欠測値補完と因果効果推定 による特定健康診査の 医療費への影響の理解

星野崇宏研究会 5期 柿迫栞緒 棚村しほり 西田康平 松永将志 山崎有紗

目次

本研究の概要

実証分析

結論・まとめ

- a. 社会背景
- b. 先行研究とその課題

- a. 使用データについて
- b. 分析方法
- c. 研究目的とその意義 c. 実証分析の結果と考察

- a. 本研究のまとめ
- b. 課題と今後の展望
- c. 社会的意義

本研究会

本研究の概要

実証分析

結論・まとめ

- a. 社会背景
- b. 先行研究とその課題
- c. 仮説の提示

- a. 使用データについて
- b. 分析方法
- c. 実証分析の結果と考察 c. 社会的意義
- a. 本研究のまとめ
- b. 課題と今後の展望

a. 社会背景

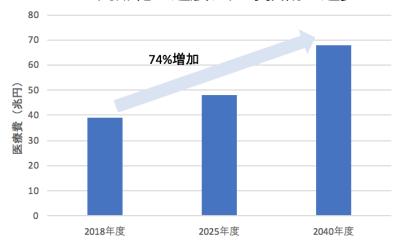
社会背景

平成20年4月

特定健康診査と特定保健指導の実施が医療保険者に義務づけ

1.医療費の増大

● 高齢化の進展、医療技術の進歩



2. 生活習慣病の有病者・予備群増加

- 疾病全体に占める生活習慣病の割合が増加傾向
- 死因の約6割、国民医療費の約3割を占める

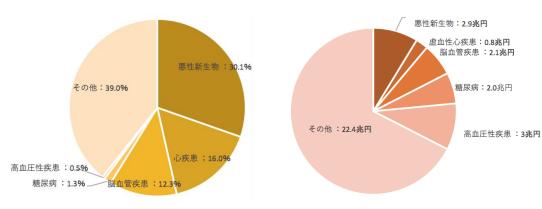


図1:医療費の増加の見通し(2018年)

図2:生活習慣病の死亡数割合と医療費

a. 社会背景

概要と現状

特定健康診査

生活習慣病の 発症リスク高いと…

保健指導

<u>対象者</u>

- 1. 公的医療保険加入者のうち40-74歳
- 2. 当該実施年度の一年間を通じて加入

検査項目

身長体重、血糖、尿検査など

生活習慣病の発症リスク

高い人→積極的支援

低い人→動機付け支援

a. 社会背景

概要と現状

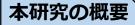
•		受診率	特定健診	保健指導		
	特定健康語	政府目標	70%	45%	保健指導	
	<u>対象者</u>	2017年度	53.1%	19.5%	習慣病の発症リスク	7

- 1. 公的医療保険加入者のう
- 2. 当該実施年度の一年 目標との大きな乖離

低い人→動機付け支援

高い人→積極的支援

<u>検査項目</u> 身長体重、血糖、尿検虫な



a. 社会背景

政府の対策





● 保険者努力支援制度



の付与

保険者

受診率向上に向けた受診勧奨を促進

- 例)特定健診の実施率に応じて
- ①後期高齢者医療制度への財政負担軽減
- ②付与金額の変化

a. 社会背景

政府の対策





保険者

- 後期高齢者支援金の見直し
- 保険者努力支援制度

● 受診率向上に向けた 受診勧奨を促進

- 例)特定健診受診率のシーケップ
- ①後期高齢者医療制作保険者努力制度
- ②付与金額の変化

1000億円

必要経費: 年間230億円弱

b. 先行研究とその課題

特定健診は本当に医療費を減少させるのか

厚労省の主張	受診率の目標達成により 2023年時点で 約200億円の医療費が適正化 される
根拠	厚労省(2016)の調査 保健指導の積極的支援参加者と不参加者では、 不参加者の方が入院外医療費高い
問題点	研究の対象者 積極的支援対象者:最も生活習慣病のリスクの高い上位9% 修了者はそのうち11% →特定健診受診群と非受診群との比較と 対象者の性質が大きく異なる

健診の受診により医療費は増加するという主張も…

- ・受診勧奨時の精密検査等に伴う医療費増加(渡邉, 2017)
- ・医療費のかかるタイミングの先延ばし(康永, 2017)

>b. 先行研究とその課題

先行研究

健診の受診有無が医療費に与える効果についての国内研究は非常に希少 特に個人レベルの医療費を用いて非受診群と比較したものは2件のみ(検索可能な範囲で)

