

日本医療研究開発機構
予防・健康づくりの社会実装に向けた研究開発基盤整備事業
ヘルスケア社会実装基盤整備事業 事後評価報告書



I 基本情報

研究開発課題名: (日本語) 予防・健康づくりサービスの平均効果と異質効果の推計デザインとその実装に関する研究
(英語) Developing and Implementing the Statistical Procedures Evaluating the Average and Heterogeneous Effects of Prevention and Health Promotion Services

研究開発実施期間: 令和4年9月26日～令和7年3月31日

研究開発代表者 氏名: (日本語) 近藤尚己
(英語) Kondo Naoki

研究開発代表者 所属機関・部署・役職:
(日本語) 京都大学・大学院医学研究科・教授
(英語) Kyoto University, Graduate school of Medicine, Professor

II 研究開発の概要

健康行動の変容やそのための社会環境整備を目指す「予防・健康づくり」のプログラムやツールが急速に発展している。しかし、医薬品等と異なり、こういったプログラムにはターゲットとなる利用者の属性による効果の異質性やアドヒアランスの低さ、対象者の属性が多様といった特徴があり、RCTのみに依存した効果評価はなじまない。そこで、予防・健康づくりプログラムの統計的評価方法に関して、文献レビュー（文献研究）、既存の介入・観察研究データの二次分析（分析研究）を行い、これらの成果をもとに政策に資するツール提供（政策研究）を行った。

文献研究: 予防・健康づくりプログラムの評価に適した研究デザインの概要、仮定、メリット・デメリットを体系的に整理・解説した。クラスターRCTや回帰不連続デザイン、差分の差分法などの疑似実験的手法に加えてTarget Trial Emulationや機械学習を用いた効果の異質性評価の詳細を、文献レビューをもとに整理した（図1）。

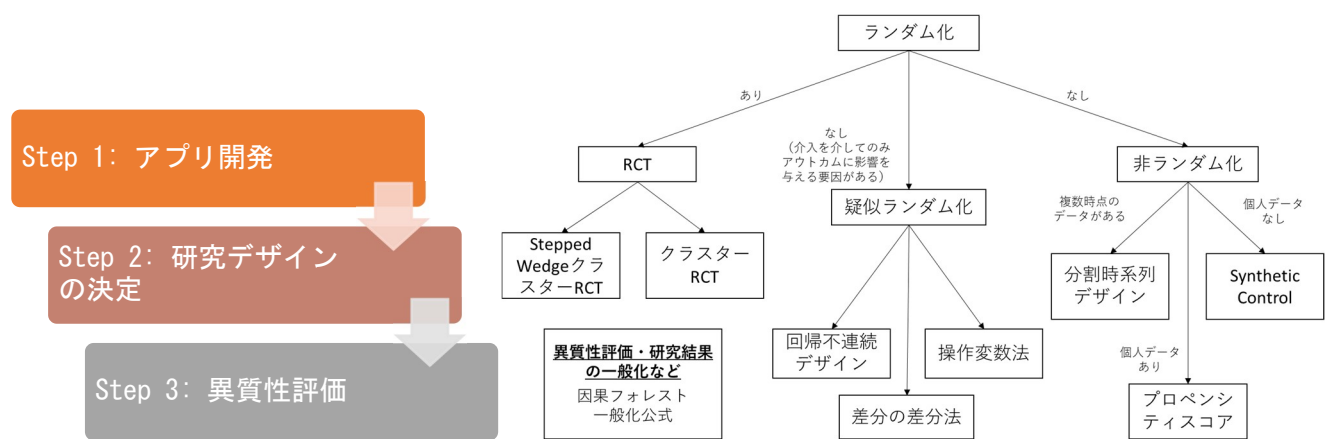


図1 予防健康づくりプログラムの評価ステップと評価手法選択フローチャート（レビュー結果の抜粋）

分析研究：A 社による行動経済学やゲーミフィケーションを応用したエンターテインメント型の健康づくりプログラム「健診戦」の因果効果評価を行い、社員の属性（職位・健康関心度など）によらず一定の効果があることが明らかになった。異質性評価をフォーマルに行うことがサービスの効果の公平性評価にも貢献しうることの証左となり、異質性評価の重要性が確認された。

