

令和 5 年度厚生労働科学研究費補助金(循環器疾患・糖尿病等生活習慣病対策総合研究事業)
『栄養・食事関連メディア情報の科学的評価及び国民への影響の分析のための研究』
分担研究報告書

動画共有サイト(YouTube)に存在する糖尿病に関する動画の質の評価

研究分担者 片桐 諒子¹、松本 麻衣¹

研究協力者 篠崎 奈々²

研究代表者 村上健太郎³

¹ 医薬基盤・健康・栄養研究所国立健康・栄養研究所栄養疫学・食育研究部

² 東京大学大学院医学系研究科栄養疫学・行動栄養学

³ 東京大学大学院医学系研究科社会予防疫学分野

【研究要旨】

動画共有サイトである YouTube から抽出した日本語と英語の糖尿病における食事に関する視聴回数の多い動画について動画の質を評価し、これらの動画を医療の専門家でない人が視聴した場合にも正しい情報が伝えられているか、またどのように判断すれば正しいと考えられる情報を見分けることができるかについて検討した。専門の評価者が糖尿病に関する動画を **useful** (糖尿病の予防や治療につながる食事について、科学的根拠を基にしたなんらかの情報が得られるか) かそうではないかを評価し、さらに **modified DISCERN**, **JAMA**, **GQS** といった客観的な指標に基づくスコアリングを行った。結果、専門家が **useful** と考えた動画は **not very useful** と考えた動画と比較して客観的な情報の質を評価する指標でも得点差があり、専門的な知識を持たない視聴者においても、これらの項目を指標として動画を判断することで、ある程度有用と考えられる動画を判別することができる可能性が示唆された。一方、投稿者が医療関係者かどうかだけでは、日本の動画については、医療関係者の動画は現在各指標のスコアが高かったものの英語の動画にはその傾向が認められず、投稿者の背景だけで動画を判断することは、誤った動画を質が高いと考えてしまうことにもつながる可能性があり多角的に判断する必要性を示唆していると考えられた。指標に含まれるような目的の明示、信頼できる情報源、提示情報のバランス、その他の情報源を提示する、情報の不確実性に言及する、投稿者の属性、資金・利益相反開示、投稿日や更新日などの項目に基づき動画の情報を判断することで、医療の専門家でない人が視聴した場合にも有用と考えられる情報を見分けることができ、質の高いと考えられる動画を視聴できる可能性が高くなると考えられた。

A. 背景と目的

International Telecommunication Union によると 2023 年には世界人口の約 67%、54 億人がインターネットを利用するとされる¹⁾。また別の調査会社によると、2021 年末の国内インターネットユーザーは 10078 万人、そのうち約

80%がソーシャルネットワークサービス(SNS)を利用していると推定しており、同会社の 2022 年時点の調査では SNS ユーザーの 62%が動画共有サイトである YouTube を利用していると回答している²⁾。

YouTube は動画共有サイトであり、プラットフォーム

フォームの性質上、誰でもクリエイターになることができるという特徴がある。このため、クオリティの異なる動画が存在することは避けられない。YouTube は質の高いコンテンツを増やすことを目指しており、4つのRの原則に基づいて誤った情報に対処している。4つのRとは、YouTubeのポリシーに違反するコンテンツを削除する(Remove)、ボーダーラインにあるコンテンツがレコメンドとして広まるのを抑える(Reduce)、信頼できるソースからのニュースや情報を見つけやすくする(Raise)、信頼できるクリエイターに還元する(Reward)という4つの原則の頭文字をとったものである³⁾。また、信頼できる医療・健康情報を見つけやすくする仕組みも作っている、とのことである。しかし、質の様々な動画が混在している現状がある。このような背景を踏まえ、YouTube上の医療情報・動画の質を科学的に評価し、視聴者の意識を高めようとする論文がある。例えば、COVID-19⁴⁾、骨粗鬆症⁵⁾などに関する論文がある。これらの論文では、動画の質を測定するために、DISCERN⁶⁾、Global Quality Score (GQS)⁷⁾、Journal of the American Medical Association (JAMA) score⁸⁾など、開発され検証されたいくつかのツールを組み合わせで使用している。

こうした背景を踏まえ、本研究では日本語および英語での糖尿病における食事に関する視聴回数の多い動画について動画の質を評価し、これらの動画を医療の専門家でない人が視聴した場合にも正しい情報が伝えられているか、またどのように判断すれば正しいと考えられる情報を見分けることができるかについて検討した。

