

明治大学総合数理学部

2021 年度

卒 業 研 究

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

学位請求者 先端メディアサイエンス学科

進藤 翔太

目次

第 1 章	はじめに	2
1.1	研究背景	2
1.2	研究目的	2
1.3	先行研究	2
1.4	研究方法	3
第 2 章	統計手法	4
2.1	χ^2 検定	4
2.2	ロジスティック回帰分析	4
第 3 章	歩数と生活習慣病のリスク分析	6
3.1	データ概要	6
3.2	データの前処理	7
3.3	分析手法	7
3.4	ユーザの罹患率とリスク比	7
3.5	ロジスティック回帰分析	8
3.6	先行研究との比較	8
3.7	考察	9
第 4 章	歩数と医療費の関係	11
4.1	データ概要	11
4.2	データの前処理	11
4.3	分析手法	11
4.4	歩数と医療費のロジスティック回帰	12
4.5	歩行時間と医療費のロジスティック回帰分析	12
4.6	単回帰分析	13
4.7	歩数と歩行時間の関係, 先行研究との比較	14
4.8	考察	15
第 5 章	おわりに	17
	参考文献	19

第 1 章

はじめに

1.1 研究背景

日本人の死因のうち、約 5 割が生活習慣病によるものであり、医療費の約 3 割を占める [1]. 生活習慣病を引き起こす要因は喫煙、飲酒、食事、ストレス、運動不足などの生活習慣によるものであるとされており、生活習慣を改善することによって生活習慣病の罹患者を減らすことは我が国にとって大きな課題となっている。問題点として、どこからが不健康で生活習慣が悪いとされるのかを理解するのが難しく、実際に自身の生活習慣が悪いのに、悪いと認識すること自体が難しい可能性がある。特に運動不足などはどこからが運動不足なのか、具体的にどのくらい罹患のリスクが上がるのかがわからないのが問題である。

1.2 研究目的

本研究では、運動不足によって生活習慣病のリスクがどの程度上がってしまうのかを具体的な数値で示すことを目的とする。運動は人間がほぼ毎日行う歩行運動に絞って調査を行い、1 日の平均歩数が少ないことにより、生活習慣病のリスクがどれくらい上がったのか、医療費に与える影響を明示する。

1.3 先行研究

1.3.1 歩数と死亡リスク

歩数と死亡リスクには相関があることがわかっている。Pedro らは 2003~2006 年に米国国民健康栄養調査に参加した 40 歳以上の男女のうち 4840 人に、最長 7 日間歩数を計測を行った。その後、約 10 年間の死亡の有無を追跡した。その結果、1 日 4000 歩の人たちに比べ、8000 歩の人々は死亡リスクが半減したことを報告している [2].

1.3.2 生活習慣と医療費の関係

生活習慣と医療費には関連がある。栗山は、40 歳から 79 歳の男女を対象に、運動不足 (1 日の歩行が 1 時間以下)、肥満、喫煙といった生活習慣関連因子が医療費に及ぼす影響を調査し、これらの因子に多く当てはまるほど医療費が高くなることを示した [3]. 辻らは歩行時間の長さやと医療費の関係を明らかにし、一部の参加者の歩

数を計測することで、歩数と医療費の回帰式*を算出した [4].

1.4 研究方法

あるヘルスケアサービス事業者から法律に従って適切に措置され、提供されたヘルスケアデータの匿名加工情報 [5] を使用する。先行研究は、計測した歩数の日数が短く、信憑性に欠けることが問題点として挙げられる。そこで、本研究ではウェアラブルデバイスに着目し、調査を行う。ウェアラブルデバイスから得られた1日の歩数の記録が年間で330日以上あるユーザを対象に調査を行う。それぞれのユーザの1日の平均歩数を計算し、レセプトにおける生活習慣病の罹患記録と照合してリスクを算出する。医療費に関しても、レセプトの記録と照合し、歩数と医療費に関連があるかを調査する。

* $y = -1.3408x + 28205$ ($y = 1$ カ月あたりの医療費, $x =$ 歩数)

