

食品臨床試験において風邪症状の 予防を検証する研究デザインの提案

馬場 亜沙美 (BABA Asami)^{1*}, 鈴木 直子 (SUZUKI Naoko)¹, 野田 和彦 (NODA Kazuhiko)¹,
波多野 絵梨 (HATANO Eri)¹, 高橋 德行 (TAKAHASHI Noriyuki)¹, 中村 駿一 (NAKAMURA Shunichi)¹,
LIU XUN¹, 柿沼 俊光 (KAKINUMA Toshihiro)¹, 山本 和雄 (YAMAMOTO Kazuo)¹

Key Words: ヒト臨床試験, ヒト試験, 特定保健用食品, 機能性表示食品, アウトカム

Proposed study design prevention of cold-like symptoms in food clinical trials

Authors: Asami Baba^{1*}, Naoko Suzuki¹, Kazuhiko Noda¹, Eri Hatano¹, Noriyuki Takahashi¹, Shunichi Nakamura¹,
Xun Liu¹, Toshihiro Kakinuma¹, Kazuo Yamamoto¹

* **Corresponding author:** Asami Baba¹

Affiliated institution:

¹ ORTHOMEDICO Inc.

[2F Sumitomo Fudosan Korakuen Bldg., 1-4-1 Koishikawa, Bunkyo-ku, Tokyo, 112-0002, Japan.]

Key Words: clinical trials, Foods for Specified Health Uses (FOSHU), Foods with Function Claims, outcome

はじめに

免疫力に関連した機能性表示食品について、「機能性表示食品の届出情報検索」¹⁾ というデータベースを利用して調査すると、近年、その届出数が増加していることがわかる。「表示しようとする機能性」の検索ボックスにキーワードとして「免疫」と入力すると、「最終製品を用いたヒト試験（ヒトを対象とした試験）により、機能性を評価している。」と「最終製品に関する研究レビュー（一定のルールに基づいた文献調査（システマティックレビュー））で、機能性を評価している。」では0件であるが、「最終製品ではなく、機能性関与成分に関する研究レビューで、機能性を評価している。」では104件の食品が抽出された（2024年6月14日時点）。「免疫」を表示した機能性表示食品は2020年から存在しており、2023年度には32件であった。

Cochrane Reviews において Zhao Y, *et al.* (2022)²⁾ の研究は、プラセボまたは無治療と比較した場合にプロバイオティクスが、上気道感染症（Upper Respiratory Tract Infection; URTI）と1回以上診断された人数を約24%減少させる可能性、URTと3回以上診断された人数を約41%減少させる可能性、急性のURTの発生率（特定期間中の新しい症例の数）を約18%減少させる可能性、急性のURTの1回の平均症状持続期間を約1.22日減少させる可能性、URTに対する抗生物質の使用人数を約42%減少させる可能性および副作用（何らかの害）を発生した人数を増加させない可能性を示した。これらのエビデンスは、それぞれ低度または中程度の確実性があると結論付けられている。したがって、食品がURTを予防する可能性は十分にある。

機能性表示食品制度では、予防効果を暗示させる

* **責任著者:** 馬場 亜沙美 (Asami Baba)¹

¹ 株式会社オルトメディコ

〒112-0002 東京都文京区小石川1-4-1 住友不動産後楽園ビル2階

表現が認められないため³⁾、「免疫機能の維持」などの表示にとどまっている。しかし、このような制度とは切り離して、風邪症状の予防効果を検証することは機能性を有した食品の可能性を拡大する上で重要である。しかし、我々の経験的に、現在実施されている食品を介入とした臨床試験（ヒト臨床試験）の多くは、機能性表示食品の届出のために行われおり、表示が困難なアウトカムを検証する試験は少ない。そこで、本項では風邪症状の予防効果を検証するための研究デザインとして、Satomura K, *et al.* (2005)⁴⁾の研究を紹介し、食品での臨床試験において適応可能かを検証する。

1. 風邪症状の予防を検証する研究デザイン

1-1. Satomura K, *et al.* (2005)⁴⁾の研究デザイン

我々は、風邪症状の予防を検証するための研究デザインとして Satomura K, *et al.* (2005)⁴⁾の研究が参考になると考えた。この研究デザインを表1に要約した。

1-2. 風邪症状の定義と測定方法

Satomura K, *et al.* (2005)⁴⁾は、URTIの初回発症率を主要なエンドポイントとしている。また、URTIの定義として、「(1)鼻症状および咽頭症状の両方がある」、「(2)少なくとも1つの症状の重症度が2段階以上悪化する」、「(3)1段階以上の症状の悪化が3日間続く」のすべての条件を満たすものとしている。