B. 方法

対象動画の抽出手順は下記の通り実施した。

①Google trendにて「糖尿病 食事」を検索用語としてYouTubeを対象とした3年間の関連キーワードを人気順に検索

- ②糖尿病、食事に関連するキーワードを選択
- ③各キーワードでYouTubeを検索、この際に「視聴回数」で並び替え
- ④1万回以上の視聴回数の動画(ショート動画は除く)を抽出
- ⑤不適格な動画を除外
- ⑥検索キーワードの上位の動画のリストに、下位の検索キーワードのリストをマージすることで1つの候補リストを作成

Google trendにて「糖尿病 食事」を検索用語としてYouTubeを対象とした2023年2月時点までの3年間の関連キーワードを人気順に検索した結果、以下を検索のキーワードとした。
日本語:・糖尿病 食事 ・糖尿病 食事 レシピ ・糖尿病の食事 ・糖尿病 予防 食事 ・糖尿病 食事 療法 ・糖尿病 食事 ダメなもの 英語:・diet for diabetes ・diabetes diet plan ・diet in diabetes ・diet plan for diabetes ・diet and diabetes ・diet for type 2 diabetes ・diabetic diet ・keto diet diabetes ・diabetes patient diet ・diabetes diet chart

これらの単語でYouTubeを検索し、言語が対象の言語ではない、タイトルから糖尿病に関連していないと考えられる、30分以上の長い動画、5年以上前にアップロードされた動画を除外し、日本語英語各々上位50本ずつを抽出した。抽出については、作業員1名が上記の明らかにタイトルや動画の長さ等が不適格な動画を除外して、上位の検索ワードのリストに下位の動画をマージし除外の可能性を考慮して50よりやや多い(60-90程度)数の動画のタイトルリストを作成した。そこからさらに複数名の作業員で不適格か判断の分かれる動画について含めるかの判断をし、50動画を確定した。

抽出した動画については、動画情報の記録日、動画URL、チャンネル登録者数、投稿日、動画時間、視聴回数、いいね数、コメント数、評価者、評価日、投稿者の職業、肩書、動画は特定の栄養療法を勧めているか?勧めてい

る・勧めていない食品、栄養素、**useful/not very useful** (糖尿病の予防や治療につながる食事について、科学的根拠を基にしたなんらかの情報が得られるか)と **modified DISCERN**、**JAMA**、**GQS** の 3 種類の国際的に動画の質の評価に使用されている指標について情報を収集した。**DISCERN** は健康情報の質を評価するために開発された指標で、今回は 5 項目で評価する改良版の **modified DISCERN** を使用した⁹⁾。項目としては D1. 目的が明確で、その目的が達せられる内容か D2. 信頼できる情報源が用いられているか D3. 提示されている情報はバランスがとれていて、偏りがないか D4. 視聴者が参照できるように、その他の情報源が示されているか D5. 情報の不確実性に言及しているかについて、糖尿病に関連する食事について専門的な知識を持っていると考えられる評価者 3 名で評価し、はいまたはいいえの人数が多い方をスコアとして合計 5 点までの評価をおこなった。**JAMA** は **Journal of American Medical Association score** で、上記の **modified DISCERN** と同様人数の多い方をそのスコアとした。項目としては、J1. 投稿者や協力者、所属先、関連する資格情報が提示されているか J2. すべての内容の参考文献や情報源、関連するすべての著作権情報が明記されているか J3. 動画の所有権、後援、広告、金銭的支援、営利目的の資金調達と支援、起りうる利益相反が、誰の目にも明らかかつ完全な形で開示されているか J4. 動画の投稿日や更新日が明記されているかを評価した。**GQS** については、1. 動画の質が低く、流れが悪い。ほとんどの情報が欠けており、視聴者にとってまったく有用でない。2. 概して動画の質が低めで流れが悪い。いくつかの情報は示されているが、多くの重要なトピックが欠けていて、視聴者にとっての有用性がかなり限られている。3. 動画の質は中程度であり、流れに改善の余地がある。いくつかの重要なトピックについては十分に論じているが、そのほかの重

要な情報については不十分である。視聴者にとっていくらかは有用である。4. 動画の質が良く、流れも概して良好である。適切な情報のほとんどが示されているものの、一部取り上げられていないトピックがある。視聴者にとって有用である。5. 動画の質が非常に高く、流れも非常に良い。視聴者にとって非常に有用である。の中から 3 人の評価者がそれぞれ一つを選択しその平均値をとった。

