

「CONSORT クロスオーバー Extension」の 紹介と解説

—CONSORT 2010 Statement: Extension to Randomized Crossover Trials(Japanese Translation)—



上岡 洋 晴 (東京農業大学大学院環境共生学専攻)

津谷喜一郎 (東京有明医療大学保健医療学部)

折笠 秀 樹 (富山大学大学院医学薬学研究部バイオ統計学・臨床疫学)

はじめに

ランダム化比較試験 (randomized controlled trial: RCT) はグレードの高いエビデンスを「つくる」ものだ。だが論文中の一定数の要素の記載が不十分で適切に「つたわら」ないことがある。するとその論文を「つかう」ものは結果の解釈が困難になり、また意思決定はバイアスを含んだものとなりうる。

RCT 報告の質を改善させるため、1996年に21項目のチェックリストとフローチャートからなる CONSORT 声明 (statement) の第1版が発表された。報告ガイドライン (reporting guideline: RG) の嚆矢となるものであった。2001年の第2版は22項目、2010年の第3版は25項目となった。この間、項目数の増加だけでなく内容も改訂されてきた。第3版は“CONSORT 2010”として世界に知られている¹⁻³⁾。

CONSORT 声明は世界中の多くの臨床系の医学雑誌の投稿規定に取り入れられ広く使われている。実際に、CONSORT 声明が公表されて以来、RCT 論文の質の改善に関連したことを示唆する論文が発表⁴⁻⁶⁾されており、人を対象とした研究分野全体においてもその貢献は大きい。日本の学会誌などにおいても普及が進んできている。

今回紹介する「CONSORT クロスオーバー Exten-

sion」は世界中の関係者からその作成が待ち望まれたものである。本稿訳著者らもそれらを期待していた。本文 (p.1884) 「5. CONSORT Extension の開発で用いられた方法」に述べられているように、この開発は、第2版の CONSORT 声明が発行された2001の翌年、2002年に始まる。昨年2019年に発行されたのであるから17年かかったこととなる。CONSORT 声明や EQUATOR Network の創始者のひとりであり、当初本プロジェクトに参加していた Douglas G Altman (1948-2017) はもはやこの世にいない。本稿訳著者らは彼・彼女らの偉業に祝意を表すものである。

この開発に長い時間がかかった理由は、クロスオーバー法はそのデザインと解析の双方の面で統計学的に複雑な問題を含むからである。今回の「CONSORT クロスオーバー Extension」はベースとする CONSORT 2010 の25項目のうち14項目に Extension (拡張) がついた。これらを正しく読むにはそれなりの素養も必要となろう。一方で、難しい統計学をそう深く理解しなくとも、チェックリストとフローチャートに必要事項を記すれば済むのが研究報告ガイドラインのよいところでもある。

表 1a ランダム化クロスオーバー試験を報告するときに含むべき情報の CONSORT チェックリスト

章/トピック (Section/Topic)	項目番号 (Item No)	チェックリスト項目 (CONSORT 2010 Item)	ランダム化クロスオーバー チェックリスト項目	報告頁 (Reported on page No) *
タイトル (Title) †	1a	タイトルにランダム化比較試験であることを記載。	タイトルにランダム化クロスオーバー試験であることを記載。	
抄録 (Abstract) †	1b	試験デザイン (trial design), 方法 (method), 結果 (result), 結論 (conclusion) の構造化抄録 (詳細は「雑誌および会議録でのランダム化試験の抄録に対する CONSORT 声明」を参照)。	クロスオーバーデザインであることを明記し, 表 2 で要約してあるすべての情報を報告。	
はじめに (Introduction)				
背景 (Background) ‡	2a	科学的背景と論拠 (rationale) の説明。		
目的 (Objective) ‡	2b	特定の目的または仮説 (hypothesis)。		
方法 (Method)				
試験デザイン (Trial Design) †	3a	試験デザインの記述 (並行群間, 要因分析など), 割付け比を含む。	クロスオーバーデザインであることの論拠。割付け比を含むデザインの特徴を記載。とくに時期の数と期間, ウォッシュアウトの期間, 持ち越し効果についての考慮。	
プロトコールの変更 (Change from protocol) ‡	3b	試験開始後の方法上の重要な変更 (適格基準 eligibility criteria など) とその理由。		
参加者 (Participant) ‡	4a	参加者の適格基準 (eligibility criteria)。		
セッティング (Settings and locations) ‡	4b	データが収集されたセッティング (setting) と場所。		
介入 (Intervention) †	5	再現 (replication) 可能となるような詳細な各群の介入。実際にいつどのように実施されたかを含む。	再現 (replication) 可能となるような詳細な介入。実際にどのように, いつ実施されたかを含む。	
アウトカム (Outcome) ‡	6a	事前に特定され明確に定義された主要・副次的アウトカム評価項目。いつどのように評価されたかを含む。		
アウトカムの変更 (Change of outcomes) ‡	6b	試験開始後のアウトカムの変更とその理由。		
症例数 (Sample size) †	7a	どのように目標症例数が決められたか。	どのように目標症例数が決められたか。参加者内変動を説明する。	
中間解析と中止のガイドライン (Interim analyses and stopping guidelines) ‡	7b	あてはまる場合には, 中間解析と中止基準の説明。		
ランダム化 (Randomisation)				
順番の作成 (Sequence generation) ‡	8a	ランダムな割振り (allocation) 順番を作成 (generate) した方法。		
	8b	割振りのタイプ: 制限の詳細 (ブロック化, ブロックサイズなど)。		
割振りの隠蔽機構 (Allocation concealment mechanism) ‡	9	ランダム割振り順番の実施に用いられた機構 (番号付き容器など), 各群の割付けが終了するまで割振り順番が隠蔽されていたかどうかの記述。		
実施 (Implementation) †	10	誰が割振り順番を作成したか, 誰が参加者を組入れ (enrollment) たか, 誰が参加者を各群に割付けた (assign) か。	誰がランダムな割振り順番を作成したか, § 誰が参加者を組入れ (enrollment) たか, 誰が参加者をその介入の順序に割付けた (assign) か。	
ブラインディング (Blinding) ‡	11a	ブラインド化されていた場合, 介入に割付け後, 誰がどのようにブラインド化されていたか (参加者, 介入実施者, アウトカムの評価者など)。		
介入の類似性 (Similarity of interventions) ‡	11b	関連する場合, 介入の類似性の記述。		
統計学的手法				
統計学的手法 (Statistical method) †	12a	主要・副次的アウトカムの群間比較に用いられた統計学的手法。	クロスオーバーデザインに適するように, 主要・副次的アウトカムの群間比較に用いられた統計学的手法 (つまり, 参加者内比較に基づく)。	
追加的解析 (Additional analyses) ‡	12b	サブグループ解析や調整解析のような追加的解析の手法。		

表 1a ランダム化クロスオーバー試験を報告するときに含むべき情報の CONSORT チェックリスト (つづき)

