

VO₂max を主要アウトカムとした 臨床試験は成立するか？

馬場 亜沙美 (BABA Asami)^{1*}, 鈴木 直子 (SUZUKI Naoko)¹,
田中 瑞穂 (TANAKA Mizuho)¹, 山本 和雄 (YAMAMOTO Kazuo)¹

Key Words : ヒト臨床試験, 特定保健用食品, 機能性表示食品, VO₂max, 主要アウトカム

Is a clinical trial with VO₂max as the primary outcome feasible?

Authors: Asami Baba ^{1*}, Naoko Suzuki ¹, Mizuho Tanaka ¹, Kazuo Yamamoto ¹

Corresponding author: Asami Baba

Affiliated institutions:

¹ ORTHOMEDICO Inc.

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

Key Words: Human clinical trials, Foods for Specified Health Uses (FOSHU), Foods with Function Claims (FFC), VO₂max, Primary outcome (endpoint)

1. はじめに

ヒト臨床試験における主要アウトカムの設定は、その試験の「科学的価値」を決定づける核心である。一般に、死亡や再発といったハードエンドポイント、臨床症状スコア、あるいは特定のバイオマーカーなどが主要アウトカムとして選択される。

一方で、最大酸素摂取量 (VO₂max) や呼吸商 (Respiratory quotient; RQ) といった運動生理学的指標が、主要アウトカムとして採用される例は決して多くない。VO₂max は以下の式で定義される生理学的に明確な指標である。

$$VO_2 \text{ max (mL} \cdot \text{kg}^{-1} \cdot \text{min}^{-1}) = \frac{\text{Absolute VO}_2 \text{ max (L} \cdot \text{min}^{-1}) \times 1000}{\text{Body weight (kg)}}$$

VO₂max は、明確な定義を持つ指標であるにもかかわらず、主要アウトカムとしては「測定値のばらつきが大きい」「統計解析や評価の難易度が高い」といった理由で敬遠されがちである。

そこで本稿では、VO₂max を主要アウトカムとして位置づけた稀有なランダム化比較試験である「Energy Balance and Breast Cancer Aspects-II (EBBA-II) 試験」の統計解析計画書 (SAP)¹⁾ を読み解きながら、以下の3点を整理・解説する。

- なぜ VO₂max が主要アウトカムとして成立し得たのか
- 臨床統計学的にどのように扱われているのか
- 今後の機能性表示食品における試験設計にどのような示唆を与えるのか

* 責任著者: 馬場 亜沙美 (Asami Baba)¹

所属機関: ¹ 株式会社オルトメディコ (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1 EBBA-II 試験の概要¹⁾

項目	内容
試験名	Energy Balance and Breast Cancer Aspects-II (EBBA-II)
試験登録	ClinicalTrials.gov (NCT02240836)
試験責任期間	Oslo University Hospital (ノルウェー)
試験デザイン	ランダム化, 評価者盲検, 並行群間比較, 優越性試験
割付	1:1 (層別因子: 閉経状態)
対象集団	新規診断乳がん患者 (術後補助療法中)
介入	12 か月間の運動プログラム
対照	標準治療 (身体活動の制限なし)
主要アウトカム	VO ₂ max のベースラインから 12 か月後までの変化量
評価時点	ベースライン, 6 か月, 12 か月 (+ 長期追跡: 2・3・5・10 年)

2. EBBA-II 試験とは？

EBBA-II 試験は、乳がんの術後補助療法を受ける患者を対象として、12 か月間の運動介入と標準治療を比較する、ランダム化評価者盲検並行群間比較試験である (表1)¹⁾。

本試験における介入の目的は、単なる体力向上にとどまらず、運動介入が心肺機能、代謝プロファイル、腫瘍微小環境に及ぼす影響を包括的に検証することにある。

3. VO₂max が主要アウトカムへ変更・採用された背景

興味深いことに、EBBA-II 試験において VO₂max は当初から主要アウトカムとして設定されていたわけではない。ClinicalTrials.gov への初期登録時点では、主要アウトカムとして以下の代謝プロファイル指標が設定されていた。

- 脂質・リポ蛋白
- 血糖
- 血圧
- 体組成
- インスリン
- 炎症・血栓関連マーカー

一方、VO₂max を含む心肺機能指標 (Cardiopulmonary exercise test; CPET) 関連指標は、副次的アウトカムとして位置づけられていた。しかし試験の進行とともに、乳がん補助療法中の患者において以下の2点が明確になった。

- 心肺機能の低下が臨床的に重要な問題であること
- 運動介入の直接的な効果を最も統合的に反映する

指標が VO₂max であること

その結果、多様な代謝指標を個別に評価するよりも、心肺機能の包括的指標である VO₂max の変化量を評価する方が、臨床的意義および解釈の一貫性に優れると判断された。この判断を受けて、データ固定前に以下の一連の手続きが実施された。

- 主要アウトカムを「VO₂max のベースラインから 12 か月後までの変化量」へ再定義
- 統計解析計画書 (SAP) の改訂
- ClinicalTrials.gov の登録内容の更新
- VO₂max を主要アウトカムとした統計的検出力および必要症例数の再計算