	研究内容・結果	問題点
竹内 (2002)	結果:健診受診群のほうが医療費が有 意に低い	①研究対象:基本健診 →特定健診の医療費抑制効果を支持するものでない ②性・年齢以外の特性の調整業 →自己選択バイアス、healthy screenee bias が除去されていない
渡邊 (2017)	特徴: ①12 年間という長期的な効果測定②自己選択バイアス, healthy screenee bias の除去 結果: 2000年前後の基本健診受診群の医療費は有意に1 人当たり積算医療費が低額	①研究対象:基本健診 →特定健診の医療費抑制効果を支持するものでない ②自記式質問紙によるベースライン調査 →誤分類の可能性 ③傾向スコアマッチングによりバイアスが生じている可能性 ④社会経済的因子の調整が不十分である可能性

b. 先行研究とその課題

先行研究

費が低額

健診の受診有無が医り 特に個人レベル 先行研究の課題: ①特定健診の受診有無が医療費に与える影響 についての研究は過去になされていない ②適切な手法で効果測定されていない 渡。 (2017)可能性 200 療費は有に

本研究の概要 c. 研究目的とその意義

本研究の特徴

目的	特定健診の受診有無が医療費に与える効果を検証する		
研究手法	多重代入を行ったうえで因果効果推定を行 う →バイアスの除去		
研究意義	L) 「受診群の方が医療費 <mark>低い</mark> 」場合		
	①医療費適正化対策としての特定健診の有効性を確認できる		
	②異質性の推定		
	→医療費抑制効果が大きく表れるのはどのような層かを分析		
	→保険者が行っている受診勧奨の効率化が可能に		
	2) 「受診群の方が医療費 <mark>高い</mark> 」場合		
	①特定健診以外の方法での医療費適正化を推進する必要		
	②次世代に負担を付け回さない、 持続可能なシステムを築く重要性 12		

本研究の概要

- a. 社会背景
- b. 先行研究とその課題
- c. 仮説の提示

実証分析

- a. 使用データについて
- b. 分析手法
- c. 実証分析の結果と考察

結論・まとめ

- a. 本研究のまとめ
- b. 課題と今後の展望
- c. 社会的意義

a. 使用データについて

使用データについて

提供元:兵庫県芦屋市

データ収集期間:2013~2018年度

対象者:2016年度を基準に満43歳以上満72歳以下の国民健康保険加入者

なお、データは提供時点で匿名化処理がなされており、個人の特定はできない。

レセプトデータ	患者が受けた保険診療について医療機関等が保険者に請求する月単位の医療報酬明細書 (縄田他, 2019)
特定健診受診券 発行者データ	各年度はじめに特定健診の対象となる被保険者に送付される受診券が発行された人の生年月日、性別等の基本情報が各対象年度ごとにまとめられたデータ
健康診断データ	特定健診受診者の受診結果について各年度ごとにまとめられているデータ
世帯所得データ	11 区分された世帯所得について各年度ごとにまとめられたデータ 14

a. 使用データについて

使用データについて

Xd:対象者の 2016 年度時点の年齢、性別、2013 年度から 2015 年度の世帯所

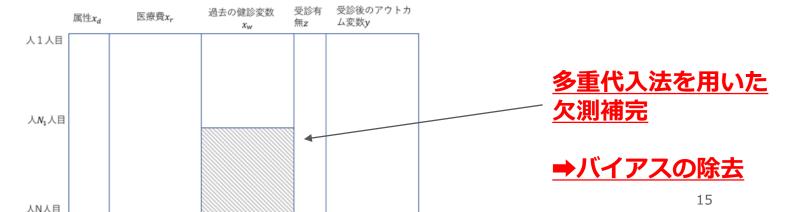
得

Xr: 2013 年度から2015 年度の外来、入院、 調剤別の診療報酬点数

Xw:2015 年度の健康診断データの診断結果

z:2016 年度の健康診断の受診有無

y:2017,2018 年度の外来と調剤の医療報酬点数の合算、入院の医療報酬点数



b. 分析手法

欠測データについて

欠測データメカニズム:データの欠測がどのようなメカニズムで生じているかを 統計的推測を行う際に有用な3区分に分類したもの

分類	ランダムな欠測	完全にランダムな欠測	ランダムでない欠測
	(missing at random;	(missing completely at	(not missing at random;
	MAR)	random; MCAR)	NMAR)
欠測の発生	観測値には依存する が、 欠測値には依存しな い	完全にランダムに発生 →MARのうち 欠測の発生が観測値に も依存しない場合	欠測値に依存する

