

厚生労働科学研究費補助金（厚生労働科学特別研究事業）  
分担研究報告書

テーマB：大規模言語モデルを用いた研究説明文書の平易化による理解力向上の検証

研究分担者 川崎良 大阪大学大学院医学系研究科社会医学講座公衆衛生学・教授

研究協力者 中島悠太 大阪大学 データビリティフロンティア機構 教授

研究協力者 大谷まゆ 株式会社サイバーエージェント

研究協力者 張培楠 株式会社サイバーエージェント

**研究要旨**

**目的:**現代の医療の重要な要素として、患者が自ら受ける医療について理解し、判断し、納得することが重要である。臨床研究の研究説明文書を題材として、大規模言語モデルによる平易化されたインフォームド・コンセント用説明文書を平易化すること、そして、文章の平易化によって患者視点で説明内容の理解度、信頼度が維持・向上するかを評価する研究を行なった。

**方法:**大規模言語モデルを用いて、大阪大学医学部附属病院で用いている研究説明用の文書（3447字）を平易化（701字～1108字）した。その文書を10～80歳代までの一般ネットパネルに提示し、理解度や同意への意向について調査した。その結果を集計し、年齢、性別、社会経済背景要因で調整をした上で平易化によって原文に比べて理解度および同意への意向において向上あるいは低下が見られたかを理解度スコアの平均値の差を持って評価した。

**結果:**研究への同意の意向については原文と平易化文章の間で差が見られなかった。その一方で、平易化文章では概ね全体の理解度が向上し、他の人に内容を説明できるとの回答が多くなった。年代別に詳細を見ると、平易化文書によって50代から70代ではかえって理解度が低下する傾向があることが示唆された。

**結論:**研究協力におけるインフォームド・コンセント文書は一般に長く、難解であることが多い。大規模言語モデルによる文書平易化によって理解度を向上させつつ、同意の意向には大きな影響を及ぼさないことが示された。また、年代によって平易化文書による理解度への影響が異なる可能性も示され、年代ごとに提示する文書の型を変えるなどの個別化によってよりきめ細やかな理解度向上を得る可能性もあると思われた。

**A. 研究目的**

現代の医療の重要な要素として、患者が自ら受ける医療について理解し、判断し、納得することが重要である。そのプロセスは、患者が医療を受ける際に、① まず検査

や治療についての情報を十分に知ること、② その情報（医療の内容、期待される利益と不利益）を適切に理解すること、③ その上で自ら判断し決断を下すこと、から構成され、ヘルス・リテラシーの目指すべき形と言える。このプロセスは、現代の臨床に

においてはインフォームド・コンセント (IC, informed consent) のプロセスとして実装されていると言える。

例えば、検査や治療について、あるいは臨床研究への参加について、医療従事者は文書を用いて患者に対してその内容を説明し、十分な理解を得たことを確認し、本人の同意を確認して医療や研究を進めていく。そのためにはICで用いる説明文書が重要な役割を担っている。その一方、現場で用いられている説明用文書で説明を受けた患者の60%はその内容が十分に理解できていないという報告や、文書のわかりにくさがその後の医療や研究参加に対する不安感の助長、信頼感の低下につながっていることが報告されている。このことは、医療の現場で限られた時間に十分な医療・研究について説明することがいかに難しいかを物語っている。近年、人工知能 (AI, artificial intelligence) 技術の著しい発展と応用によりその応用範囲が医療にも及んでいる。自然言語を解析し、生成するいわゆる大規模言語モデル (LLM, Large Language Model) が開発され、さまざまな用途で用いられている。医療における種々の文書の生成、情報抽出、要約、言い換え、平易化もその応用の一つである。ICで用いる説明文書についてもLLMを用いて、その内容を分かり易くすること、その際には同時に内容の理解度を維持あるいは向上させることができるか、また、根拠に乏しい情報、誤った情報が混入することがないか、流動的な改変文書について誰がどのように責任を取ることができるか、技術的な議論と倫理的・法的・社会的課題の議論の双方が活発化している。

今回、医療現場で用いられるIC説明用文書を、LLMで平易化することの効果を経験的な評価とヘルス・リテラシーの醸成への寄与の評価を目的として研究を行った。具体的には、大阪大学医学部附属病院で行われ公開されている臨床研究の研究説明文書を題材として、LLMによる平易化されたIC説明用文書を平易化すること、そして、文章の平易化によって患者視点で説明内容の理解度、信頼度が維持・向上するかを評価する研究を行った。

## B. 研究方法

- (1) 説明文書の選定：本研究では患者説明用の文書として「全ゲノム情報等の高精度かつ迅速な患者還元および新たな創薬等の創出を通じた高度化がんプレジジョン医療の実践」(大阪大学医学部附属病院観察研究等倫理審査委員会 承認番号:22195-6; 研究責任者:大阪大学大学院医学系研究科消化器外科学・教授 土岐祐一郎) (2023年7月8日 第5版) を元文書として用いた。選定の理由としては、複数の診療科に関連する内容であること、治療に関する内容では患者当事者でなければ判断できない要素が関与するため研究的な内容であること、そして、内容の理解が比較的難しい研究協力の依頼であること、から選定した。この説明文書全文は23ページに及ぶため、本検証では「7.様々な目的での利用」(3頁)を用いることとした(資料1)。
- (2) LLMを用いた説明文書の平易化：元の説明文書をChatGPT4 (OpenAI, USA)

および cyberagent/calm2-7b-chat (サイバ  
ーエージェント, 日本) の二つの LLM モデ  
ルを用いて平易化した。平易化において  
は、{対象：一般人、10代、若い人、高齢  
者、お年寄り、医学的知識のない方向け  
に、}, {平易化指示：分かりやすく、簡潔  
に、要点をまとめて、理解しやすいよう  
に}, {文字数：1000 字程度で} の指示語の組  
み合わせでプロンプトを作成し、それぞ  
れのプロンプトを用いて平易化文書を 100 編  
作成した。作成された平易化文書につい  
て、医学生 3 名が以下の観点から 5 段階評  
価を行い、6 編に絞り込んだ。

A) 明らかに誤った情報が含まれていない  
か？

1. 明らかな誤りがある (1 点)
2. 誤りの可能性がある (2 点)
3. 誤りはない (3 点)
4. 概ね誤りはない (4 点)
5. 全く誤りはない (5 点)

B) 理解しやすいか？

1. 全くわからない (1 点)
2. ややわかりにくい (2 点)
3. わかりやすい (3 点)
4. とてもわかりやすい (4 点)
5. 大変わかりやすい (5 点)

C) 難しい語句があるか？

1. たくさんある (1 点)
2. 結構ある (2 点)
3. ある (3 点)
4. 少しある (4 点)
5. 一つもない (5 点)

