個人であるのに対して、日本の国民健康・栄養調査では単位区であるので、世帯あるいは個人単位での抽出確率は単位区の抽出確率に等しい (図 1-2 : n/N)。なお、世帯数は単位区によって異なるので、抽出確率の逆数による重みの合計は必ずしも母集団の性年齢構成に一致しない。母集団の性年齢構成への調整は poststratification によって行う。

2. 応答バイアスの調整

まず、非回答者の特徴について十分に整理する必要がある。国民健康・栄養調査単独で得られる情報としては、世帯の性・年齢構成と人数および調査地区がある。また、国民健康・栄養調査は国民生活基礎調査からのサブサンプルに対して行われるので、国民生活基礎調査とのレコードリンクが可能になれば、健康状態等に由来する応答バイアスの可能性についても検討することが可能になるであろう。Hot-deck、回帰補完、多重補完その他のいずれを適用するかは、非回答者の特徴について整理した後に数

種類の方法を試みて検討する必要がある。また、非回答は原則的に世帯単位で生ずるので、世帯の補完値の推定と個人の補完値の推定の2段階を踏む必要があるかもしれない。

3. 標本の性年齢等の構成を母集団の構成に調整する

NHANES III は国勢調査区と全国レベルの2段階の poststratification を行うが、日本の国民健康・栄養調査は単位区内では全数が抽出されているので、最終抽出段階の層（都道府県市区町村等）の人口に対する1段階のみの poststratification を行えばよいだろう。これによって、層全体の推計値が得られる。

最終的に日本全国あるいはブロック別の推定値を得るためには、各層人口で重み付けをした平均値等を計算する。

4. 推定値の標本誤差について

複雑なサンプリングにおける標本誤差の推定のためのソフトウェアパッケージとして SUDAAN 等が開発されており、NHANES III でもこれを用いている。日本の国民健康・栄養調査は、近似的に層化クラスター抽出とみなすことができるので、SAS の SURVEYMEANS プロシージャ等も利用可能であろう。（テーマ2参照）

【テーマ2】誤差率とサンプルサイズの計算 抽出方法について

B2. 方法

現行の国民健康・栄養調査の標本抽出の概要を図2-1に示す。まず、国民生活基礎調査のために、直近の国勢調査での国勢調査区から、都道府県等を層とした層化クラスター抽出を行い、これを親標本とする。次に、親標本の国勢調査区を分割して約20世帯程度からなる単位区を作り、単位区を抽出単位とした無作為抽出を行い、選ばれた300単位区を国民健康・栄養調査の対象とする。実際には、親標本のうち他の調査のために抽出された単位区は除外して、残った単位区の中から無作為抽出を行っているが、

他の単位区が無作為抽出されている限り、親標本全体から単位区を無作為抽出したとみなしてよい。

また、この抽出方法は厳密には、国勢調査区を第1次抽出単位、単位区を第2次抽出単位とした2段階抽出とみなすべきである⁷⁾。ただし、国勢調査区1個あたりの単位区数のばらつきがそれほど大きくなければ、層全体に単位区を設定してそれを単純無作為に抽出した場合（1段階抽出）でも2段階抽出した場合でも、任意の単位区が選ばれる確率はほぼ等しい。また、全体の抽出率が非常に小さければ1段階抽出の場合に同一の国勢調査区から複数の単位区が抽出される確率は非常に小さいため、2段階抽出の場合でも同一の国勢調査区から複数の単位区が抽出される確率が十分に小さければ（例えばそうならないように操作すれば）、近似的に1段階抽出（単位区を抽出単位とした層化クラスター抽出）とみなして差し支えないであろう。以下の推計は、その前提のもとで行う。

点推定の方法

栄養素等の連続型変数の場合、全国平均の推定値 \hat{Y} は次式で得られる⁸⁾。

$$\hat{Y} = \left(\sum_{h=1}^H \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij} y_{hij} \right) / w_{\dots} \quad \dots \text{(式1)}$$