第2章

統計手法

2.1 χ^2 検定

本研究では独立性の χ^2 検定を用いる。独立性の χ^2 検定は分類基準間に関連があるかを検定する。例えば、歩数群と生活習慣病で、集計表が表 2.1 のようになったとする。このとき、帰無仮説は H_0 :(歩数群と生活習慣病には関連がない) と仮定される。 i 行の合計を $f_{i\cdot}$, j 列の合計を $f_{\cdot j}$ とすると、 i 列, j 行目の理論値は

$$\frac{f_{i\cdot} \times f_{\cdot j}}{N} \quad (2.1)$$

で表され、表 2.1 から各成分の理論値を計算したものが、表 2.2 となる。 χ^2 値は各成分において、表 2.1 の実測値から表 2.2 の理論値の差を 2 乗したものを理論値で割った総和となる。今回の例では χ^2 値は 7.2747 で p-value は 0.007 となり、p-value が有意水準 0.05 を下回ったことから、帰無仮説 H_0 は棄却され、歩数群と生活習慣病には関連があると結論づけられる。

表 2.1 歩数群と生活習慣病の集計表

歩数	生活習慣病である	生活習慣病ではない	計
4000 歩未満	109	157	266
6000~7999 歩	212	462	674
計	321	619	940

表 2.2 各成分の理論値

歩数	生活習慣病である	生活習慣病ではない
4000 歩未満	90.8	175.2
6000~7999 歩	230.2	443.8

2.2 ロジスティック回帰分析

ロジスティック回帰分析は医療統計で多く用いられる統計手法である。生活習慣病に罹患する確率を y とすると、

$$y = \frac{1}{1 + e^{-(a_1 x_1 + a_2 x_2 + \dots + a_i x_i + a_0)}} \quad (2.2)$$

で表される. 変数 x_i には, 今回調査する歩数のみだけでなく, 年齢, BMI などの交絡因子を含む. a_i は各変数の係数, a_0 は定数である. ロジスティック回帰分析を行う際に, 歩数のみでなく, 他の因子を変数を加えるのは, それらの影響を取り除いた分析を行うとともに, 他の因子と影響を比較するためである. ロジスティック回帰の結果から各因子の影響力を測る指標として, オッズ比 OR(Odds Ratio) が存在する. ロジスティック回帰分析において, 交絡因子の影響を調整した OR は

$$OR = e^{a_i} \quad (2.3)$$

で定義される. 表 2.3 の集計表において, 6000~7999 歩を基準とした 4000 歩未満の OR は,

$$OR = \frac{a/b}{c/d} \quad (2.4)$$

で定義され, 相対リスク RR(Relative Risk) は

$$RR = \frac{a/a+b}{c/c+d} \quad (2.5)$$

で定義される. $a \ll b, c \ll d$ の場合, $OR \approx RR$ となり, オッズ比は相対リスクと同じとみなせる.

表 2.3 歩数群と生活習慣病の集計表

歩数	生活習慣病である	生活習慣病ではない	計
4000 歩未満	a	b	$a + b$
6000~7999 歩	c	d	$c + d$
計	$a + c$	$b + d$	N

第3章

歩数と生活習慣病のリスク分析

3.1 データ概要

ヘルスケアサービス事業者から提供されたヘルスケアデータの匿名加工情報のうち、使用した匿名加工データの統計量を表 3.1 に示す。レセプトデータの属性、歩数データの属性、健康診断データの属性を表 3.2 に示す。

表 3.1 匿名加工データの統計量

データ名	レセプトデータ	歩数データ	健康診断データ
属性数	15	4	49
期間	2012–2018	2014–2018	2008–2018
レコード数	3,936,877	24,936,120	964,636
対象者	288,568	58,655	198,740

表 3.2 各匿名加工データの属性

データ名	属性	例
レセプト	仮個人 id	user_A
	診療開始日	2010/11/10
	icd10 コード 1_基本疾患_細分類	I10
歩数	仮個人 id	user_A
	年月日	2015/6/11
	歩数	5602
	歩数データソース	スマートフォン自動連携
健康診断	仮個人 id	user_A
	健診受診月	2010/11/10
	BMI	21.2
	喫煙	はい

