

包絡分析法の予防医学への応用の挑戦

中村 翔, 小笠原 悠, 成松 宏人

包絡分析法では会社など意思決定を行う事業体の効率性を対象とするが、事業体を個人の身体として効率性スコアを算出することもできる。われわれは肥満や高血圧・脂質異常症に注目し、入力/出力情報に対して得られた効率性を個人のもつ発症リスクと解釈することで、包絡分析法を予防医学に対して応用することを試みた。ここでは摂取カロリーや消費カロリーなどの情報を入力情報、BMI や血圧などの測定系の指標を出力情報としている。得られた効率性スコアを従来の一次予防のようなポピュレーションアプローチではなく個別医療に向けた提案を実現するため、介入研究を通して行動変容への有用性を示す必要がある。

キーワード：包絡分析法, DEA, 予防医学, ゲノム医療

1. はじめに

日本人の3大死因は老衰を除くと悪性新生物、心疾患、脳血管疾患であり [1]、本邦において3大疾患として広く認識されている。この3大疾患は生活習慣病とも呼ばれ、その名のとおりわれわれの生活習慣がその発症に大きく寄与することがわかっている。たとえば、多くの人にとってより身近な高血圧、糖尿病、脂質異常症（高コレステロール血症）、肥満症などの有病率の高い生活習慣病も、直接または続発・併存する疾患を通して、3大疾患をはじめとする life-threatening な病気の罹患リスクを増加させる。

生活習慣病は各個人が生活習慣に留意することで発症リスクを低減することができる。生活習慣病に至っていない未発症者の発症を予防することを一次予防という。健康日本 21 などこの一次予防を念頭においており [2, 3]、留意すべき生活習慣として健康的な食事・栄養摂取、身体活動量、飲酒、喫煙といった生活習慣が挙げられる。

これらの留意すべき生活習慣が広く知られている一方で、普段から健康に気をつけている、健康を意識している人は全年齢で約半分くらいであるとの報告もある [4]。何もやっていない人のうち約2割は健康的な生活習慣といわれているものがあまりに一般的で個別最適

化されていないために「何をどのようにやったらよいかわからない」状況であり、「特に理由はない」あるいは「健康なので特に何もする必要はない」ので健康を意識していないと回答している人が4割にのぼる [4]。健康診断などでは正常範囲であると判断され健康に関して専門職などから一般的な情報提供にとどまる可能性が高い集団はその時点では健康であるが故に、現状では何も改善する必要がないと感じやすい。重要なのは、このような集団が一次予防の対象集団であり、一次予防が目的とする健康な人が健康なままでいることを実現することが必要な集団であるということである。

現在の一次予防はいわゆる健康とされる集団全体に介入を行うポピュレーションアプローチが一般的である。ポピュレーションアプローチの難しさの一つに、集団が大きいことと、対象者の行動変容を惹起するための動機付けに、たとえば何かしらの疾患に罹患している健康に気を付けるインセンティブを自己で保有する人と比較して、多大なエフォートを要することである。介入に充てられるリソースは限られていることが多いため、集団における介入対象の規模と介入の程度のトレードオフが発生し、集団全体で見たときに効果的かつ効率的に一次予防を提供する手法が未だ確立されていない。

われわれは現状の健康診断の検査項目のカットオフでは健康と判定される集団を対象とし、現状のリスク評価法では健康と判定されてしまうが潜在的な将来の疾患リスクを有する集団を、包絡分析法 (Data Envelopment Analysis; 以下 DEA) によって判別できるのではないかと考えこれまで研究を行ってきた。本稿では、われわれのこれまでの研究を紹介するとともに、今後の展望について紹介する。

なかむら しょう, なりまつ ひろと
 神奈川県立保健福祉大学ヘルスイノベーション研究科
 〒210-0081 神奈川県川崎市殿町 3-25-10 Research Gate Building 2-A
 神奈川県立がんセンター臨床研究所がん予防・情報学部
 〒241-8515 神奈川県横浜市旭区中尾 2-3-2
 research@nakasho.org
 hiroto-narimatsu@umin.org
 おがさわら ゆう
 東京都立大学都市環境学部
 〒192-0397 東京都八王子市南大沢 1-1
 ogayu@tmu.ac.jp

