

レセプトの匿名加工情報を用いた歩数による
生活習慣病罹患, 医療費の関係について

菊池研究室 4年

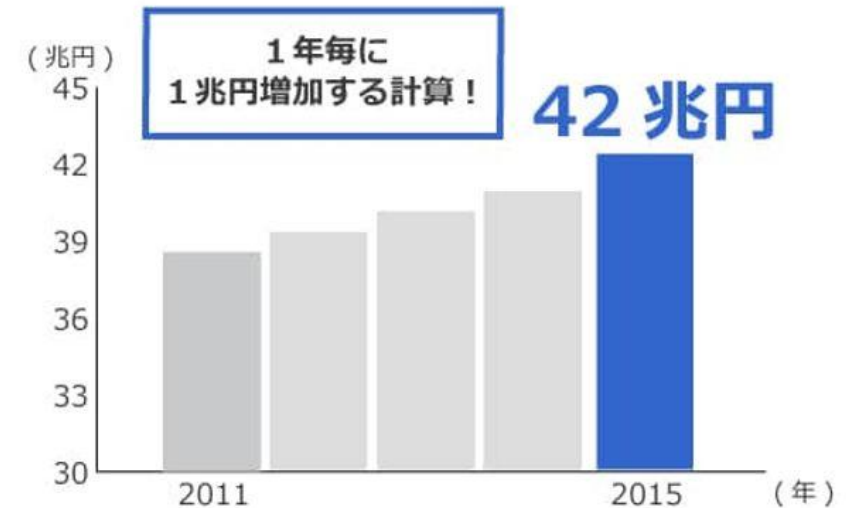
進藤 翔太

研究背景と目的

高齢者の増加に伴って、医療費は毎年増加傾向

毎日、運動するのは大変なので、歩行方法を変えることで医療費を抑えたりすることはできないだろうか？

国民医療費の年次推移



出典: 厚生労働省

健康状態を維持し、医療費を抑えるベストな歩行方法を明らかにする



本研究の新規性

	本研究	先行研究 [Pedro 2020]
計測器	ヘルスケアデバイス	歩数計
計測期間	330日以上	7日
使用データ	ヘルスケア事業者から提供された匿名加工情報	独自に収集したデータ (個人情報を含む)



データ収集が容易に
(ビッグデータ)

+

個人情報を保護

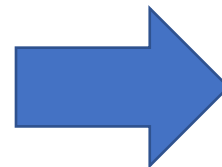
匿名加工情報と活用例(購買データ)

氏名	生年月日	性別	住所
山田太郎	1975/10/11	男	東京都中野区中野~
鈴木花子	1960/12/20	女	千葉県千葉市中央区~

氏名	日付	購入品
山田太郎	2021/1/10	タバコ
鈴木花子	2021/1/11	洗剤

個人情報

匿名加工



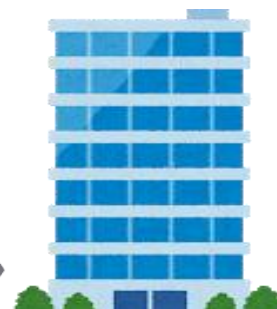
個人を特定できないように

仮id	生年	性別	住所
00001	1975	男	東京都
00002	1960	女	千葉県

仮id	日付	購入品
00001	2021/1/10	タバコ
00002	2021/1/11	洗剤

匿名加工情報

マーケティング企業
(第三者)



個人情報として扱う必要がなくなる
第三者に提供が可能



消費者属性と購買傾向を分析し、新商品
を開発

データ概要

ヘルスケアサービス事業者から提供された匿名加工情報

データ名	レセプトデータ	歩数データ	健康診断データ	医療費データ
属性数	15	4	49	49
期間	2012-2018	2014-2018	2008-2018	2008-2018
レコード数[行]	3,936,887	24,936,120	964,636	964,636
対象者[人]	288,568	58,655	198,740	198,740

グループ1

40～59歳 男性2159人
・歩数 330日以上
・健康診断の記録あり

歩数(客観的)

グループ2

40～59歳 男性39,724人
・健康診断の記録あり

歩行時間(主観的)