B 社が提供する AI 健康管理アプリ「カロママ プラス」の効果に関する RCT の結果を Target Trial Emulation (TTE) を応用して観察データで再現することに成功した（表 1）。TTE が一部有益である一方、限界点も明らかとなった。さらに、B 社及び C 社とともに、高血圧リスク遺伝情報を活用したカロママ プラスの介入効果に関する RCT を実施・完了した。カロママ プラスの効果が出やすい集団を同定するための要因分析を実施して国際誌に報告した。また、特定保健指導の動機づけ支援のデータセットを使用し、運動習慣を獲得しやすい集団を同定するための要因分析を実施し国際誌で報告した。

RCTの介入群における体重減少 (95%信頼区間)	TTE [重み付けあり] (95%信頼区間)	TTE [重み付けなし] (95%信頼区間)
-2.4kg (-3.35, -1.45)	-2.93kg (-3.00, -2.86)	-2.33kg (-2.39, -2.27)

表 1 TTE 評価結果に基づき、Nakata et al のランダム化比較試験データをカロママ プラスのログデータ（観察データ）に当てはめて因果効果を推計した結果の比較

加えて、予防・健康づくりサービスに特有の介入評価における課題に対応した「拡張型実験デザイン」の開発と検証を進めた：①ランダム割けが困難な状況において疑似実験デザインでの介入評価、②介入の利用行動がコントロール困難な状況において、Nested RCT で介入を実装し、受療後の行動変化がアウトカムに与える影響を検証し、適切な結果が算出できることを確認した(図 2)。複数の論文で成果公表した。

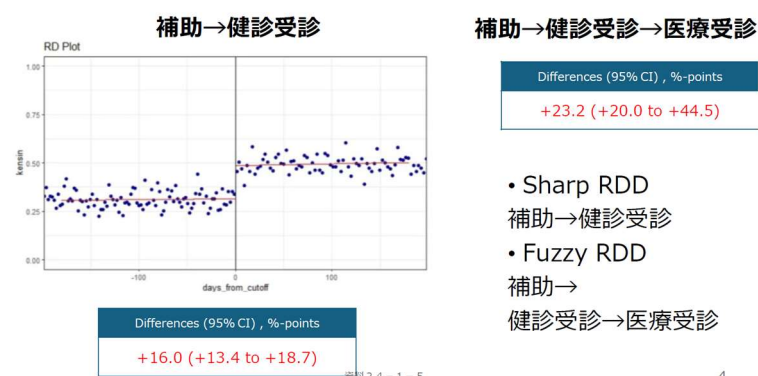


図2 外生的な割り付け変数（生年月日）を用いた非ランダム割り付け介入データを用いた疑似実験デザインによる因果効果の推計結果

令和5年度から追加した、健康管理アプリ「カロママ プラス」利用の異質効果推計の分析においては、家族歴や普段の歩行習慣等により効果が異なることが明らかとなり、予防・健康づくりプログラムの設計と評価の際には異質効果を想定することの重要性をあらためて浮き彫りにした（表3）。

変数	オッズ比 (95%信頼区間)	p 値
歩行習慣	0.248 (0.079, 0.786)	0.018
歩行速度	0.324 (0.105, 1.004)	0.051
家族歴	4.269 (0.972, 18.746)	0.055

調整変数：性別、年齢、減量体験、体重変動、運動習慣、既往歴、教育歴、年収、同居状況、婚姻状況、最初の1週間の各項目のアプリ入力回数

表3 カロママ プラスの効果が特に強い可能性のある属性変数の抽出結果（3%以上の減量成功の多変量調整オッズ比）

政策研究：文献研究・実証研究各班の知見を参考にしながら、各ステークホルダーへのインタビュー調査・コンソーシアムでのワークショップ、産業医や第三者（健康教育学会学術大会ラウンドテーブル）からのフィードバック（図3）を経て作成・改訂した「保健事業導入のための購入者向けチェックリスト」案を実装コンソーシアム内で検討・開発し、出版した（図4）。開発者版（図5）も開発した。これらを活用するためのマニュアルについて、産業保健の専門家によるエキスパートレビューや健康教育・ヘルスプロモーションの専門家による助言を受けつつ作成した（予防・健康づくりサービスの品質チェックリスト利用の手引き）。また、これまでに記録された実装コンソーシアムや専門家からの意見を整理し、合意形成プロセスを視覚化し、またこれらツールのフィージビリティを評価し、ツールの改善を行った。