2. 肥満に対する DEA の応用

われわれが初めて DEA を個人に応用することを試みたのは、2015 年 5 月に出版された PLOS ONE の論文であった [5]. 帝京大学大学院公衆衛生学研究所の中田 善規教授との共同研究によって decision making unit (DMU) を個人として扱い、DEA を適用するアイデアが生まれた。ここでは最も基本的なモデルの一つである CCR モデル (入力指向) を使用している。ある事業体 o に対する効率性スコア θ を表す CCR モデルは以下で表される [6].

$$\begin{aligned} & \min_{\theta, \lambda_1, \dots, \lambda_n} \theta \\ \text{subject to} & \\ & \sum_{j=1}^n x_{ij} \lambda_j \leq \theta x_{io}, \quad i = 1, \dots, m; \\ & \sum_{j=1}^n y_{rj} \lambda_j \geq y_{ro}, \quad r = 1, \dots, s; \\ & \lambda_j \geq 0, \quad j = 1, \dots, n. \end{aligned}$$

ここで、 x_{ij} と y_{rj} は事業体 j における入力 i と出力 r の数量を表している。ただし、これらの入力と出力それぞれの要素の少なくとも一つは正の値を取ることが仮定されている。制約式中の $\lambda_j, j = 1, \dots, n$ は CCR モデルにおける生産可能集合を表す変数であり、 $\lambda = (\lambda_1, \dots, \lambda_n)$ は intensity vector と呼ばれる。CCR モデルの詳細や生産可能集合の定義についてはたとえば文献 [6] を参照してほしい。

前述の論文 [5] における CCR モデルの入力変数は摂取カロリーの逆数と消費カロリー、出力変数には BMI の逆数を適用している。この効率性スコアは、摂取カロリーが多く、消費カロリーが少なくても BMI が低い状態を実現するような個人のもつ特徴、すなわち個人が有する肥満リスクに対応する評価指標と解釈することが可能である。個人が先天的にもっている肥満リスクの一つとしては遺伝的特徴、つまり、生まれつき肥満になりやすい・なりにくい、というものが挙げられる。そこでわれわれは肥満の遺伝学的リスク指標を比較対象とし、結果として、CCR モデルによって得られた効率性スコアは個人に対する肥満のリスクを評価する際に用いることができる可能性を示唆した。

2.1 肥満の遺伝学的リスク指標

ここでは、肥満の遺伝学的リスク指標について簡単に説明する。人の保有する全ゲノムの配列を決定するための国際共同研究プロジェクトであるヒトゲノム計画は 2000 年に完了した [7]. それ以降、人のさまざま

な形質について、遺伝学的な要因がどの程度影響しているかに関する研究が行われてきた。その一つにゲノムワイド関連研究 (genome-wide association study; GWAS) がある。GWAS は、興味対象の形質に関連している一塩基多形 (single nucleotide polymorphism; SNP) を明らかにすることが可能であり、これまでに数多くの SNPs と形質との関連がジェノタイプング技術の向上やコストの低下などによって報告されてきた [8].

われわれが用いた肥満の遺伝的リスク指標は、この SNPs の計算に基づいている。肥満の遺伝学的リスク指標は BMI との関連が明らかとなっている SNPs を対象に、先行 GWAS で明らかとなっている各 SNPs の回帰係数で加重和を取ったものであり、この値は遺伝学的リスクスコアと呼ばれる。この肥満の遺伝学的リスクスコアと BMI が関連することはこれまでの研究で明らかとなっているが、将来の BMI 変化量との関連はまだ十分明らかとなっていないことや、遺伝学的リスクスコアによって実際に肥満となるかを予測するには精度が十分ではないことに注意しなくてはならない [9].

