

国民健康・栄養調査の各調査項目における欠損値によるバイアスの検討

研究分担者 横山 徹爾（国立保健医療科学院生涯健康研究部）

研究要旨

近年、国民健康・栄養調査の協力率は低下傾向にあることから、代表性が損なわれることが危惧される。そこで本分担研究では、国民健康・栄養調査とその親標本である国民生活基礎調査とを個人単位でレコードリンケージして多重代入法を適用し、未回答・未測定および非協力によるバイアスがどの程度生じているのかを推定することを目的とする。多重代入法では、第1段階で国民健康・栄養調査の情報のみで欠損値の補完を行ったうえで、第2段階で国民生活基礎調査の情報を用いて非協力者の補完を行う。初年度は、第1段階として、国民健康・栄養調査の身体状況調査、栄養摂取状況調査、生活習慣調査の3つから、多重代入法による欠損値の補完とバイアスの推定を行った。その結果、身体状況調査、栄養摂取状況調査、生活習慣調査のいずれでも、未回答・未測定によるバイアスは非常に小さかった。次年度は国民生活基礎調査とのレコードリンケージにより第2段階の分析を行う。

A. 研究目的

国民健康・栄養調査の通常調査年の対象地区は、直近の国民生活基礎調査の単位区から無作為抽出により選定され、国民を代表するデータとして長年にわたって活用されている。しかし近年、国民健康・栄養調査の協力率は低下傾向にあることから、その代表性が損なわれることが危惧される。西ら¹⁾は国民生活基礎調査に協力した世帯の20歳以上の世帯員を対象として、個人単位でレコードリンケージを行い、世帯や個人の特徴により国民健康・栄養調査の協力率に差がみられたことから、代表性が損なわれてきている可能性を示唆した。しかし、そのような協力率の差により、どの程度の大きさのバイアスが生じているのかは不明である。また、国民健康・栄養調査は、身体状況調査、栄養摂取状況調査、生活習慣調査の3つから成り、各調査により協力率が異なるだけでなく、調査項目によって未回答・未

測定（欠損値）の頻度が異なるため、これによりさらにバイアスが加わる可能性も否定できない。

このような調査への非協力および調査項目毎の欠損を補完し、生じているバイアスの大きさを推定可能な統計学的方法として、多重代入法(multiple imputation)が提案されている²⁾。多重代入法は、既知の（測定されている）対象者の属性や調査項目の値から、欠損値の値を推定して補完するというのが基本的な考え方である。Okuboら³⁾は、乳幼児栄養調査と、その親標本である国民生活基礎調査とをレコードリンケージして、国民生活基礎調査で得られた情報から、多重代入法により乳幼児栄養調査の欠損値を補完し、非協力によるバイアスはかなり小さいことを示した。

国民健康・栄養調査（通常調査年）の対象地区も、親標本である国民生活基礎調査の単位区から無作為抽出されているため、国民生活基礎調査とレコードリンケージすれ

ば、国民生活基礎調査で把握された属性や調査項目の値から、国民健康・栄養調査の非協力者の値を推定して補完し、バイアスの大きさを評価することが可能と考えられる。この補完には2段階を踏む³⁾。第1段階では、国民健康・栄養調査のみで欠損値の補完を行う。例えば欠損値の少ない生活習慣状況調査の情報を用いて、欠損値の多い身体状況調査の値を補完する。同一調査内で補完する方が多くの情報を利用できるという利点がある。第1段階の補完を行ったうえで、第2段階で国民生活基礎調査の情報を用いて、国民健康・栄養調査の非協力者の補完を行う。これらにより最終的に、未回答・未測定・非協力によるバイアスを推定する。

本分担研究では、国民健康・栄養調査と国民生活基礎調査とを個人単位でレコードリンクして多重代入法を適用し、未回答・未測定・非協力によるバイアスがどの程度生じているのかを推定することを目的とする。初年度は、国民生活基礎調査データが入手できていないため、第1段階として、国民健康・栄養調査の身体状況調査、栄養摂取状況調査、生活習慣調査の3つから、一部項目の欠損値を多重代入法により補完し、バイアスの推定を行った。

B. 方法

平成30年国民健康・栄養調査データを目的外利用申請したうえで使用した。20歳以上の男女を分析対象とし、身体状況調査、栄養摂取状況調査、生活習慣調査データを個人ごとに突合し、多重代入法により未回答・未測定(欠損値)の値を補完した。用いた項目は、性・年齢、世帯員の人数、および表1に示した身体状況調査7項目、生活習慣調査11項目、栄養摂取状況調査27項目であり、男女別に分析した。多重代入法の計算に

は、統計ソフトSASのPROC MIを使用し、連続変数の補完には回帰法を、カテゴリー変数の補完には判別関数法を用いた。補完前と補完後では年齢構成が変わる可能性があるため、各項目の平均値と割合を年齢調整したうえで比較した。

C. 結果

表1に、分析に使用した項目の欠損値の割合(何れかの調査データがある人数に対する割合)を示す。生活習慣調査の欠損値は1.4~2.0%と比較的少なく、栄養摂取状況調査の食品群・栄養素等摂取量の欠損値は約13~14%である。身体状況調査では、BMIと歩数の欠損値が20%前後なのに対して、血圧と血液検査は約50~60%が欠損値である。