b. 分析手法

伝統的な欠測データ処理とその問題点

①欠測値に関する部分を無視して解析を行う→完全ケース分析 / リストワイズ削除→ 一部でも変数に欠測があるユニットを削除

特徴 MCARのもとで真のパラメータ値に対する一致性を持つ

問題点① MCARでない場合 推定値にバイアス

問題点② サンプルサイズが小→推定値の分散が過大評価

b. 分析手法

伝統的な欠測データ処理とその問題点

②欠測値に値を代入し、擬似完全データを作成 単一代入法:1つの欠測値に1つの値を代入

ex1) 平均値代入:観測値から推定した平均値を代入

MCAR→平均値や回帰分析の係数が不偏性を持つ

ex2) 回帰代入:欠測値を観測値で説明する回帰分析モデルによる予測値を代入

MAR→回帰代入の結果から得られた平均値は一致性がある

問題点 標本分散・不偏分散が過小評価→推定値の分散も過小評価 回帰モデルが正しくないと推定量が一致性を持たない(回帰代入)



解決策:多重代入法

b. 分析手法

多重代入法(Multiple Imputation; MI(Rubin, 1987))

複数の擬似完全データを作成し、各データごとに計算した複数の推定値の統合を 行う

代入ステージ

単一代入法などを用いて欠測値を補完した複数の擬似完全データセットを作成

解析ステージ

作成した擬似完全データを用いて分析を行い、各データセットごとに推定値とその共分散行列の推定値を得る

統合ステージ

Rubinのルール(Rubin, 1987)により、各推定値とその共分散行列の推定値を統合

b. 分析手法

代入ステージ

【一般的な多重代入法】

補完値の生成アルゴリズムの仮定: 1つの変数のみに欠測が生じる

⇔特定の1つの変数のみ欠測が生じることは**ほとんどない**



ギプスサンプラー(マルコフ連鎖モンテカルロ法の一種)への類推から データ拡大アルゴリズムを行う

【連鎖式による多重代入法(multiple imputation by chained equation; MICE)】

統合ステージ

Rubinのルール

パラメータ θ に対して、擬似完全データを利用した際のD個の推定値 $(\hat{\theta}_1, \cdots, \hat{\theta}_D)$ とその共分散行列の推定値 (W_1, \cdots, W_D) が得られているとき、統合された推定値 $\bar{\theta}_D$ は、算術平均より

$$ar{m{ heta}}_D = rac{1}{D} \sum_{d=1}^D \hat{m{ heta}}_d$$

となる。

統合ステージ

推定量の共分散行列の推定値 T_D は、代入値内共分散行列の平均

$$\bar{W}_D = \frac{1}{D} \sum_{d=1}^D W_d$$

および推定量の代入値間共分散行列の推定値

$$B_D = rac{1}{D-1} \sum_{d=1}^D \left(\hat{oldsymbol{ heta}}_d - ar{oldsymbol{ heta}}_D
ight)^2$$

を用いて

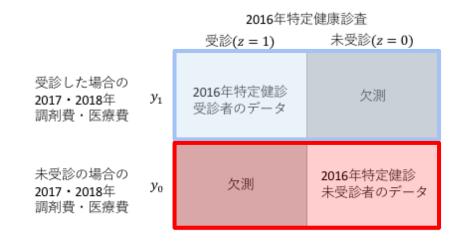
$$T_D = \bar{W}_D + \left(1 + \frac{1}{D}\right) B_D$$

となる。

b. 分析手法

因果効果とは

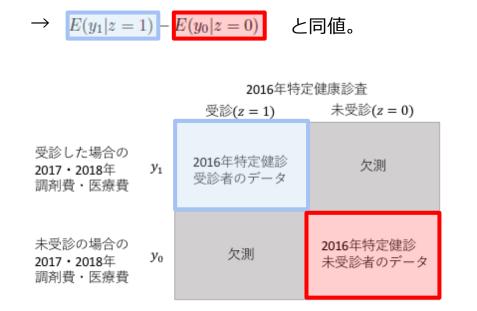
本来は $E(y_1 - y_0) = E(y_1) - E(y_0)$ を因果効果としたい \rightarrow 欠測により不可能。(Rubinの因果効果モデル)



b. 分析手法

単純な差

zが無作為割当である場合→先ほどみた $E(y_1 - y_0) = E(y_1) - E(y_0)$



健診を受診する層と受診しない層 は異なるため、

単純比較は因果効果でない

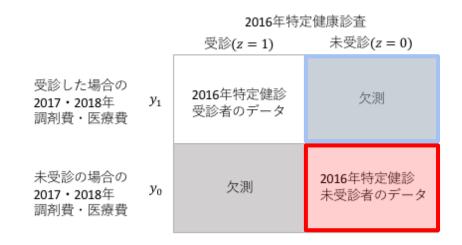
年齢や所得等、受診有無以外の 共変量(交絡要因)の調整が必須 →**傾向スコアの導入**

b. 分析手法

ATU(対照群の平均処置効果)