3.2 データの前処理

ウェアラブル端末の歩数の記録が2016年、330日以上ある40～59歳のユーザー2159人を対象とする。2016年の健康診断データから、BMI、喫煙の有無の問診票の項目、2016年～2018年のレセプトデータから、表3.3に示す5つの生活習慣病のICD-10コードを取り出す。ICD-10コードは、世界保健機関(WHO)が作成している疾病の分類である。ICD-10コードによる生活習慣病の分類は[6]を参考にした。

表 3.3 ICD-10 コードによる生活習慣病の分類

生活習慣病	ICD-10 コード
高血圧症	I10
糖尿病	E11, E12, E13, E14
脂質異常症	E780, E781, E785
虚血性心疾患	I209, I219, I259
高尿酸血症	E790

3.3 分析手法

5つの生活習慣病について、歩数が4000未満、4000～5999、6000～7999、8000～9999、10000以上の5つのグループに分け、次の2つの分析を行う。

(1) 6000～7999歩を基準にして、各疾病に3年以内に罹患するリスク比 Relative Risk(RR)を求め、独立性の χ^2 検定をRのchisq.testで行う。

(2) Rのglmを用いて、ロジスティック回帰分析を行う。目的変数 y を3年以内の罹患の有無、説明変数を歩数を2000歩ずつに分け、6000～7999歩を基準にして、相対的に比較する歩数群 x_1 、年齢 x_2 、BMI x_3 、喫煙の有無 x_4 としたロジスティックモデル

$$y = \frac{1}{1 + e^{-(a_1 x_1 + a_2 x_2 + a_3 x_3 + a_4 x_4)}}$$

を用いて、交絡因子の影響を考慮した各疾病のオッズ比 Odds Ratio(OR)を求める。

3.4 ユーザの罹患率とリスク比

ユーザーの特徴、罹患率、リスク比、 χ^2 検定の結果を表3.4に示す。4000未満のグループは全疾病について罹患率が高く、6000～8000歩の基準群と比べるとリスクが高血圧症が1.86倍、高尿酸血症が1.8倍、脂質異常症で1.63倍、糖尿病で1.36倍、生活習慣病全体としても1.3倍とリスクが上がり、統計的に有意な水準であった(p値が0.05未満)。4000～5999歩ではリスクが高血圧症で1.63倍、脂質異常症1.35倍で有意差が見られたものの、生活習慣病全体では1.12倍で有意差が見られなかった。8000歩ではすべての項目で差が見られなかった。10000歩以上では糖尿病のリスクが1.52倍と高くなったが、高血圧症が0.88倍、高尿酸血症が0.76倍と低くなった影響で、生活習慣病全体では1.06倍とリスクがほとんど変わらない結果となった。

表 3.4 平均歩数による各疾患の罹患率

平均歩数	<4000	4000-5999	6000-7999	8000-9999	≥10000
全体	266	679	674	363	177
平均年齢	48.9	49.0	50.0	48.8	49.9
BMI の平均	24.2	24.0	24.0	23.9	24.0
喫煙の有無 (%)	70(26.3)	169(24.9)	146(21.7)	82(22.6)	36(20.3)
罹患数 (%)	109(40.1)	240(35.3)	212(31.5)	121(33.3)	59(33.3)
生活習慣病全体 RR(95%CI)	1.30(1.09-1.56)	1.12(0.97-1.31)	1	1.06(0.88-1.27)	1.06(0.84-1.34)
p-value	0.007	0.144	1	0.583	0.699
罹患数 (%)	41(15.4)	92(13.5)	56(8.3)	37(10.2)	13(7.34)
高血圧症 RR(95%CI)	1.86(1.27-2.70)	1.63(1.19-2.23)	1	1.22(0.83-1.82)	0.88(0.49-1.58)
p-value	0.002	0.003	1	0.369	0.792
罹患数 (%)	63(23.7)	146(21.5)	117(17.4)	75(20.7)	47(26.6)
糖尿病 RR(95%CI)	1.36(1.04-1.79)	1.24(1.00-1.54)	1	1.19(0.92-1.54)	1.52(1.13-2.06)
p-value	0.033	0.063	1	0.222	0.008
罹患数 (%)	53(19.9)	112(16.5)	82(12.2)	45(12.4)	27(15.3)
脂質異常症 RR(95%CI)	1.63(1.19-2.24)	1.35(1.04-1.77)	1	1.02(0.72-1.43)	1.25(0.84-1.87)
p-value	0.003	0.028	1	0.993	0.333
罹患数 (%)	25(9.4)	51(7.5)	35(5.2)	20(5.5)	7(3.4)
高尿酸血症 RR(95%CI)	1.80(1.10-2.96)	1.45(0.95-2.19)	1	1.06(0.62-1.81)	0.76(0.34-1.69)
p-value	0.026	0.10	1	0.943	0.630
罹患数 (%)	25(9.4)	41(6.0)	39(5.8)	20(5.5)	17(9.6)
虚血性心疾患 RR(95%CI)	1.62(1.00-2.63)	1.04(0.68-1.60)	1	0.95(0.56-1.61)	1.66(0.96-2.86)
p-value	0.066	0.94	1	0.966	0.098