3. 高血圧・脂質異常症に対する DEA の応用

肥満に対する DEA の応用を行った後、個人を DMU とした DEA によって疾患リスクの評価を試みた形質が高血圧と脂質異常症である [10]. ここでは、この高血圧と脂質異常症に対して DEA の応用を行って得られた結果を紹介する。この研究は、山形大学医学部メディカルサイエンス推進研究部が実施している山形県コホート研究 (高島) のデータを用いて実施された [11]. 山形県コホート研究は 1979 年に舟形町で始まった糖尿病健診に端を発し、21 世紀 COE プログラム (2004–2008) およびグローバル COE プログラム (2008–2012) に採択されたことをきっかけに開始された分子疫学コホート研究である [12]. 研究の目的は、人々の生活習慣と体質である遺伝学的情報がどのように組み合わせるとがんといった生活習慣病を引き起こすのかを明らかにすることである。また、2005 年に名古屋大学医学部を中心として開始されたゲノムコホート研究である日本多施設共同コホート研究 (Japan Multi-Institutional Collaborative Cohort [J-MICC] Study) と共通のプラットフォームで実施されている [13]. 山形県コホート研究 (高島) は、この山形県コホート研究のうち、山形県東置賜郡高島町にて 2004 年から 2006 年に実施されたベースライン調査で研究協力者となった方々を対象としており、2011 年に二次調査が実施されている。

表1 高血圧と脂質異常症モデルの入出力変数

変数名	対象疾患	
	高血圧	脂質異常症
入力変数		
食塩摂取量*	○	—
総エネルギー摂取量*	○	○
身体活動量	○	○
出力変数		
収縮期血圧*	○	—
拡張期血圧*	○	—
中性脂肪*	—	○
HDL コレステロール	—	○
LDL コレステロール*	—	○

*これらの変数は逆数が用いられた。

われわれはこの山形県コホート研究（高島）のベースライン調査時のデータに DEA を適用し、効率性スコアが 2011 年の二次調査までの高血圧および脂質異常症の新規発症と関連するかを見ることで、効率性スコアをこれらの疾患のリスクスコアとして用いることができるかを検証した。

この研究で解析した高血圧と脂質異常症の DEA モデルに用いられた入出力変数を表 1 に示す。いずれも、入力指向型の CCR モデルで解析した。なお、効率性スコアおよび発症データがともに使用可能なデータ数は高血圧と脂質異常症でそれぞれ 520 人分と 584 人分であった。

得られた効率性スコアは先ほどの肥満に対する DEA の適用の研究 [5] と同様、効率性スコアが高いほど疾患に対する耐性がある（発症リスクが低い）と考えることができることから、それぞれの対象疾患に対する個人のもつ発症リスクと関連した指標と解釈できる。そこでわれわれは得られた効率性スコアが実際に発症リスクを表しているかどうかを検証した。検証には疾患の 5-8 年以内の発症を表したダミー変数を被説明変数とし、効率性スコアを説明変数、そして、年齢、性別、BMI を調整変数としたロジスティック回帰分析を用いた。その結果、ベースライン調査時点でいずれの疾患も有していなかった集団において、効率性スコア 0.10 点上昇による調整後オッズ比は高血圧に関して 0.66（90% 信頼区間 0.55-0.78）、脂質異常症に関して 0.84（90% 信頼区間 0.75-0.94）であった。これは、効率性スコアが 0.1 上昇すると、発症有無の比（発症/非発症）が高血圧で 0.66 倍、脂質異常症で 0.84 倍と低下することを表している。すなわち、被説明変数の疾患の有無でオッズ比とその信頼区間が 1.0 より低