測定方法は、試験参加者が日誌に直接記入する形式を採用している。質問項目として、鼻症状（鼻漏、くしゃみ）、咽頭症状（痛み、搔痒感）、気管支症状（咳、痰）、全身症状（発熱、関節痛、倦怠感）などのURTIに関連した項目が設計されている。これらの項目に対して、Jackson method⁵⁾に従い「なし」、「軽度」、「中等度」、「重度」の4段階に分類される。「軽症」は忙しいときには症状が気にならない程度の場合、「中等症」は常に不快感がある場合、「重症」は日常生活に支障をきたす場合を指す。

1-3. 解析方法

上述のようなエンドポイントでは、ある一定の期間の中で個人ごとの関心のある事象（例えば、死亡

表1 Satomura K, *et al.* (2005)⁴⁾の研究デザイン

対象者	選抜基準	(1) 健康な成人ボランティア (2) 18歳から65歳の男女
	除外基準	(1) 頻繁にうがいをする習慣がある (2) 長年にわたって上気道感染症 (URTI) に感染したことがほとんどない (3) 免疫力が低下している (例: コントロール不良の糖尿病患者やステロイド使用者) (4) ポビドンヨードを使用する資格がない (例: 甲状腺疾患患者)
介入	水うがい群	約20 mLの水で約15秒間うがいを3回続け、少なくとも1日3回実施
	ポビドンヨードうがい群	約20 mLの7%ポビドンヨードを15~30倍に希釈したもので約15秒間うがいを3回続け、少なくとも1日3回実施
対照	対照群	各参加者ごとのうがい習慣
主要アウトカム		初回URTI発症率
介入期間		60日間
実施時期		2002~2003年の冬季
デザイン		非盲検・ランダム化・並行群間比較
サンプルサイズ	水うがい群	122
	ポビドンヨードうがい群	133
	対照群	132
主解析		ITTに対するlog-rank test

ITT: Intention to treat

や再発)と事象発現までの時間が記録される。このとき、関心のある事象の発現率や事象発現までの平均的な時間の計算、さらには計算した発現率を介入群間で比較することにおいて、途中で試験を辞退した参加者のデータを適切に扱わなければ、発現率や発現までの時間を過少(あるいは過大)に評価する恐れがある。本稿では、関心のある事象と事象発現までの解析において用いられる代表的な手法として、「Kaplan-Meier法」と「ログランク検定」を概

表2 試験参加者5名の生存時間模擬データ

試験参加者 ID	発現時間(日)	発現・途中打ち切り
1	$t_{(1)}$	6 発現 $\delta_{(1)}=1$
2	$t_{(2)}$	10 発現 $\delta_{(2)}=1$
3	$t_{(3)}$	15 打ち切り $\delta_{(3)}=0$
4	$t_{(4)}$	27 発現 $\delta_{(4)}=1$
5	$t_{(5)}$	36 発現 $\delta_{(5)}=1$

$t(i)$: 昇順の第 i 番目の生存時間の観測値, $\delta(i)$: 昇順の第 i 番目の観測値が中途打ち切りであるときに 0, 非中途打ち切りのときに 1 をとる打ち切り指標関数

表3 Kaplan-Meier法による発現率の推定

時間区間	発現率
$0 \leq t < 6$	$t(i) \leq t$ を満たす観測値はないため $\hat{S}(t) = 1 - 1 = 0$
$6 \leq t < 10$	$t_{(1)}$ が $t(i) \leq t$ であるから $\hat{S}(t) = 1 - \left(\frac{5-1}{5-1+1}\right)^1 = 0.2$
$10 \leq t < 15$	$t_{(1)}, t_{(2)}$ が $t(i) \leq t$ であるから $\hat{S}(t) = 1 - \left(\frac{5-1}{5-1+1}\right)^1 \left(\frac{5-2}{5-2+1}\right)^1 = 0.4$
$15 \leq t < 27$	$t_{(1)}, t_{(2)}, t_{(3)}$ が $t(i) \leq t$ であるから $\hat{S}(t) = 1 - \left(\frac{5-1}{5-1+1}\right)^1 \left(\frac{5-2}{5-2+1}\right)^1 \left(\frac{5-3}{5-3+1}\right)^0 = 0.4$
$27 \leq t < 36$	$t_{(1)}, t_{(2)}, t_{(3)}, t_{(4)}$ が $t(i) \leq t$ であるから $\hat{S}(t) = 1 - \left(\frac{5-1}{5-1+1}\right)^1 \left(\frac{5-2}{5-2+1}\right)^1 \left(\frac{5-3}{5-3+1}\right)^0 \left(\frac{5-4}{5-4+1}\right)^1 = 0.7$
$36 \leq t$	$t_{(1)}, t_{(2)}, t_{(3)}, t_{(4)}, t_{(5)}$ が $t(i) \leq t$ であるから $\hat{S}(t) = 1 - \left(\frac{5-1}{5-1+1}\right)^1 \left(\frac{5-2}{5-2+1}\right)^1 \left(\frac{5-3}{5-3+1}\right)^0 \left(\frac{5-4}{5-4+1}\right)^1 \left(\frac{5-5}{5-5+1}\right)^1 = 1$