3.5 ロジスティック回帰分析

歩数の項目のロジスティック回帰分析の結果を表 3.5 に、高血圧症における全項目のロジスティック回帰分析の結果を表 3.6 に示す。表 3.5 でロジスティック回帰分析の交絡因子を考慮した分析を行っても、4000 歩未満であると、3 年以内の短期間に高血圧症で 2.01 倍、高尿酸血症で 1.82 倍、脂質異常症で 1.78 倍、糖尿病で 1.47 倍、生活習慣病全般でもリスクを 1.5 倍に上げ、p 値が 0.05 未満で統計的に有意であった。表 3.6 で歩数、年齢、BMI が p 値が統計的に有意な水準で、生活習慣病のリスクを上げる因子であるとわかったが、喫煙の項目に有意差は見られなかった。

3.6 先行研究との比較

先行研究の 4000 歩を基準とした約 10 年以内の死亡リスクと本研究の 6000~8000 歩を基準とした 3 年以内の生活習慣病の罹患リスクの比較を表 3.7 に示す。先行研究は歩数に比例して死亡リスクが減ったのに対

表 3.5 ロジスティック回帰分析の結果 (歩数)

平均歩数	x_1	<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

表 3.6 高血圧症におけるロジスティック回帰分析

説明変数 x_i	a_i	OR	p-value
歩数 (1:<4000 0:6000–7999)	0.698	2.01	0.002
年齢	0.087	1.08	0.000
BMI	0.123	1.13	0.000
喫煙 (1: はい 0: いいえ)	-0.032	0.97	0.903

表 3.7 先行研究とのリスク比較

平均歩数 (先行研究)	本研究 (生活習慣病)	先行研究 (死亡)
10000 以上 (12000)	1.06	0.35
6000–8000(8000)	1	0.49
4000 未満 (4000)	1.30	1

して、本研究では、10000 歩以上では 1.06 倍と罹患リスクが再び増加する非単調な振る舞いをする結果となった。歩数が少ない場合にはどちらも共通してリスクが上がっていたが、先行研究では死亡リスクは 8000 歩を基準とすると 4000 歩は約 2 倍上がったのに対して、本研究では 6000~8000 歩を基準として、4000 歩未満の生活習慣病のリスクが 1.3 倍に上がったため、リスクの上がり方には違いが生じた。

3.7 考察

歩数は BMI、年齢と同様に生活習慣病をリスクを上げる要因となることが判明した。喫煙の項目に有意差が見られなかったのは、過去の喫煙歴を考慮せず、現在の喫煙の有無のみで分析を行ってしまったことが原因の 1

つとして挙げられる。

3年以内の短期間で、4000歩未満の群は生活習慣病全体のリスクを1.3倍に上げてしまうが、4000歩以上の4つの群では有意差が見られなかったことから、歩数が極端に少ない場合は短期間で、健康に悪影響を及ぼしてしまうと考えられる。本研究では3年以内の罹患の影響を見たが、先行研究の様に長期的に見たら、影響を及ぼすことが示唆される。

第 4 章

歩数と医療費の関係

4.1 データ概要

本研究では、あるヘルスケアサービス事業者から法律に従って、適切に措置され、提供されたヘルスケアデータの匿名加工情報 [5] を使用する。表 4.1 に使用した匿名加工データの統計量を示す。