い（高い）ことは、効率性スコアの点数が上昇するほど疾患を発症しづらい（しやすい）ことを意味する。

同論文ではさらに特定健康診査で評価されるリスク（従来リスク評価）と効率性スコアを組み合わせた層別化解析を実施している。従来リスク評価には、特定保健指導の区分判定に用いられている以下の 11 項目を取り上げ、その個数で 4 群に層別化している。11 項目は、収縮期血圧・拡張期血圧 ($\geq 130/85$ mmHg)、高血圧症治療中、空腹時血糖 (≥ 100 mg/dl)、HbA1c ($\geq 5.6\%$)、糖尿病治療中、中性脂肪 (≥ 150 mg/dl)、HDL コレステロール (< 40 mg/dl)、脂質異常症治療中、BMI (≥ 25 kg/m²)、喫煙歴有りである。この従来リスクの一つも当てはまらない人を低リスク、一つ当てはまる人を中間リスク、二つ当てはまる人を高リスク、三つ以上当てはまる人を最高リスク群としている。

高血圧に関しては、低リスクから高リスク群までは、層別前の解析と同様に、将来の発症リスクに関するオッズ比は、効率性スコアの点数が高いほど低いという関連が見られた。一方で、最高リスク群においては、効率性スコア 0.10 点上昇による調整後オッズ比が 0.82（90% 信頼区間 0.59-1.13）であり、この関連が観察されなかった。ここで重要なのは、われわれが DEA を予防医療に適用しようと試みた一番の目的が、現在は特定健康診査をはじめとした健康診断でリスクが検知されない集団において、将来の疾患リスクが高い集団を判別することであった点である。この点で、最高リスク群はこれまでもたとえば「黄色信号群」、「予備軍」といった形で介入がなされてきた可能性が高い集団であり、一定程度リスクの可視化ができていた集団である。対して、低リスクから高リスク群、特に従来リスクの一つも有していない低リスク群はこれまでリスクの可視化ができず、それが故に自治体の保健事業の介入対象にもなりにくかった集団である。このような集団において、DEA による効率性スコアが将来の疾患発症リスクと関連していることを明らかにした点において、一次予防の実現に向けて一歩進んだと思える結果であった。

脂質異常症については、先に述べたとおり、集団全体では効率性スコアは将来の脂質異常症の発症と関連していたものの、従来リスクでの層別化解析ではどのサブグループにおいても疾患発症との関連が示されなかった。また、従来リスクの 4 群において低リスク群を参照群としたときに、高血圧に関しては中間リスクから最高リスク群になるにつれて疾患発症のオッズ比が増加していくのに対して、脂質異常症については、従来リスクが増加してもどの群においてもオッズ比が

1.8 程度であった。原因として考えられる理由の一つは、脂質異常症という疾患概念のまとめ方である。脂質異常症には中性脂肪が高い状態、LDL コレステロールが高い状態、HDL コレステロールが低い状態が考えられ、これら複数の病態をまとめてしまっていることが原因となっている可能性がある。そのほかには、従来リスク評価に含まれる耐糖能、BMI、喫煙歴と血清コレステロール値は続発性脂質異常症との関連が知られているとおり、従来リスクの評価項目自体が直接的に血清コレステロール値に影響を与えることがある。高血圧であれば、従来リスクに食塩摂取量が含まれていることと同等の影響を及ぼしている可能性がある。このような理由で、脂質異常症では層別化解析で高血圧と同じような結果が観察されなかった可能性がある。

われわれはこの山形県コホート研究によって得られた結果を基に、DEA による効率性スコアを予防介入に応用する前向き介入研究を実施した。

4. 高血圧 DEA モデルの予防介入研究

前節で紹介した山形県コホート研究の成果を用いた研究を基に、山形県東置賜郡高島町にて介入研究を実施した [14]。町の特定健康診査を受診した方で、特定保健指導区分が情報提供となった方、かつ、研究参加時点で高血圧の基準を満たさず、既往歴もない方を対象にした介入研究である。介入は先ほどの高血圧の変数を採用した DEA の CCR モデルから得られた効率性スコアを用いた将来の高血圧発症リスクの通知であった。研究結果が有用だと判明したのちに、社会実装が可能な介入とするため、介入群のうち高血圧発症リスクがおおよそ集団の 50% より高い方にリスクを郵送にて通知するという介入を行った。この結果は既に取りまとめ済みであり、本原稿執筆時点において査読付き英文論文誌での査読結果にもとづき、論文を修正中であることから、近日中に広く結果が公開できることを予定している。