インタビュー結果

A社+B社+C社

抽出した複合語：

- ・行動変容
- ・ヘルスサービス
- ・コアデザイン原則
- ・モデル理論
- ・認知行動療法
- ・保健指導
- ・厚労省
- ・産業医
- ・アウトカム

…サービス開発に際して留意していること

…開発したサービスの説明

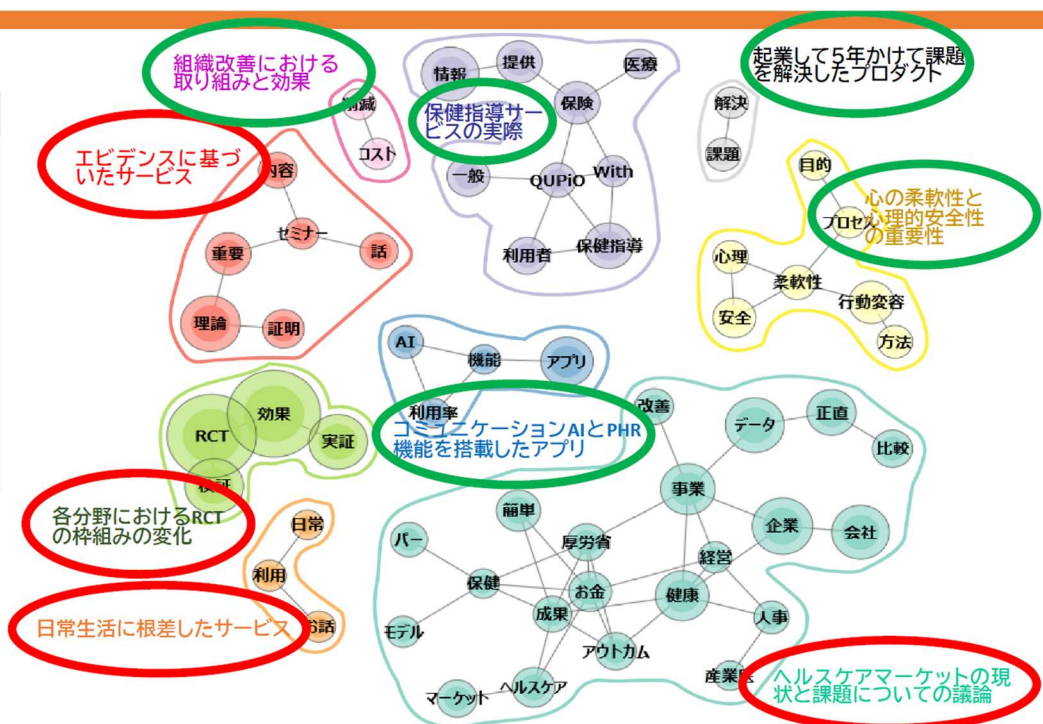


図3 予防健康づくりサービス開発に関する各ステークホルダーが留意している要因や開発物への認識に関するインタビューデータの分析結果

IV. チェックリスト

安全で効果のある予防・健康づくりサービスって何？

予防・健康づくりサービス[※]の品質チェックリスト

導入を検討している予防・健康づくりサービスについて№1～9の項目をチェックしてみましょう。（№10はオプション項目です。）

※ ここでの「予防・健康づくりサービス」とは、法令に基づく特定保健指導やストレスチェックのようなものから、企業や保険者が加入者・従業員のリスクの状況に合わせて行う事業（例、肥満対策や睡眠習慣改善、栄養指導など）があります。学術的理論をベースに開発されたもの、AI など最新のテクノロジーによるもの、優れた専門職によるカウンセリングやコーチングなどの人的サービス、スマートフォンや ICT を活用したものなどがあり、対象となる顧客層も組織（保険者や企業）や個人をターゲットにしたものまであります。