4.2 データの前処理

2016 年に健康診断結果とウェアラブルデバイスの歩数の記録が 330 日以上ある 40~59 歳の男性 2159 人と、2016 年に健康診断の結果がある 40~59 歳の男性 39,724 人を対象とする。健康診断データから、BMI、現在の喫煙の有無、1 日 1 時間の歩行の有無、1 回 30 分以上の軽く汗をかく運動を週 2 回以上、1 年以上しているかの有無の項目を取り出す。レセプトデータから、翌年 2017 年の医療費を計算する。

4.3 分析手法

2016 年にウェアラブルデバイスに歩数の記録がある 2159 人と健康診断の結果のみがある 39,724 人を次の 2 つの分析を行う。

(1) 歩数の記録がある 2159 人では、目的変数 y を翌年 2017 年の医療費が 80,000 円以上 (約上位 50%), 170,000 円以上 (約上位 25% 以上), 500,000 円以上 (約上位 5% 以上) になるそれぞれの確率とし、説明変数を歩数群 x_1 (6000~7999 歩を基準として、相対的に比較する、年齢 x_2 , BMI x_3 (18.5 以上かつ 25 未満の普通体重を基準として、低体重 BMI 18.5 未満, 普通 BMI 18.5 以上かつ 25 未満, 肥満 BMI 25 以上), 喫煙の有無 x_4 とし

表 4.1 匿名加工データの統計量

データ名	医療費データ	歩数データ	健康診断データ
属性数	10	4	49
期間	2004-2018	2014-2018	2008-2018
レコード数	19,053,822	24,936,120	964,636
対象者	288,623	58,655	198,740

表 4.2 歩数についてのロジスティック回帰分析

医療費	80,000 円以上	170,000 円以上	500,000 円以上			
N 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	0.98	0.954
10000 歩以上	0.99	0.971	1.15	0.474	1.11	0.787
年齢	1.07	0.000	1.08	0.000	1.07	0.000
低体重	0.51	0.167	0.69	0.478	2.25	0.287
肥満	1.60	0.000	1.95	0.000	1.81	0.002
喫煙	0.84	0.105	0.93	0.558	1.22	0.368

たロジスティックモデル

$$y = \frac{1}{1 + e^{-(a_1 x_1 + a_2 x_2 + a_3 x_3 + a_4 x_4)}} \quad (4.1)$$

を用いて、オッズ比 Odds Ratio(OR) を求める。

(2) 健康診断結果のみがある 39,724 人を対象に、(1) のモデルから、歩数 x_5 を 1 日 1 時間以上の歩行の有無に変更し、新たに週 2 回以上の運動の有無 x_6 を追加したモデル 2 を用いて、OR を求める。

$$y = \frac{1}{1 + e^{-(a_2 x_2 + a_3 x_3 + a_4 x_4 + a_5 x_5 + a_6 x_6)}} \quad (4.2)$$

(3) 歩数と 1 カ月の医療費の関係について、単回帰分析を行う。

4.4 歩数と医療費のロジスティック回帰

歩数の記録がある 2159 人のロジスティック回帰分析の結果を表 4.2 に示す。表 4.2 から、ロジスティック回帰分析では、すべての医療費の区切りで、歩数と 1 年後の医療費には統計的に有意な差がなく (p-value が 0.05 以上)、歩いた歩数が多いほど、医療費が下げているわけではなかった。年齢と肥満のみは有意な説明変数であり、年齢が高かかったり、肥満であると、医療費が上がる傾向があった。

4.5 歩行時間と医療費のロジスティック回帰分析

健康診断結果のみがある 39,724 人のロジスティック回帰分析の結果を表 4.3 に示す。表 4.3 から、1 日 1 時間以上の歩行をしている人は、すべての医療費の区切りで、OR が 1 未満になっていること、医療費が高くなるにつれ、OR が下がっていることが示されている。すなわち、1 時間以上の歩行をしていない人に比べ、医療費が全体的にかからない上、医療費が高額になる確率も低い。年齢が高く、肥満であると、医療費が上がる傾向があった。喫煙者、週 2 回の運動をしている人は医療費が下がる傾向もあった。

