

医学論文における臨床参加患者群データである“Patient characteristics”のl-多様性による匿名性の評価

北村 憲太^{1,a)} IRVAN MHD¹ 山口 利恵¹

概要：薬の治療効果などの医療の効果を確認するには臨床研究に参加した患者群における人種の比率や併存疾患の割合や年齢分布などの評価が必要である。そのため、医学論文では“Patient characteristics”と呼ばれる、臨床研究に参加した患者群の人種の比率や年齢分布などの統計情報が記載される。しかし、Patient characteristics 内の特定疾患の比率が高い場合などには、臨床試験に参加した人の疾患が推測されてしまうというプライバシー侵害のリスクが生じうる。匿名加工情報の分野では、公開された情報内の属性値の比率の偏りによるプライバシー侵害を行う攻撃は同種攻撃や背景知識攻撃として知られており、このような攻撃への脆弱さを測る指標としてl-多様性が提案されている。本研究では具体的な医学論文の中の Patient Characteristics の匿名性をl-多様性を用いて評価した。

キーワード：プライバシー, l-多様性, 医療

Anonymity evaluation of “Patient characteristics (feature data of clinical trial participating patients in medical papers)” by l-diversity

Abstract: To confirm the effectiveness of medical treatment, such as the therapeutic effects of the drug, it is necessary to evaluate the racial proportion, the proportion of complications, and the age distribution of the patients who participated in the clinical study. For this reason, medical papers include statistical information called “Patient characteristics”, which includes the racial proportion and age distribution of the patient group participating in the clinical study. However, if the ratio of specific diseases in the patient characteristics is high, there is a risk of privacy violation because the diseases of the participants in the clinical study may be inferred. In the field of anonymized information, privacy-violating attack based on the bias of the ratio of attribute values in the disclosed information is known as homogeneous attack or background knowledge attack, and l-diversity is proposed as a measure of vulnerability to such attacks. In this study, we evaluate the anonymity of Patient Characteristics in actual medical papers by l-diversity.

Keywords: Privacy, l-diversity, Healthcare

1. はじめに

1.1 背景

昨今、プライバシーを保護した上で公的機関・企業・医療データを公開するプライバシー保護データ公開 (PPDP, Privacy-Preserving Data Publishing) と呼ばれている分野が盛んに研究されている。この分野の一般的な問題として、データの有用性を保持しながらどの程度プライバシーを保

護した情報公開が可能かという問題がある [1]。有用性を保ったまま匿名化する手法としてはk-匿名化 [2] や差分プライバシー [3] などの手法が提案されている。また、l-多様性 [4] やt-近接性 [5] といった指標を用いることで公開情報時の匿名性と有用性のバランスを定量的に評価することが提案されている。医療においても、医療データの公表における匿名化と有用性の関係はトレードオフの関係が指摘されており、どの程度の匿名加工でどの程度の有効性が得られるかは定かではなく、k-匿名化などの匿名加工技術の適用が議論されている [6], [7]。

医学論文で公開される医療データ形式には Patient char-

¹ 東京大学大学院情報理工学系研究科
Graduate School of Information Science and Technology,
The University of Tokyo

^{a)} kitamura@yamagata.ic.i.u-tokyo.ac.jp

表 1 Patient characteristics. Covid-19 に関する臨床研究論文より適宜修正して掲載 [8].

Table 1 Patient characteristics. From clinical research paper on Covid-19, modified as appropriate [8].

Characteristics	Convalescent (N = 228)	Placebo (N = 105)
Median age (IQR) — yr	62.5 (53–72.5)	62 (49–71)
Age category — no. (%)		
<65 yr	126 (55.3)	54 (51.4)
≥65 to <80 yr	75 (32.9)	43 (41)
≥80 yr	27 (11.8)	8 (7.6)
Female sex — no. (%)	67 (29.4)	41 (39.0)
Median time to onset of symptoms (IQR) — days	8 (5–10)	8 (5–10)
Coexisting conditions — no. (%)		
No other conditions	80 (35.1)	37 (35.2)
Body-mass index >30	104 (45.6)	52 (49.5)
Hypertension	111 (48.7)	48 (45.7)
Diabetes	40 (17.5)	21 (20)
Chronic obstructive pulmonary disease	23 (10.1)	2 (1.9)
Asthma	9 (3.9)	5 (4.8)
Chronic renal failure	10 (4.4)	4 (3.8)
Hematologic cancer	4 (1.8)	3 (2.9)
Solid tumors	23 (10.1)	11 (10.5)
Current tobacco use	6 (2.6)	6 (5.7)
Previous tobacco use	101 (44.3)	37 (35.2)
Congestive heart failure	8 (3.5)	3 (2.9)
Thromboembolic disease	5 (2.2)	2 (1.9)
Previous medications used — no. (%)		
ACEI or ARB 2	69 (30.3)	32 (30.5)
Frequent or recent use of NSAID	37 (16.2)	13 (12.4)
Anticoagulation	14 (6.1)	6 (5.7)
Corticosteroids	7 (3.1)	2 (1.9)
Immunosuppressants	6 (2.6)	3 (2.9)
Statins	61 (26.8)	21 (20)
Laboratory values		
Median total SARS-CoV-2 antibody titer (IQR)	1/50 (0–1:800)	1:50 (0–1:1600)
Negative total SARS-CoV-2 antibody titer — no./total no. (%)	65/145 (44.8)	34/70 (48.6)
Median d-dimer level (IQR) — ng/ml	697 (470–1150)	797 (550–1224)
Median ferritin level (IQR) — ng/ml	939 (441–1634)	645 (362–1180)
Severity inclusion criteria — no. (%)		
Oxygen saturation <93% at FiO2 0.21	224 (98.2)	100 (95.2)
mSOFA or SOFA ≥2	32 (14)	17 (16.2)
Hospitalization area at enrollment — no. (%)		
Emergency department	11 (4.8)	3 (2.9)
General ward	150 (65.8)	77 (73.3)
Critical care unit	67 (29.4)	25 (23.8)
Use of oxygen supplementation devices (n=299) — no. (%)		
Low-flow nasal cannula	146 (64.0)	70 (66.7)
Venturi or nonrebreather mask	49 (21.5)	16 (15.2)
High-flow nasal cannula	11 (4.8)	7 (6.7)
Noninvasive ventilatory support	0	0
Treatments during trial — no. (%)		
Supplemental oxygen	206 (90.4)	93 (88.6)
Glucocorticoids	209 (91.7)	101 (96.2)
Lopinavir–ritonavir	7 (3.1)	3 (2.9)
Tocilizumab	6 (2.6)	8 (7.6)
Ivermectin	4 (1.8)	1 (1)
Hydroxychloroquine	1 (0.4)	0