$\hat{S}(t)$: Kaplan-Meier法に基づく生存率の推定値, $t(i)$: 昇順の第 i 番目の生存時間の観測値

説する。一般に、関心のある事象は、死亡に限らず上述のエンドポイントである URTI の初回発現や疾患の増悪といった場合も考えられるが、ここでは、このような場面で得られるデータの総称として生存時間データと呼ぶ。

Kaplan-Meier法は、生存時間データに基づいて所与の時間 t における生存率を求めるための1つの方法である。Kaplan-Meier法に基づく生存率の推定値 $\hat{S}(t)$ は、標本サイズを n 、昇順の第 i 番目の生存時間の観測値を $t(i)$ とすると以下の式で与えられる。

$$\hat{S}(t) = \prod_{t(i) \leq t} \left(\frac{n-i}{n-i+1} \right)^{\delta(i)} \quad (1)$$

$\delta(i)$ は昇順の第 i 番目の観測値が中途打ち切りであるときに 0, 非中途打ち切りのときに 1 をとる打ち切り指標関数である。なお、上記の式は生存率を計算する式であるため、事象が起こるたびに 1 から 0 にかけて推定値は低下していくが、発現率の場合は、1 から計算された推定値を差し引くことにより 0 から 1 にかけて上昇していくように算出できる。例として、表2の模擬データを用いて Kaplan-Meier法に基づき計算し、その結果を表3にまとめた。

また、推定値を y 軸、時間を x 軸としてグラフ化することにより、発現率の推移を可視化できる(図1)。

2つの Kaplan-Meier 曲線を比較する代表的な解析手法として、ログランク検定がある。ログランク検定は、以下の過程により計算される。

- ①事象が観測された各時点で治療群・対照群 × 事象の有・無の 2×2 分割表を生成する。
- ②1セルに着目し、観測事象数と期待事象数の差を計算する。
- ③全イベント時点で算出した観測事象数と期待事象数の差を合計する(合計値は「ログランクスコア」と呼ばれる)。
- ④ログランクスコアの2乗をスコア