導入を検討するサービス名（又は概要）：

チェック日： 年 月 日（ ） チェック者名：

No.	チェック項目	はい	いいえ	不明/ 該当しない
事業計画・事業目的と導入するサービスの整合性を確認しましょう				
1	サービス導入が保健事業の目的に合致しているかどうかを確認しましょう 【すべて「はい」にチェックできれば OK】			
	1-1. サービスの目的と保健事業の目的が合致している	☐	☐	☐
	1-2. サービスの有効性を示す具体的な数値目標を設定できる（利用者数、成功率、利用期間など）	☐	☐	☐
	1-3. そのサービスを利用する対象者の特性とサービス開発時の有効性検証に参加した集団の特性は完全に、もしくは部分的に一致している	☐	☐	☐
2	サービスの提供が中止された際の対応方法を検討しましょう 【「はい」にチェックできれば OK】			
	2-1. サービスの提供が中止された際の対応についての説明がある（サービスの提供が中止された際には、対象者への通知や、責任の範囲を明らかにするなど、様々な課題や問題が発生する）	☐	☐	☐
サービスの品質を確認しましょう				
3	サービスの対象者は保健事業のターゲットと一致しているかどうかを確認しましょう 【3-1、3-2のいずれかが「はい」にチェックできれば OK】			
	3-1. 特定の状態にある人や要件を満たす人を対象としたサービスである（例、腰痛のある人、40 歳以上、喫煙者、心筋梗塞の既往歴があるなど）	☐	☐	☐
	3-2. 対象者の状態や要件を問わないサービスである	☐	☐	☐
4	サービスの提供が保証される期間を確認しましょう 【すべて「はい」にチェックできれば OK】			
	4-1. サービスの提供が保証される期間が提示されている	☐	☐	☐
5	4-2. サービスのサポート体制や保証される期間が提示されている	☐	☐	☐
	サービスを導入する目的（目指す成果）は明確かどうかを確認しましょう 【3-1～3-4のいずれかが「はい」にチェックできれば OK】			
	5-1. 情報提供を目的とするサービスである	☐	☐	☐
	5-2. 記録や管理を目的とするサービスである（例、食事や体重の記録や管理など）	☐	☐	☐
	5-3. 健康状態の改善を目的とするサービスである（例、生活習慣の改善により心筋梗塞の再発を予防するなど）	☐	☐	☐
	5-4. 行動変容を促すサービスである（例、サービス利用者の行動変容、その積み重ねによる健康の改善など）	☐	☐	☐

図 4・5 「チェックリスト」：「購入者版」（左）、「開発者 β 版」（右） いずれも 1 ページ目イメージ抜粋

これら学術成果を、19 編の学術論文、18 回のセミナーや学会発表（うち国際イベント 7 回）、公開データベース 1 件、公開ツール 2 件として公表した。より広い対象に向けて 30 回のイベントに登壇して普及に努めた。複数の公的研究費によるプロジェクトへ引き継ぎ、学術的な今後の展開の筋道を立てた。

実施にあったってはポストドクや大学院生、若手事業者へと積極的な分析や論文執筆、コンソーシアム運営等の機会を提供し、それぞれ成果を上げる支援をした。

今後の社会実装については、AMED で令和 7 年度から始まったヘルスケアサービス実用化事業・ヘルスケアサービス加速化事業が開始されている。本研究班で得た知見やツール（開発者・購入者チェックリスト）の活用を積極的に提案していき、実社会への普及やさらなるツールの改善を目指す。本研究成果を基盤としたヘルスケア事業者との共同研究も複数開始されており、さらなる推進を期待している。

成果資料掲載サイト

E-LIFE ヘルスケアナビ | ヘルスケアサービス事業者・利用者向け情報サイト(<https://healthcare-service.amed.go.jp/>)

掲載ページ：予防・健康づくりの新たな研究手法の開発 | E-LIFE ヘルスケアナビ | ヘルスケアサービス事業者・利用者向け情報サイト：

- サービス導入時のチェックリスト利用の手引き（https://healthcare-service.amed.go.jp/assets24/pdf/checklist_guide.pdf）
- サービス導入時のチェックリスト（<https://healthcare-service.amed.go.jp/assets24/pdf/checklist.pdf>）