acteristics と呼ばれる 表 1 のようなものがある [8]. 薬の治療効果は患者の併存疾患や年齢によって左右されるため、臨床試験に参加した患者群の併存疾患の割合や年齢分布を確認できるように Patient characteristics は公開される。Patient characteristics はこのような有用性のあるデータであるが、持病などのセンシティブな情報について記載されるためプライバシーの侵害のリスクとなりうる。そのため、Patient characteristics においてどの程度の匿名性が保たれるかの指標が必要となる。情報公開における匿名性指標の評価では、匿名性の侵害を試みる攻撃者を想定し、そのような攻撃者の再構成や再識別などの攻撃をどの程度防

表 2 脆弱性のある Patient characteristics の例。

Table 2 Examples of vulnerable Patient characteristics.

	治療群 (N = 228)	プラセボ群 (N = 105)	治療群+プラセボ群 (N = 333)
例 ① 糖尿病	0	1	1
例 ② 高血圧	225	100	325
例 ③ 喫煙	3	100	103

ぐことが可能かを匿名性の指標にすることが提案されている [9]. そのため、Patient characteristics においても、匿名性の侵害を試みる攻撃を想定し、そのような攻撃をどの程度防ぐことが可能かを検討することが必要となる。

1.2 目的

本論文の目的は、プラセボ群と治療群の 2 群に分かれるような Randomized controlled trial (RCT) と呼ばれる臨床研究における Patient characteristics に対する 3 つの再識別攻撃「Patient Detect Placebo (PDP) 攻撃」、「Patient Family Detect on Overall Category (PFDOC) 攻撃」、「Patient Family Detect on Placebo and Treatment Category (PFDPTC) 攻撃」を想定し、それぞれの攻撃に対する定量的匿名性指標を提案する点にある。なお、本論文における「再識別」とは「公表データセットにおける匿名化された個人の記録を個人情報まで辿ること」と考える [9].

PDP 攻撃は、患者自身が攻撃者であり、「患者が属するのは治療群かプラセボ群か」を再識別する攻撃である。PDP 攻撃の例として、表 2 の例 ①の糖尿病のカテゴリーでは、糖尿病患者が Patient characteristics に 1 人であるため、臨床試験に参加した糖尿病の患者は自分がプラセボ群だと再識別可能であり、後述するプラセボ効果がなくなるという健康被害を受けうる。PFDOC 攻撃は、患者家族が攻撃者であり、「患者があるカテゴリーに属するかどうか」を再識別する攻撃である。PFDPTC 攻撃の例として、表 2 の例 ②の高血圧のカテゴリーでは、治療群とプラセボ群の合計人数 (333 人) と高血圧群の合計人数 (325 人) が近いので、臨床試験参加者は高い確率で高血圧だと判明する。PFDPTC 攻撃は、患者家族が攻撃者であり、「患者が治療群かプラセボ群かを推定した上で、患者があるカテゴリーに属するかどうか」を再識別する攻撃である。PFDPTC 攻撃の例として、表 2 の例 ③の喫煙のカテゴリーでは、ある患者がプラセボ群だと知られた場合に、プラセボ群 (105 人) とプラセボ群の喫煙群 (100 人) の人数が近いので、その参加者は喫煙の可能性が高いと判明する。

上記の 3 つの攻撃はいずれも攻撃者が持つ予備情報と Patient characteristics において属する患者の多様性が低いカテゴリーをリンクさせることで患者情報を特定してい

る。このような攻撃者の予備情報と公開情報の中の多様性が低いカテゴリーをリンクさせる攻撃に対する耐性の指標として、k-匿名化 [2] においてはl-多様性 [4] が提案されている。本論文では、上記3つの攻撃に対して、l-多様性の考え方を応用した定量的な匿名性指標をそれぞれの攻撃に対し提案する。また、この匿名性指標を用いて具体的な論文における Patient characteristics の匿名性を検討する。

1.3 本論文の構成

本論文の構成は次のとおりである。第2章では関連研究としてk-匿名化とその評価指標であるl-多様性、プラセボ効果と二重盲検を紹介する。第3章では準備として、本研究における再識別攻撃と提案する定量的匿名性指標について説明する。第4章では再識別を行う具体的な Patient characteristics [8] の説明と結果の検討方法を説明する。第5章では結果をまとめる。第6章では考察を行う。第7章では本研究のまとめと今後の展望を述べる。