D) 原文を正確に表現しているか？

1. 全く正確ではない (1 点)
2. あまり正確ではない (2 点)
3. 正確である (3 点)
4. とても正確 (4 点)
5. 非常に正確 (5 点)

E) 内容を人に説明できるか？

1. 全く説明できない (1 点)
2. あまり説明できない (2 点)
3. 説明できる (3 点)
4. 概ね説明できる (4 点)
5. 十分に説明できる (5 点)

日本語文章の難易度を示す指標として、日  
本語文章難易度判別システム

(<http://jreadability.net/>) を用いた指標化  
を行なった。

(3) 医療従事者による平易化説明文書の  
内容評価：上記で選別された 6 編について、  
医師 7 名、看護師 4 名が医学的な内容の誤  
りがないか、そして、上記の評価項目に従  
って採点を行なった。

(4) 一般人による平易化された説明文書の  
評価：ネット調査を手掛ける委託業者 (イ  
ンテージ社) の一般人からなるパネル (対  
象者年齢群：15-19 歳、20 代、30 代、40  
代、50 代、60 代、70 代、80 代；各 500 名  
程度を目標にリクルートを行なった。用意  
した原文、平易化された IC 説明用文書 1-  
6 の計 7 編を年齢、性別ごとにランダムに  
提示した。回答時間に制限は設けなかつ  
た。本文を提示後、資料 2 に示す項目につ  
いて回答を求めた。

(5) 説明文書生成を可能にするチャットボットの試作：個別化された説明を提供する形のプロトタイプとして、患者向けの説明文書をもとに質問に答えるチャットボットを試作した。大阪大学医学部附属病院で用いている胃がん手術に関する情報を外科医として回答に答えるもので、外部情報の検索を制限し、大阪大学で用いている説明文書のみを参照しつつ回答を生成するように設定した。

## C. 研究結果

原文書をもとに作成した 100 編の生成文章のうち、医学生による予備評価によって選定された 6 編を資料 3 から資料 8 に示す。

表 2 には、使用した LLM、文字数、日本語文章難易度判別システムスコアと文章難易度、文章の形式を示す。文字数は 800–1000 字程度で A4 一枚に収まる程度の分量となった。文章の体裁としては、原文のような説明文型に加え、箇条書き型、Q&A 型、会話型などが生成された。これら 6 編について、まず医療従事者が採点した評価の分布を図 1 に示す。平易化 4 が他の文章に比べてスコアが低かった。ネットパネル調査においては、年齢・性の層毎にランダムに原文か平易化文書 6 編のいずれかが提示されることとした。表 1 には性別・年齢層別の回答者人数の分布を示す。性別及び年齢層については均等に回答が得られた。資料 3 から資料 8 を提示した上での回答をもとに 5 段階評価を 1 点から 5 点として平均スコアを算出したものを図 2 に

示す。粗スコア値においては、原文に比べて、平易化 2、平易化 3 については有意な差はなく、平易化 5 は「目的」について、平易化 6 は「国家戦略」についての理解度が悪い方向に、平易化 4 と平易化 7 は「全体の理解度」と「他人へ説明できる」については理解度の向上の方向に差がみられた。

文章の理解においては、年齢や性別、社会経済的要素として居住都道府県、職業（学生含む）、世帯収入などが影響する可能性を考え、これらの変数で調整した上での原文に対する平易化文章の理解度スコアの差を算出したものが表 3 である。その結果、「国家戦略についての理解度」については原文には明記されているものの、平易化の過程で省略されているか明記されているとは言えない内容であることを反映して、理解度が低くなっていた。また、研究への同意の意向については原文と平易化文章の間で差が見られなかった。その一方で、平易化 5 以外では、全体の理解度、他の人に内容を説明できるなどのスコアが向上していた。逆に言えば、平易化 5 については、原文に比べてかえって理解度が低い結果となった。平易化 5 については、文章の日本語文章難易度判別システムスコアが他の平易化文書に比べて低いこと、また、文書形式が説明文形式であることなどが違いとして示唆された。

図 3 には年代別にスコアを文字数順、日本語文章難易度判別システムスコア順、そして、文書形式別に示した。暖色系は理解度スコアが低下する方向に、寒色系は理解度スコアが向上する方向にあったことを示している。年代別に見ると平易化文

書によって 50 代から 70 代ではかえって理解度が低下する方向にあることが示唆された。

このように、年代によって平易化文書による理解度への影響が、理解度を上げる方向と下げる方向とに分かれているとすれば、全体としての研究の理解度を向上させるためには、年代ごとに提示する文書を変える、個別化することで全体の理解度を助けるかは今後検証が必要であると考えている。

個別化された研究説明あるいは臨床情報提供の例として、大阪大学医学部附属病院が提供している情報をもとに回答を提供するチャットボットのプロトタイプを作成した（大阪大学医学部附属病院患者向けの胃がん手術説明文書に基づいたチャットボット：<https://chat.openai.com/g/g-TL4Ks53mu-ban-da-hua-zi> (ChatGPTs を用いて作成)；<https://huggingface.co/spaces/peinan/pdfchat> (cyberagent/calm2-7b-chat))。これらのチャットボットで用いた題材は胃がんについての手術説明書であり、外部検索機能を用いず、提供された手術説明書に基づいて回答をすることを指示したものとしている。今回の研究期間中には検証には至らなかったが、このチャットボットについての評価も今後検討していきたい。

#### D. 考察

今回、医療現場で用いられる研究に関する IC 説明用文書を、LLM を用いて平易化することで理解度の向上、あるいは、同意取得の

意向への影響などを評価した。今回は小規模な技術的評価にとどまっているが、文章の平易化によって同意への意向に影響を及ぼすことなく理解度を向上させることができる可能性を明らかにすることができた。

医療従事者による評価において評価が低かった平易化 5 についてはネットパネル調査での理解度において原文に比べて理解度が向上することはなかった。かえって目的の理解度においては理解度が低い方向に影響している可能性もあり、これが文章の難易度によるものか、あるいは、文章の体裁（説明文）によるものかについてはさらに検討する余地がある。

また、年代によって好まれる文書に特徴があることが示され、特に 10 歳代から 40 歳代と 80 歳代は平易化文書が理解度を高める方向にある傾向であるのに対して、50 歳から 70 歳代においてはかえって平易化が理解度を低下させている可能性もあり、年代に合わせて好まれる平易化文書の型がある可能性も示された。本来であれば患者に対して丁寧な説明を行う上で、個別にその理解を助けることができれば理想であるが、それが LLM の助けを借りて情報提供を行う可能性が示唆された。

#### E. 結論

研究に関する IC 説明用文書を、LLM を用いて平易化することで理解度の向上、あるいは、同意取得の意向への影響などを評価した。今回は小規模な技術的評価にとどまっているが、文章の平易化によって同意への意向に影響を及ぼすことなく理解度を向上させる