章/トピック (Section/Topic)	項目番号 (Item No)	チェックリスト項目 (CONSORT 2010 Item)	ランダム化クロスオーバー チェックリスト項目	報告頁 (Reported on page No)*
結果 (Results)				
参加者の流れ (Participant flow) (フローチャートを強く推奨) †	13a	各群について、ランダム割付けされた人数、意図された治療を受けた人数、主要アウトカムの解析に用いられた人数の記述。	それぞれの順序と時期に分けて、ランダム割付けされた人数、意図された治療を受けた人数、主要アウトカムの解析に用いられた人数の記述。	
追跡不能と除外 (Losses and exclusions) †	13b	各群について、追跡不能例とランダム化後の除外例を理由とともに記述。	それぞれの段階で除外した参加者数 (それぞれの順序と時期に分けて) と、その理由。	
募集 (Recruitment) ‡	14a	参加者の募集期間と追跡期間を特定する日付。		
試験終了 (Trial end) ‡	14b	試験が終了または中止した理由。		
ベースライン・データ (Baseline data) †	15	各群のベースラインにおける人口統計学的 (demographic)、臨床的な特性を示す表。	各順序と時期別にベースラインにおける人口統計学的 (demographic)、臨床的な特性を示す表。	
解析された人数 (Number analyzed) †	16	各群について、各解析における参加者数 (分母)、解析が元の割付け群によるものであるか。	各解析における参加者数 (分母) と解析が元の割付け群によるものであるか。	
アウトカムと推定 (Outcome and estimation) †	17a	主要・副次的アウトカムのそれぞれについて、各群の結果、介入のエフェクト・サイズの推定とその精度 (95%信頼区間など)。	主要・副次的アウトカムのそれぞれについて、介入のエフェクト・サイズの推定とその精度 (95%信頼区間など) を含む結果は参加者内比較に基づくべきである。¶加えて、各時期における各介入結果が推奨される。	
2 項アウトカム (Binary outcomes) ‡	17b	2 項アウトカムについては、絶対エフェクト・サイズと相対エフェクト・サイズの両方を記載することが推奨される。		
補助的解析 (Ancillary analysis) ‡	18	サブグループ解析や調整解析を含む、実施した他の解析の結果。事前に特定された解析と探索的解析を区別する。		
害 (Harm) †	19	各群のすべての重要な害 (harm) または意図しない効果 (詳細は「ランダム化試験における害のよりよい報告: CONSORT 声明の拡張」を参照)。	すべての重要な害 (harm) や、そのデザインが原因となる意図しない効果 (詳細は「ランダム化試験における害のよりよい報告: CONSORT 声明の拡張」を参照) の詳細。	
考察 (Discussion)				
限界 (Limitation) †	20	試験の限界、可能性のあるバイアスや精度低下の原因、関連する場合は解析の多重性の原因を記載。	試験の限界、可能性のあるバイアスや精度低下の原因、関連する場合は解析の多重性の原因を記載。潜在的な持ち越し効果を考察する。	
一般化可能性 (Generalisability) ‡	21	試験結果の一般化可能性 (外的妥当性、適用性)。		
解釈 (Interpretation) ‡	22	結果の解釈、有益性と有害性のバランス、他の関連するエビデンス。		
その他の情報 (Other information)				
登録 (Registration) ‡	23	登録番号と試験登録名。		
プロトコル (Protocol) ‡	24	可能であれば、完全なプロトコルの入手方法。		
資金提供者 (Funding) ‡	25	資金提供者と他の支援者 (薬剤の供給者など)、資金提供者の役割。		

CONSORT=Consolidated Standards of Reporting Trials

*ページ番号は雑誌の投稿規定において任意となる。

†元の CONSORT 項目の修正

‡修正なしの CONSORT 項目

§ランダムな順序はここではランダムな順番のリストを指す。典型的にはコンピュータプログラムにより生成されたものである。これは、ランダム化クロスオーバー試験における介入の順序と混同されないようにすべきである。たとえば、個々の試験参加者が介入 B を受ける前に介入 A を受けるということ。

¶参加者内の比較では、測定値が独自のコントロールとして作用するため、各参加者の測定値の間の相関が考慮される。したがって、測定値は独立していない。

表 1b ランダム化クロスオーバー試験を報告するときに含むべき情報の CONSORT チェックリスト (1つのカラムにまとめたもの)

章/トピック (Section/Topic)	項目番号 (Item No)	ランダム化クロスオーバー: チェックリスト項目	報告頁 (Reported on page No)*
タイトル (Title) †	1a	タイトルにランダム化クロスオーバー試験であることを記載。	
抄録 (Abstract) †	1b	クロスオーバーデザインであることを明記し、表2で要約してあるすべての情報を報告。	
はじめに (Introduction)			
背景 (Background) ‡	2a	科学的背景と論拠 (rationale) の説明。	
目的 (Objective) ‡	2b	特定の目的または仮説 (hypothesis)。	
方法 (Method)			
試験デザイン (Trial Design) †	3a	クロスオーバーデザインであることの論拠。割付け比を含むデザインの特徴を記載。とくに時期の数と期間、ウォッシュアウトの期間、持ち越し効果についての考慮。	
プロトコルの変更 (Change from protocol) ‡	3b	試験開始後の方法上の重要な変更 (適格基準 eligibility criteria など) とその理由。	
参加者 (Participant) ‡	4a	参加者の適格基準 (eligibility criteria)。	
セッティング (Settings and locations) ‡	4b	データが収集されたセッティング (setting) と場所。	
介入 (Intervention) †	5	再現 (replication) 可能となるような詳細な介入。実際にどのように、いつ実施されたかを含む。	
アウトカム (Outcome) ‡	6a	事前に特定され明確に定義された主要・副次的アウトカム評価項目。いつどのように評価されたかを含む。	
アウトカムの変更 (Change of outcomes) ‡	6b	試験開始後のアウトカムの変更とその理由。	
症例数 (Sample size) †	7a	どのように目標症例数が決められたか。参加者内変動を説明する。	
中間解析と中止のガイドライン (Interim analyses and stopping guidelines) ‡	7b	あてはまる場合には、中間解析と中止基準の説明。	
ランダム化 (Randomisation)			
順番の作成 (Sequence generation) ‡	8a	ランダムな割振り (allocation) 順番を作成 (generate) した方法。	
割振りの隠蔽機構 (Allocation concealment mechanism) ‡	8b	割振りのタイプ: 制限の詳細 (ブロック化、ブロックサイズなど)。	
実施 (Implementation) †	9	ランダム割振り順番の実施に用いられた機構 (番号付き容器など)、各群の割付けが終了するまで割振り順番が隠蔽されていたかどうかの記述。	
	10	誰がランダムな割振り順番を作成したか、§誰が参加者を組入れ (enrollment) たか、誰が参加者をその介入の順序に割付けた (assign) か。	
ブラインディング (Blinding) ‡	11a	ブラインド化されていた場合、介入に割付け後、誰がどのようにブラインド化されていたか (参加者、介入実施者、アウトカムの評価者など)。	
介入の類似性 (Similarity of interventions) ‡	11b	関連する場合、介入の類似性の記述。	
統計学的手法 (Statistical method) †	12a	クロスオーバーデザインに適するように、主要・副次的アウトカムの群間比較に用いられた統計学的手法 (つまり、参加者内比較に基づく)。	
追加的解析 (Additional analyses) ‡	12b	サブグループ解析や調整解析のような追加的解析の手法。	

1. CONSORT クロスオーバー Extension のスコープ

第1は、クロスオーバー試験における重要な方法論的特徴をまとめることである。第2は、クロスオーバー試験の経験的エビデンス (empirical evi-

dence, 逸話的エビデンスとも称する) を考察、すなわち、これまで公表されたクロスオーバー試験論文の「質」についての研究をレビューすることである。その後で、クロスオーバー試験に適合した CON-