2. 関連研究

k-匿名化とその評価指標であるl-多様性、プラセボ効果と二重盲検について述べる。

2.1 k-匿名化 [2] とl-多様性 [4]

k-匿名化とは、データセットの中で準識別子が同一属性の人がk人以上にすることで個人を特定する確率が $1/k$ 以下となり匿名性を保てるという匿名化手法である。

k-匿名化においては同種攻撃と呼ばれる再識別手法に対する脆弱性が指摘されており、同一属性の人がk人以上であっても、例えば全員が糖尿病である場合にその属性に含まれている人は全員が糖尿病だとわかってしまう。また、k-匿名化においては背景知識攻撃と呼ばれる再識別手法に対する脆弱性も指摘されており、例えばある人種の一般的な集団の特性として心筋梗塞の確率が非常に低いとわかっている場合に、心筋梗塞が含まれる属性にその人種の人が含まれている場合、心筋梗塞以外の疾患であることが推定される。l-多様性はこのような攻撃に耐性を持つ匿名化手法であり、同一属性の人の病名などのセンシティブ情報の種類を複数にすることによって、同種攻撃や背景知識攻撃による再識別を難しくする。l-多様性の定量的評価指標としては、単純l-多様性、エントロピーl-多様性、再起的(c, l)-多様性などが提案されている。

エントロピーl-多様性は、以下のように定義される。

定義 1. (エントロピー l-多様性): q^* -ブロックは準識別子が同一の値の組み合わせを持つタプルであり、データテーブルは q^* -ブロックの集団であり、 q^* -ブロックの中のセンシティブ属性を s とする。全ての q^* -ブロックに対して下記の式が満たされる時、データテーブルはエントロピー l-多様性を持つとする。

$$-\sum_{s \in S} p(q^*, s) \log(p(q^*, s)) \geq \log(l)$$

(q^* -ブロックの中の注目属性の値 s の母集団を S とし、 $p(q^*, s)$ は q^* -ブロックのセンシティブ属性 s の割合)

2.2 プラセボ効果 [10] と二重盲検 [11]

プラセボ効果とは、薬理学的な効果が期待できないような偽薬であっても臨床的な効果が得られることである。

臨床試験において、プラセボ群と治療群に患者をランダム割賦する際に、患者も医者もプラセボ群か治療群かわからないようにする二重盲検という臨床研究の手法がとられる。治験内容によって各患者がプラセボ群か治療群かの通知が行われるかどうかや通知時期は異なる。

3. 準備

以下では、本研究における再識別攻撃と提案する定量的匿名性指標について説明する。本研究では Patient characteristics の匿名性を評価するにあたり Patient characteristics を再識別する「PDP 攻撃・PFDOC 攻撃・PFDPTC 攻撃」という3つのタイプの再識別攻撃を想定する。なお、Patient characteristics には治療群とプラセボ群が無い形式や併存疾患群や併存疾患に対する服薬群が無い形式もある [12] が、本研究ではプラセボ群と非プラセボ群が存在するタイプのもので、併存疾患群や併存疾患に対する服薬群が存在する形式を研究の対象とする。

PDP 攻撃は、臨床試験に参加した患者が自身の参加した臨床試験の論文で Patient characteristics を見る場合に、患者自身が治療群なのかプラセボ群なのかを再識別するという攻撃である。関連研究で紹介したように、プラセボ群にも医学的効果がありうるため、もし患者自身がプラセボ群であることを知った場合にはプラセボ効果が失われてしまい健康被害を受ける可能性がある。PDP 攻撃では攻撃者は患者自身であるため、プラセボ群であるかどうか以外の Patient characteristics の全カテゴリーに対する患者自身の情報を補助情報として持つ。

攻撃者はまず、Patient characteristics の全カテゴリーにおいて以下に定義される Patient Detect Placebo エントロピー (PDP エントロピー) を計算する。

定義 2. (Patient Detect Placebo エントロピー (PDP エントロピー)): Patient characteristics のカテゴリーの1つにおいて、プラセボ群が A 人、非プラセボ群が B 人の時に、PDP エントロピーを以下のように計算する。

$$-(A/(A+B)) \log(A/(A+B)) \\ -(B/(A+B)) \log(B/(A+B))$$

なお、 $A=0$ もしくは $B=0$ の場合、PDP エントロピーは 0 と考えるものとする。

PDP エントロピーが低いカテゴリーでは、攻撃者である

患者自身が当該カテゴリーに当てはまる場合に、治療群なのかプラセボ群なのかを判別してしまう可能性が高くなる。そのようなリスクに対し、以下のように定義される安全性の定量的指標 Patient Detect Placebo エントロピー- l -多様性 (PDP エントロピー- l -多様性) を考える。

定義 3. (*Patient Detect Placebo* エントロピー- l -多様性 (PDP エントロピー- l -多様性)): *Patient characteristics* のあるカテゴリー群に対して PDP エントロピー- l -多様性が保たれているとは、カテゴリー群の全ての PDP エントロピーに対して、「PDP エントロピー $\geq \log(l)$ 」となることである。

PFDOC 攻撃は、臨床試験に参加した患者の家族が臨床試験の論文で Patient characteristics を見る場合に、患者の並存疾患や並存疾患に対する服薬の情報を再識別するという攻撃である。患者家族は患者の治療対象に関する説明を医療者から受けたり患者から報告されることがあるため、臨床試験の対象疾患に対する治療内容や臨床試験に参加しているといった情報は補助情報として持つ。一方、患者の併存疾患である性感染症などは患者本人が家族にカミングアウトしていない場合など、患者が患者家族に知られていない状況がある。そのため、患者家族は患者の並存疾患や併存疾患に対する服薬状況については知らないものとする。PFDOC 攻撃では攻撃者は患者家族であるため、再識別対象の患者がプラセボ群であるかどうか、患者の併存疾患、併存疾患に対する服薬状況以外の Patient characteristics における全カテゴリーの患者の情報を補助情報として持つ。

攻撃者はまず、Patient characteristics の併存疾患、併存疾患に対する服薬のカテゴリーにおいて以下に定義される Patient Family Detect on Overall Category エントロピー (PFDOC エントロピー) を計算する。