ただし、

$h = 1, 2, \dots, H$: 層番号（全体で H 層）

$i = 1, 2, \dots, n_h$: 層 h 内での単位区番号

$j = 1, 2, \dots, m_{hi}$: 層 h , 単位区 i 内での個人番号

n : 全体の人数合計

w_{hij} : 層 h , 単位区 i , 個人 j の基本の重み

y_{hij} : 層 h , 単位区 i , 個人 j の栄養素摂取量等

また、

$$w_{\dots} = \sum_{h=1}^H \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij} \quad \dots \text{(式2)}$$

である。

抽出単位は単位区なので、単位区の抽出確率の逆数とその単位区に属する個人の基本の重み w_{hij} となる。層 h において、単位区総数 L_h は、国勢調査区総数を N_h 、国民生活基礎調査の対象となる国勢調査区数を k_h 、国民生活基礎調査の対象となる国勢調査区に設定された単位区総数を M_h とすると、

$$L_h = N_h M_h / k_h \quad \dots (式3)$$

により推計される⁷⁾。従って、

$$w_{hij} = L_h / n_h \quad \dots (式4)$$

とする。この基本の重みを決めるために必要な情報を整理すると、表2-1のようになる。

標準誤差の推定法

全国推定値 \hat{Y} の分散は、Taylor 展開に基づき得られた次式により推定される⁸⁾。

$$\hat{V}(\hat{Y}) = \sum_{h=1}^H \hat{V}_h(\hat{Y}) \quad \dots (式5)$$

ここで、 $n_h > 1$ の場合、

$$\hat{V}_h(\hat{Y}) = \frac{n_h}{n_h - 1} \sum_{i=1}^{n_h} (e_{hi} - \bar{e}_{h..})^2 \quad \dots (式6)$$

$$e_{hi} = \left(\sum_{j=1}^{m_{hi}} w_{hij} (y_{hij} - \hat{Y}) \right) / w_{hi..} \quad \dots (式7)$$

$$\bar{e}_{h..} = \left(\sum_{i=1}^{n_h} e_{hi} \right) / n_h \quad \dots (式8)$$

であり、 $n_h = 1$ の場合、

$$\hat{V}_h(\hat{Y}) = \begin{cases} \text{計算不能 (全ての層で } n_h = 1 \text{ の場合)} \\ 0 \text{ (それ以外の場合)} \end{cases}$$

とする。

経時変化を比較するための年齢調整

栄養素等摂取量の長期的な経年推移や、健康日本21の評価等のために異なる年の調査結果を比較するためには、性・年齢等の構成の違いを補正する必要が生ずる場合もあるだろう。

そのためには、直接法による年齢調整を行うのが一つの方法である。例えば、poststratification (後述) で調査年の国勢調査人口に調整する代わりに、共通の基準年の国勢調査人口に調整すればよい。あるいは、2回の標本調査間での差を比較することのみに興味があるのならば (例えば健康日本21のベースライン値と5年後の比較を国民栄養調査結果によって行う場合など)、初回調査時の標本の性年齢構成を基準として、2回目調査時の標本の性年齢構成を調整するのも一つの方法と思われる。

2回の独立な調査の有意差検定を行うためには、それぞれの調査における推定値を \hat{Y}_1, \hat{Y}_2 、その分散を $\hat{V}(\hat{Y}_1), \hat{V}(\hat{Y}_2)$ とすると、

$$\hat{V}(\hat{Y}_1 - \hat{Y}_2) = \hat{V}(\hat{Y}_1) + \hat{V}(\hat{Y}_2) \quad \dots (式9)$$

であるから、

$$Z = \frac{\hat{Y}_1 - \hat{Y}_2}{\sqrt{\hat{V}(\hat{Y}_1) + \hat{V}(\hat{Y}_2)}} \quad \dots (式10)$$