表2に、元の値の(補完前の)平均値、補完値の平均値(欠損値だった者の予測値の平均)、補完後の全体平均値、および補完前後の平均値の差(バイアス)を示す。身体状況調査ではバイアスは非常に小さく、男性の歩数の+0.6%が最大だった。男性の歩数の補完値の平均は215歩少ないが、欠損値の割合が19%と多くないため、全体へのバイアスは小さいと考えられる。血圧と血液検査は約半数以上が未測定だったにもかかわらず、バイアスはほとんど認められなかった。

栄養摂取状況調査のうち、栄養素等摂取量では、補完値の平均にずれは認められたが、欠損値の割合が小さいためバイアスはほとんど認められず、男性のカルシウムの+0.4%が最大だった。食品群別摂取量も同様で、もともと摂取量の少ない種実類ときのご類が男性で1%を超えたほかは、やはりバイアスは非常に小さかった。生活習慣調査は欠損値が2%以下で非常に少ないため、補

完しても全体の割合に変化はなかった（表には示していない）。

D. 考察

国民健康・栄養調査の身体状況調査、栄養摂取状況調査、生活習慣調査の何れかのデータがある20歳以上の男女を対象として、多重代入法により欠損値の補完を行い、未回答・未測定によるバイアスの大きさを推定したが、バイアスは極めて小さかった。これは、回答・測定ができた者とそうでない者とで背景因子が類似しているか、欠損値の頻度が少ないためと考えられる。用いる項目によっても補完性能は異なるため、他の項目を用いた場合には結果は変わってくる可能性もある。生活習慣調査の項目で用いた、「食品選択で重視すること」、「主食・主菜・副菜を組み合わせる頻度」、「歯の状況」などは、食事への影響が大きいと思われるので、補完の性能は高いはずだが、推定された食品群・栄養素等の摂取量のバイアスは非常に小さかった。

今回補完に用いたのは、国民健康・栄養調査の何れかの項目を調査できた人のデータであり、調査に協力的な比較的均質な人たちとも考えられるので、未回答・未測定項目があっても大きなバイアスは生じなかったのかも知れない。国民生活基礎調査とレコードリンケージを行った西らの報告では¹⁾、世帯人員、世帯構造、世帯業態によって協力率が大きく異なることが示されており、国民生活基礎調査のこれらの項目を用いて補完を行った場合には、より大きなバイアスが推定される可能性もある。この検証は次年度の課題である。

E. 結論

国民健康・栄養調査の身体状況調査、栄養摂取状況調査、生活習慣調査を用いて多重代入法により欠損値の補完を行った範囲では、未回答・未測定によるバイアスは小さかった。次年度は国民生活基礎調査とのレコードリンケージにより同様の分析を行い、非協力によるバイアスの大きさを検討する。

【参考文献】

- 1) 西信雄, 他. 国民健康・栄養調査の協力率とその関連要因. 厚生の指標 2012;59(4):10-15.
- 2) Sterne JA, White IR, Carlin JB, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ. 2009;338:b2393.
- 3) Okubo H, Yokoyama T. Sociodemographic Factors Influenced Response to the 2015 National Nutrition Survey on Preschool Children: Results From Linkage With the Comprehensive Survey of Living Conditions. J Epidemiol. 2020;30(2):74-83.

F. 健康危機情報

なし

G. 研究発表

なし

H. 知的財産権の出願・登録状況

なし

表1. 平成30年国民健康・栄養調査における欠損値の頻度(20歳以上, 一部項目のみ)

	男 n=3095※	女 n=3547※
身体状況調査		
腹囲	31.7%	29.1%
歩数	19.0%	19.2%
HbA1c	61.2%	51.1%
血清総コレステロール	61.1%	50.9%
BMI	22.5%	21.3%
収縮期血圧	58.4%	48.0%
拡張期血圧	58.4%	48.0%
生活習慣調査		
問1 食品選択で重視すること	1.4%	1.5%
問2 主食・主菜・副菜を組み合わせる頻度	1.5%	1.5%
問3 喫煙の状況	1.4%	1.3%
問5 睡眠時間	1.5%	1.4%
問6 睡眠で休養が充分とれているか	1.4%	1.4%
問7 飲酒の状況	1.4%	1.4%
問8 自分の歯の本数	1.9%	1.8%
問9-1 歯ぐきが腫れている	1.8%	1.8%
問9-2 血が出る	1.8%	2.0%
問10 健診等の受診状況	1.4%	1.4%
問11 就業状況	1.5%	1.4%
栄養摂取状況調査		
摂取食品群・栄養素等	14.0%	13.2%

※何れかの調査データがある人数に対する欠損値の割合

表2. 平成30年国民健康・栄養調査における欠損値によるバイアスの推定(20歳以上, 一部項目のみ)