定義 4. (*Patient Family Detect on Overall Category* エントロピー (PFDOC エントロピー)): *Patient characteristics* における治療群全体の人数が N_a 人、プラセボ群全体の人数が N_b 人であり、治療群とプラセボ群の全体の人数を合計した人数が $N_c (= N_a + N_b)$ 人として、*Patient characteristics* の対象のカテゴリーにおいて、プラセボ群が A 人、非プラセボ群が B 人の場合に、PFDOC エントロピーを以下のように計算する。

$$-((A+B)/N_c)\log((A+B)/N_c) \\ -((N_c-A-B)/N_c)\log((N_c-A-B)/N_c)$$

なお、 $A+B=0$ の場合、PFDOC エントロピーは 0 と考えるものとする。

PFDOC エントロピーが低いカテゴリーは、そのカテゴリーの確率 $(A+B)/N_c$ が高い場合に、臨床試験に参加した全員がそのカテゴリーに当てはまる確率が高くなり、そのカテゴリーの確率が低い場合には、臨床試験に参加した全員がそのカテゴリーに当てはまる確率が低くなる。そ

のため、再識別対象の患者も同様にそのカテゴリーに当てはまる可能性が低いあるいは高いことが示唆される。そのようなリスクに対し、以下のように定義される安全性の定量的指標 Patient Family Detect on Overall Category エントロピー- l -多様性 (PFDOC エントロピー- l -多様性) を考える。

定義 5. (*Patient Family Detect on Overall Category* エントロピー- l -多様性 (PFDOC エントロピー- l -多様性)): *Patient characteristics* のあるカテゴリー群に対して PFDOC エントロピー- l -多様性が保たれているとは、カテゴリー群の全ての PFDOC エントロピーに対して、「PFDOC エントロピー $\geq \log(l)$ 」となることである。

PFDPCTC 攻撃は、臨床試験に参加した患者の家族が臨床試験の論文で Patient characteristics を見る場合に、患者がプラセボ群であるのか治療群であるかを確信している条件のもとで、患者の並存疾患や並存疾患に対する服薬の情報を再識別するという攻撃である。患者家族は PFDOC 攻撃と同様に臨床試験の対象疾患に対する患者の治療内容や患者が臨床試験に参加しているといった情報は補助情報として持つ。そのため、そのような補助情報を持つカテゴリーに対し PDP 攻撃を行うことにより、プラセボ群であるか治療群であるかの確信を得ているとする。一方、患者家族は患者の並存疾患や併存疾患に対する服薬状況については知らないものとする。

攻撃者はまず、Patient characteristics の併存疾患、併存疾患に対する服薬のカテゴリーにおいて以下に定義される Patient Family Detect on Placebo and Treatment Category エントロピー (PFDPCTC エントロピー) を計算する。

定義 6. (*Patient Family Detect on Placebo and Treatment Category* エントロピー (PFDPCTC エントロピー)): *Patient characteristics* の対象のカテゴリーにおいて、プラセボ群が A 人、治療群が B 人として、患者がプラセボ群であると再識別者が確信している場合に、PFDPCTC エントロピーを以下のように計算する。

$$-(A/(A+B))\log(A/(A+B))$$

なお、 $A=0$ の時に PFDPCTC エントロピー = 0 とする。また、患者が治療群であると再識別者が確信している場合に、PFDPCTC エントロピーは、以下のように計算する。

$$-(B/(A+B))\log(B/(A+B))$$

なお、 $B=0$ の時に PFDPCTC エントロピー = 0 とする。

また、攻撃者は、PFDOC 攻撃で定義した PFDOC エントロピーも計算し、得られた PFDOC エントロピーと PFDPCTC エントロピーの差分の絶対値である Patient Family Detect on Placebo and Treatment Category 差分エントロピー (PFDPCTC 差分エントロピー) を計算する。

定義 7. (*Patient Family Detect on Placebo and Treatment Category* 差分エントロピー (*PFDPCTC* 差分エントロピー)): *PFDPCTC* 差分エントロピー = $|PFDPCTC$ エントロピー - *PFDOC* エントロピー |

PFDPCTC 差分エントロピーが低いカテゴリーに関しては、そのカテゴリーの全患者の割合 $(A + B) / (Na + Nb)$ と比べた時に、患者が属すると再識別者が確信しているプラセボ群あるいは治療群でのそのカテゴリーの全患者の割合 A / Na もしくは B / Nb との差が大きいことが示唆される。そのため、患者の併存疾患や併存疾患に対する服薬情報に対するプライバシーの侵害となりうる。そのようなリスクに対し、以下のように定義される安全性の定量的指 *Patient Family Detect on Placebo and Treatment Category* 差分エントロピー 1-多様性 (*PFDPCTC* 差分エントロピー 1-多様性) を考える。

定義 8. (*Patient Family Detect on Placebo and Treatment Category* 差分エントロピー 1-多様性 (*PFDPCTC* 差分エントロピー 1-多様性)): *Patient characteristics* のあるカテゴリー群に対して *PFDPCTC* 差分エントロピー 1-多様性が保たれているとは、カテゴリー群の全ての *PFDPCTC* 差分エントロピーに対して、「*PFDPCTC* 差分エントロピー $\leq \log(l)$ 」となることである。

4. 実験方法

以下では、本研究における再識別を行う具体的な *Patient characteristics* [8] の説明と結果の検討方法を述べる

4.1 マテリアル [8]