が、帰無仮説 $H_0: \bar{Y}_1 = \bar{Y}_2$ のもとで、近似的に基準正規分布 $N(0, 1)$ に従うことを利用する。

計算プログラム

統計ソフトウェアパッケージ SAS (ver. 9.1) を用いて以上の計算を行い、全国平均の推定値 \hat{Y} と標準誤差を得るためには、例えば以下のようにする。

```
PROC SURVEYMEANS;
  STRATA KEN /LIST; /* 層番号 h */
  CLUSTER TANIKU; /* 単位区番号 i */
  VAR EIYOSO; /* 栄養素等 y_hij */
  WEIGHT WGT; /* 基本の重み w_hij */
```

RUN;

層別に推計する場合には、層別クラスター抽出では各層で独立に標本抽出を行っていると考えられるので、単に別々に解析すればよい。例えば BY ステートメントを用いて、

```
PROC SURVEYMEANS;
  CLUSTER TANIKU; /* 単位区番号 */
  VAR EIYOSO; /* 栄養素等 */
  WEIGHT WGT; /* 基本の重み */
  BY KEN; /* 層番号 */
```

RUN;

とする。ただし、単位区数 $n_h = 1$ の層では標準誤差を推定することはできない。

また、性年齢階級等のドメイン別の推計を行うためには、各ドメインから独立に抽出しているわけではないのでBYステートメント等を用いて単に別々に解析するのは不相当であり、次のように DOMAIN ステートメントを用いる。(計算式は文献8参照)

```
PROC SURVEYMEANS;
  STRATA KEN /LIST; /* 層番号 */
  CLUSTER TANIKU; /* 単位区番号 */
  VAR EIYOSO; /* 栄養素等 */
  WEIGHT WGT; /* 基本の重み */
  DOMAIN SEX*AGE10; /* 性・年齢階級 */
RUN;
```

経験ベイズ推定

笹瀬らの方法⁹⁾を用いて、都道府県別に経験ベイズ推定を試みた。この方法は、各層内では個人単位での単純無作為抽出を前提としているため、クラスター抽出の場合には厳密には適用することはできないと思われるが、人口規模の違いの影響を検討するために仮に試行した。

C2. 結果

平成14年国民栄養調査の全国推計

(1) 基本推計

基本の重みを計算するための基本情報(表2-1)が得られていないので、仮に全ての個人の基本の重みが等しい($=1$)と仮定して推計を行った結果を表2-2に示す。男女ともに、標準誤差率1%以下の栄養素が多い。また、性年齢階級の標準誤差率を表2-3に示す。主栄養素で

は、標準誤差率2%以下となっている性年齢階級が多い。

(2) 非回答者の補完

調査に回答が得られた世帯と得られなかった世帯とで、性・年齢・その他の背景因子(究極的には食習慣)に違いがある場合、集計結果に偏りが生ずる可能性がある。そのため、可能ならば何らかの方法で非回答者のデータの補完を行うことが望まれる。国民健康・栄養調査で、非回答者の背景因子として把握可能なのは、性・年齢・世帯員数である。また、国民健康・栄養調査は国民生活基礎調査のサブサンプルであるため、両調査をリンクすることが可能であれば、非回答者の特性をより細かく把握することができるので、補完の際の偏りをより小さくできる可能性がある。補完には以下の方法が考えられる。

- Hot-Deck: 非回答者と同じ特性を有するグループの者(同じ性・年齢等)の値で置き換える⁴⁾。
- 線形回帰: 栄養素を目的変数、特性(性・年齢等)を説明変数として回答者のデータを用いて重回帰式を作成し、この式を用いて非回答者の値を予測する⁴⁾。
- Propensity score: 回答か非回答かを目的変数、背景因子を説明変数として多変量ロジスティック回帰を行い、非回答の確率を回帰式で表す。その確率(propensity score)によって複数の群を作り、非回答者の値を、同じ群から無作為に復元抽出した者の値で置き換える。