ことができる可能性、また、年代によって好まれる文書に特徴がある可能性を明らかにすることができた。

- <https://huggingface.co/spaces/peinan/pdfchat> (cyberagent/calm2-7b-chat を用いて作成)

## F. 健康危険情報

なし

## G. 研究発表

### 1. 論文発表

準備中

### 2. 学会発表

準備中

## H. 知的財産権の出願・登録状況

(予定を含む。)

### 1. 特許取得

なし

### 2. 実用新案登録

なし

### 3. その他

- 大阪大学医学部附属病院患者向けの胃がん手術説明文書に基づいたチャットボット作成例：  
<https://chat.openai.com/g/g-TL4Ks53mu-ban-da-hua-zi> (ChatGPTs を用いて作成)

図 1. 生成された平易化文書に対する医療従事者による評価結果

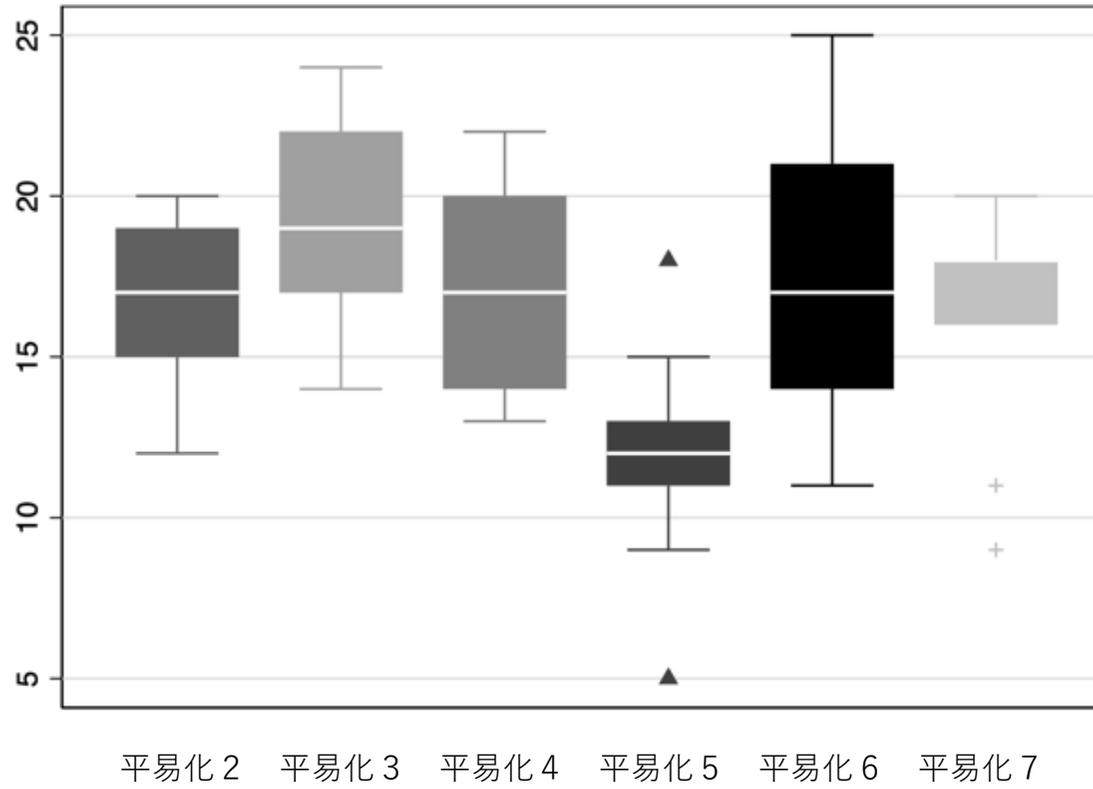


図2. ネットパネル調査での理解度スコア平均値(粗集計)

(\* :原文に対して理解度低い; #:原文に対して理解度高い;  $p < 0.05$  with One-way ANOVA with Bonferroni correction)