Covid-19 の重症肺炎に対する回復期血漿療法の治療効果に関する論文の *Patient characteristics* を本研究では検討する。なお、この *Patient characteristics* は、治療群：プラセボ群 = 2 : 1 の人数比となるように患者群がランダム割賦されている。本研究の再識別攻撃では、平均値などのバリュー変数のカテゴリーは採用せず、カテゴリカル変数のカテゴリーを解析対象とした。カテゴリカル変数でないために解析対象から外したカテゴリーは、{Median age, Median time to onset of symptoms, Laboratory values の Median total SARs-CoV-2 antibody tier, Laboratory values の Median D-dimer level, Laboratory values の Median ferritin level} である。なお、Laboratory values の {Negative total SARS-CoV-2 antibody titer — no./total no. (%) } のカテゴリーに関しては、「Sampling していないもの・Sampling した上で陰性であったもの・Sampling した上で陽性であったもの」の 3 種類が考えられるため、{No sampling, Sampling & Negative, Sampling & Positive} として解析対象とした。

4.2 再識別者モデルごとの再識別結果の検討方法

PDP 攻撃においては、ランダム割賦である 2:1 の場合に

は $PDP \text{ エントロピー} = - (1/3) \log(1/3) - 2/3 \log(2/3) = 0.918$ となる。*PDP* エントロピーが 0.918 より低いカテゴリーは平均的な場合より再識別のリスクが高いカテゴリーと考えられる。これを *Patient Detect Placebo* リスキーカテゴリー (*PDP* リスキーカテゴリー) と定義する。

定義 9. (*Patient Detect Placebo* リスキーカテゴリー (*PDP* リスキーカテゴリー)): 治療群とプラセボ群へのランダム割賦が $a : b$ の時、 $-(a / (a + b)) \log((a / (a + b))) - (b / (a + b)) \log((b / (a + b)))$ より *PDP* エントロピーが低いカテゴリーを *PDP* リスキーカテゴリーとする。

PDP リスキーカテゴリーに属するカテゴリー数の、全体のカテゴリー数に対する割合を非匿名性の指標として検討する。また、*PDP* エントロピー 1-多様性の 1 の値に対して、保全できるカテゴリー数の全体のカテゴリー数に対する割合も匿名性の評価指標とする。また、全カテゴリーに対して *PDP* エントロピー 1-多様性を保全する 1 の値も検討する。

PFDOC 攻撃においては、患者の併存疾患あるいは併存疾患に対する薬の服薬のあるカテゴリーの *PFDOC* エントロピーが全カテゴリーの *PFDOC* エントロピーの平均より低いものは、*Patient characteristics* の中での比較において再識別のリスクがあると考えられる。これを *Patient Family Detect on Overall Category* リスキーカテゴリー (*PFDOC* リスキーカテゴリー) と定義する。

定義 10. (*Patient Family Detect on Overall Category* リスキーカテゴリー (*PFDOC* リスキーカテゴリー)): *PFDOC* エントロピーが全カテゴリーの *PFDOC* エントロピーの平均より低いカテゴリーを *PFDOC* リスキーカテゴリーとする。

PFDOC リスキーカテゴリーに属するカテゴリー数の、全体のカテゴリー数に対する割合を非匿名性の指標として検討する。また、*PFDOC* エントロピー 1-多様性の 1 の値に対して、保全できるカテゴリー数の全体のカテゴリー数に対する割合も匿名性の評価指標とする。また、全カテゴリーに対して *PFDOC* エントロピー 1-多様性を保全する 1 の値も検討する。

PFDPCTC 攻撃においては、患者の併存疾患あるいは併存疾患に対する服薬におけるカテゴリーの *PFDPCTC* 差分エントロピーが全カテゴリーの *PFDPCTC* 差分エントロピーの平均値より高いものは、*Patient characteristics* の中での比較において再識別のリスクがある。これを *Patient Family Detect on Placebo and Treatment Category* 差分エントロピーリスキーカテゴリー (*PFDPCTC* 差分エントロピーリスキーカテゴリー) と定義する。

定義 11. (*Patient Family Detect on Placebo and Treatment Category* 差分エントロピーリスキーカテゴリー (*PFDPCTC* 差分エントロピーリスキーカテゴリー)): *PFDPCTC* 差分エントロピーが全カテゴリーの *PFDPCTC* 差分エントロピーの平均値より低いカテゴリーを *PFDPCTC*

差分 リスキーカテゴリーとする。

PFDPTC 差分 リスキーカテゴリーに属するカテゴリー数の、全体のカテゴリー数に対する割合を匿名性の指標として検討する。また、PFDPTC 差分エントロピー 1-多様性の 1 の値に対して、保全できるカテゴリー数の全体のカテゴリー数に対する割合も匿名性の評価指標とする。また、全カテゴリーに対して PFDPTC 差分エントロピー 1-多様性を保全する 1 の値も検討する。