非回答による偏りがどの程度あるかを把握するためには、これら複数の方法を試みる。国民生活基礎調査とのリンクが可能であれば、回帰(線形回帰および propensity score)による方法が特に有用であろう。これは今後の課題である。

(3) Poststratificationによる人口調整

非回答者の補完を行った上で、抽出確率の逆数を基本の重みとして集計を行えば、集計結果として得られた性・年齢構成は、実際の日本全国（あるいは各都道府県）の人口構成に近いものになることが期待される。ただし、台帳の不完全さ等の理由により、人口構成に少しずれが生ずる可能性がある。これを調整するのがPoststratificationによる人口調整である。

非回答者の補完を行うために必要な情報は整理されていないが、人口構成の違いの影響をおおまかに把握するために、非回答者の補完をせずに、全国人口を用いてPoststratificationによる人口調整を試みた。その結果を表2-4に示す。年齢階級によって摂取量の差が大きいと思われる栄養素において、調整の有無で最大4%程度のずれが認められた。

都道府県別推計のための標本数

国民健康・栄養調査は毎年実施されているので、標本数の設計に必要な情報が事前にかなり詳細に得られている。そこで、都道府県別に一定の精度で（ある誤差率で）推定を行うために必要な単位区数を、平成14年国民栄養調査データに基づくMonte Carlo simulationにより以下の手順で推定した。

(1) 平成14年国民栄養調査データ300単位区の分布を、ある都道府県内での近似的な母分布とみなす。

(2) これら300単位区から $K(=2\sim 40)$ 個の単位区を無作為に復元抽出して誤差率を計算するという作業を1000回ずつ行い、 $K=2\sim 40$ のそれぞれの場合ごとに誤差率の分布を作成する。

(3) 誤差率の分布の中央値が目標の誤差率となる単位区数を“5割の県で所定の誤差率を達成するために必要な単位区数”とし、同様に分布の80点に対応する単位区数を“8割の県で所定の誤差率を達成するために必要な単位区数”とする。

Simulationの結果は図2-2および表2-6にまとめた。例えば、8割の県で食塩摂取量の平均値を誤差率5%で推定するために必要な単位区数は10個である。実際の調査で予想通りの誤差率が得られるかを確認するために、平成12～14年の国民栄養調査で計算された都道府県別単位区数と誤差率も同じ図にプロットした。図から分かるように、Simulationにより予測した標本数別の誤差率は、実際の調査で得られた結果とおおむね一致しており、この方法は標本数の設計に有用であると考えられた。

さらに、性・年齢階級別に推定する場合についても、同様の手法を用いて必要な単位区数を予測した結果を表2-5にまとめた。

経験ベイズ推定の応用

現行の国民健康・栄養調査では、人口の小さな県の標本数がかかなり小さく標本誤差が大きい。そのため、単年度のデータを都道府県別に集計して標本平均を算出すると、人口の小さな県では栄養素の摂取量が極端に多い、または少ないという現象が生じやすい。これは、単なる偶然によるバラツキの影響であるから、誤った解釈に陥らないように注意が必要である（図2-3の●）。このような問題を避けるためには、経験ベイズ推定を行うことを考慮した方がよい。クラスター抽出の場合にはかなり複雑と思われるので、仮に単純無作為抽出を仮定して経験ベイズ推定⁹⁾を行った結果を図2-3に示す。標本サイズが小さい県でも、極端なばらつきが生じにくくなっていることがわかる。今後、層化クラスター抽出の場合の経験ベイズ推定法の応用が望まれる。

D2. 考察

都道府県別に高い精度で推計を行うためには、ほとんどの県で国民健康・栄養調査の対象者だけでは誤差率が大きすぎて標本数が不十分である。独自に都道府県民（健康・）栄養調査を行

うにあたっては、本研究で明らかにした誤差率と標本数との関係を参考にするとよいだろう。一般に、栄養素等の平均値に比べて、割合の推定（例：メタボリックシンドローム有病率）ではより多くの標本数が必要である。調査の際には、重視する変数として何を採用するかで必要なサンプルサイズが大きく異なるので、計画段階で十分な検討が必要である。