リスク分析の誤謬

相対リスクを普通に計算した場合の問題点

例	喫煙者	非喫煙者
生活習慣病である	100	50
生活習慣病でない	100	150

生活習慣病の割合 喫煙者 $100/200=1/2$ 非喫煙者 $50/200=1/4$

非喫煙者を基準とした喫煙者が生活習慣病になる相対リスクは

$1/2 / (1/4) = 2$ 倍

これは本当に正しい？

喫煙者は運動不足や肥満な人が多いかも

だとしたら、喫煙のみでなく、交絡を含んでしまっている



ロジスティック回帰について

そこで、医療統計などではロジスティック回帰が多く用いられる

ロジスティック回帰では、生活習慣病になる確率を p として、

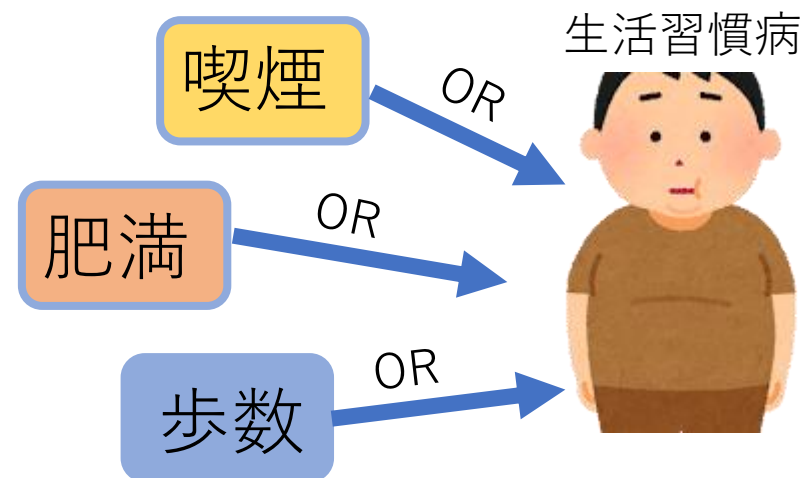
$$p = \frac{1}{1+e^{-ax}} \quad (a = a_1x_1 + a_2x_2 + \dots + a_ix_i + \beta)$$

(x = 喫煙, 肥満かどうか, 歩数などの健康因子, β = 係数)

で与えられる。

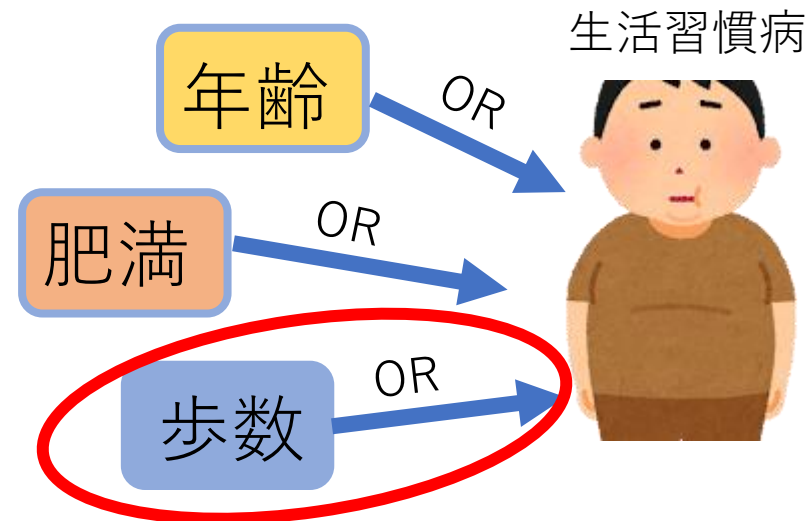
オッズ比 OR(各因子の影響力の指標)

$$OR = e^{a_i}$$



生活習慣病リスク 分析手法

・ 40~59歳男性,2159人を対象(グループ1)に目的変数を3年以内の生活習慣病の罹患の有無として,歩数について,交絡因子の影響を取り除いたロジスティック回帰を行う.