表 1b ランダム化クロスオーバー試験を報告するときに含むべき情報の CONSORT チェックリスト (1つのカラムにまとめたもの) (つづき)

章/トピック (Section/Topic)	項目番号 (Item No)	ランダム化クロスオーバー: チェックリスト項目	報告頁 (Reported on page No) *
結果 (Results)			
参加者の流れ (Participant flow) (フローチャートを強く推奨) †	13a	それぞれの順序と時期に分けて, ランダム割付けされた人数, 意図された治療を受けた人数, 主要アウトカムの解析に用いられた人数の記述。	
追跡不能と除外 (Losses and exclusions) †	13b	それぞれの段階で除外した参加者数 (それぞれの順序と時期に分けて) と, その理由。	
募集 (Recruitment) ‡	14a	参加者の募集期間と追跡期間を特定する日付。	
試験終了 (Trial end) ‡	14b	試験が終了または中止した理由。	
ベースライン・データ (Baseline data) †	15	各順序と時期別にベースラインにおける人口統計学的 (demographic), 臨床的な特性を示す表。	
解析された人数 (Number analyzed) †	16	各解析における参加者数 (分母) と解析が元の割付け群によるものであるか。	
アウトカムと推定 (Outcome and estimation) †	17a	主要・副次的アウトカムのそれぞれについて, 介入のエフェクト・サイズの推定とその精度 (95%信頼区間など) を含む結果は参加者内比較に基づくべきである。¶ 加えて, 各時期における各介入結果が推奨される。	
2 項アウトカム (Binary outcomes) ‡	17b	2 項アウトカムについては, 絶対エフェクト・サイズと相対エフェクト・サイズの両方を記載することが推奨される。	
補助的解析 (Ancillary analysis) ‡	18	サブグループ解析や調整解析を含む, 実施した他の解析の結果。事前に特定された解析と探索的解析を区別する。	
害 (Harm) †	19	すべての重要な害 (harm) や, そのデザインが原因となる意図しない効果 (詳細は「ランダム化試験における害のよりよい報告: CONSORT 声明の拡張」を参照) の詳細。	
考察 (Discussion)			
限界 (Limitation) †	20	試験の限界, 可能性のあるバイアスや精度低下の原因, 関連する場合は解析の多重性の原因を記載。潜在的な持ち越し効果を考察する。	
一般化可能性 (Generalisability) ‡	21	試験結果の一般化可能性 (外的妥当性, 適用性 (applicability))。	
解釈 (Interpretation) ‡	22	結果の解釈, 有益性と有害性のバランス, 他の関連するエビデンス。	
その他の情報 (Other information)			
登録 (Registration) ‡	23	登録番号と試験登録名。	
プロトコル (Protocol) ‡	24	可能であれば, 完全なプロトコルの入手方法。	
資金提供者 (Funding) ‡	25	資金提供者と他の支援者 (薬剤の供給者など), 資金提供者の役割。	

CONSORT=Consolidated Standards of Reporting Trials

* ページ番号は雑誌の投稿規定において任意となる。

†元の CONSORT 項目の修正

‡修正なしの CONSORT 項目

§ランダムな順序はここではランダムな順番のリストについて言及する。典型的にはコンピュータプログラムにより生成されたものである。これは, ランダム化クロスオーバー試験における介入の順序と混同されないようにすべきである。たとえば, 個々の試験参加者が介入 B を受ける前に介入 A を受けるということ。

¶参加者内の比較では, 測定値が独自のコントロールとして作用するため, 各参加者の測定値の間の相関が考慮される。したがって, 測定値は独立していない。

SORT チェックリストの, アブストラクトのチェックリスト, 改訂版フローチャート, さらに優れた報告例を示した。

このガイドラインでは, すべての試験参加者 (participant) が 2 つの順序 (2×2 または AB/BA デザインとして知られている) のいずれかで 2 つの介入を

受けるランダム化クロスオーバー試験のもっとも単純で, もっとも一般的な形式に焦点を当てる。推奨事項のほとんどが, より複雑なデザイン (3 つ以上の介入, 時期, または順序) にも適用される。別のセクション「7. さらに複雑な研究デザイン」(p.1889) では, 3 つ以上の介入を比較した試験で生じる特有

Box 1 用語解説

Period (時期) : 1つの治療を受ける時間の長さ。

Sequence (順序) : 治療順序 (AB, BA), AB アームに割振られた参加者は最初に治療 A, 次いで治療 B を受ける。BA アームはその逆になる。

Within participants variability (参加者内変動) : 試験参加者内における差の推定標準偏差。

Washout (ウォッシュアウト, 狭義に薬剤介入での場合には「休業」とも称する) : 治療効果を持ち越させないようにするための、2つの治療間における治療のない時間の長さ。

Carry over effect (持ち越し効果) : 最初の介入の効果が2番目の時期に持ち越すこと。

Period effect (時期効果) : アウトカムの結果が、治療効果とは無関係に時とともに変化すること。

Within participant comparison (参加者内比較) : 試験参加者内の比較では、測定値が独自のコントロールとして作用するため、各参加者の測定値の間の相関が考慮される。したがって、測定値は独立ではない。

の問題についても概説した。

2. ランダム化クロスオーバー試験の方法論的特徴

並行群間 (パラレル) 試験とは対照的に、クロスオーバー試験の各個人はランダムな順序で複数の介入を受ける。試験参加者は一連の介入にランダム化される。こうして各参加者は自分をコントロールとしても機能する。このように事前に設定されたデザインにおいて、一部の個人がコンプライアンス違反または救急薬の使用を通じてクロスする試験、またはコントロール群すべての参加者に最後に実験的治療にクロスする機会が与えられるような試験とは混同すべきではない。Zeng ら⁷⁾は、「クロスオーバー割振り」と名付けられた研究のほぼ 1/4 ($n=17/72$) が、ランダム化クロスオーバーデザインを使用せず、試験参加者が試験の途中で介入を変えることを認めていた。

ランダム化クロスオーバー試験には特有の問題がある。1つ目は、「持ち越し効果」(carry over effect)の可能性がある。最初の介入の効果は2番目の時期まで持続するため、観察された治療の違いは、治療を受けた順序によって異なる (Box 1 の「用語解説」を参照)。持ち越し効果にはさまざまな原因が考えられる。効果が残っている薬物やその他の治療の問題に加えて、試験参加者の後から生じる反応は、以前の副作用や以前の治療に対する影響を受ける可能性がある。したがって、最初の介入の終了と2番目の介入の開始との間に十分な「ウォッシュアウト」(washout)を含めることが推奨される。これにより、

最初の介入の影響が2番目の介入時期のアウトカムの測定に「引き継がれない」ようになる。

もう1つの問題は「時期効果」(period effect)であり、これは対象のアウトカムが治療効果に関係なく時間とともに変化するときが発生する。たとえば、患者の状態が安定していない場合や、治療の効果が季節的である場合などである。

さらなる問題は、最初の介入が大きく成功したり、あるいは失敗したりした場合に、試験参加者が試験から脱落する可能性である。これらの参加者の結果を分析に含めることはできない。

デザイン

単純な AB/BA クロスオーバーデザインのおもな長所は、両方の介入が同じ試験参加者を対象に評価されることである。これにより、群レベルではなく個人での比較が可能になる。さらに、試験参加者は2つの介入の経験を比較することで、嗜好 (介入に対する好み) を表すことができるが、1つの介入しか受けないパラレルデザインではそれは不可能である⁸⁾。