5. 実験結果

PDP 攻撃に関して、図 1 に表 1 から作成した PDP エントロピーを示す。なお、点線は 2:1 割賦時の平均的な PDP エ

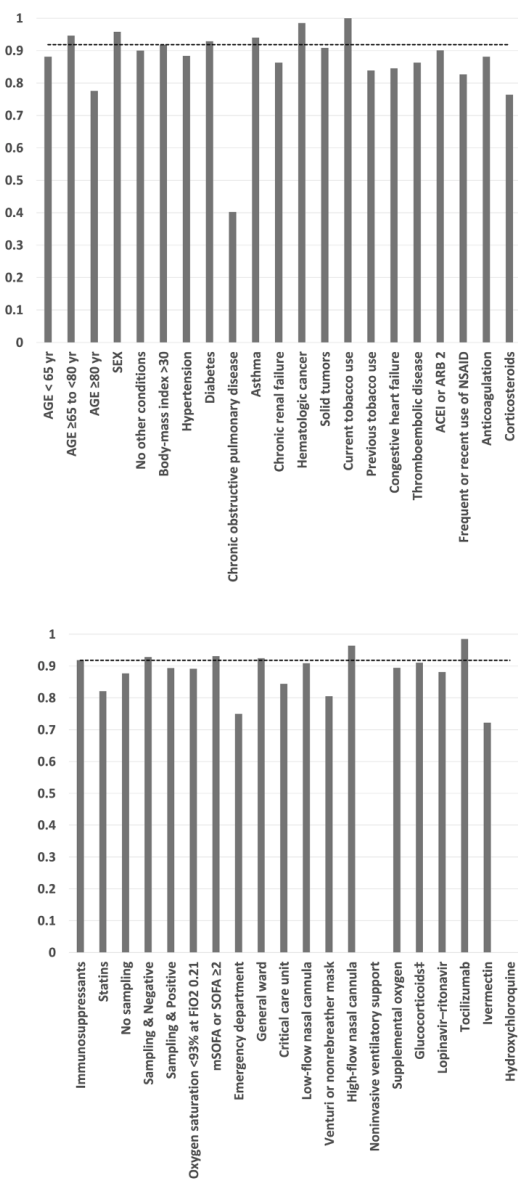


図 1 表 1 の Patient characteristics のカテゴリーごとの PDP エントロピー。点線は 2:1 割賦時の PDP エントロピー。

Fig. 1 PDP entropy for each category of patient characteristics in Table 1. The dotted line is the PDP entropy for the 2:1 allocation.

ントロピー (= 0.918) である。対象とした 41 カテゴリーの内 PDP リスキーカテゴリーは 28 カテゴリーであり 68%の割合であった。PDP エントロピー 1-多様性の 1 値ごとに保全されるカテゴリー数は図 2 に示す。特に低い値である 0.5 以下の 3 カテゴリーは、「Noninvasive ventilatory support, Hydroxychloroquine, Chronic obstructive pulmonary disease」であり、それぞれの PDP エントロピーは 0.0, 0.0, 0.4 であった。それぞれのカテゴリーの治療群：プラセボ群の人数比は、Noninvasive ventilatory support が 0:0, Hydroxychloroquine が 1:0, Chronic obstructive pulmonary disease が 23:2 の人数比であった。Noninvasive ventilatory support のカテゴリーは該当する患者はおらず、当該カテゴリーに属する患者がプラセボ群や治療群と判明するリスクは無い。しかし Hydroxychloroquine のカテゴリーに当てはまる患者は、治療群であることがわかってしまう。さらに、Chronic obstructive pulmonary disease のカテゴリーに当てはまる患者は、高い確率で治療群であることがわかってしまう。人数に関しては、特に下位 3 カテゴリーのものは、カテゴリー間の重複が無い場合 26 人、重複が最大の場合 25 人となり、全体の人数から見た割合は、7.8 - 8.1% となった。全カテゴリーを保全する PDP エントロピー 1-多様性の 1 の値は 1 = 1 であった。

PFDOC 攻撃に関して、図 3 に表 1 から作成した PFDOC エントロピーを示す。なお、点線は対象とした 19 カテゴリーの PFDOC エントロピーの平均値 (= 0.57) である。対象とした 19 カテゴリーの内 PFDOC リスキーカテゴリーは 11 カテゴリーであり 58%の割合であった。PFDOC エントロピー 1-多様性の 1 値ごとに保全されるカテゴリー数は図 4 に示す。特に低い値である 0.2 以下の 4 カテゴリーは、「Hematologic cancer, Thromboembolic disease, Corticosteroids, Immunosuppressants」であり、それぞれの人数は、Hematologic cancer 7 人, Thromboembolic disease 7 人, Corticosteroids 9 人, Immunosuppressants 9 人

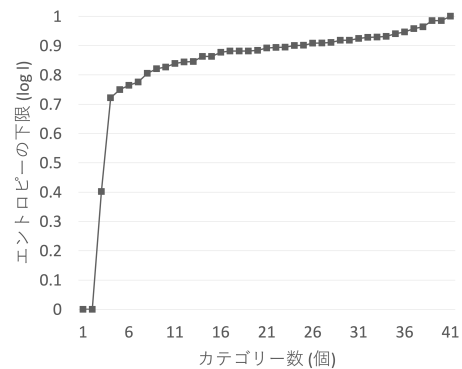


図 2 PDP エントロピー 1-多様性の下限 vs 匿名性が保全されるカテゴリー数。

Fig. 2 PDP entropy 1-diversity lower bound vs number of anonymity-preserving categories.

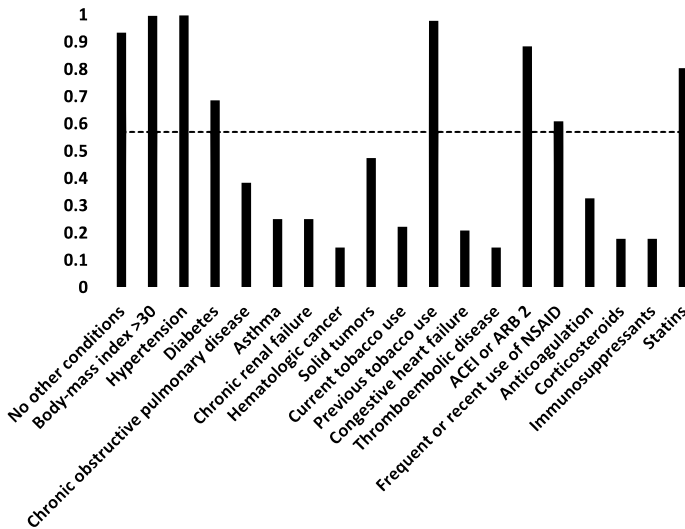


図 3 表 1 の Patient characteristics のカテゴリーごとの PFDOC エントロピー. 点線は対象とした 19 カテゴリーの PFDOC エントロピーの平均値の 0.57.

Fig. 3 PFDOC entropy for each category of Table 1 Patient characteristics. The dotted line (= 0.57) is the average of the PFDOC entropy of the 19 targeted categories.