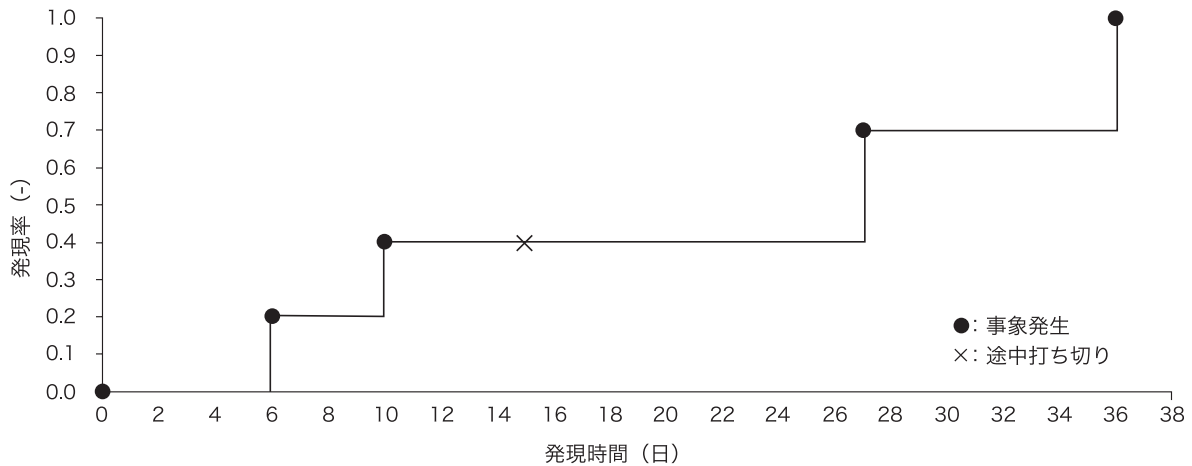


図1 Kaplan-Meier 法に基づく発現率の推定値のプロット

の分散で除すことにより、ログランク検定統計量を算出する。

⑤ログランク検定統計量は自由度1のカイ二乗分布に従うため、カイ二乗分布よりP値が算出される。

計算の過程からもわかるように、ログランク検定は、事象発生時点ごとの生存曲線のズレを合計して曲線全体の差を比較しているの、ズレを全時点で等しく扱うことができる。

1-4. 例数設計

ログランク検定を用いる場合の例数設計の計算にはいくつか種類があるが、今回は要求事項が少なく簡単に計算できるFreedman (1982)⁶⁾の方法を紹介する。

Freedmanの方法は生存時間が指数分布に従うと仮定することを条件とし、要求事項は治療群の生存率 P_c 、対照群の生存率 P_e 、有意水準 α に対応する正規分布 $N(0, 1)$ の分位点 Z_α 、第二種の過誤 β に対応する正規分布 $N(0, 1)$ の分位点 Z_β のみである。そして、以下の式により全体のサンプルサイズ N は求められる。

$$N = \frac{2d}{2 - P_c - P_e} \quad (2)$$

ここで、 d は全体の総事象数を正規変数化したものであり、

$$d = (Z_{\alpha/2} + Z_\beta)^2 \left(\frac{1 + \theta}{1 - \theta} \right)^2 \quad (3)$$

であらわされ、ハザード比 θ は、生存時間が指数分

布の場合にはどの時点でも一定になることから、

$$\theta = \frac{\ln(P_e)}{\ln(P_c)} \quad (4)$$

であらわせる。

Satomura K, et al. (2005)⁴⁾の研究での例数設計は、Freedmanの方法が採用されたかは定かではないが、結果が近似するかを検証してみる。

Satomura K, et al. (2005)⁴⁾の研究では、先行研究の治療群のURTIおよびインフルエンザの累積発症率61.8%、対照群の累積発症率85.8%のデータを用いて、検出力を0.9、有意水準を0.05として例数設計を行っており、サンプルサイズは各群70名と算出されていた。これらの情報を上記の式に当てはめてみる。

$$\theta = \frac{\ln(0.858)}{\ln(0.618)} \approx 0.318 \quad (5)$$

$$d = (-1.96 - 1.28)^2 \left(\frac{1 - 0.318}{1 - 0.618} \right)^2 \approx 39.206 \quad (6)$$

$$N = \frac{1 \times 39.206}{2 - 0.618 - 0.858} \approx 150 \quad (7)$$

となることから各群のサンプルサイズは75名となり、Satomura K, et al. (2005)⁴⁾の研究と結果が近似していることが分かった。

まとめ

本稿では、風邪症状の予防を検証するための研究

デザインとして Satomura K, *et al.* (2005)⁴⁾ の研究を参考とし、症状の発現率とその解析方法について解説した。症状の発現は、ある一定の期間の中で個人ごとに発症するタイミングが異なるため、そのタイミングも重要なデータとなる。そのため、最終的な発現率の差を *t* 検定などで群間比較する方法ではなく、試験期間中を通じた群間の累積発症率のズレを解析できるログランク検定の方が適切だと考えられる。本稿で提案したように厳密なランダム化比較試

験を設計することで、より信頼性の高いエビデンスの取得に寄与できる。

現状、機能性表示食品は制度上、疾病の予防効果を暗示する表現が認められていないため、風邪症状の予防効果を示唆するような表示は不可能である。したがって、機能性表示食品の制度に則った報告ができるかは検証する必要がある、機能性表示に適用できるかは議論の余地がある。

参考文献

1. 消費者庁：機能性表示食品の届出情報検索。(2024年6月11日アクセス可能：<https://www.fld.caa.go.jp/caaks/cssc01/>)
2. Zhao Y, Dong BR, Hao Q. Probiotics for preventing acute upper respiratory tract infections. *Cochrane Database Syst Rev.* **8**(8): CD006895. 2022. (PMID: 36001877)
3. 消費者庁：機能性表示食品の届出等に関するガイドライン(改正令和6年4月1日(消食表第217号))。(2024年6月11日アクセス可能：https://www.caa.go.jp/policies/policy/food_labeling/foods_with_function_claims/assets/notice_240401_01.pdf)
4. Satomura K, Kitamura T, Kawamura T, Shimbo T, Watanabe M, Kamei M, Takano Y, Tamakoshi A: Great Cold Investigators-I. Prevention of upper respiratory tract infections by gargling: a randomized trial. *Am J Prev Med.* **29**(4): 302-307. 2005. (PMID: 16242593)
5. Jackson GG, Dowling HF, Spiesman IG, Boand AV: Transmission of the common cold to volunteers under controlled conditions. I. The common cold as a clinical entity. *AMA Arch Intern Med.* **101**(2): 267-278. 1958. (PMID: 13497324)
6. Freedman LS: Tables of the number of patients required in clinical trials using the logrank test. *Stat Med.* **1**(2): 121-129. 1982. (PMID: 7187087)