重要な方法論的問題は、クロスオーバーデザインの使用が妥当であるかどうかである。少なくとも調査のすべての期間にわたって、慢性的または比較的安定している状態や疾患 (たとえば、多発性硬化症や関節リウマチ) の症状に対する治療 (つまり、疼痛などの症状の治療) にもっとも適している。そして、治療効果が可逆的かつ短期間であるような場合である。関心のある状態が治癒する可能性がある場

合、または試験参加者が試用期間中に死亡する可能性のあるようなものにはクロスオーバーデザインは適さない。

しかしながら、適切ではない状態でも広く使用されている。たとえば、不妊治療を意図したアウトカムにおいてである。試験の最初の時期で妊娠した女性は、その後の試験から除外される。それにもかかわらず、クロスオーバーデザインが当該研究分野⁹⁾で保護されている。妊娠は統計的にランダムな欠測 (missing at random) と取り扱われるという批判¹⁰⁾があるにもかかわらず普及したままである¹¹⁾。試験の症例数の計算は、患者の反応における参加者内変動に基づいている。クロスオーバーデザインは、異なる治療に対する参加者の反応間に高い正の相関がある場合、パラレルデザインよりもはるかに効率的である。パラレルデザインと比較して、目標とする効果量・タイプ1エラー (第1種の過誤)・パワーが同じだと仮定すると、必要とする試験参加者数は少なくなる。

クロスオーバー試験にはいくつかの弱点がある。とくに前述のように持ち越し効果が生じる可能性がある。参加者は最初の治療後に脱落する可能性があり、そのため2番目の治療を受けられないことになる。それは副作用に関連しているかもしれない。

解析

クロスオーバー試験の解析は、ペアのデータに基づいている必要がある¹²⁻¹⁴⁾。推定アプローチは、同じ個人で繰り返される測定の間隔を考慮に入れるべきである。有意性の検定では、持ち越し効果や時期効果がないと仮定した場合、連続変数のアウトカムにおいては参加者内での差に基づき、対応のある t 検定、2項アウトカムでは Mainland-Gart 検定のような手法を用いるべきである^{15,16)}。

クロスオーバー試験を解析するために以前は推奨されていたが、批判されることになった方法は、持ち越し効果を検定し、これが統計的に有意である場合は、第2時期のデータを破棄して第1時期のデータのみを分析するものである。つまり、最初の時期のデータが、パラレルデザインのデータであるかのように解析するものである。Freeman¹⁷⁾は、この解析戦略には欠陥があり、バイアスのある結果につな

がることを示した(ただし、2つの解析間の選択が、事前の仮説の結果に基づいている場合には一般化できる)。Senn¹³⁾は、2つの時期/2つの治療のクロスオーバーデザインの使用は、持ち越し効果が最小限であるという仮定に基づいたうえで有効になると述べている。

クロスオーバーデザイン特有の他の統計的課題は、起こりうる時期効果 (period effect) と持ち越し効果 (carry over effect) の調整が必要なことである。統計モデルには、持ち越し効果のパラメーターを含める。

時期効果は、解析において調整が可能である。AB/BA デザインでは、各順序に等しい数の試験参加者が割振られているとき、時期効果は主たる治療効果の推定にバイアスをもたらさない。しかし、時期効果は、治療効果がどれほどランダムな分散に起因するかによるため、分散の推定に影響をもたらす。時期効果の程度を理解させるためにデータを示すとともに、解析で時期効果が調整されたかどうか、そのような決定が事前に行われたかどうかを明確に情報提供することが重要である。

3. クロスオーバー試験はどれほど行われているか？

2000年12月に公開されたすべてのPubMedが付与 (index) したRCTの詳細なレビューでは、74% (383/519) の試験がパラレルデザイン、22% (116/519) がクロスオーバーデザインであったことを報告している¹⁸⁾。2000年12月にMEDLINEで検索した試験のうち、22% (116/526) はクロスオーバー試験であり、そのなかでほとんどが2つの治療 (72%)、2つの時期 (64%) で実施していた¹⁹⁾。2006年12月に公開されたすべてのPubMedが付与 (index) したRCTのレビューでは、77% (477/616) がパラレルデザイン、16% (100/616) がクロスオーバーデザインだったことを報告している²⁰⁾。2007年から2010年の間にClinicalTrials.govに登録された介入研究のレビューによると、11.2% (4,351/38,969) がクロスオーバーデザインであった²¹⁾。2012年12月のPubMedでのレビューでは、RCTのなかの8.7% (98/1,122) がクロスオーバーデザインであった²²⁾。

4. ランダム化クロスオーバー試験の報告の質はどうか？

CONSORT に関連する RCT の報告の質についての論文は比較的存在するが、クロスオーバー試験の報告の質を具体的に検討した論文は少ない。Mills ら¹⁹⁾は、2000 年 12 月に MEDLINE で検索されたランダム化クロスオーバー試験においては、デザイン、解析、および解釈の詳細について高頻度で省略されていたことを報告している。しかし、ほとんどの試験では、ウォッシュアウト期間を報告して擁護し (69%, 87/127)、解析におけるペアのデータの使用を報告していた (95%, 121/127)。

Gewandte ら²³⁾は、1993 年から 2013 年の間に公表された慢性疼痛に対する薬物治療のクロスオーバー臨床試験 124 編を調べた結果、28% (35/124) がベースラインとウォッシュアウト後の疼痛のレベルを報告し、31% (23/75) だけが参加者内変動に基づいていることを具体的に示した症例数の計算を報告していたことを明らかにした。

Straube ら²⁴⁾は、1990 年から 2014 年の間に PubMed で公開された慢性疼痛に関するクロスオーバー試験 98 編を検討した結果、抄録における有害事象の報告が不十分であり、本文中にもまれにしか報告されておらず、治療時期ごとの内訳を示していたのはわずか 23% (23/98) だったことを報告している。

Zeng ら⁷⁾は、2014 年 9 月に ClinicalTrials. gov を用いてランダム化クロスオーバー試験 54 編を調べ、約 2/3 が単純な AB/BA デザインであり、ほとんどの試験 (87%, 47/54) が試験参加者の流れに関して十分な情報を示していたことを報告している。ベースライン特性は、すべての試験参加者について単一群として報告されることが多く (59%, 32/54)、主要アウトカムと有害事象は「介入ごと」に報告 (それぞれ、81%, 44/54, 83%, 45/54) されていた。しかし、ベースライン特性、アウトカム測定、および有害事象の結果の記載は、クロスオーバーデザインを完全に反映していないようであった。

いくつかの研究²⁵⁻²⁷⁾は、メタアナリシスに関連するランダム化クロスオーバー試験の報告を検討している。メタアナリシスに含まれるように、かなり頻繁にデータが不適切に報告されていたことを示して