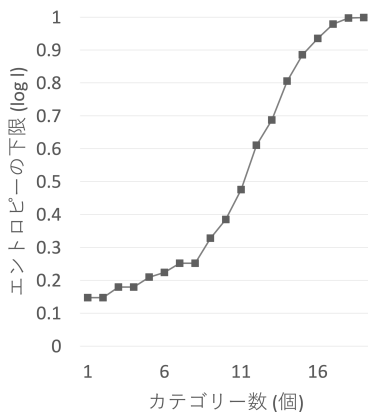


図 4 PFDOC エントロピー l-多様性の下限 vs 匿名性が保全されるカテゴリー数.

Fig. 4 PFDOC entropy l-diversity lower bound vs number of anonymity-preserving categories.

であった。これらはいずれも、「病名の確率が低い」カテゴリーのために PFDOC エントロピーが低くなったものであり、病名を確定できないカテゴリーであった。人数に関しては、0.2 以下のものは、カテゴリー間の重複が無い場合 32 人、重複が最大の場合 9 人となり、全体の人数から見た割合は、2.7 - 9.6% となった。全カテゴリーを保全する PFDOC エントロピー l-多様性の 1 の値は 1.107 であった。

PFDPCTC 攻撃に関して、図 5 に表 1 から作成した PFDPCTC 差分エントロピーを示す。点線は対象とした 19 カテゴリーの PFDPCTC 差分エントロピーの平均値 (= 0.029) である。対象とした 38 カテゴリーの内、PFDPCTC 差分 リスクカテゴリーは 12 カテゴリーであり 32% の割合であった。PFDPCTC 差分エントロピー l-多様性の 1 値ご

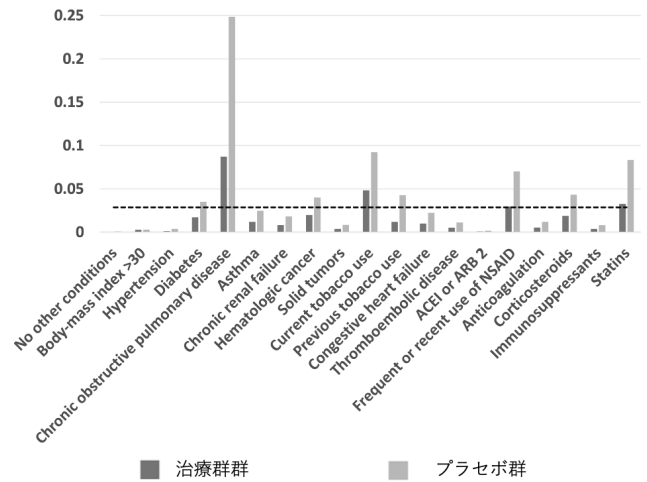


図 5 表 1 の Patient characteristics のカテゴリーごとの PFDPCTC 差分エントロピー. 点線は対象とした 19 カテゴリーの PFDPCTC 差分エントロピーの平均値の 0.029.

Fig. 5 PFDPCTC differential entropy for each category of Table 1 Patient characteristics. The dotted line (= 0.029) is the average PFDPCTC difference entropy of 19 categories.

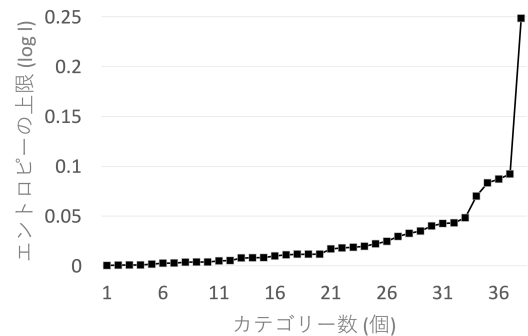


図 6 PFDPCTC 差分エントロピー l-多様性の上限 vs 匿名性が保全されるカテゴリー数

Fig. 6 PFDPCTC differential entropy l-diversity upper bound vs number of anonymity-preserving categories.

とに保全されるカテゴリー数は図 6 に示す。PFDPCTC 差分エントロピー l-多様性が特に高い値である 0.05 以上の 5 カテゴリーは、「Frequent or recent use of NSAID のプラセボ群, Statins のプラセボ群, Chronic obstructive pulmonary disease のプラセボ群, Chronic obstructive pulmonary disease の治療群, Chronic obstructive pulmonary disease のプラセボ群」であった。それぞれの「治療群とプラセボ群の合計人数」との人数比は、「15 人/105 人 vs 82 人/333 人 (14% vs 25%)」, 「21 人/105 人 vs 82 人/333 人 (20% vs 25%)」, 「2 人/105 人 vs 25 人/333 人 (2% vs 7.5%)」, 「6 人/105 人 vs 12 人/333 人 (5.7% vs 3.6%)」, 「23 人/228 人 vs 25 人/333 人 (10.1% vs 7.5%)」であった。人数に関しては、PFDPCTC 差分エントロピーが 0.05 以上のものは、カテゴリー間の重複が無い場合 67 人 (プラセボ群 44 人, 治療群 23 人), 重複が最大の場合 44 人 (プラセボ群 21 人,

治療群 23 人) となり, 全体の人数から見た割合は, 13.2 - 20.1%となった. 全カテゴリーを保全する PFDPTC 差分エントロピー 1-多様性の 1 の値は 1.188 であった.

6. 考察

本研究の対象論文 [8] にて, PDP エントロピー 1-多様性は $l = 1$ でしか匿名性を保てないため匿名性が保たれていないことが指摘された. 実際に, PDP エントロピーが特に低い Hydroxychloroquine は確実に, Chronic obstructive pulmonary disease は高い確率で当該カテゴリーの患者は治療群だと判明する. 治療群の場合はプラセボ効果がなくなることによる健康被害はないが, プラセボ群と判明するリスクもあった. PDP 攻撃に対して脆弱性のあるカテゴリーの指摘ができるため, PDP エントロピー 1-多様性による匿名性評価の機能が示されたと言える.