各疾病におけるロジスティック回帰分析(歩数)



年齢, BMI, 喫煙の
影響を排除した

平均歩数		<4000	4000-5999	6000-7999	8000-9999	≥10000
生活習慣病 全体	OR	1.50	1.19	1	1.10	1.12
	p-value	0.008	0.141	1	0.464	0.52
高血圧症	OR	2.01	1.74	1	1.30	0.89
	p-value	0.002	0.002	1	0.248	0.727
糖尿病	OR	1.47	1.30	1	1.28	1.82
	p-value	0.033	0.064	1	0.147	0.003
脂質異常症	OR	1.78	1.42	1	1.04	1.35
	p-value	0.003	0.025	1	0.854	0.220
高尿酸血症	OR	1.82	1.46	1	1.08	0.86
	p-value	0.029	0.096	1	0.792	0.734
虚血性心疾患	OR	1.68	1.04	1	0.96	1.74
	p-value	0.054	0.854	1	0.891	0.071

歩数の影響が強い

※赤は統計的に有意

高血圧における歩数の影響力

高血圧症におけるロジスティック回帰分析

説明変数	a	OR	p-value
歩数 (1:<4000, 0:6000 – 7999)	0.69796	2.01	0.002
年齢	0.07692	1.08	0.000
BMI	0.12277	1.13	0.000
喫煙(1: はい 0:いいえ)	-0.03155	0.97	0.903

簡単に言うと

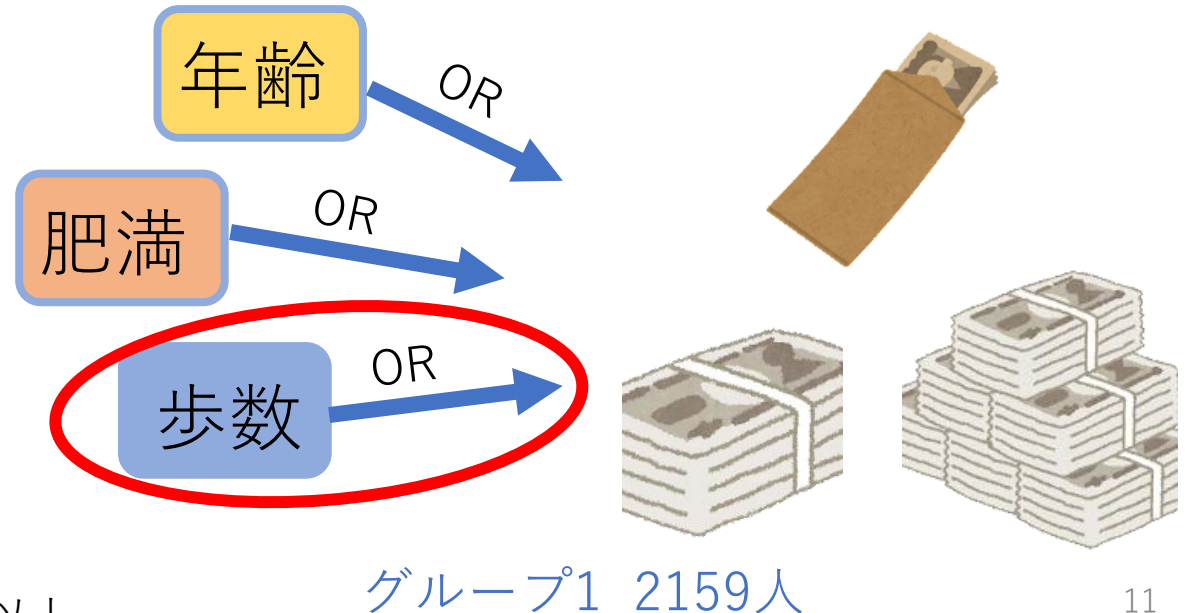
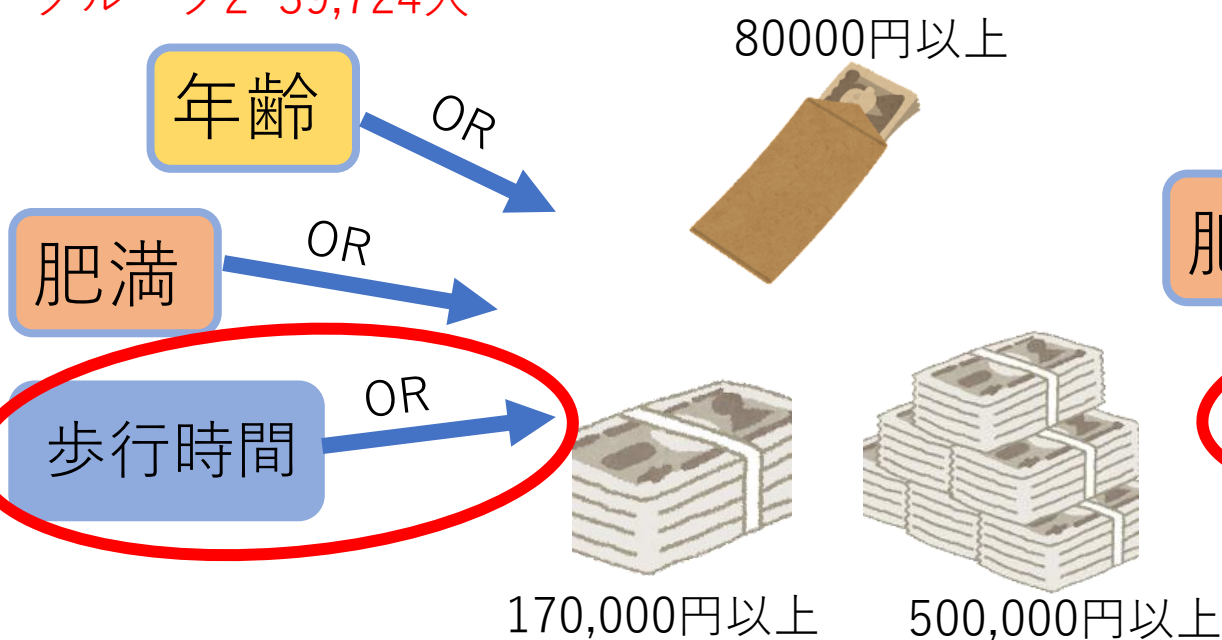
高血圧症で同等の悪影響