いる。

これらの研究は、問題が数年以上にわたって改善されていないことを意味しており、ランダム化クロスオーバー試験の報告に関するガイドラインを要望するものであった。

5. CONSORT Extension の開発で用いられた方法

2002 年 5 月に、多様な研究デザインについて 2001 年版 CONSORT 声明の拡張 (extension) を検討するために、CONSORT グループの数人の著者が米国バージニア州アーリントンで会合を開いた。クロスオーバー試験へと拡張する論文の最初の草案は、2002-2003 年の間に Altman と Elbourne によって開発された。そして、2010 年に CONSORT 声明が更新された。この拡張に関する作業は、2014 年に Dwan と Li がグループに参加したときに前進した。2014 年から 2018 年の期間に、著者間で多数のテレビ会議を通じて完了した。CONSORT グループのガイダンスに従って、CONSORT グループのエグゼクティブで、EQUATOR 運営グループの議長でもある Altman をメンバーに含めた。草稿は、CONSORT グループ内に広く、また他の選ばれた個人に配布され、それらのフィードバックを受けて改訂、そしてエグゼクティブによって最終承認された。

6. CONSORT クロスオーバー Extension の CONSORT チェック項目

チェックリストの項目について説明し、ランダム化クロスオーバー試験における CONSORT 2010 のチェックリスト項目との変更の焦点を当てる。背景を説明し、優れた報告について 1 つ以上の例を示す (本稿では、詳細例示は省略し、該当する文献の紹介に留めた)。また、変更しない他のチェック項目についても説明するが、実施にはクロスオーバーデザインについての特有の配慮が必要である。

表 1a, 1b は、CONSORT 2010 のチェックリストに、ランダム化クロスオーバー試験としての推奨される変更点を示している。表 1a には、CONSORT 2010 とランダム化クロスオーバー RCT とを併記、表 1b には両者を 1 つのカラムにまとめている。読者が使いやすいほうを参考にするとよい。

表 2 ランダム化クロスオーバー試験報告の抄録に含むべき情報: 抄録チェックリストのための CONSORT Extension^S

項目 (Item)	チェックリスト項目 (Description)
タイトル (Title)*	タイトルにランダム化クロスオーバー試験であることを記載。
試験デザイン (Trial Design)*	試験デザインの記載 (クロスオーバー試験と時期の数)。
方法 (Methods)	
参加者 (Participant) †	参加者の適格基準 (eligibility criteria) とデータが収集された場所 (setting)。
介入 (Intervention)*	すべての参加者に行われた介入内容。
目的 (Objective) †	特定の目的や仮説。
アウトカム (Outcome) †	この報告での明確に定義された主要アウトカム評価項目。
ランダム化 (Randomisation)*	参加者がどのようにしてその順番に割振られたか。
ブラインディング (Blinding, masking)*	参加者, 介入実施者, アウトカムの評価者が介入に対してブラインディングされたか。
結果 (Results)	
ランダム化された人数 (Numbers of randomised)*	それぞれの順序にランダム化された人数。
募集 (Recruitment)	試験の状況。 †
解析された人数 (Numbers analysed)*	解析された参加者数。
アウトカム (Outcome)*	主要アウトカムについて, 参加者内比較に基づくエフェクト・サイズの推定とその精度。
害 (Harms) †	重要な有害事象や副作用。
結論 (Conclusions) †	結果の全体的な解釈。
試験登録 (Trial registration) †	登録番号と試験登録名。
資金提供者 (Funding) †	資金提供源。

CONSORT=Consolidated Standards of Reporting Trials

*元の CONSORT 項目の修正

†修正なしの CONSORT 項目

‡これは学会抄録に適用する。

§ Hopewell S, Clarke M, Moher D, et al, CONSORT Group. CONSORT for reporting randomised trials in journal and conference abstracts. *Lancet* 2008; 371: 281-3. doi: 10.1016/S0140-6736(07)61835-2

(1) タイトルと抄録

項目 1a 「タイトル」: タイトルにランダム化クロスオーバー試験であることを明記する。

その主要な理由は、読者が見分ける助けになるからである。そこから読者は症例数や解析に関連してデザインの詳細を考慮しはじめることになる (書き方が参考になる先行研究^{29,30)})。

項目 1b 「抄録」: クロスオーバーデザインであることを明記し、表 2 で要約してあるすべての情報を報告する。

明確で、透明性があり、十分な内容を記載した抄録であることが重要である。抄録だけしか読まない読者もいるだろうし、さらに先に読むべきかを判断するのに抄録を読む人も多いことだろう。よく書かれた抄録は、データベースからの関連報告の検索の助けにもなる。2008 年には、「ランダム化比較試験の抄録報告における拡張版」³¹⁾が公表され、CON-

SORT 2010 とも結合している。また、抄録では、研究デザイン、すなわち順序のランダム化と参加者内比較を考慮した解析を明確にすべきである。表 2 にクロスオーバー試験の抄録で含むべき情報を示した (書き方が参考になる先行研究³²⁾)。

(2) 方法

項目 3a 「試験デザイン」: クロスオーバーデザインであることの論拠。割付け比を含むデザインの特徴を記載。とくに時期の数と期間, ウォッシュアウトの期間, 持ち越し効果についての考慮。

方法には、クロスオーバーデザインで実施したことの論拠を含めるべきである。とくに、2×2 クロスオーバーデザインでは、持ち越し効果を十分なパワーで検出できず、調整できないため、持ち越し効果は無視できるという妥当性を示す必要がある。デザインの記述には、治療時期, 試用期間, ウォッシュアウト期間 (ある場合) などの情報を含めて、

何回の時期を通して何回の介入がテストされたかを明確にすべきである（書き方が参考になる先行研究³³⁻³⁶⁾）。

項目 3b「プロトコールの変更」: 試験開始後の方法上の重要な変更（適格基準: eligibility criteria など）とその理由（CONSORT 2010 からの変更はなし）。

持ち越しに対する検定は推奨されない。しかし、もし検定を行って第 1 期のデータしか用いなければ、そのことを報告すべきである。検定の実施については、項目 12a（統計学的手法）でも論じられなければならない。持ち越しが生じた理由についても議論されなければならない。

項目 5「介入」: 再現（replication）可能となるような詳細な介入。実際にどのように、いつ実施されたかを含む。

この項目においては、2×2 ランダム化クロスオーバー試験としての拡張のため、「各群において」が取り除かれた。すべての試験参加者が両方の介入を受けるからである。

項目 7a「症例数」: どのように目標症例数が決められたか。参加者内変動を説明する。

クロスオーバーデザインの主たる利点は、有意水準とパワーとエフェクトサイズが与えられたとき、各参加者は 1 つの治療だけを受けるパラレルデザインよりも小さい症例数ですむことにある。各試験参加者が対照にもなっている（両介入を受ける）ので、参加者内変動が除去されるためである。著者は通常の有意水準とパワーを含む症例数の計算を報告するとともに、連続変数の場合、参加者内変動を報告することが重要である。連続変数をアウトカムとしたクロスオーバー試験では、参加者内群間差の標準偏差の推定値を症例数設定の際に見積もらなければならない。実際には、多くの試験において、この推定値が存在することはほとんどない。ただし、これを無視すると、クロスオーバー試験の症例数を多めに見積もってしまう可能性がある³⁷⁾。参加者内群間差に関する標準偏差を推定するために、参加者内相関を含め何らかの試みをすべきだろう。

同じように、2 項アウトカムにおいては、ペアデータであるという特性を考慮せず、ペアデザインに起因する参加者内比較を考慮しないと、不要に多くの症例数になってしまうだろう。症例数の計算が再現されうるように、適切な記載をすることが望まれる。追跡不能者の許容範囲も報告されるべきである（書き方が参考になる先行研究³⁸⁾）。