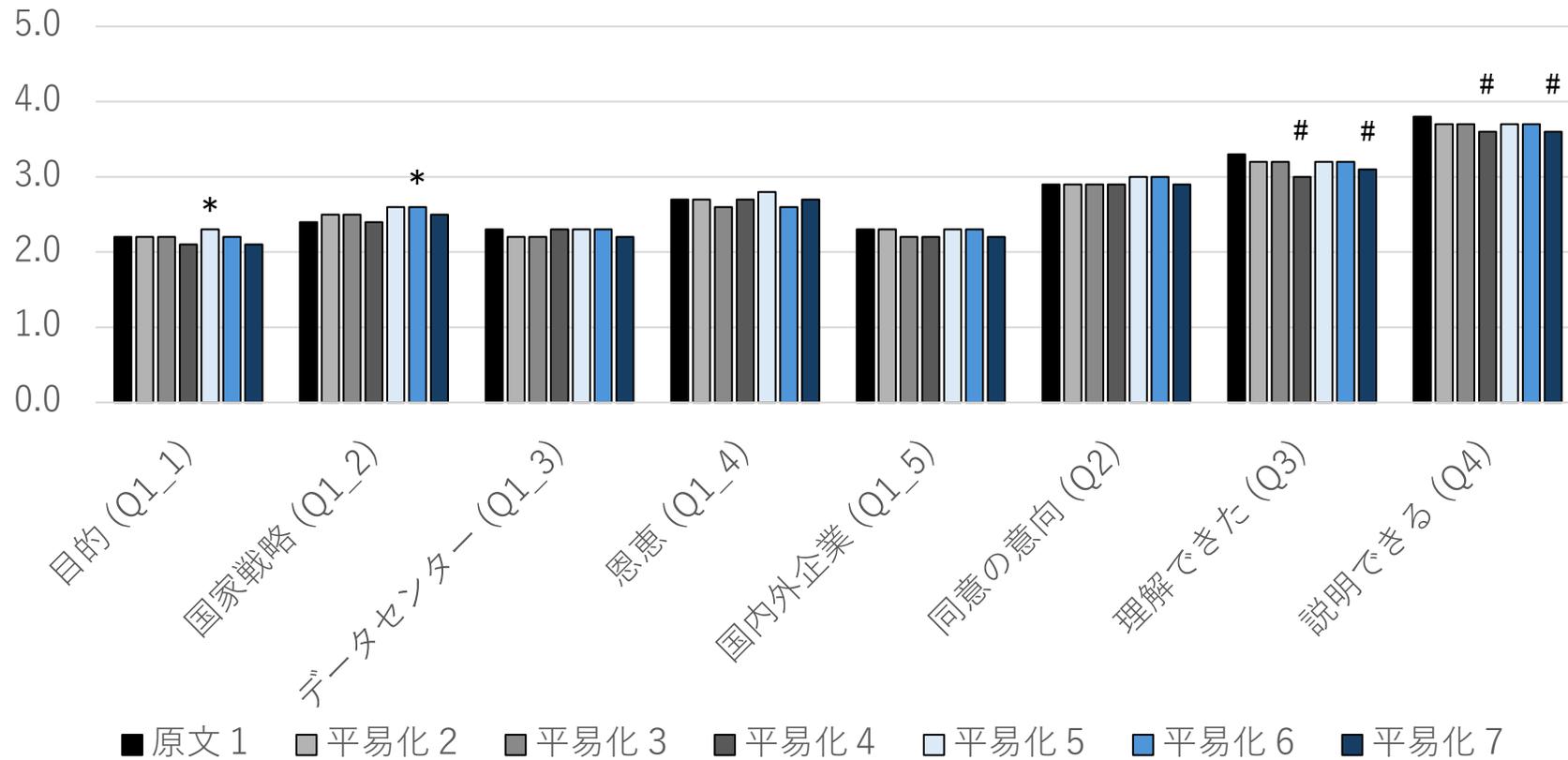


図3.年代別かつ文字数、リーダビリティ、文書形式順に見た理解度スコア平均値のヒートマップ  
 (暖色系は原文に比べて理解度が低い方向に、寒色系は原文に比べて理解度が高い方にあることを示す。)

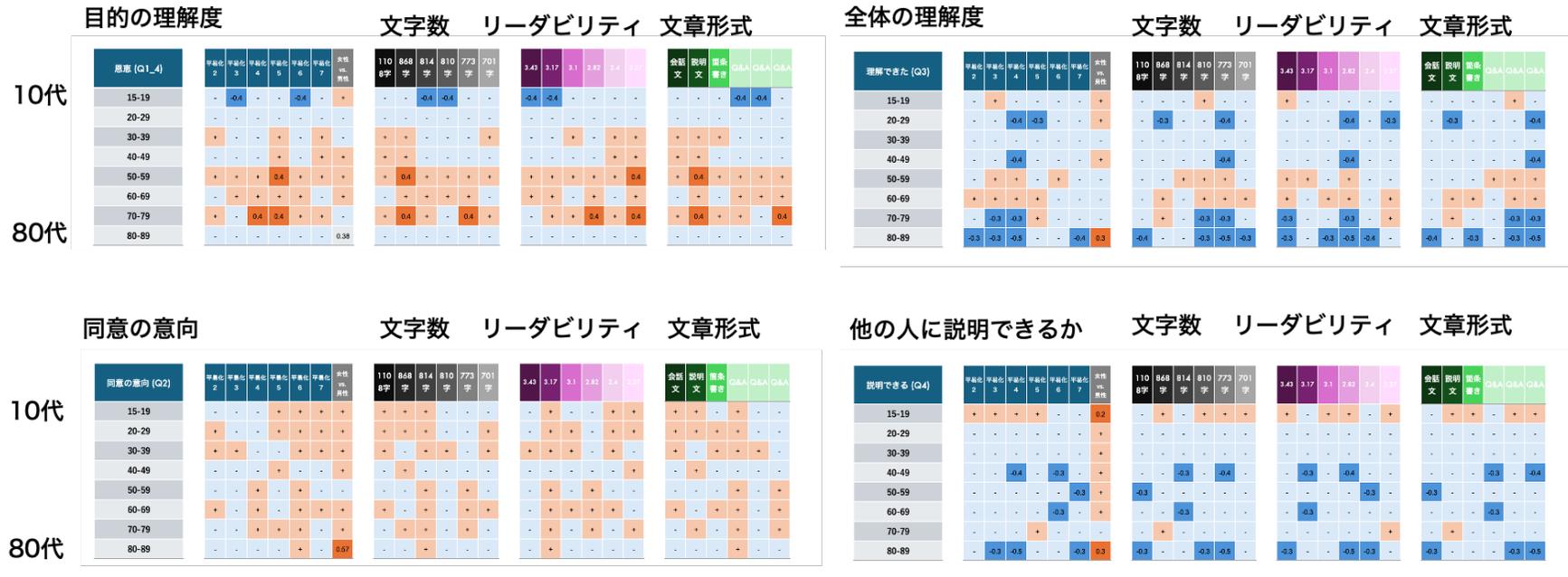


表1. 対象者年齢・性別の分布

		原文1	平易化2	平易化3	平易化4	平易化5	平易化6	平易化7	p-value
<b>N</b>	<b>4193</b>	601	596	603	597	606	598	592	
<b>性別</b>	<b>男性</b>	314 (52.2%)	304 (51.0%)	288 (47.8%)	310 (51.9%)	327 (54.0%)	312 (52.2%)	326 (55.1%)	0.26
	<b>女性</b>	287 (47.8%)	292 (49.0%)	315 (52.2%)	287 (48.1%)	279 (46.0%)	286 (47.8%)	266 (44.9%)	
<b>年代</b>	<b>15-19</b>	72 (12.0%)	72 (12.1%)	72 (11.9%)	73 (12.2%)	74 (12.2%)	72 (12.0%)	73 (12.3%)	1.00
	<b>20-29</b>	72 (12.0%)	71 (11.9%)	73 (12.1%)	71 (11.9%)	73 (12.0%)	71 (11.9%)	71 (12.0%)	
	<b>30-39</b>	75 (12.5%)	78 (13.1%)	78 (12.9%)	76 (12.7%)	79 (13.0%)	75 (12.5%)	72 (12.2%)	
	<b>40-49</b>	76 (12.6%)	75 (12.6%)	78 (12.9%)	75 (12.6%)	76 (12.5%)	76 (12.7%)	75 (12.7%)	
	<b>50-59</b>	77 (12.8%)	75 (12.6%)	78 (12.9%)	76 (12.7%)	76 (12.5%)	75 (12.5%)	73 (12.3%)	
	<b>60-69</b>	76 (12.6%)	76 (12.8%)	74 (12.3%)	78 (13.1%)	78 (12.9%)	76 (12.7%)	78 (13.2%)	
	<b>70-79</b>	78 (13.0%)	75 (12.6%)	75 (12.4%)	75 (12.6%)	74 (12.2%)	76 (12.7%)	76 (12.8%)	
	<b>80-89</b>	75 (12.5%)	74 (12.4%)	75 (12.4%)	73 (12.2%)	76 (12.5%)	77 (12.9%)	74 (12.5%)	