PFDOC エントロピー 1-多様性は l が 1.107 より小さければ匿名性を保てることが指摘された. PFDOC エントロピーが特に低い 0.2 以下のカテゴリーは, 「疾患ではない」という確率が高いものであり, 「疾患でない」という情報は当該論文では秘匿しにくいことが示された. PFDOC 攻撃に対して脆弱性のあるカテゴリーの指摘ができるため, PFDOC エントロピー 1-多様性による匿名性評価の機能が示されたと言える. 一方で, このような情報公開により, 一般の人口と比較した有病率などの解離も人権侵害になりうるが, このような医療研究で出された論文自体がエビデンスとして世間の知識となり「相関特定」となる [9]. そのため, PFDOC 攻撃はプライバシーの侵害に当たらない可能性もある.

PFDOC 差分エントロピー 1-多様性は l が 1.188 より大きければ匿名性を保てることが指摘された. PFDPTC 差分エントロピーが特に高い 0.05 以上のカテゴリーにおいて, 実際に「治療群 + プラセボ群」での併存疾患の有病率あるいは併存疾患に対する服薬率が, 「治療群もしくはプラセボ群」において確率が高くなる例が指摘された. PFDPTC 攻撃に対して脆弱性のあるカテゴリーの指摘ができるため, PFDPTC 差分エントロピー 1-多様性による匿名性評価の機能が示されたと言える.

7. おわりに

医療データの中で特に Patient characteristics の 3 つの定量的な匿名性指標として PDP エントロピー 1-多様性, PFDOC エントロピー 1-多様性, PFDPTC 差分エントロピー 1-多様性を提案した. また, 具体的な Patient characteristics において, それぞれの匿名性指標による匿名性を評価し, 匿名性評価の機能を確認した. 今後の課題としては, 本論文では RCT と呼ばれる医学研究の手法での, 併存疾患や併存疾患に対する服薬のカテゴリーが存在する Patient characteristics を対象としたが, スコアリング研究

や症例報告などの研究手法における様々な形での Patient characteristics [12] における匿名性の評価指標を作成することが考えられる.

参考文献

- [1] B. C. M. Fung, K. Wang, R. Chen, and P. S. Yu, "Privacy-preserving data publishing: A survey on recent developments," *ACM Computing Surveys*, vol.42, no.4, pp.1-51, 2010.
- [2] L. Sweeney, "k-anonymity: A Model for Protecting Privacy," *International Journal on Uncertainty, Fuzziness and Knowledge-based Systems*, vol.10, no.5, pp.557-570, 2002
- [3] C. Dwork, "Differential privacy: A survey of results," *International conference on theory and applications of models of computation*, pp1-19, 2008.
- [4] A. Machanavajjhala, J. Gehrke, D. Kifer, and M. Venkatasubramanian, "l-diversity: Privacy beyond k-anonymity," *ACM Transactions on Knowledge Discovery from Data (TKDD)*, vol.1, no.1, article 3, 2007.
- [5] N. Li, T. Li, and S. Venkatasubramanian, "t-closeness: Privacy beyond k-anonymity and l-diversity," *2007 IEEE 23rd international conference on data engineering*, pp.106-115, 2007.
- [6] Khaled El Emam (著), Luk Arbuckle (著), 木村 映善 (監修), 魔狸 (監修), 笹井 崇司 (翻訳). データ匿名化手法 —ヘルスデータ事例に学ぶ個人情報保護. オライリージャパン, 2015.
- [7] K. Rajendran, M. Jayabalan, and M. E. Rana, "A study on k-anonymity, l-diversity, and t-closeness techniques," *International Journal of Computer Science and Network Security*, vol.17, no.12, pp.172-177, 2017.
- [8] V. A. Simonovich, L. D. Burgos Prats, P. Scibona, M. V. Beruto, M. G. Vallone, C. Vázquez, N. Savoy, D. H. Giunta, L. G. Pérez, M. L. Sánchez, A. V. Gamarnik, D. S. Ojeda, D. M. Santoro, P. J. Camino, S. Antelo, K. Rainero, G. P. Vidiella, E. A. Miyazaki, W. Cornistein, O. A. Trabadelo, F. M. Ross, M. Spotti, G. Funtowicz, W. E. Scordo, M. H. Losso, I. Ferniot, P. E. Pardo, E. Rodriguez, P. Rucci, J. Pasquali, N. A. Fuentes, M. Esperatti, G. A. Speroni, E. C. Nannini, A. Matteaccio, H. G. Michelangelo, D. Follmann, H. C. Lane, and W. H. Belloso, for the PlasmAr Study Group, "A randomized trial of convalescent plasma in Covid-19 severe pneumonia," *New England Journal of Medicine* vol.384, no.7, pp.619-629, 2021.
- [9] C. Dwork, A. Smith, T. Steinke, and J. Ullman, "Exposed! a survey of attacks on private data," *Annual Review of Statistics and Its Application*, vol.4, pp61-84, 2017.
- [10] J. K. Ted, and G. Franklin, "Placebo effects in medicine," *New England Journal of Medicine*, vol.373, no.1, pp.8-9, 2015.
- [11] A. Bhatia, P. S. Appelbaum, and K. L. Wisner, "Unblinding in Randomized Controlled Trials: A Research Ethics Case," *Ethics & Human Research*, vol.43, no.2, pp.28-34, 2021.
- [12] R. Pisters, D. A. Lane, R. Nieuwlaat, C. B. De Vos, H. J. Crijns, and G. Y. Lip, "A novel user-friendly score (HAS-BLED) to assess 1-year risk of major bleeding in patients with atrial fibrillation: the Euro Heart Survey," *Chest*, vol.138, no.5, pp.1093-1100, 2010.