項目 8a「順序の作成」: ランダムな割振り（allocation）順番を作成（generate）した方法（CONSORT 2010 からの変更はなし）。

クロスオーバー RCT においては、割振り順番は介入を受ける順序を意味する。割振りは順序 1 の場合、参加者は A から B の介入、順序 2 の場合、参加者は B から A の介入を受けるとようになる（書き方が参考になる先行研究^{39,40)}）。

項目 10「実施」: 誰がランダムな割振り順番を作成したか、誰が参加者を組入れ（enroll）たか、誰が参加者をその介入の順序に割付けた（assign）か。

この項目では、「その介入の順序に」という文言が追加された。

項目 12a「統計学的手法」: クロスオーバーデザインに適するように、主要・副次的アウトカムの群間比較に用いられた統計学的手法（つまり、参加者内比較に基づく）。

国際医学雑誌編集者会議（International Committee for Medical Journal Editors: ICMJE）と CONSORT グループは、解析方法は、「精通している読者が報告された結果を検証するためにオリジナルデータにアクセスしうる十分な情報の記載」がなされるべきだとしている。（<http://www.icmje.org/recommendations/browse/manuscript-preparation/preparing-for-submission.html#d>）

クロスオーバーデザインだということと、使用された統計的手法がわかれば、読者がそれを評価できるようになる。この解析には、比較のなかで参加者内の特性に重きを置くべきである。

手法の部分では、解析にどの手法が用いられたかを示すべきである。試験参加者内の解析がどのような方法で行われたのかを明確に示す。たとえば、参

加者内群間差で t 検定を用いたとか、参加者、時期、治療効果を含む分散分析を用いたなどである。時期効果や持ち越し効果をモデル化したのであれば、それを報告すべきである。同じように、2項アウトカムにおいては、条件付きロジスティック回帰が Mainland-Gart 検定の代替法となる。参加者内比較を考慮しないと、治療効果の分散を過大に評価する可能性がある。

一部のクロスオーバー試験では、試験参加者は両方の時期の最初と最後にアウトカム変数を測定し、治療効果は各時期の変化スコアを使用して推定される。このアプローチは、持ちだし効果を取り除くとされている。しかしながら、それは治療効果の正確さやバイアスのある推定値を生み出す可能性があるため^{41,42)} 推奨できない。

欠測データは、他の研究デザインと同じようにクロスオーバー試験でも一般的な問題になるが、その特殊性はより複雑である。欠測データがない場合、解析モデルではそれを明記すべきである。ベースライン・データの役割は、しばしばベースライン調整が標準誤差を増加させるので注意深く検討される必要がある。すべての利用可能なデータの混合効果モデル（たとえば、ここでは固定効果と変量効果の混合）は、この条件下で被検者内従属に対する適切な調整を含み、典型的な望ましい第一段階であり、Rubin の「ランダムな欠測 (missing at random)」仮説において適切である。大雑把に言うと、データが設定されるか否かにかかわらず、治療順序と初期データが与えられると後期のアウトカムデータの分布は同一であるという意味である。データが設定されようとされまいと、ベースラインと治療順序が与えられたときのアウトカムデータの分布が同じと仮定すれば、欠測データを除外した解析により、正しい ITT 推定が得られる（つまり、ランダムな欠測ということ）。

データを多重補完し、アウトカムの分布をベースラインと治療順序を与えられた観察されたアウトカムと異なるものにするにより、この検定不可能な仮定に対する結論の頑健性を探ることができる。患者のサブセットからの多重補完の実施は、補完データが流れとして妥当であり、変動を適切に反映するため推奨される⁴⁷⁾ (書き方が参考になる先行研

究^{35,44-46)})。

(3) 結果

項目 13a 「参加者の流れ」: それぞれの順序と時期を分けて、ランダム割付けされた人数、意図された治療を受けた人数、主要アウトカムの解析に用いられた人数の記述(フローチャートが強く推奨される。図 1 を参照)。

フローチャートは、CONSORT 声明の必須要素であり、広く採用されている。クロスオーバー試験においては、全期間を通しての試験参加者の流れを理解するのに重要である。研究を通しての試験参加者の流れを把握できるフローチャートが推奨されるが、書式や内容はその試験の特殊性に合わせて変更することができる。縦型で時間経過を含むことを推奨する(書き方が参考になる先行研究^{47,48)})。

項目 13b 「追跡不能と除外」: それぞれの段階で除外した参加者数(それぞれの順序と時期を分けて)と、その理由。

試験の途中で脱落する試験参加者は、1つの介入だけアウトカムは評価されないことになる。脱落は有益な情報を示すことになるかもしれない。たとえば、実施された治療に不満だったので他の治療はもう望まないかもしれない。そうであれば、結果にバイアスをもたらす。それぞれの介入に関して、順序と時期とに分け、できればフローチャート内に理由も入れて試験参加者減少を示すべきである。

不完全データを取り扱う統計解析もある(項目 12a) (書き方が参考になる先行研究^{49,50)})。

項目 15 「ベースライン・データ」: 各順序と時期別にベースラインにおける人口統計学的 (demographic)、臨床的な特性を示す表。

個人別のランダム割振りにより、ベースライン時の群の特徴における差異は系統誤差というよりは偶然の結果とみなせる⁵¹⁾)。ランダム化クロスオーバー試験においては、介入によって影響しうるベースラインの特徴が第2時期の最初で、第1期の最初の状態に戻っているかどうかを把握することが望ましい。順序ごとの情報は、ランダム化により試験開始時の重要な変数に対して順序間均衡性の達成を評価

するのに必要である。時期ごとの情報は、次の時期における治療効果と試験参加者の特性の時間的変動との交絡があるかどうかの理解に役立つ。性別や年齢などは、2つの時期の開始時ともに同じ値を取る。ただし、不安定な予後因子と主要アウトカムのベースライン値は、各時期の初めに確認する必要がある。もし、アウトカムの特性が時間とともに変化するなら、順序ごとにベースラインを表にするだけでは、時期間差（つまり、治療間差）を推測できない（参考になる先行研究^{39,52}）。

項目 16 「解析された人数」: 各解析における参加者数（分母）と解析が元の割付け群によるものであるか。

解析された試験参加者数は結果を解釈するうえで重要である。クロスオーバーの解析は、ペアの特性を考慮しなければならない。それぞれのアウトカムの解析人数は、参加者内群間差あるいは比の人数と等しいはずである。しかし、すべての試験参加者がそれぞれのアウトカムの解析に寄与するわけではない。試験参加者が1つの時期の解析に役立てられないときには、それに対応する期間は失われる。持ち越し効果や時期効果がないと仮定する。補完するデータは復活し、補完をしないとデータは消失する。これにより、パワーが問題となる。症例数、つまりパワーは、すべての試験参加者が情報を提供するという前提で計算している。試験参加者数を報告しないと、統計学的パワーがどれくらい低下したかを評価できない。持ち越し効果あるいは時期効果があるときには、欠測データによってバイアスをもった推定値となる。加えて、CONSORT 2010 ガイドライン¹⁻³で詳細の説明があるように、per protocol 解析なのか、それとも ITT 解析を実施したのかを明らかにすべきである。

項目 17a 「アウトカムと推定」: 主要・副次的アウトカムのそれぞれについて、介入のエフェクトサイズの推定とその精度（95%信頼区間など）を含む結果は参加者内比較に基づくべきである。加えて、時期ごとの各介入結果が推奨される。