表 2.平易化されたファイルの文字数、難易度

	原文1	平易化2	平易化3	平易化4	平易化5	平易化6	平易化7
平易化に使用した LLM	-		ChatGPT4			calm2-7b-chat	
総文字数	3447 字	701 字	810 字	773 字	868 字	814 字	1108 字
文章難易度	上級前半	中級後半	中級後半	中級後半	上級前半	中級後半	上級前半
リーダビリティ値 [日本語文章難易度判別システム (jreadability.net) ]	1.70	3.10	3.43	2.82	2.27	3.17	2.40
文書形式	説明文	箇条書き	Q&A	Q&A	説明文	Q&A	会話文

表3. 原文に対する平均スコア(年齢、性別、居住都道府県、職業(学生含む)、世帯年収で調整した多変量解析)

	目的についての理解度 (Q1_1)				国家戦略についての理解度 (Q1_2)				データセンターについての理解度 (Q1_3)				恩恵についての理解度 (Q1_4)			
	$\beta$	95%CI	p 値		$\beta$	95%CI	p 値		$\beta$	95%CI	p 値		$\beta$	95%CI	p 値	
原文 1	(基準)	-	-		(基準)	-	-		(基準)	-	-		(基準)	-	-	
平易化 2	0.04	-0.06	0.13	0.48	<b>0.11</b>	<b>0.01</b>	<b>0.21</b>	<b>0.03</b>	-0.02	-0.12	0.08	0.70	-0.01	-0.12	0.10	0.91
平易化 3	0.04	-0.06	0.14	0.40	0.09	-0.01	0.19	0.09	-0.05	-0.15	0.06	0.37	-0.10	-0.21	0.01	0.07
平易化 4	0.01	-0.09	0.11	0.88	0.05	-0.05	0.15	0.31	0.02	-0.09	0.12	0.76	0.02	-0.09	0.13	0.72
平易化 5	<b>0.19</b>	<b>0.09</b>	<b>0.29</b>	<b>&lt;0.01</b>	<b>0.16</b>	<b>0.06</b>	<b>0.26</b>	<b>&lt;0.01</b>	0.00	-0.10	0.11	0.95	0.10	-0.01	0.21	0.08
平易化 6	0.06	-0.04	0.16	0.25	<b>0.21</b>	<b>0.11</b>	<b>0.31</b>	<b>&lt;0.01</b>	0.05	-0.06	0.15	0.39	-0.10	-0.21	0.01	0.07
平易化 7	-0.02	-0.12	0.08	0.70	<b>0.13</b>	<b>0.03</b>	<b>0.23</b>	<b>0.01</b>	-0.08	-0.18	0.03	0.15	-0.02	-0.13	0.09	0.74
	国内外企業利用について(Q1_5)				同意の意向 (Q2)				全体の理解ができた (Q3)				他の人に説明できる (Q4)			
	$\beta$	95%CI	p 値		$\beta$	95%CI	p 値		$\beta$	95%CI	p 値		$\beta$	95%CI	p 値	
原文 1	(基準)	-	-		(基準)	-	-		(基準)	-	-		(基準)	-	-	
平易化 2	0.00	-0.11	0.10	0.96	0.00	-0.11	0.12	0.95	<b>-0.14</b>	<b>-0.23</b>	<b>-0.04</b>	<b>0.01</b>	<b>-0.12</b>	<b>-0.22</b>	<b>-0.02</b>	<b>0.02</b>
平易化 3	-0.07	-0.17	0.04	0.20	-0.06	-0.17	0.06	0.32	<b>-0.13</b>	<b>-0.23</b>	<b>-0.03</b>	<b>0.01</b>	<b>-0.15</b>	<b>-0.25</b>	<b>-0.05</b>	<b>&lt;0.01</b>
平易化 4	-0.04	-0.14	0.07	0.46	-0.05	-0.16	0.07	0.42	<b>-0.24</b>	<b>-0.33</b>	<b>-0.14</b>	<b>&lt;0.01</b>	<b>-0.19</b>	<b>-0.28</b>	<b>-0.09</b>	<b>&lt;0.01</b>
平易化 5	0.02	-0.09	0.12	0.76	0.04	-0.07	0.16	0.44	-0.07	-0.16	0.03	0.19	-0.05	-0.15	0.05	0.29
平易化 6	0.08	-0.03	0.18	0.15	0.05	-0.06	0.16	0.38	<b>-0.11</b>	<b>-0.20</b>	<b>-0.01</b>	<b>0.03</b>	<b>-0.14</b>	<b>-0.24</b>	<b>-0.04</b>	<b>0.01</b>
平易化 7	-0.10	-0.20	0.01	0.07	-0.04	-0.15	0.07	0.46	<b>-0.18</b>	<b>-0.28</b>	<b>-0.08</b>	<b>0.00</b>	<b>-0.20</b>	<b>-0.30</b>	<b>-0.10</b>	<b>&lt;0.01</b>

$\beta$ : 原文の平均スコアに対する差(負の方向が理解度が高いことを示す) 95%CI: 95%信頼区間

## (資料1) 原文1

「全ゲノム解析等実行計画」では、ご提供いただいた試料・情報を用いて「解析・データセンター」に全ゲノム等解析データおよび臨床情報のデータベースを構築し、国内外の研究者や企業の間でデータを広く共有して診療や研究・創薬等に活用するための仕組みを構築します。この仕組みは、日本の患者さんや市民の皆さんのゲノム医療を日本に住む方々のデータを使って実現するための基盤であるとお考えください。そのため、以下の目的で多様な疾患に関する研究・創薬等にデータを活用させていただきます。①全ゲノム解析等の成果を患者さんに還元すること②新たな個別化医療等を実現し、日常診療への導入を目指すこと③全ゲノム解析等の結果を研究・創薬などに活用すること④上記①～③に関わる人材の育成や保健医療政策の検討を行うこと本研究では、データの適正な利用を確保するため、関連法令・指針、AMED との契約および「全ゲノム解析等実行計画」に基づいてデータ共有ルールを定め、第三者も加わったデータ利活用審査委員会が公正な立場からデータの利用を審査・監督する仕組みを設けます。データの利用者には、関連法令・指針、データの利用条件、およびセキュリティに関するガイドラインの遵守が義務づけられます。ゲノム研究は急速に進歩しており、今後さらに研究が進むと、現時点では想定されていない新たな研究に試料や情報を活用させていただくことが必要となる可能性があります。その場合には、新たに研究計画を作成し、関連法令・指針に基づき、データ利活用審査委員会や研究機関の倫理審査委員会による審査で改めて承認を受けた上で研究に活用させていただきます。なお、AMED 研究開発データ統合利活用プラットフォームを通じてデータが利活用される場合があります。その場合も、上記に準じて取り扱われます。「全ゲノム解析等実行計画」ではデータの利活用に関する状況をできる限り公開し、透明性の確保に努めます。本計画で解析・収集されたデータは大変に貴重なものです。そのため、様々な目的で使わせていただきたいと考えております。以下では、その用途として考えられる事項をご説明します。① 将来の研究目的での利用医療機関や研究機関の努力だけで、新しい医薬品や診断技術等を開発することは困難です。また従来、こうした製品・技術の開発に携わる企業が、日本の患者さん等のゲノムデータを収集・解析する環境が整備されておらず、課題とされてきました。そのため、「全ゲノム解析等実行計画」では、研究機関・医療機関だけでなく、医薬品や医療機器等の研究開発に携わる国内外の企業が、研究、医薬品等の開発、科学的なエビデンスに基づく予防等の目的で、データを利用できる環境を整備します。これにより、国内での研究開発を促進することが期待されます。現在はまだ計画されていませんが、今後、さらに研究が進むと、新たな研究が立案される可能性があります。その場合に、あなたからいただいた試料や情報を活用させて頂く可能性があります。その際、国内外の機関（民間企業を含む）との共同研究になる可能性があるほか、共同研究先にあなたから頂いた試料や情報を提供する可能性があります。具体的には、以下のような形でデータを利用させていただきます。a) 「全ゲノム解析等実行計画」に参画する研究機関・医療機関およびこれらの機関との共同研究を行う機関による利用本