このプロセスは ICH E9 (臨床試験のための統計的原則) や CONSORT 声明の考え方に照らしても許容される、方法論的に正当な手続きである^{2,3)}。その根拠として以下の要素が満たされていることが挙げられる。

- データ固定前に変更が決定・文書化されていること
- SAP の改訂履歴および変更の医学的妥当性が明示されていること
- 主要アウトカム変更に伴い、検出力・必要症例数の再計算が適切に行われていること

【ポイント】

EBBA-II 試験のプロセスから学べるのは、代謝指標 (脂質・血糖など) が運動介入の「間接的結果」であるのに対し、VO₂max は運動介入の「直接的かつ統合的な効果」であると再整理された点である。これは機能性表示食品の臨床試験設計にも重要な示唆を与える。食品の作用機序がエネルギー代謝や運

動耐容能に関与する場合、複数の代謝マーカーを個別に追うよりも、VO₂max を主要アウトカムに据える方が、介入の直接的な有効性を高感度かつ包括的に捉える指標となり得る。

4. EBBA-II 試験で VO₂max はどのように解析されたか？

4-1. 主要アウトカムの定義

EBBA-II 試験における主要アウトカムは、「VO₂max のベースラインから 12 か月後までの変化量」を連続変数として評価すると定義されている。VO₂max を連続変数として扱うことで、以下の統計学的・臨床的な利点が確保される。

- 生理学的情報量を最大限に保持できる
- 二値化に伴う統計的検出力の低下を回避できる
- 介入による群間差の大きさを定量的に解釈可能となる

4-2. 統計モデルの設計思想

EBBA-II 試験の主解析には、線形混合モデル (Linear Mixed Model) が用いられている。これは、VO₂max が縦断的に繰り返し測定される連続指標であることを前提とした設計である。本モデルには表 2 の固定効果が含まれている。

また、ランダム効果 (Random effects) として、被験者ごとのランダム切片が導入されている。これにより、被験者間のベースライン VO₂max のばらつきと、同一被験者内での測定値相関が適切にモデル化されている。

さらに、本解析では、時間を「6 か月」を結節点 (Knot) とする区分線形モデル (Piecewise Linear Model) として組み込んでおり、介入初期 (0~6 か月) と介入後期 (6~12 か月) で異なる変化勾配を許容する設計となっている。これは、運動介入による心肺機能の変化率が一定ではないという生理学的背景を反映した、合理的な仮定である。

4-3. 解析設計の統計的意義

この解析設計により、以下の利点が得られる。

- 欠測データを含む縦断データを自然に

扱える

- 全ランダム化症例を解析に含める modified Intention to Treat 原則を維持できる
- 中間時点 (6 か月) の情報もパラメータ推定に活用できるため、推定効率が高い

これは、いわゆる Mixed Model for Repeated Measures (MMRM) や constrained longitudinal data analysis (cLDA) に近い形である。

4-4. VO₂max を主要アウトカムとしたサンプルサイズ設計

SAP においては、統計的検出力を担保するため表 3 の仮定が明示されている。

この前提に基づき、1 群あたり約 253 例で 80% の検出力が確保できると算定され、1 群あたり約 300 例という目標症例数が設定された (300 例時の検出力を逆算すると 86%)。

【重要な示唆】

VO₂max の標準化効果量は決して大きくない。しかし、二値化せず連続変数として扱い、縦断情報を最大限活用する線形混合モデルを用いることで、現実的な症例数で検証可能な設計が成立している。これは、VO₂max は「ばらつきが大きいから主要アウトカムにできない」のではなく、「適切なモデルと設計を前提にすれば成立する」ことを示している。

表 2 主解析に用いた線形混合モデルの固定効果¹⁾

固定効果	内容
群	運動介入群 vs 標準治療群
時点	0, 6, 12 か月
群×時点の交互作用	介入効果の時間的推移を評価
閉経状態	ランダム化時の層別因子 (閉経前 / 閉経後)

表 3 VO₂max を主要アウトカムとしたサンプルサイズ設計の前提条件¹⁾

設計パラメータ	設定値
ベースライン VO ₂ max (両群平均)	約 32 mL/kg/min
運動介入群の変化量	0 (変化なし)
標準治療群の変化量	- 0.5 mL/kg/min (低下)
群間差	0.5 mL/kg/min
変化量の標準偏差	2.0
有意水準	両側 5%
検出力	80~86%
目標症例数	1 群あたり約 300 例

5. 呼吸商 (RQ) はなぜ主要アウトカムに採用されなかったのか？

EBBA-II 試験では、RQ も測定されているが、あくまで副次的・探索的アウトカムという位置づけに留まっている。RQ はエネルギー基質の燃焼比率(糖質と脂質の利用割合)を示す生理学的に重要な指標であるが、以下の要因による影響を強く受ける。

- 測定直近の食事内容
- 測定時の環境条件や心理的緊張
- 測定直前の身体活動量

これらの変動要因が存在するため、RQ は VO₂max と比較して日内・日差変動が大きく、臨床的意義の解釈が曖昧になりやすい。すなわち、長期間の介入効果を評価する上で不可欠な再現性が担保しにくいいため、主要アウトカムとしては不適格であると判断されたと考えられる。