ランダム化クロスオーバー試験の結果では、信頼区間を含む点推定が主要・副次的アウトカムに対し

て報告すべきである。これは、参加者内の適切な分析部分を除き、CONSORT ガイドラインと同じである。パラレルデザインの結果のように、または参加者を二重に数えて提示すべきではない。理想的には、(時期間) 相関が研究のパワーに影響を与えるので、解析する各主要アウトカムの相関係数も、後のクロスオーバーデザインを計画するときに役立つように提供する必要がある。

2 値アウトカムについては、マッチングした表で示すのが望ましい。これにより、読者は一致したペアと不一致なペアを確認できるようになる。マッチングした表があると、今後のメタアナリシスがしやすくなる。というのは、参加者内相関を考慮すれば、適切な式で治療間差を下方修正できるからだ⁵³⁻⁵⁵。パラレルデザインと同じ様式で 2×2 の表で提示すると、治療間分散を適切に調整できない¹⁰。ペアごとの表示は、その後のサンプルサイズの計算にも役立つ。

介入ごとに、そして時期ごとに結果を示すことを推奨する²⁶。というのは、治療と時期の交互作用の有意差に気づくからだ。これは解析担当者がどのように解析したかによらない。理想的には、試験参加者がどちらを好んだかという好みのアウトカムも参加者レベルで報告すべきである。たとえば、参加者ごとに介入 A を好む人と介入 B を好む人に分けて、McNemar 検定を用いて分析する。時期効果を補正するなら、Mainland-Gart 検定または Prescott 検定を行う（参考になる先行研究^{34,35,50,52}）。

項目 19 「害」: すべての重要な害 (harm) や、そのデザインが原因となる意図しない効果（詳細は「ランダム化試験における害のよりよい報告: CONSORT 声明の拡張」²⁸）を参照）の詳細。

それぞれの介入における有害事象の種類と全体の頻度が記載されるべきである。加えて、有害事象の一致と不一致や、効果量の推定と精度（群間比較がなされる場合）を示すことは、その介入の相対的な安全性の情報を与えることになる。

(4) 考察

項目 20 「限界」: 試験の限界、可能性のあるバイアスや精度低下の原因、関連する場合は解析の多重

表 3 医中誌 Web におけるランダム化比較試験の詳細デザイン: 2015-2020 年の期間*に公表された原著論文

公表全体・キーワード	パラレルデザイン	クロスオーバーデザイン
論文全体	4,712 (71.4%)	1,886 (28.6%)
「薬理と治療」**	239 (66.8%)	119 (33.2%)
健康食品**	375 (81.5%)	85 (18.5%)
薬物療法**	1,705 (83.5%)	338 (16.5%)
検 索 式		件 数
#1	((ランダム化比較試験/TH or ランダム化比較試験/AL)) and (DT=2015:2020 PT=原著論文)	6,598
#2	((クロスオーバー研究/TH or クロスオーバー研究/AL)) and (DT=2015:2020 PT=原著論文)	2,057
#3	#1 and #2	1,886
#4	(([薬理と治療]/JN or 薬理と治療/JN)) and (DT=2015:2020 PT=原著論文)	701
#5	(([薬理と治療]/JN or 薬理と治療/JN)) and (DT=2015:2020 PT=原著論文)	701
#6	#3 and #5	119
#7	#1 and #5	358
#8	((薬物療法/TH or 薬物療法/AL)) and (DT=2015:2020 PT=原著論文)	57,123
#9	((食事療法/TH or 食事療法/AL)) and (DT=2015:2020 PT=原著論文)	2,232
#10	(健康食品/TH or 健康食品/AL)	19,722
#11	((健康食品/TH or 健康食品/AL)) and (DT=2015:2020 PT=原著論文)	1,197
#12	#1 and #11	460
#13	#3 and #11	85
#14	#1 and #8	2,043
#15	#3 and #8	338

*2020年8月18日検索。2020年1月1日から2020年8月18日までの期間に公表された原著論文が対象。

**キーワードは任意に選んだ。

性の原因を記載。潜在的な持ち越し効果を考察する。

クロスオーバーデザインの限界は、第1時期の治療が第2時期の結果に影響を及ぼすことであり、もう一方の治療のアウトカムを改善させたり、効果を抑制したりする。持ちこし効果は潜在的に試験を無効にする。そうした限界の報告は試験結果を無効であることを知らせることになる。報告の必要がある限界には、2番目の介入が適用される前の追跡不能者、および適用された順序が試験参加者のランダム化された順序とならないように介入を混合することが含まれる。試験期間中の疾患の安定性の観点からのクロスオーバーデザインの妥当性についても議論しなければならない(書き方が参考になる先行研究^{35,44,57)})。

7. さらに複雑な研究デザイン

これまで、2×2の単純な試験デザインについて説明してきた。クロスオーバーデザインでは、3つ以上の介入での比較(多群比較試験のためのCONSORT拡張版)⁵⁸⁾やクラスタークロスオーバーRCTもある。このデザインにおいては、それぞれの集団

がランダム化された順序で複数の介入を受ける⁵⁹⁾。最近のレビューによると、適切な解析が行われたかを確認することや、報告を改善することが必要なようである⁶⁰⁾。クラスタークロスオーバー試験のCONSORT拡張版の開発は現在、第一に進められている(McKenzie Jら)。

クロスオーバー試験(例として、試験参加者のなかで両目が測定された場合)には、反復測定(多時点)や試験参加者内の多重性の問題もある。他にクロスオーバー試験において低頻度で用いられるかたちでは、生物学的同等性研究、Balaamのデザイン、追加期間デザイン、n-of-1デザイン、不完全ブロックデザイン¹³⁾もある。

おわりに

本稿ではCONSORTランダム化クロスオーバーExtensionの紹介を行った。日本における学術雑誌において、どの程度の割合でクロスオーバーデザインの原著論文が公表されているのか気になるところである。そこで、医中誌Webを用い、「研究デザイン」(research design: RD)としての簡易検索を2020