OR	歩数	年齢	BMI
2.01倍	4000歩未満	9歳年を取る	5.68上がる

医療費 分析手法

1. 40~59歳男性 39,724人を対象(グループ2)に目的変数を翌年の医療費が80,000円以上, 170,000円以上, 500,000円以上かどうかとし, 歩行時間について, ロジスティック回帰を行う.
2. 1と同様の内容をグループ1で, 歩数についてロジスティック回帰を行う.

グループ2 39,724人



グループ1 2159人

歩行時間, 歩数が医療費に及ぼす影響(ロジスティック回帰)

歩行時間(問診表の回答)と医療費

医療費	80,000円以上		170,000円以上		500,000円以上	
人数(割合)	18,569(46.7%)		10,469(26.4%)		2,554(6.4%)	
項目	OR	p-value	OR	p-value	OR	p-value
1日1時間以上歩行	0.89	0.000	0.85	0.000	0.76	0.000

オッズ比が1未満
医療費を下げる
効果がある

歩数と医療費

医療費	80,000円以上		170,000円以上		500,000円以上	
人数(割合)	1023(47.4%)		539(25.0%)		117(5.4%)	
項目	OR	p-value	OR	p-value	OR	p-value
4000歩未満	0.96	0.792	1.14	0.456	1.74	0.059
4000-5999歩	1.13	0.258	1.07	0.606	1.30	0.293
8000-9999歩	1.06	0.669	1.11	0.488	1.18	0.954
10000歩以上	0.99	0.971	1.15	0.474	1.11	0.787

p-valueが全部
0.05以上
歩数は医療費
に影響していない

まとめ

生活習慣病リスク

- 4000歩未満だと3年以内の高血圧症, 糖尿病, 脂質異常症, 高尿酸血症のリスクが1.36~1.86倍上がった
- 4000歩未満は, 高血圧症において9歳衰えた, BMIが5.68上がるのと同じくらい不健康

医療費

- 歩行時間が1時間以上だと医療費が下がっていたが, 歩数は医療費に影響していなかった.

本研究のベストな歩行法

生活習慣病のリスクを下げ、医療費を抑えるには

1日6000～8000歩で、1時間以上の歩行

の適度な歩行を毎日行うのがベスト

みなさんも正しい生活習慣を心掛けましょう