【テーマ3】習慣的摂取量の分布の推定

B3. 方法

(1) 正規分布を仮定した分散分析に基づく簡便な方法

1日の食事調査で得られた栄養素摂取量等は、習慣的な摂取量の個人差（個人間変動：分散 σ_b^2 ）に、個人内の摂取量の日間変動（個人内変動：分散 σ_w^2 ）が加わって、分布幅（分散 σ^2 ）が広いものとなる。すなわち、一般に、個人間変動と個人内変動が独立ならば、

$$\sigma^2 = \sigma_b^2 + \sigma_w^2$$

という関係がある。

複数日（2日以上）の食事調査によって、個人間変動 σ_b^2 および個人内変動 σ_w^2 を推定するためには、個体を要因とした一元配置分散分析を行う^{10, 12)}。個体の効果の平均平方和を s_b^2 、誤差の平均平方和を s_w^2 とすると、

$$\hat{\sigma}_b^2 = (s_b^2 - s_w^2) / n,$$

$$\hat{\sigma}_w^2 = s_w^2 \quad \dots (式11)$$

で推定される。習慣的な摂取量の分布が正規分布、個人内変動は全ての個体に共通で正規分布、個人間変動と個人内変動が独立と仮定できるならば、これにより、2日以上食事調査から、習慣的な摂取量の分布を推定可能である¹⁰⁾。習慣的な摂取量の分布を推定するためには、まず1日調査のデータを、分散が $\hat{\sigma}_b^2$ となるように調

整する。すなわち、個体 i の j 日目の摂取量を W_{ij} 、調整した摂取量を Y_{ij} とすると、

$$Y_{ij} = (W_{ij} - \bar{W}_{ij}) \frac{\hat{\sigma}_b}{\hat{\sigma}} + \bar{W}_{ij}$$

ここで、 $\hat{\sigma}^2$ は W_{ij} の分散、 \bar{W}_{ij} は W_{ij} の平均である。一部の対象者のみ複数日調査を行う場合には、一部の対象者の個人間変動 $\hat{\sigma}_b'^2$ と個人内変動 $\hat{\sigma}_w'^2$ を式11より計算し、

$$\hat{\sigma}_b'^2 = \hat{\sigma}^2 \frac{\hat{\sigma}_b'^2}{(\hat{\sigma}_b'^2 + \hat{\sigma}_w'^2)}$$

により推定する。以下、これをANOVA法と呼ぶ。

ビタミンAなど、分布型が強く歪んだ栄養素の場合には、正規分布に近似させるために、べき変換を用いる。べき指数 γ は、1, (1.5)⁻¹, (2.0)⁻¹, ..., (10)⁻¹, または対数変換（0に相当）のうちから、誤差分散

$$\sum_{i=1}^n \sum_{j=1}^k (U_{ij} - \beta_0 - \beta_1 W_{ij}^\gamma)^2$$

が最小になるもの選ぶ¹¹⁾。 W_{ij} の最小値が0で対数変換の場合には、 $W_{ij} > 0$ となる最小の W_{ij} の0.5倍を加算してから計算を行う。 β_0, β_1 は最小二乗法によって決める。 U_{ij} は、全データをプールしたときのBlom正規スコア、つまり、

$$U_{ij} = \Phi^{-1}[(s_{ij} - 3/8) / (nk + 1/4)]$$

ここで、 Φ^{-1} は標準正規分布の分布関数の逆関数、 s_{ij} は W_{ij} の順位である。

(2) Musser SMらのセミパラメトリックな分布変換を用いた方法

前記のANOVA法では、習慣的な摂取量（べき変換した場合も含む）の分布が正規分布で、個人内変動は全ての個体に共通という仮定をおいている。Musser SMらは¹¹⁾、セミパラメトリックな変換方法を用いて、栄養素摂取量の分布を正規分布に近似させ、また、個人内変動が全て