そこで、今回は傾向スコアとIPW推定量を用いて 「未受診の人がもし健診を受けた場合の効果(ATU)」を解析

ATU:
$$E(y_1 - y_0 \mid z = 0) = E(y_1 \mid z = 0) - E(y_0 \mid z = 0)$$



b. 分析手法

因果効果とは

$$E(y_1 \mid z = 0)$$
 の求め方

第 i 人が2016年の健診を受ける確率(傾向スコア) $e_i = p(z_i = 1 \mid x_i)$ を

とおくと、適切な重みwを用いて表した $\frac{1}{W_1}\sum_{i=1}^N\frac{(1-z_i)y_i}{e_i}$

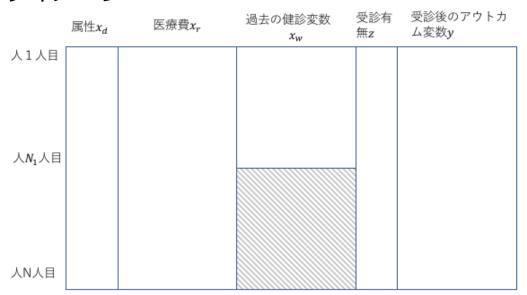
が $E(y_1 \mid z = 0)$ が一致推定量となる。

このwが傾向スコアを用いていることから、 「**傾向スコアによる重み付け推定法**」と呼ばれる

b. 分析手法

分析の枠組み

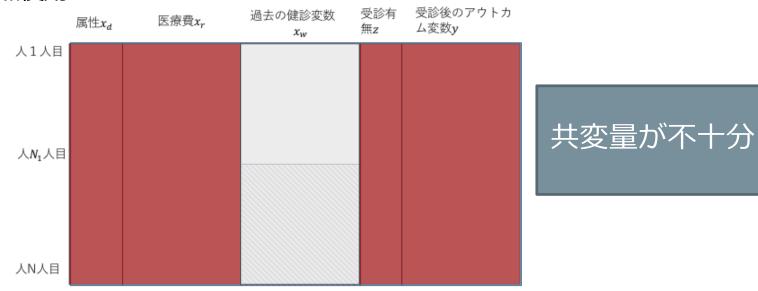
そもそものデータイメージ



b. 分析手法

分析の枠組み

①健診情報無使用



b. 分析手法

分析の枠組み

②完全ケースデータ

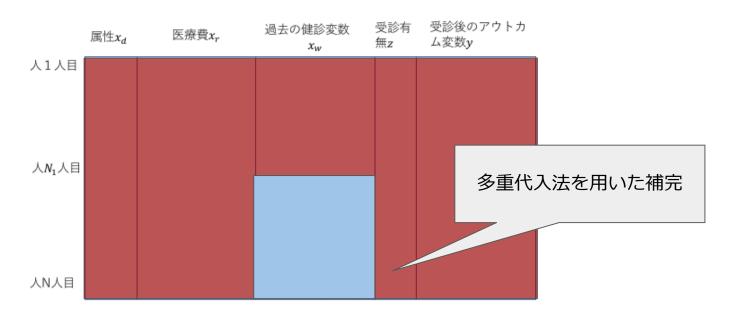


2015年の健診受診者に限定

b. 分析手法

分析の枠組み

③多重代入後



b. 分析手法

分析の枠組み

名称	健康診断 情報	標本数(人)	受診群	割合(%)	特徴
①健診情報無 使用	×	8149	3244	39.809	健診情報がな く不十分
②完全ケース データ	0	2931	2349	80.143	年齢等の分布 に偏り有
③多重代入後	0	8149	3244	39.809	全対象者の 健診情報有

c. 実証分析の結果と考察

単純な差

I 2群の差をみると、健診未受診群のほうが医療費が高いII しかし、共変量について考慮した結果ではないIII完全ケースの対象者は必ず2015年の健診に参加しており、標本の偏り等の問題がある。(選択バイアス)

$E(y_1 z=1) - E(y_0 z=0)$	外来・調剤医療費(点)		入院医療費(点)	
	推定值標準偏差		推定値	標準偏差
1 • 3	-2764	152989	-9740	142126
② (完全ケース)	-13958	201225	-6350	111200

c. 実証分析の結果と考察

ATU(対照群における平均処置効果)