表 4.3 歩行時間についてのロジスティック回帰

医療費 N 人数 (割合)	80,000 円以上 18,569(46.7%)	170,000 円以上 10,469(26.4%)	500,000 円以上 2554(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
年齢	1.06	0.000	1.07	0.000	1.07	0.005
低体重	0.88	0.685	1.00	0.985	1.18	0.230
肥満	1.65	0.000	1.83	0.000	1.90	0.000
喫煙	0.83	0.000	0.88	0.000	0.95	0.249
週 2 回の運動 (30 分以上)	0.83	0.000	0.93	0.025	0.85	0.005

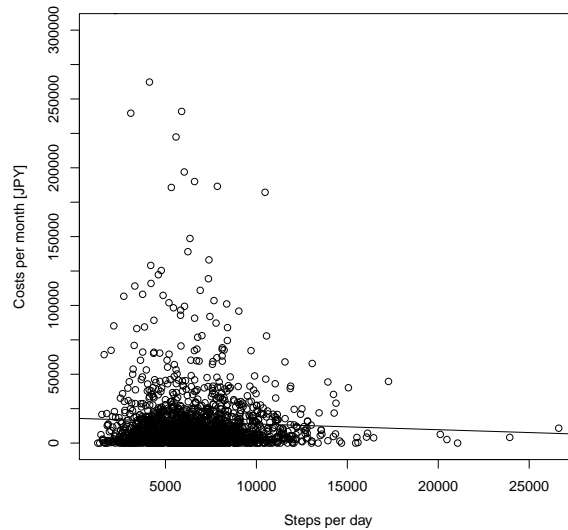


図 4.1 1 日の平均歩数と 1 カ月の医療費の散布図

4.6 単回帰分析

1 日の平均歩数 x と 1 カ月の医療費の平均 y についての回帰モデルは $y = -0.4146x + 18004$ である。歩数 (X 軸) と医療費 (Y 軸) の散布図, モデルの回帰直線を図 4.1 に示す。モデル (3) の回帰式は先行研究の $y = -1.3408x + 28205$ と, 傾きが異なっていた。本研究と先行研究の回帰結果から, 1000 歩あたりで削減できる医療費を整理した結果を表 4.4 に示す。先行研究の結果では, 1000 歩あたり, 1340.8 円医療費が削減できるのに対し, 本研究では 414.6 円削減であり, 効果が小さい。

表 4.4 1000 歩あたりで削減できる医療費

	対象者	削減できる医療費 (1000 歩あたり)
本研究	40-59 歳の男性	414.6 円
Tuji[4]	40-79 歳の男女	1340.8 円

表 4.5 歩数ごとの 1 日 1 時間以上の歩行の有無 (問診表) の回答

歩数	はい	いいえ	無回答	割合 (はい/回答あり)%
4000 歩未満	42	170	54	24.1 %
4000-5999 歩	149	402	128	27.0 %
6000-7999 歩	217	346	111	38.5 %
8000-9999 歩	163	141	59	53.6 %
10000 歩以上	92	53	33	63.4 %

表 4.6 [3] との医療費の比較

	歩数	人数	1 月あたりの医療費	比率
本研究	10000 歩以上	58	13,121	1.30
	8000-9999 歩	124	11,359	1.13
	6000-7999 歩	218	10,061	1.00
	4000-5999 歩	218	12,325	1.22
	4000 歩未満	77	10,208	1.01
Kuriyama[3]	1 日の歩行時間	人数	1 月あたりの医療費	比率
	1 時間以上	4772	20,376	1.00
	1 時間未満	4986	21,877	1.10

4.7 歩数と歩行時間の関係, 先行研究との比較

表 4.5 に歩数の記録がある 2159 人のうち, 1 日 1 時間以上の歩行しているか, 歩数ごとの集計を示す. 歩数が多いほど, 1 日 1 時間以上の歩行をしている. ロジスティック回帰分析の結果から, 歩行時間が 1 時間以上だと医療費が下がる傾向にあった. しかし, 歩数が多いほど, 歩行時間が 1 時間以上の割合が多いにもかかわらず, 医療費が下がるわけではなかった.