評価者は互いの評価を知ることなく評価を実施した。これらを日本語および英語の動画各々について実施した。評価者は糖尿病の食事および科学的根拠の判断に関して専門的知識を持っている者(研究者および医師、管理栄養士)であり、これらの評価者が **useful** かどうかを判断し、それと動画の質に関する客観的指標が一致する傾向であれば、専門的知識を持っている評価者と類似の評価を、専門知識を持たない人でも客観的指標を用いることで判断することが可能となる、ととらえて解析を実施した。

C. 結果

C-1. 対象の動画の特性

今回、タイトルでは明らかでなかったが動画の内容を検討した結果糖尿病に関する動画でないものが日本語の動画で 3 つ、英語の動画で 2 つあったため、日本語動画 47 本、英語動画 48 本が評価の対象となった。(表 1)

日本語の動画の約半数、英語動画の 3 割程度が専門家の中で **useful** と判断されたが、専門家の中でも、**not very useful** の動画の方が意見が一致しやすい傾向にあった。**Useful** の動画は投稿者が医療関係者と明らかにしている動画が多かった。英語の動画は視聴回数が多い動画については非常に登録者数が多い動画が視聴回数上位に含まれていたが、そうした動画は **not very useful** に分類されることが多かった。また、いいねやコメントの数も **not very useful** に分類される動画に多い傾向があ

った。

C-2. 各指標における動画の質の検討

結果は表2の通りである。日本語の動画で **modified DISCERN** の項目（目的が明確、信頼できる情報源、提示情報のバランス、その他の情報源、情報の不確実性）すべての項目において **useful** と判定された動画で **not very useful** とされた動画よりも「はい」と専門家が回答していた。英語の動画においては、その他の情報源の提示と情報の不確実性への言及については、**useful** と **not very useful** で統計学的有意差は認めず、**not very useful** の英語動画においても比較的「はい」と答えた評価者が多かった。日本語の動画においては **not very useful** の動画においてこの2項目を含め全体的に **not very useful** の動画は **modified DISCERN** の項目を含んでいない傾向にあったといえる。

JAMA については、投稿者等の所属、資格情報、文献や情報源、資金・利益相反開示、投稿日や更新日の項目が含まれ、日本語、英語の動画ともに **useful** の動画については **not very useful** よりも「はい」の項目が多かった。ただし、情報源の提示や利益相反の開示については英語の動画において **useful** と **not very useful** において統計的な有意差がなく、日本語の動画との性質の違いが認められた。

全体的な質を評価する **GQS** については英語の動画では統計的な有意差を認めなかったが日本語の動画では認めており、専門家の有用か否かの判断においては、動画の質や流れの良さの評価において差が生じることが確認された。

医療者か否かで各指標を検討した結果、日本語の動画では医療関係者の動画はそうでない動画と比較して、各指標のスコアが有意に高かったものの英語の動画ではその傾向が認められなかった。

その他、勧められているもしくは勧められて

いない食品や栄養素については、日本語の動画では日本で摂取される米や大豆などの食品が含まれており、英語の動画では日本では多くは摂取されない食品（種子類など）も多く含まれた。

D. 考察

専門の評価者が糖尿病に関する動画を評価し、**useful** と考えた動画は **not very useful** と考えた動画と比較して客観的な情報の質を評価する指標でも得点差があり、専門的な知識を持たない視聴者においても、これらの項目を指標として動画を判断することである程度有用と考えられる動画を判別することができる可能性が示唆された。一方、投稿者が医療関係者かどうかだけでは、日本の動画については、医療関係者の動画は現在各指標のスコアが高かったものの英語の動画にはその傾向が認められず、投稿者の背景だけで動画を判断することは、誤った動画を質が高いと考えてしまうことにもつながる可能性があり多角的に判断する必要性を示唆していると考えられた。動画内で勧められているもしくは勧められていない食品や栄養素については、日本語の動画では日本でよく摂取されている食品があげられやすいことから、情報を得たい日本人の視聴者は日本語の動画を見る可能性が高いことが示唆され、日本語の動画に対する評価方法を知ることが重要と考えられた。

modified DISCERN の項目（目的が明確、信頼できる情報源、提示情報のバランス、その他の情報源、情報の不確実性への言及）については、日本語の動画においてはすべての項目で **useful** の動画でスコアが高いことからこれらの項目は判断として有用と考えられる。一方で英語の動画では日本語の動画よりも情報源の提示と情報の不確実性への言及（**not very useful** の動画でも）が行われている傾向にあった。これらから、日本語英語