I いずれの結果も、未受診群の方が**医療費が高い**という結論 II 多重代入法を用いた結果が最も妥当だと考えられる III その結果、外来+薬局調剤費は**6万8千円**、入院費は**10万円**抑制

結果変数	外来・調剤医療費	賃 (点)	入院医療費(点)		
ATU	推定値	標準誤差	推定値	標準誤差	
1	-2220	2742	-9584	2574	
2	-11745	8947	-5974	9508	
③(多重代入)	-6880	2842	-10386	2207	

c. 実証分析の結果と考察

異質性の推定

世帯所得によって、健診受診による医療費抑制効果に違いがあるのか推定

→所得を標準化しwとした上で、結果変数を単回帰で説明 ただし、今回は分析の枠組み①を使用

外来医療費+薬局調剤費

$$E(y_1 - y_0|w = c) = \hat{\beta}_{10} + \hat{\beta}_{11}w - (\hat{\beta}_{00} + \hat{\beta}_{01}w) = -1618.3 + 1238.4w$$

入院医療費

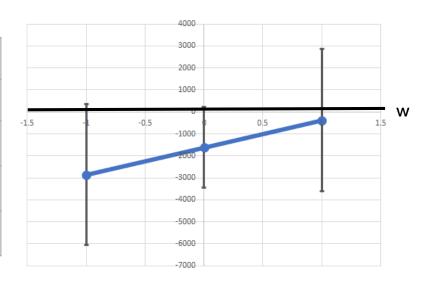
$$E(y_1 - y_0|w = c) = \hat{\beta}_{10} + \hat{\beta}_{11}w - (\hat{\beta}_{00} + \hat{\beta}_{01}w) = -9419 + 3283w$$

c. 実証分析の結果と考察

異質性の推定

外来・調剤医療費

w = 所得	外来・調剤医療費(点)		
所得との差	推定値	標準誤差	
w = - 1	-2857	3209	
w = 0	-1619	1840	
w = 1	-380	3241	

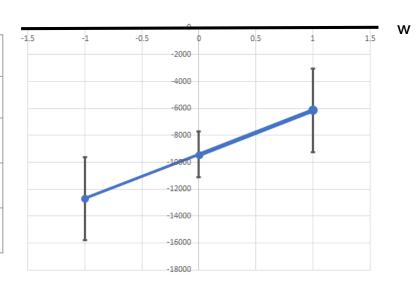


c. 実証分析の結果と考察

異質性の推定

入院医療費

w = 所得	入院医療費(点)	
所得との差	推定値	標準誤差
w = - 1	-12702	3089
w = 0	-9419	1722
w = 1	-6136	3123



c. 実証分析の結果と考察

異質性の推定

- I 両結果変数とも、所得が低い人ほど抑制効果が大きい
- Ⅱ 所得が高い人は、私費で人間ドックを受けている等の可能性

本研究の概要

- a. 社会背景
- b. 先行研究とその課題
- c. 仮説の提示

実証分析

- a. 使用データについて
- b. 分析方法
- c. 実証分析の結果と考察

結論・まとめ

- a. 本研究のまとめ
- b. 課題と今後の展望
- c. 社会的意義

a. 本研究のまとめ

結論

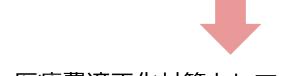
「特定健診の受診行為には医療費抑制効果がある」

- 1. IPW推定法により、**特定健診を受診しなかった人は受診により外来・調剤医療** 費と入院医療費のいずれも抑制できることが確認された。
- 1. 異質性の推定により、**外来・調剤医療費、入院医療費ともに所得が低い人ほど 抑制の効果が高く、その差は大きい**ことがわかった。

c. 社会的意義

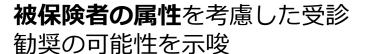
今回の研究の意義

医療費適正化対策としての特定健診には有効性がある



- ・医療費適正化対策としての特定健診の正当性を支持する根拠を初めて示した
- ・特定健診を**受診する際のインセンティブ** として受診勧奨に利用可能

異質性の推定: 所得の低い人のほうが医療 費抑制効果が高い



b. 課題と今後の展望

今後の展望

1. 効果測定が**短期**にとどまる

2. 2016年単年の健診受診のみ

3. 異質性の分析内容がやや希薄



1. 中長期的な医療費抑制効果の測定

2. 継続的な特定健診受診の効果測定

3. 異質性の詳細な分析