主な論文（抜粋）

- Fukuma S, Mukaigawara M, Iizuka T, Tsugawa Y. Impact of the national health guidance intervention for obesity and cardiovascular risks on healthcare utilization and healthcare spending in working-age Japanese cohort: regression

安全で効果のある予防・健康づくりサービスって何？

予防・健康づくりサービス[※]の品質チェックリスト（開発側）β 版[※]

サービス提供している予防・健康づくりサービスについて№1～8の項目をチェックしてみましょう（№10はオプション項目です。）[※]

※ ここでの「予防・健康づくりサービス」とは、法令に基づく特定保健指導やストレスチェックのようなものから、企業や保険者が加入者・従業員のリスクの状況に合わせて行う事業（例、肥満対策や睡眠習慣改善、栄養指導など）をさします。学術的理論をベースに開発されたもの、AI など最新のテクノロジーによるもの、優れた専門職によるカウンセリングやコーチングなどの人的サービス、スマートフォンや ICT を活用したものなどがあり、対象となる顧客層も組織（保険者や企業）や個人をターゲットにしたものまであります。[※]

※このちゅくりすと[※]はサービス提供側（開発者側）がサービスの品質や安全性を上げ、担保するためにサービス提供側・中に関与するチェックリストです[※]

サービス提供するサービス名（又は概要）：

チェック日： 年 月 日（ ） チェック者名：

No.	チェック項目 [※]	はい [※]	いいえ [※]	不明/ 該当しない
提供するサービスと市場・顧客ニーズの整合性を確認できる情報を明示しましょう [※]				
1 [※]	サービス提供の目的や成果が明確か確認しましょう 【すべて「はい」にチェックできれば OK】 [※]			
	1-1. サービス提供の目的や成果が明確になっている [※]	☐	☐	☐
	1-2. サービスの有効性を示す具体的な数値目標を導入先の顧客にも提供できる（参加率、プログラムの完了率、利用期間など） [※]	☐	☐	☐
	1-3. サービス開発時の有効性検証に参加した集団の特性やサービス利用に関する想定ケースが公開され、導入先の顧客がサービス選定時に参照できる [※]	☐	☐	☐
2 [※]	サービスの提供を中止する際の対応方法を明示しましょう 【「はい」にチェックできれば OK】 [※]			
	2-1. サービスの提供が中止された際の対応を明示している（サービスの提供が中止された際には、対象者への通知や、責任の範囲を明らかにするなど、様々な課題や問題が発生する） [※]	☐	☐	☐
サービスの品質を確認しましょう [※]				
3 [※]	サービスの対象者が明確になっているか確認しましょう [※]			
	3-1. 特定の状態にある人や要件を満たす人を対象としたサービスである（例、腰痛のある人、40 歳以上、喫煙者、心筋梗塞の既往歴があるなど） [※]	☐	☐	☐
	3-2. 対象者の状態や要件を問わないサービスである [※]	☐	☐	☐
4 [※]	サービスの提供が保証される期間を確認しましょう 【すべて「はい」にチェックできれば OK】 [※]			
	4-1. サービスの提供が保証される期間を提示している [※]	☐	☐	☐
5 [※]	4-2. サービスのサポート体制や保証される期間を提示している [※]	☐	☐	☐
	サービスを導入する目的（目指す成果）が明確か確認しましょう [※]			
	5-1. 情報提供を目的とするサービスである [※]	☐	☐	☐
	5-2. 記録や管理を目的とするサービスである（例、食事や体重の記録や管理など） [※]	☐	☐	☐
	5-3. 健康状態の改善を目的とするサービスである（例、生活習慣の改善により心筋梗塞の再発を予防するなど） [※]	☐	☐	☐
	5-4. 行動変容を促すサービスである（例、サービス利用者の行動変容、その積み重ねによる健康の改善など） [※]	☐	☐	☐

discontinuity design. *BMJ Open*.2022.12: e056996 DOI: 10.1136/bmjopen-2021-056996