の動画ともに、D2.信頼できる情報源が用いられているか D3. 提示されている情報はバランスがとれていて、偏りがないかが差がつきやすく、質の判断として重要と考えられた。

JAMA の項目については、投稿者等の所属、資格情報、文献や情報源、資金・利益相反開示、投稿日や更新日の項目が含まれている。YouTube の特徴上、投稿日は明示されるためスコアとしては差がつきにくい、古い動画であることなどが明らかになるため、情報源として新しいかの判断に有用と考えられた。一方、利益相反については英語の動画の投稿者には、日本語の動画の投稿者よりも浸透していると考えられ *useful* と *not very useful* の動画で差がつきにくい項目となっていたが日本語の動画においてはこうした資金源や根拠となる情報が開示されているか否かも質の判断の根拠になると考えられた。

本研究は、動画の本数としては 50 本ずつという本数を対象としており、評価者も各動画 3 名ずつで評価をつけるなど十分と考えられる方法で日本語および英語の糖尿病に関する動画の質を評価した。糖尿病に関する動画であり、他の疾患では異なる可能性があるものの、すでに海外で健康情報の質の評価に使用される項目は日本語の糖尿病に関する食事についての動画を専門的な知識を持たない人が判断する際の一助になると考えられた。

E. 結論

糖尿病に関する日本語および英語の動画を YouTube から抽出し、客観的指標としてすでに動画の質の評価で用いられている *modified DISCERN*, *JAMA*, *GQS* を使用して評価した。その結果、専門家が有用であると判断した動画はこれらの指標でもスコアが高く、指標に含まれるような目的の明示、信頼できる情報源、提示情報のバランス、その他の情報源を提示する、情報の不確実性に言及する、投稿者の

属性、資金・利益相反開示、投稿日や更新日などの項目に基づき動画の情報を判断することで、医療の専門家でない人が視聴した場合にも有用と考えられる情報を見分けることができ、質の高いと考えられる動画を視聴できる可能性が高くなると考えられた。一方で投稿者の属性のみで判断するのは避けた方がよい可能性が高い。投稿者が自由に動画を発信できる動画共有サイトにおいても、健康情報を判断するのに既存の客観的指標の項目が使用できる可能性が示唆された。

F. 健康危険情報

なし

G. 研究発表

1. 論文発表

なし

2. 学会発表

なし

H. 知的所有権の出願・登録状況

1. 特許取得

なし

2. 実用新案登録

なし

3. その他

なし

I. 参考文献

1) International Telecommunication Union Statistics

URL:<https://www.itu.int/en/ITU-D/Statistics/Pages/stat/default.aspx>

2) ICT 総研 2022 年度 SNS 利用動向に関する調査 URL: <https://ictr.co.jp/report/20220517-2.html/>

3) YouTube YouTube の取り組み URL:

<https://www.youtube.com/howyoutubeworks/>

- 4) Calvin C et al. The Reliability and Quality of YouTube Videos as a Source of Public Health Information Regarding COVID-19 Vaccination: Cross-sectional Study. *JMIR Public Health Surveill.* 2021;7:e29942.
- 5) Mustafa EO et al. Quality of English-language videos available on YouTube as a source of information on osteoporosis. *Arch Osteoporos.* 2022;17:19.
- 6) Charnock D et al. DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. *J Epidemiol Community Health.* 1999;53:105–111.
- 7) André Bernard et al. A systematic review of patient inflammatory bowel disease information resources on the World Wide Web. *Am J Gastroenterol.* 2007;102:2070-7.
- 8) Silberg WM et al. Assessing, controlling, and assuring the quality of medical information on the Internet: Caveant lector et viewor--Let the reader and viewer beware. *JAMA* 1997;277:1244-5.
- 9) Singh AG et al. YouTube for information on rheumatoid arthritis—a wakeup call? *J Rheumatol.* 2012;39:899–903.