表 4.6 に, 先行研究 [3] と本研究の比較を示す. [3] では, 喫煙の有無 (過去の喫煙歴を含む), BMI は 25 以上の肥満, 22 以上かつ 25 未満の普通を比較, 歩行時間は 1 時間未満で運動不足としたリスクの組み合わせで, 医療費の平均を求めている. これと比較するために, BMI の項目はまったく同じ条件, 喫煙は現在の喫煙の有無, 歩数によって分け, 6000~7999 歩を基準として, 1 カ月の平均医療費の比率を計算した. また, 表 4.6 ではすべてのグループが非喫煙者, BMI が普通体重である. 表 4.6 の非喫煙者かつ普通体重の人たちにおける歩数との 1 カ月の医療費を図 4.3 に示す. [3] では, 1 日 1 時間以上歩く人たちと比較して, 1 日 1 時間未満の人たちは医療費が 1.1 倍高くなっている. 一方, 本研究の歩数では, 6000~7999 歩の医療費が 1 番安く, 10000 歩以上が 1 番高く, 医療費の比率が 1.3 倍という不規則な振る舞いであった.

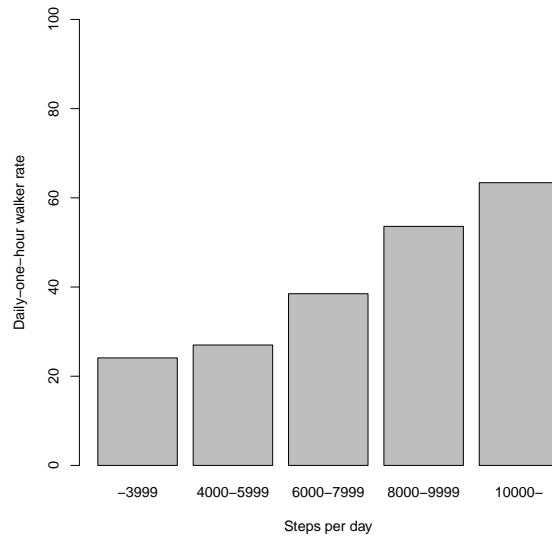


図 4.2 1日の平均歩数と歩行時間1時間以上の割合

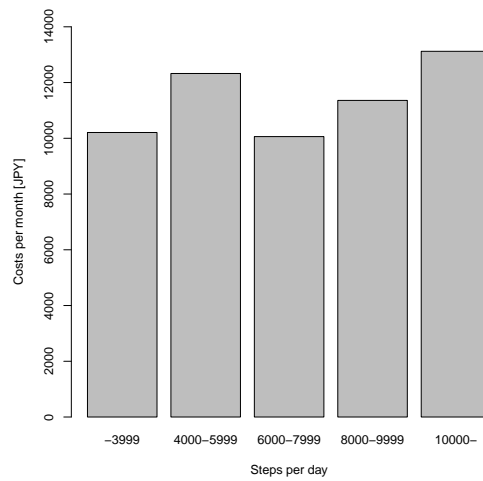


図 4.3 非喫煙者、普通体重の人々における1日の平均歩数と1カ月の医療費

4.8 考察

歩行時間が長いと翌年の医療費は下がる傾向にあり、医療費が高額になる確率も低いことがわかった。しかし、歩数が多いほど医療費が下がる傾向があるわけではなかった。歩数が多い人は、1時間以上歩いている割合が多くなるにもかかわらず、医療費を下げるわけではなかったため、歩行した距離よりも、歩行による運動の時間の方が医療費を抑える上では重要であると示唆される。

表 4.3 で、喫煙者だと、医療費が下がるという結果になったのは、先行研究 [3] では、過去に喫煙歴がある人も

喫煙者に含んで分析を行っていたが、本研究では、過去の喫煙歴は考慮せず、現在の喫煙の有無についてのみで分析を行ったことが原因のひとつとして挙げられる。

表 4.4 において、先行研究 [4] と本研究で削減できる医療費に差があったのは、対象とする年齢、性別に違いがあったことが大きな原因であると思われる。先行研究は 40～79 歳と年齢の幅が広く、年齢が高いほど医療費がかかり、歩数も少なくなる傾向があるため、純粋な歩数の影響だけでなく、年齢が交絡因子となり、その影響を大きく反映させてしまっている可能性がある。それに比べ、本研究では 40～59 歳と年齢の幅を絞り、性別も男性のみを対象としたため、40～59 歳の男性において、先行研究に比べ、信憑性の高い歩数と医療費の関係を示せていると考えられる。しかし、他の年齢、性別では歩数と医療費の関係を示せていないのが課題である。