	男性					女性				
	元の値	補完値の	補完後全体	バイアス		元の値	補完値の	補完後全体	バイアス	
	平均※1	平均※2	平均※3	平均※4	%※5	平均※1	平均※2	平均※3	平均※4	%※5
身体状況調査										
腹囲	86.4	85.6	86.3	0.1	0.1%	80.9	80.2	80.8	0.1	0.1%
歩数	6775.6	6560.6	6735.0	40.6	0.6%	5844.2	5405.1	5825.7	18.5	0.3%
HbA1c	5.79	5.73	5.79	0.00	0.0%	5.71	5.65	5.69	0.02	0.3%
血清総コレステロール	199.6	200.2	199.8	-0.2	-0.1%	208.3	206.6	208.2	0.1	0.0%
BMI	23.8	23.8	23.8	0.0	-0.1%	22.4	22.4	22.5	0.0	-0.1%
収縮期血圧	132.2	130.3	132.2	0.1	0.0%	126.6	125.0	126.6	0.1	0.1%
拡張期血圧	82.2	82.4	82.4	-0.2	-0.3%	76.8	76.6	77.0	-0.2	-0.2%
栄養摂取状況調査										
<i>栄養素等摂取量</i>										
総エネルギー(kcal/日)	2165.0	2134.5	2159.7	5.3	0.2%	1726.4	1729.6	1728.0	-1.5	-0.1%
総たんぱく質(g/日)	78.4	77.2	78.2	0.1	0.2%	66.0	65.9	66.0	0.0	-0.1%
総脂質(g/日)	66.8	67.0	66.7	0.1	0.1%	56.0	56.2	56.1	0.0	-0.1%
脂肪エネルギー比率, %	27.4	27.8	27.4	0.0	-0.1%	28.8	28.8	28.8	0.0	0.0%
炭水化物(g/日)	281.6	276.3	280.8	0.8	0.3%	230.0	230.1	230.2	-0.2	-0.1%
カリウム(mg/日)	2446.4	2337.2	2438.1	8.3	0.3%	2277.0	2268.7	2280.5	-3.5	-0.2%
カルシウム(mg/日)	502.5	475.8	500.4	2.1	0.4%	498.5	500.1	499.8	-1.3	-0.3%
鉄(mg/日)	8.25	7.98	8.23	0.02	0.2%	7.49	7.50	7.50	-0.01	-0.2%
ビタミンC(mg/日)	98.2	91.5	97.9	0.2	0.2%	104.1	101.9	104.1	0.0	0.0%
食塩相当量(g/日)	11.0	10.9	11.0	0.0	0.1%	9.3	9.4	9.3	0.0	-0.1%
<i>食品群別摂取量</i>										
穀類(g/日)	493.2	492.9	492.1	1.1	0.2%	355.9	356.3	355.8	0.1	0.0%
いも類(g/日)	52.8	49.5	52.4	0.4	0.8%	49.1	49.3	49.2	-0.1	-0.2%
砂糖・甘味料類(g/日)	6.64	6.40	6.63	0.02	0.2%	6.66	6.97	6.71	-0.05	-0.8%
豆類(g/日)	66.5	61.4	66.1	0.4	0.6%	65.8	66.4	66.1	-0.2	-0.3%
種実類(g/日)	2.40	2.04	2.37	0.03	1.4%	2.91	2.94	2.92	-0.01	-0.5%
野菜類(g/日)	290.1	278.6	289.2	0.9	0.3%	272.8	276.2	273.7	-0.9	-0.3%
果実類(g/日)	88.3	79.2	88.2	0.1	0.1%	110.1	107.3	110.2	-0.1	-0.1%
きのこ類(g/日)	17.1	14.4	16.8	0.3	1.8%	16.9	17.7	17.0	-0.1	-0.8%
海藻類(g/日)	9.54	9.15	9.53	0.01	0.1%	8.46	8.52	8.48	-0.02	-0.2%
魚介類(g/日)	76.0	72.1	75.9	0.1	0.1%	64.4	61.9	64.3	0.1	0.2%
肉類(g/日)	124.5	128.9	124.4	0.1	0.1%	87.6	89.7	87.6	-0.1	-0.1%
卵類(g/日)	44.5	43.2	44.3	0.1	0.3%	40.5	40.6	40.6	-0.1	-0.1%
乳類(g/日)	99.8	92.1	99.2	0.6	0.6%	116.4	115.1	116.6	-0.2	-0.2%
油脂類(g/日)	12.9	12.9	12.8	0.0	0.2%	10.1	10.2	10.1	0.0	-0.2%
菓子類(g/日)	21.8	21.2	21.7	0.1	0.6%	28.2	29.6	28.4	-0.2	-0.8%
嗜好飲料類(g/日)	770.2	749.3	768.2	1.9	0.3%	618.3	616.0	619.0	-0.7	-0.1%
調味料・香辛料類(g/日)	68.1	67.1	68.0	0.1	0.2%	59.2	59.1	59.2	-0.1	-0.1%

※1 何れかの調査データがある者の年齢構成(※3の年齢構成に等しい)に直接法で調整。

※2 多重代入法により補完した者のみの平均値(年齢調整なし)

※3 多重代入法により欠損値を補完した全体の平均値 ※4=※1-※3 ※5=※4/※3