この RQ と VO₂max の対比は、ヒト臨床試験において「何を主要アウトカムに据えるべきか」を考えるうえで重要な視点を提供している。

6. 機能性表示食品の臨床試験への示唆

EBBA-II 試験は乳がん術後患者を対象とした研究であり、疾病に罹患していない健常者(あるいは境界域の者)を対象とする機能性表示食品のヒト臨床試験とは前提となる集団が異なる。しかし、VO₂max を主要アウトカムとして成立させた数少ない成功例として、今後の試験設計に以下の示唆を与えている。

- ① VO₂max は「条件付きで」主要アウトカムとして成立する

EBBA-II 試験のプロセスから、VO₂max を主要アウトカムに据えるためには、少なくとも以下の3条件を試験計画段階で満たす必要があると考えられる。

- 対象集団において、心肺機能の維持・向上が明確な健康課題となっていること
- 介入がエネルギー代謝や心肺機能の改善と直接的に結びつく科学的合理性を有すること
- 統計解析計画において、連続変数としての縦断的評価や線形混合モデルの適用が前提とされていること

これらの条件が厳密に設計・文書化されていれば、VO₂max は単なる参考値ではなく、主要アウトカムとして十分に機能し得る。

- ②機能性表示食品の試験設計で VO₂max が敬遠される理由とその突破口

一方で、機能性表示食品の届出を企図したヒト臨床試験において、VO₂max が副次的アウトカムに留まりやすい理由も、EBBA-II 試験との対比から明確になった。主な障壁は以下の2点である。

- 自覚症状や健康イベントとの直接的関連の弱さ：日常生活における「疲れにくさ」や「身体活動量の維持」といった主観的指標と、VO₂max の変動をどう関連づけるかが課題となる。
- 「臨床的意義」の説明の難しさ：VO₂max の微小な数値改善が、健常者の健康にどのようなベネフィットをもたらすのか、消費者庁の手引き等に沿った明確なロジック構築が求められる。

総じて、機能性表示食品のヒト臨床試験において VO₂max を評価項目として活用するには、単なる「測定項目の追加」ではなく、成分の作用機序の立証から適切な解析モデルの選定に至るまで、科学的かつ一貫したストーリー構築が不可欠であると言える。

7. まとめ

本稿では、VO₂max を主要アウトカムとして正式に採用した稀有なランダム化比較試験である「EBBA-II 試験」を題材に、「VO₂max は主要アウトカムとして成立し得るのか？」という問いについて検討してきた。

EBBA-II 試験の最大の特徴は、VO₂max を主要アウトカムとして「選んだ」という結果そのものではなく、その選択に至るプロセスの妥当性と、それを支える統計解析計画が極めて精緻であった点にある。試験開始当初は代謝プロファイルが主要アウトカムとして想定されていたが、試験進行に伴う医学的再評価を経て、VO₂max が介入の直接的かつ統合的な効果を示す主要な関心変数として再定義された。この変更は、データ固定前に実施され、SAP 改訂および必要症例数の再計算を伴う、ICH E9 等のガイドラインに準拠した正当な手続きとして行われている。

さらに、特筆すべきは、その解析手法である。VO₂max は連続変数として扱われ、縦断データを前提とした線形混合モデル（MMRM に準じる設計）を適用することで、効果量がさほど大きくないという前提においても、現実的なサンプルサイズで十分に検証可能な臨床試験デザインが構築されていた。

EBBA-II 試験は乳がん術後患者を対象とした研究であり、その結果や設定をそのまま機能性表示食品などの健常者を対象としたヒト臨床試験に適用できるわけではない。しかし、介入の作用機序とアウトカムの論理的対応、多指標評価と単一統合指標のト

レードオフ、そして主要アウトカムとして成立させるための統計学的要件という観点において、臨床研究者および試験計画立案者に強力な参照枠組みを提供している。

結論として、VO₂max はどのような試験にも適合する万能な主要アウトカムではない。しかし、対象集団における臨床的文脈、介入成分の合理的な作用機序、そして縦断的データ解析を前提とした堅牢な統計設計という3条件が満たされれば、主要アウトカムとして十分に成立し得る指標であることを、EBBA-II 試験は示している。

参考文献

1. Energy Balance and Breast Cancer Aspects (EBBA-II). ClinicalTrials.gov identifier: NCT02240836. Available at: <https://clinicaltrials.gov/study/NCT02240836> (Accessed 2026-06-9).
2. 厚生労働省: ICH E9 臨床試験のための統計的原則 (2026 年 6 月 9 日アクセス可能 : <https://www.pmda.go.jp/files/000156112.pdf>)
3. Sally Hopewell, An-Wen Chan, Gary S Collins, Asbjørn Hróbjartsson, David Moher et al: CONSORT 2025 explanation and elaboration: updated guideline for reporting randomised trials. *BMJ* **389**: e081124, 2025 (DOI: <https://doi.org/10.1136/bmj-2024-081124>)