5. さまざまな DEA モデルの適用の可能性と問題

これまで使用したモデルはすべて入力指向の CCR モデルという DEA における最も基本的なモデルの一つを用いたものだが、ほかにも DEA には BCC モデルや SBM モデル、Network DEA、Fuzzy DEA などさまざまなモデルや手法が提案されている。先に述べた脂質異常症など、CCR モデルで従来手法とそれほど変わらない結果が得られた疾患であっても、ほかの

モデルを適用することで、新たな傾向が得られる可能性がある。

その一方で、肥満や高血圧、脂質異常症の人に実際に介入する主体は、特定健康診査では保健師や管理栄養士の方の役割が中心となることや、予防医学は特定の地域に限らず広い適用範囲をもつことから、社会実装を考慮した場合には、実装先の組織あるいは業務担当者がリスク評価の結果を理解でき、納得したうえで、正しい解釈に基づいた介入がしやすいシンプルなモデルであることが望ましい。

さらに、使用できるデータの制限とエビデンスの問題がある。Network DEA のように入力と出力の関係を精緻にしていく方向性も当然存在する。しかし、高血圧や脂質異常症などの一次予防のデータは基本的には特定健康診査などの健診データが元になるため、一般的なデータ項目を変数としてモデルを構築することが必要であり、従来の区分判定の基準も存在することから、複雑な構造を取り入れるには注意が必要である。これは構築したモデルの汎用性に関わる問題であることから、一般的にさまざまな分野で広く見られる問題であろう。もう一つのエビデンスの問題とは、入力と出力の関係を複雑にしていくほど、入力と出力の関係性に対してより細かな医学側のエビデンスが必要になるということである。たとえば、先述した肥満の遺伝学的リスク指標は BMI に関連した 75 個の SNPs の情報から算出されているが、たとえば肥満遺伝子として有名な *FTO* 遺伝子のような遺伝子上の SNPs など、具体的にどのように肥満と関連するかが明らかになっているものもあれば [15]、75 個の SNPs の中には、肥満症という表現系に関する生物学的機序がまだ明らかになっていない遺伝子や、遺伝子のコーディング領域以外の SNPs も含まれている。そのため、遺伝学的リスク指標を使用するのではなく、各 SNPs を 0-1 の固定した値として DEA のモデルに組み込むというモデル構築の方向性には注意が必要になる。このように、既にガイドラインなどの施策に反映されている内容や高いエビデンスレベルが得られているかを加味してモデルや入出力の変数を選択しなければならない。

6. まとめと今後の展望

本稿では、われわれがこれまで取り組んできた予防医学へ DEA を応用した一連の研究とその成果を紹介した。現在も社会実装に向けた検証を実施しており、本稿では紹介できなかったが、CCR 以外のモデルの適用や、遺伝情報を用いた個別化医療に向けた DEA

の適用も進めている。今後これらの結果が得られ次第、広く予防医学分野に公表していくつもりである。

一方で、ほかの疾患への応用やさまざまなモデルの適用の模索は限定的であり、まだまだ予防医学へのDEAの応用というアプローチは発展段階である。今後の発展のためには、前節で述べたとおりORと医学両方の知見が必要なことから、異なる分野の研究者間の連携が不可欠である。よって、今後より一層、ORと医学研究者が議論を行うことができる場が広がることを期待したい。