表 1 選択した動画の性質

	日本語の動画			英語の動画		
	Total	Useful	Not useful	Total	Useful	Not useful
n	47	23	24	48	18	30
意見が一致した数	20	4	16	17	3	14
投稿者の職業・背景						
医療者	24	17	7	21	10	11
患者	9	0	9	1	0	1
その他	2	1	1	5	1	4
不明	12	5	7	20	7	13
登録者数(千人)						
平均	164.9	163.5	166.3	3895.5	172.0	5200.8
標準偏差	244.8	202	283.6	6886.3	272.8	8233.6
中央値	60	40	80	2460	63.5	3375.0
四分位	30	20	30	458.5	36.7	5070.0
	200	360	145	457.0	218	5410.0
いいね数						
平均	8050.3	4804	11161.4	47472.9	34261.2	55400
標準偏差	9008	4358.9	11112	35012.1	17654.5	40381.5
中央値	3833	3430	6258	37000	35500	42000
四分位	2708	2093	3027.5	25500	26000	25000
	11000	5015	15000	66000	39000	71000
再生回数						
平均	846,908	485,246	1,193,501	2763643	1772797	3358151
標準偏差	1,045,151	387,860	1,335,903	2297959	499524.4	2727079
中央値	450947	296702	607207	1959461	1799408	2503508
四分位	279522	230105	358532	1600952	1527639	1633644
	883706	642358	1653365	2881688	2026492	3837237
コメント数						
平均	290.8	121.5	432	2190.2	1531	2563.7
標準偏差	426.6	90.6	536.4	2086.4	783.1	2483.5
中央値	173	100	227	1756	1362	1930.5
四分位	86.5	58.5	164	1040	1040	1206.0
	297.5	180	402.5	2575	2035	3156.0
動画時間 (分)						
平均	10.8	11	10.7	12.9	16.6	10.7
標準偏差	5.6	6.2	5.1	7.5	7.0	7.0
中央値	9.13	9.13	9.13	10.6	16.0	9.3
四分位	7.15	6.1	7.6	7.0	12.0	6.4
	13.2	16.1	11.2	18.8	20.4	12.3

表2 日本語・英語動画における客観的指標による質の評価

	日本語				英語			
	Total	useful	not very useful	P 値	Total	useful	not very useful	P 値
DISCERN								
平均スコア	1.66	2.51	0.85	<0.0001	1.71	2.41	1.29	<0.0001
標準偏差	1.17	1	0.61		0.85	0.75	0.59	
D1	0.62 (0.33)	0.78 (0.22)	0.46 (0.34)	0.0003	0.66 (0.32)	0.83 (0.17)	0.56 (0.34)	0.0005
D2	0.30 (0.33)	0.57 (0.27)	0.06 (0.13)	<0.0001	0.17 (0.32)	0.41 (0.39)	0.03 (0.13)	0.0008
D3	0.28 (0.30)	0.43 (0.27)	0.14 (0.26)	0.0004	0.22 (0.27)	0.43 (0.28)	0.10 (0.18)	0.0001
D4	0.35 (0.34)	0.55 (0.36)	0.15 (0.17)	<0.0001	0.49 (0.30)	0.52 (0.37)	0.47 (0.26)	0.60
D5	0.11 (0.23)	0.17 (0.30)	0.04 (0.11)	0.06	0.17 (0.19)	0.22 (0.16)	0.13 (0.21)	0.11
JAMA								
平均スコア	1.75	2.19	1.33	<0.0001	2.16	2.48	1.96	0.01
標準偏差	0.73	0.63	0.56		0.70	0.54	0.73	
J1	0.41 (0.44)	0.62 (0.44)	0.19 (0.32)	0.0005	0.49 (0.46)	0.59 (0.47)	0.42 (0.45)	0.22
J2	0.12 (0.22)	0.23 (0.27)	0.01 (0.07)	0.001	0.12 (0.29)	0.26 (0.41)	0.33 (0.13)	0.03
J3	0.23 (0.29)	0.33 (0.28)	0.14 (0.28)	0.02	0.56 (0.38)	0.63 (0.38)	0.51 (0.38)	0.30
J4	0.99 (0.07)	1	0.97 (0.09)	0.16	1	1	1	-
GQS								
平均スコア	2.29	2.94	1.67	<0.0001	2.39 (0.69)	3.07 (0.39)	1.98 (0.45)	<0.0001
標準偏差	0.85	0.57	0.55					

各指標の小項目の数値は平均(標準偏差)、p値は useful と not very useful のスコアの比較