表 4.5 で、歩数が多いほど、歩行時間が 1 時間以上の割合が多くなっているのにもかかわらず、医療費が下がるわけではなかった。このようになった原因の 1 つとして、歩行速度が医療費に関係している可能性があることが挙げられる。表 4.5 で、10,000 歩以上歩いている人で、1 時間以上歩いている割合は 63.4% だが、残りの 36.6% は 1 時間以上歩いておらず、人によって歩行速度にある程度の差があることが推測される。このことから、長距離をある程度の時間をかけて歩くのは健康的かもしれないが、長距離を短時間で速く歩くことはかえって体に悪影響を与えてしまう可能性があることが示唆される。

本研究で医療費を分析した期間は 1 年のみであり、長期間の調査はできていないのが課題である。歩数が少ない (4000 歩未満) 場合には、生活習慣病のリスクを上げてしまうことから、長期間の医療費との関係を調べると、歩数不足が、医療費を高くさせる傾向になる可能性がある。しかし、歩数が多いと生活習慣病のリスクが下がるわけではなかったため、長期的に見ても、歩数が多いほど医療費が下がるとは言い切れない。歩数が 6000～7999 歩のグループは図 4.4 において、医療費が最も低く、生活習慣病の罹患率が最も低かったことから、1 時間以上の歩行で 6000～7999 歩ほど歩くことが医療費を抑える上では適切であると考えられる。

第 5 章

おわりに

本研究では歩数と生活習慣病の罹患リスク, 医療費について, 匿名加工情報を用いて調査を行った。歩数が 4000 歩未満であると, 3 年以内の生活習慣病の罹患リスクを上げてしまうことがわかった。ただし, それ以上の歩数では罹患リスクがほとんど変わらず, 歩行による健康の向上の明確な証拠は得られなかった。医療費に関しては, 歩行時間が 1 時間以上だと医療費が下がる傾向にあったが, 歩数が多いほど医療費が下がるという傾向はなかった。本研究では, 歩数の計測期間は長い, データの都合上, 長い期間で病気を追跡できていないことが今後の課題である。

謝辞

本研究を行うにあたり，多くの方より御指導いただきました．特に明治大学総合数理学部先端メディアサイエンス学科，菊池浩明教授に深く感謝申し上げます．また，メンターとしてご指導いただいた伊藤聡志さん，研究室の皆様に深く感謝の意を表するとともに，謝辞とさせていただきます．

参考文献

- [1] 生活習慣病予防協会, <http://www.seikatsusyukanbyo.com/prevention/about.php>, 2020年12月参照
- [2] Pedro F. Saint-Maurice, et al. “Association of Daily Step Count and Step Intensity With Mortality Among US Adults”, JAMA, 323(12) pp.1151-1160, 2020.
- [3] Kuriyama S, Hozawa A, Ohmori K, et al, “Joint impact of health risks on health care charges: 7-year follow-up of National Health Insurance beneficiaries in Japan (the Ohsaki Study)”, Prev Med. 39, 1194-1199, 2004.
- [4] 辻一郎, 健康長寿社会を実現する—「2025年問題」と新しい公衆衛生戦略の展望—, 大修館書店, 2015.
- [5] 伊藤 聡志, 池上 和輝, 菊池 浩明, 健康診断データとレセプトデータの匿名加工情報を用いた疾病リスク分析, 情報処理学会論文誌, Vol.62, No.9, pp.1560-1574, 2021.
- [6] 全国健康保険協会, “生活習慣関連 10 疾患の動向に関する調査” (https://www.kenporen.com/toukei_data/pdf_chosa_r01_06_01.pdf, 2020年9月参照)