参考文献

- [1] 政府統計の総合窓口 (e-Stat), 「調査項目を調べる—人口動態調査 (厚生労働省)」, <https://www.e-stat.go.jp/> (2023年2月4日閲覧)
- [2] 厚生労働省, 「生活習慣病予防」, https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/kenkou_iryoku/kenkou/seikatsu/seikatsuyuukan.html/ (2023年2月4日閲覧)
- [3] 厚生労働省, 「健康日本21 (第二次)」, https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/kenkou_iryoku/kenkou/kenkounippon21.html (2023年2月4日閲覧)
- [4] 厚生労働省, “第2章 健康をめぐる現状と意識”, 平成26年版厚生労働白書—健康長寿社会の実現に向けて—健康・予防元年, 2014.
- [5] H. Narimatsu, Y. Nakata, S. Nakamura, H. Sato, R. Sho, K. Otani, R. Kawasaki, I. Kubota, Y. Ueno, T. Kato, H. Yamashita, A. Fukao and T. Kayama, “Applying data envelopment analysis to preventive medicine: a novel method for constructing a personalized risk model of obesity,” *PLOS ONE*, **10**, e0126443, 2015.
- [6] W. W. Cooper, L. M. Seiford and J. Zhu (eds.), *Handbook on Data Envelopment Analysis*, Springer, 2011.
- [7] International Human Genome Sequencing Consortium, “Initial sequencing and analysis of the human genome,” *Nature*, **409**, pp. 860–921, 2001.
- [8] The GWAS Catalog, <https://www.ebi.ac.uk/gwas/> (2023年2月4日閲覧)
- [9] R. J. F. Loos and G. S. H. Yeo, “The genetics of obesity: from discovery to biology,” *Nature Reviews Genetics*, **23**, pp. 120–133, 2022.
- [10] S. Nakamura, H. Narimatsu, Y. Nakata, M. Sakaguchi, T. Konta, M. Watanabe, Y. Ueno, K. Ishizawa, H. Yamashita, T. Kayama and T. Yoshioka, “Efficiency score from data envelopment analysis can predict the future onset of hypertension and dyslipidemia: A cohort study,” *Scientific Reports*, **9**, 16309, 2019.
- [11] Yamagata University Genomic Cohort Consortium and H. Narimatsu, “Constructing a contemporary gene-environmental cohort: study design of the Yamagata Molecular Epidemiological Cohort Study,” *Journal of Human Genetics*, **58**, pp. 54–56, 2013.
- [12] 山形大学医学部メディカルサイエンス推進研究所, 「山形県コホート研究」, <https://www.id.yamagata-u.ac.jp/IPMSR/project.html> (2023年2月4日閲覧)
- [13] K. Takeuchi, M. Naito, S. Kawai, M. Tsukamoto, Y. Kadomatsu, Y. Kubo, R. Okada, M. Nagayoshi, T. Tamura, A. Hishida, M. Nakatochi, T. Sasakabe, S. Hashimoto, H. Eguchi, Y. Momozawa, H. Ikezaki, M. Murata, N. Furusyo, K. Tanaka, M. Hara, Y. Nishida, K. Matsuo, H. Ito, I. Oze, H. Mikami, Y. Nakamura, M. Kusakabe, T. Takezaki, R. Ibusuki, I. Shimoshikiryō, S. Suzuki, T. Nishiyama, M. Watanabe, T. Koyama, E. Ozaki, I. Watanabe, K. Kuriki, Y. Kita, H. Ueshima, K. Matsui, K. Arisawa, H. Uemura, S. Katsuura-Kamano, S. Nakamura, H. Narimatsu, N. Hamajima, H. Tanaka and K. Wakai, “Study profile of the Japan multi-institutional collaborative cohort (J-MICC) study,” *Journal of Epidemiology*, **31**, pp. 660–668, 2021.
- [14] H. Narimatsu and S. Nakamura, “Single-blind randomized parallel group trial on the effect of primary preventive intervention using data envelopment analysis aiming the reduction of the onset of hypertension,” *UMIN CTR*, UMIN Clinical Trials Registry number: UMIN000037883, 2019.
- [15] C. Church, L. Moir, F. McMurray, C. Girard, G. T. Banks, L. Teboul, S. Wells, J. C. Brüning, P. M. Nolan, F. M. Ashcroft and R. D. Cox, “Overexpression of FTO leads to increased food intake and results in obesity,” *Nature Genetics*, **42**, pp. 1086–1092, 2010.