- Inoue K, Athey S, Tsugawa Y. Machine-learning-based high-benefit approach versus conventional high-risk approach in blood pressure management. *International Journal of Epidemiology* 2022; dyad037.DOI : 10.1093/ije/dyad037
近藤尚己, 『『自然に健康になれる環境づくり』に向けたヘルスケア産業の変革—誰も取り残されないウェルビーイングの達成に向けて—』, 企画の責任者と座長を務めた医研シンポジウム 2023 の報告論文集 2024 年 1 月
- Mori Y, Inoue K, Sato H, Tsushima T, Fukuma S. Beta-blocker initiation under dobutamine infusion in acute advanced heart failure: a target trial emulation with observational data. *European Heart Journal Open*, Volume 4, Issue 4, July 2024, oeae054
- Wan J, Wakaba K, Onoue T, Tsushita K, Nakata Y. Factors associated with acquiring exercise habits through health guidance for metabolic syndrome among middle-aged Japanese workers: A machine learning approach. *Prev Med Rep*. 2024, 48, 102915. doi: 10.1016/j.pmedr.2024.102915.
- Inoue K, Athey S, Baicker K, Tsugawa Y. Heterogeneous effects of Medicaid coverage on cardiovascular risk factors: secondary analysis of randomized controlled trial. *BMJ*. 2024; 386: e079377.
- Inoue K, Adomi M, Efthimiou O, Komura T, Omae K, Onishi A, Tsutsumi Y, Fujii T, Kondo N, Furukawa AT. Machine Learning Approaches to Evaluate Heterogeneous Treatment Effects in Randomized Controlled Trials: A Scoping Review. *J Clin Epidemiol*. 2024; 176: e111538.
- Yamada T, Sugimoto K, Nagata H, Fukuda Y, Sato K, Kondo N. Heterogeneous effects on body mass index in the “Checkup Championship”: A behavioral science-based health promotion program by health interest level. *Preventive Medicine Reports*. In press.

(英文)

Preventive and health promoting programs are rapidly developing. Unlike pharmaceuticals, however, these programs are characterized by heterogeneity in effectiveness and low adherence among the attributes of the target users, making randomized controlled trial (RCT)-based evaluation unsuitable. Therefore, we conducted a literature review on statistical evaluation methods for prevention and health promotion programs (literature study), conducted a secondary analysis of existing intervention and observational study data (analytical study), and provided tools to contribute to policy making based on these results (policy study).

Literature review: We systematically organized and described the overview, assumptions, and advantages and disadvantages of research designs suitable for evaluating prevention and health promotion programs. Details of the evaluation of heterogeneity of effects using Target Trial Emulation (TTE) and machine learning in addition to quasi-experimental methods such as cluster RCTs, regression discontinuity designs, and difference-in-differences methods were organized based on the literature review.

Analytical study: A causal effect evaluation of "Health Checkup Championship," an entertainment-type health promotion program applying behavioral economics and gamification by Company A, revealed that the program has a certain effect regardless of employee attributes (job position, health interest level, etc.). This was evidence that formal evaluation of heterogeneity can contribute to fairness evaluation of service effectiveness, and confirmed the importance of heterogeneity evaluation. The results of an RCT on the effectiveness of Company B's health management application "Calo-mama Plus" were successfully replicated in observational data by applying TTE. While TTE could be beneficial, its limitations were also identified. In addition, together with Companies B and C, we conducted and completed an RCT on the effect of the Calo-mama Plus intervention using hypertension risk genetics. We conducted an analysis to identify the population most likely to benefit from the intervention. The analysis found that the effect of Caro-mama Plus differed depending on family history, usual walking habits, etc., highlighting once again the importance of assuming heterogeneous effects when designing and evaluating prevention and health promotion programs. In addition, we developed and validated an "extended experimental design" to address issues in intervention evaluation specific to prevention and health promotion services: 1) evaluation of interventions in a quasi-experimental design in situations where random assignment is difficult, 2) implementation of interventions in nested RCTs in situations where intervention use behavior is difficult to control, and we also tested the impact of behavioral changes after receiving treatment on outcomes, and confirmed that appropriate results could be calculated.

Policy research: We developed "Purchaser's Checklist for Implementing Health Promotion Services." A developer's version was also developed, with the manuals for using these tools. We also organized the input from implementation consortia and experts, visualized the consensus-building process, and evaluated the feasibility of these tools to improve them.