研究の実施機関を含む多数の医療機関・研究機関が「全ゲノム解析等実行計画」に参画し、データの収集や解析を行います。得られたデータを効率的に解析するために、参画機関が共同でデータを利用して診療、研究・創薬等を行います。また、参画機関との共同研究を行う機関（企業を含む）がデータを利用させていただく場合もあります。利用を希望する医療機関や学術研究機関は、具体的な研究計画を立案した後、倫理審査委員会の承認を得て、データの利用審査を受けることになっています。利用に際しては、データ利活用審査委員会への報告が義務づけられます。b) 上記 a)以外の機関による利用上記 a)以外の機関は、有償でデータを利用することが可能です（おもに企業による利用が想定されています）。データの利用を希望する研究機関、医療機関、企業等は、具体的な研究計画を立案し、倫理審査委員会の承認を得るとともに、データ利用を申請し、データ利活用審査委員会による審査を受けます。審査により、適切な利用を行うと認められた場合にのみデータの利用が許可されます。ただし、医薬品や医療機器等の開発に携わる企業は、それぞれ競争をしながら開発を進めており、データ利用の詳しい目的や目標等について詳細な情報を公開することはできません。しかし、できる限りの情報を公開し、透明性を担保いたします。

② 国内外の公的データベースへの登録と活用本研究と類似した研究をする無駄を防いだり、解析するデータの量や規模を大きくして研究結果の信頼性を高めるために、公的なデータベースにデータを提供する取り組みも進んでいます。解析されたデータが、データベースを介して他の研究者に利用されることによって、新しい技術の開発が進むとともに、様々な病気について、今まで不可能であった疾患の原因の解明や治療法・予防法の確立が早まる可能性が期待できるためです。「解析・データセンター」で保管されるあなたのデータのうち、全ゲノム等解析データおよび一部の臨床情報は、公的な研究用データベース（国外のものを含む）にも提供・登録されます。公的データベースに登録されたデータは、日本国内の研究機関に所属する研究者だけではなく、国内外の製薬企業等の民間企業や海外の研究機関に所属する研究者も利用する可能性があります。将来、どの国の研究者から利用されるか、現時点ではわかりません。しかし、どの国の研究者に対しても、その国の法令に沿って作成されたデータベースのガイドライン等に準じた利用が求められます。データベースに登録されたデータの利用には、データの種類によって異なるアクセスレベル（制限公開、非制限公開）が設定されます。個人の特定につながらない情報（頻度情報・統計情報等）は非制限公開データとして不特定多数による利用が可能な形で提供され、個人ごとのゲノムデータ等は制限公開データとして、科学的観点と研究体制の妥当性に関する審査を経た上で、データの利用を承認された研究者のみに利用されます。現在、登録を予定しているデータベースには、代表的なものとして以下が挙げられます。a) 国立研究開発法人科学技術振興機構（JST）が運営する、バイオサイエンスデータベースセンター（NBDC）のヒトデータベース NBDC ヒトデータベースでは、インターネットのウェブサイトを通じてデータが公開されています。NBDC にデータを登録する際には、データを 2 種類に分けて取り扱います。（ア）多くの方のデータを集計した結果

は、研究者一般に広く公開します。(イ) 個人ごとの詳しいデータについては、科学的観点と個人情報保護のための体制などについて厳正な審査を受けて、承認された研究者にのみ、利用を許可します。NBDC では、データを登録する研究者、および、データを利用する研究者の双方に、 厳格なセキュリティ環境の整備を義務づけています。NBDC の運営状況や NBDC から公開されているデータの内容をお知りになりたい場合には、下記のウェブサイトをご覧ください。NBDC ヒトデータベース

<http://humandbs.biosciencedbc.jp/>③ 人材育成や保健医療政策など公益性の高い目的での利用ゲノム医療の発展には、大規模なデータを解析できる人材の育成が急務であり、これらのデータを使用した実習や研修などに使わせていただきたいと考えています。また、これらのデータは、国や地域が立案する保健医療の政策の検討においても、大きく役立つものだと考えられます。ぜひ行政機関やその依頼を受けた研究者に使わせていただきたいと考えています。上記以外にも、現在はまだわかりませんが、将来、医療をよりよくするために公益性の高い目的で使用させていただく可能性があります。こうした場合にも、利用を希望する者に対しては、利用目的を明らかにしたうえで審査を受けて頂くとともに、データ利用とセキュリティに関するガイドラインを遵守していただきます。なお、本研究にご協力を頂くにあたっては、本計画で解析・収集されたデータを上記①～③でお示している全ての目的で利用することにご同意頂く必要があります。

## (資料2)パネル調査で用いた理解度調査票

次の文章は「ゲノム医療の実現を目指した研究への参加のお願い」の文章の一部です。この文章を読んで以下の質問に回答してください。

### 「文章の提示」

設問1 この文章の内容について、以下の質問に回答ください。

- ① がんや難病の早期診断や治療法開発に貢献することが期待されている。
- ( 1 ) そう思う。  
( 2 ) ややそう思う。  
( 3 ) どちらでもない。  
( 4 ) あまりそう思わない。  
( 5 ) そう思わない。
- ② 国家戦略としてゲノム医療が推進されている。
- ( 1 ) そう思う。  
( 2 ) ややそう思う。  
( 3 ) どちらでもない。  
( 4 ) あまりそう思わない。  
( 5 ) そう思わない。
- ③ 協力することによってあなたの情報は解析・データセンターに保管される。
- ( 1 ) そう思う。  
( 2 ) ややそう思う。  
( 3 ) どちらでもない。  
( 4 ) あまりそう思わない。  
( 5 ) そう思わない。
- ④ 協力することによって、あなたの治療に直接的な恩恵がある。
- ( 1 ) そう思う。  
( 2 ) ややそう思う。  
( 3 ) どちらでもない。  
( 4 ) あまりそう思わない。  
( 5 ) そう思わない。
- ⑤ 協力することによって、あなたの解析結果は国内外の企業が利用する可能性がある。
- ( 1 ) そう思う。  
( 2 ) ややそう思う。  
( 3 ) どちらでもない。  
( 4 ) あまりそう思わない。  
( 5 ) そう思わない。

設問2 この文章を読んで「ゲノム医療」に協力したいと思いましたか？

- ( 1 ) 協力したいと思った。
- ( 2 ) どちらかと言えば協力したいと思った。
- ( 3 ) どちらでもない。
- ( 4 ) どちらかと言えば協力したくないと思った。
- ( 5 ) 協力したくないと思った。

設問3 この文章の内容を十分理解できたと思いますか？

- ( 1 ) とてもよく理解できた。
- ( 2 ) 大部分はよく理解できた。
- ( 3 ) 概ね理解できた。
- ( 4 ) あまり理解できなかった。
- ( 5 ) 全く理解できなかった。

設問4 この文章の内容を他の人に説明できると思いますか？

- ( 1 ) しっかり説明できると思う。
- ( 2 ) 大部分は説明できると思う。
- ( 3 ) おおむね説明できると思う。
- ( 4 ) あまり説明できそうにない。
- ( 5 ) 全く説明できそうにない。

設問5 このような文章の理解を助けるのは以下のどのような要素でしょうか？重要な要素と思われる順に上位三つ選んでください。（複数回答可）

1位 2位 3位

- ( ) ( ) ( ) 文章が短いこと。
- ( ) ( ) ( ) 細かな点まで説明がなされていること。
- ( ) ( ) ( ) Q&A形式であること。
- ( ) ( ) ( ) 箇条書き形式であること。
- ( ) ( ) ( ) 口語調であること。
- ( ) ( ) ( ) その他 ( 自由記載 )
- ( ) ( ) ( ) その他 ( 自由記載 )
- ( ) ( ) ( ) その他 ( 自由記載 )

質問は以上です。ご協力ありがとうございました。

## **(資料3) 平易化2**

### **1. 計画の概要:**

- ・日本中の人々のゲノム（DNA の情報）を解析して、大きなデータベースを作ります。
- ・このデータベースは、病気の治療や新薬の開発に役立てるために、研究者や企業と共有されます。

### **2. 目的:**

- ・ゲノム解析の成果を患者さんに還元する。
- ・新しい個別化医療（人それぞれに合った治療）を実現する。
- ・ゲノムデータを使って、病気の研究や新薬の開発を進める。
- ・関連する研究者や医療関係者の育成、政策の検討を行う。

### **3. データ利用のルール:**

- ・データの使用は法律やガイドラインに基づいて厳しく管理されます。
- ・使う前には、専門の委員会が公正にチェックします。

### **4. 将来の研究での利用:**

- ・新しい医薬品や技術を開発するために、日本だけでなく海外の企業や研究機関もデータを使うことがあります。
- ・これにより、より多くの研究が進むことが期待されます。

### **5. 公的データベースへの登録:**

- ・解析されたデータは、公的な研究用データベースにも登録され、他の研究者も使えるようになります。
- ・これにより、新しい発見が進む可能性があります。

### **6. 公益性の高い利用:**

- ・このデータは、教育や保健医療政策のためにも使われます。
- ・例えば、医療の向上や行政の政策決定に役立つ情報として活用されることがあります。

### **7. 同意が必要:**

- ・この計画に参加するためには、上記のすべての利用目的に同意する必要があります。

この計画は、将来の医療を良くするために、たくさんの人のゲノムデータを使って研究を進めるといふものです。そして、その過程で色々なルールに従ってデータを管理し、安全に利用することが大切にされています。

#### **(資料4) 平易化3**

**Q1: 全ゲノム解析等実行計画って何ですか？**

A1: この計画では、患者さんから提供されたゲノム情報（体の設計図のようなもの）と健康に関するデータを使って、データベースを作ります。このデータベースを使って、日本の研究者や会社が医療の改善や新薬の開発に役立てることができます。

**Q2: この計画でデータはどう使われますか？**

A2: データは主に4つの目的で使われます。1) 患者さんへの医療の改善、2) 新しい治療法の開発、3) 研究や新薬開発、4) 医療関係者の育成や政策の検討です。

**Q3: データはどのように守られますか？**

A3: データの安全と適切な利用を守るため、法律やルールに基づき、専門の委員会がデータの使い方を審査し、監督します。データを使う人は、これらのルールを守ることが必須です。

**Q4: どんな研究にデータが使われるか、将来変わることはありますか？**

A4: はい、研究は進歩するので、新しい研究にもデータが使われることがあります。その場合は、新しい計画を立てて、専門の委員会の承認を得てから使います。

**Q5: 企業もこのデータを使えますか？**

A5: はい、研究機関や医療機関だけでなく、医薬品や医療機器の研究開発を行う企業もデータを利用できます。ただし、適切な利用計画と承認が必要です。

**Q6: このデータはどこに登録されますか？**

A6: データは、国内外の公的な研究用データベースに登録されます。例えば、日本の科学技術振興機構が運営するヒトデータベースなどです。

**Q7: このデータは他にどんな目的で使われますか？**

A7: データは、医療従事者の育成や保健医療政策の検討にも使われます。これは、より良い医療サービスを提供するために重要です。

**Q8: データを提供する際、何に同意する必要がありますか？**

A8: 提供する際には、この計画で収集されたデータが上述の全ての目的で使われることに同意する必要があります。

## (資料5) 平易化4

### Q1: 「全ゲノム解析等実行計画」とは何ですか？

A1: 「全ゲノム解析等実行計画」とは、人々の遺伝子に関する大きなデータを集めて、それを使って病気の治療や新しい薬の開発をするための計画です。日本に住む人々の遺伝子情報を集めて、それを研究者や会社と共有して、もっといい医療を作るための基盤を作ります。

### Q2: この計画でデータはどう使われますか？

A2: この計画では、集めた遺伝子データを使って、病気の治療法を改善したり、新しい薬を作ったりします。また、このデータは研究者や医師の訓練、健康政策の考え方を作るのにも使われます。

### Q3: データの安全はどう守られますか？

A3: この計画では、データを安全に使うために、法律やルールに従って、特別な委員会がデータの使い方を審査して監督します。データを使う人は、セキュリティに関するガイドラインを守る必要があります。

### Q4: このデータは将来どう使われるかもわかりますか？

A4: 将来の研究のためにもデータが使われる可能性があります。新しい研究計画があれば、また審査を受けて承認された後に使います。これは国内外の様々な研究機関や会社が参加するかもしれません。

### Q5: 他の国の研究者もデータを使えますか？

A5: はい、公的なデータベースにデータを登録することで、日本だけでなく他の国の研究者も利用できます。ただし、どの国の研究者が使っても、その国の法律に基づいたガイドラインに従って使われます。

### Q6: この計画に参加するとどのようなメリットがありますか？

A6: この計画に参加することで、あなたのデータが医学の進歩に貢献し、将来的にはより良い医療や新しい薬の開発に役立つかもしれません。また、データは教育や公共政策の検討にも使われ、社会全体の健康向上に貢献する可能性があります。

## (資料6) 平易化5

あなたの全ゲノム解析等の結果は、医学の進歩と健康管理に役立つために研究機関・医療機関に提供されます。

本研究では、あなたに提供いただいた試料・情報を活用して「解析・データセンター」に全ゲノム等解析データと臨床情報のデータベースを構築します。このデータベースは国内外の研究者や企業の間で共有され、将来的には日本の患者さんの診療に役立つとされています。

研究機関や医療機関は、データを利用して診療や研究等を行い、データベースを利活用審査委員会などが公正に審査・監督します。研究機関・医療機関は、関連法令やAMEDとの契約に基づいてデータを利用し、その情報を公開することが求められます。本研究では、データの適切な利用を確保するため、関連法令・指針、データの利用条件、セキュリティに関するガイドラインを遵守しなければなりません。また、データの利用には、同意が必要です。

データ提供によって、あなたの健康状態を把握することができ、医学の進歩に貢献できます。ただし、データ利用により、将来的には新たな研究や治療の開発が期待されますが、その際に試料・情報を活用させていただく可能性があるため、研究計画が承認される前に改めて研究計画の審査を受ける必要があります。

研究機関や医療機関は、データの適切な利用を確保するため、関連法令やAMEDとの契約に基づいて利用し、その情報を公表します。

あなたには、本研究の参加に同意していただく必要があります。同意していただけない場合には、本研究の対象外となります。その場合でも、医学の進歩に貢献することはできません。

もし、あなたが提供した試料や情報が、新たな研究や研究結果の公開によって、あなたや他の患者さんに不利益が生じる可能性がある場合には、その提供を控えていただく必要があります。

以上、本研究の目的や利用方法などについてご理解いただいたうえで、同意いただけず場合には、同意書に署名をお願いいたします。同意書にご署名いただきました場合には、研究機関・医療機関に提出をお願いいたします。

## (資料7) 平易化6

Q: わたしのゲノムデータは、どのように利用されるのですか？

A: あなたのゲノムデータは、次世代の医療の発展に貢献するために、病気の治療や予防、新薬の開発、健康管理などの研究に使われます。「全ゲノム解析等実行計画」は、データ提供者であるあなたの同意を頂いたうえで、研究に活用します。

Q: 遺伝子検査を受けると、どんなメリットがあるのですか？

A: 遺伝子検査を受けると、将来かかりやすい病気のリスクや体質を知ることができます。また、遺伝的なリスクに基づいた健康管理もできるので、病気の予防や健康維持に役立ちます。ただし、リスクが高い場合でも、生活習慣の改善などでリスクを軽減することができますので、将来の健康管理に役立てることができます。

Q: データを使用する際、わたしの個人情報を守られるのですか？

A: 「全ゲノム解析等実行計画」では、あなたの個人情報やゲノムデータの取り扱いについて、厳重に管理します。あなたのデータは、研究に活用される前に、データ利活用審査委員会によって厳正に審査され、データの安全な利用が確保されます。データの利用に際しては、関連法令・指針に従い、あなたのプライバシーや人権が守られるよう、適切に配慮されます。

Q: わたしが協力を拒否することはできますか？

A: あなたは、提供いただいたデータを研究に活用することについて、自由意思で拒否することができます。ただし、同意をいただいた場合は、上記のようにデータを利用していただきます。

Q: わたしの健康にどんなメリットがあるのですか？

A: あなたの遺伝子情報をもとに、将来かかるかもしれない病気を予防したり、病気にかかりにくい体質に改善することができます。遺伝子検査を受けることで、生活習慣の見直しなど、健康管理にも役立ちます。また、あなたにとってより健康的な方法を選択し、健康的な生活を送ることができるようになるかもしれません。

## (資料8) 平易化7

患者：今日はどうしたのですか？

医師：今日は、あなたの全ゲノム解析等実行計画について説明したいと思います。

患者：全ゲノム解析ですか？

医師：そうです、あなたの遺伝子をすべて調べることで、将来の病気の予防や治療法の開発に役立つかもしれないのです。

全ゲノム解析等実行計画では、ご提供いただいた試料・情報を用いて「解析・データセンター」に全ゲノム等解析データおよび臨床情報のデータベースを構築し、国内外の研究者や企業の間でデータを広く共有して診療や研究・創薬等に活用するための仕組みを構築します。

この仕組みは、日本の患者さんや市民の皆さんのゲノム医療を日本に住む方々のデータを使って実現するための基盤であるとお考えください。

データの適正な利用を確保するため、関連法令・指針、AMED との契約および「全ゲノム解析等実行計画」に基づいてデータ共有ルールを定め、第三者も加わったデータ利活用審査委員会が公正な立場からデータの利用を審査・監督する仕組みを設けます。

データの利用者には、関連法令・指針、データの利用条件、およびセキュリティに関するガイドラインの遵守が義務づけられます。

患者：なるほど、では、私のデータはどのように使われるのですか？

医師：まず、将来の研究目的での利用です。具体的には、研究機関・医療機関が、医薬品や医療機器等の研究開発に携わる国内外の企業が、研究、医薬品等の開発、科学的なエビデンスに基づく予防等の目的で、データを利用できる環境を整備します。

次に、国内外の公的データベースへの登録と活用です。解析されたデータが、データベースを介して他の研究者に利用されることによって、新しい技術の開発が進むとともに、様々な病気について、今まで不可能であった疾患の原因の解明や治療法・予防法の確立が早まる可能性が期待できるのです。

患者：では、私のデータはどこで使われるのですか？

医師：公的な研究用データベース（国外のものを含む）に、解析されたデータが提供・登録されます。公的データベースに登録されたデータは、日本国内の研究機関に所属する研究者だけではなく、国内外の製薬企業等の民間企業や海外の研究機関に所属する研究者も利用する可能性があります。

患者：なるほど、私のデータは、将来的には海外の研究者にも使われるかもしれないですね。

医師：はい、そのため、将来的に医療がよりよくなる可能性が期待できます。

患者：わかりました。自分のデータが使われることで、将来の医療がよりよいものになるな

ら、協力する意義があると思います。

医師：ありがとうございます。あなたの協力が、将来の医療の発展につながります。