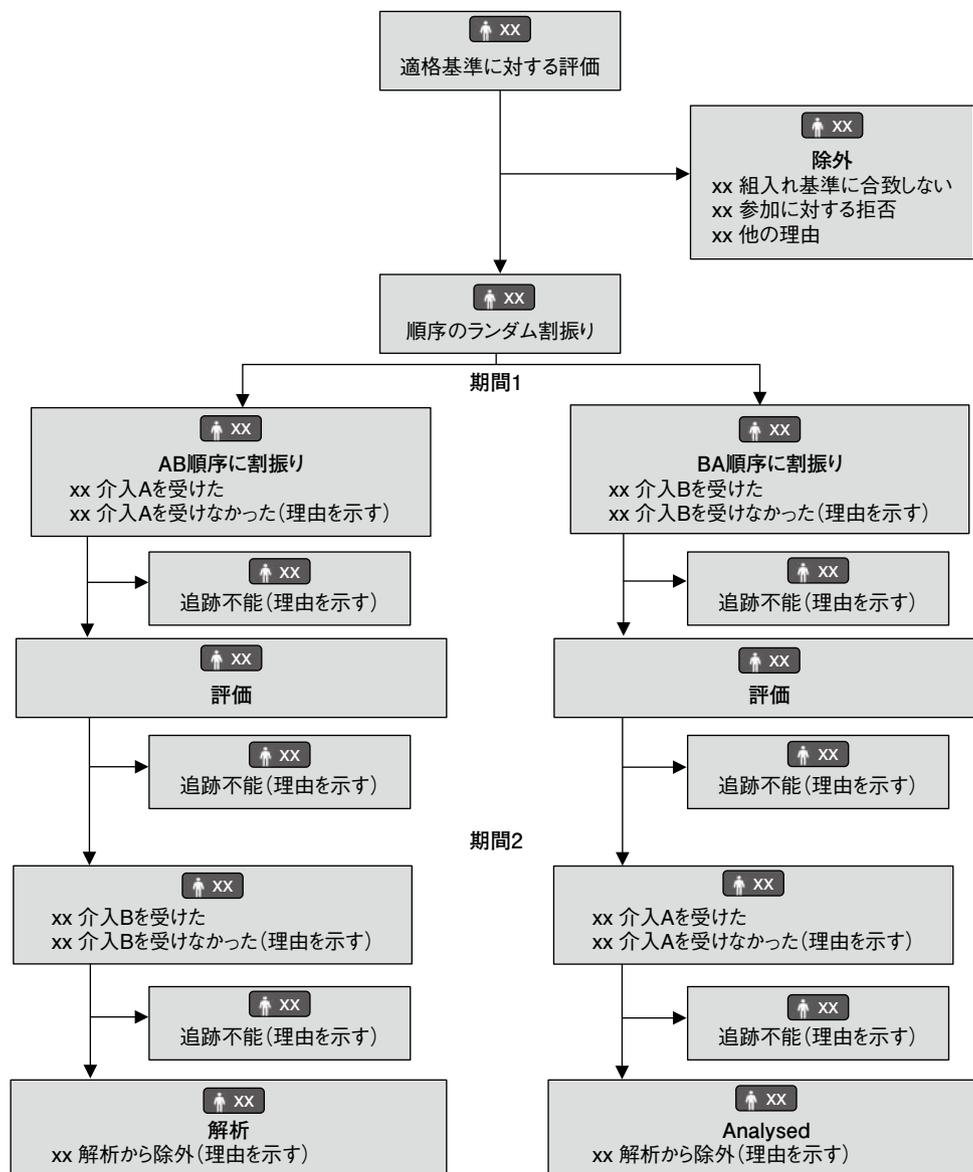


図 1 クロスオーバー試験のための CONSORT フローチャート (2つより多い群または時期なら修正が必要である)

A=介入, B=コントロールあるいは他の介入, xx=人数, CONSORT=Consolidated Standards of Reporting Trials

多くのケースではウォッシュアウト期間が示されている。

年 8 月 18 日に実施した (表 3)。

2015 年から 2020 年の検索日までの 4 年 8 ヶ月間分で、全体の RCT が 6,598 件、そのなかでクロスオーバーデザインは 1,886 件 (22.8%)、パラレルデザインは 4,712 (71.4%) だった。さらに雑誌名 (journal name: JN) の「薬理と治療」、また、一般用語としての「健康食品」と「薬物療法」の計 3 つについて、同じように割合を調べた。その結果、クロ

スオーバーデザインだけの件数 (%) で示すと、薬理と治療では 119 (33.2%)、健康食品で 85 (18.5%)、薬物療法で 81 (23.6%) であった。

全体では、約 20% ということで、PubMed で検索した先の折笠による研究における報告の 9-22% と比較するとほぼ同じ程度であった。しかし、「薬理と治療」誌に投稿される RCT は健康食品に関する論文が多く、クロスオーバーを用いたものが増えている

表 4 クロスオーバーデザインとパラレルデザインの対比

特 徴	ク ロ ス オ ー バ ー	パ ラ レ ル
Size	○小さくすむ	×大きくなりがち
Cost	○安くすむ	×高くなりがち
Duration	○短くすむ	×長くなりがち
Between-subject variation	○大きいほど利点	×大きいほど欠点
Attrition	×起こりやすい	○起こりにくい
Carryover	×入ると厄介	○入らない
Disease type	×安定した慢性疾患に限る (ぜんそく, 狭心症, 糖尿病等)	○とくに制約はない (感染症など急性疾患も可)
Response	×早期のものに向く (早く動く血糖値など)	○早期でなくてもよい (血管イベントなども可)
Reversible	×戻らないといけない (死亡や心筋梗塞はだめ)	○戻らなくてもよい (死亡や心筋梗塞も可)

両者ともに良し (○) 悪し (×) がある。Attrition (摩擦) とは欠損値や中途脱落のようにデザインからの逸脱事象を指す。

(出典: 折笠秀樹. クロスオーバー試験の計画および解析. 薬理と治療 2016; 44: 1261-76. 表 5 を転載)

という編集委員, 学術アドバイザー, 一読者としての感覚があることから, 33.2%という数字はそれを裏付けているのかもしれない。本稿はその詳細を調べることを目的としているわけではないので, 考察はここまでとする。

ランダム化クロスオーバー試験は, 臨床系においては前述したように病状が安定している, あるいは治癒しにくいような慢性疾患に適していることが知られている。臨床系のなかでも実験的介入, たとえば, 食後血糖上昇抑制試験などの早く動くマーカーで数日のウォッシュアウト期間を設けて行うような単回試験では頻繁に用いられている。マーカーは異なるにしても, 前述したように, 「薬理と治療」誌のなかで数多く掲載されているサプリメントや天然成分を用いたランダム化クロスオーバー試験はそれと似ている。さらには, 体力科学の分野においても, 一般人やスポーツ選手における体力・競技力の向上 (具体的には, たとえば筋肉量・筋出力の増加を主要アウトカム) を目的として, このデザインで実施されることが少なくない。

治療だけでなく, ヘルスケアを広義にとらえる場合において, パラレルデザインで実施されることが基本になるが, 試験参加者数や費用の削減などを理由にクロスオーバーデザインでの実施を強く望む研究者も多い。その場合, 計画段階で methodologist に相談することが肝要と思われるが, その両者を直感的に検討するうえで表 4 の「クロスオーバーデザインとパラレルデザインの対比」⁶¹⁾が参考になるだろう。もちろん, 本稿のなかで何度も言及してきた持ち越し効果と時期効果の大きな問題を回避できるか

が最大の焦点となる。

なお, このプロトコル版については EQUATOR ネットワークにおける “Reporting guidelines under development” (<https://www.equator-network.org/library/reporting-guidelines-under-development/reporting-guidelines-under-development-for-clinical-trials-protocols/>) を確認したが開発途中のものはないようだ。したがって, 研究においては報告する際, というよりも計画段階がもっとも重要なので, 表 1a, 1b のチェックリストや図 1 のフローチャートを熟読して, 先を見越したプロトコル設定が望まれる。

今後, わが国において「クロスオーバーデザイン」が用いられる雑誌では, 報告の質を高めるために投稿規定にこの「CONSORT ランダム化クロスオーバー Extension」のチェックリストを掲載すべきであろう。また投稿者は, 試験の計画段階でチェックリストが完成できそうにないときは, その論文の不備を指摘されることを避けるために, 通常のパラレルデザインを使うべきかもしれない。

邦訳対象論文

Dwan K, Li T, Altman DG, Elbourne D. CONSORT 2010 statement: extension to randomised crossover trials. *BMJ* 2019; 366: l4378. doi: 10.1136/bmj.l4378

参考文献

- 1) Schulz KF, Altman DG, Moher D; CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010; 340: c332.
- 2) 津谷喜一郎, 元雄良治, 中山健夫 (訳). CONSORT 2010

声明：ランダム化並行群間比較試験報告のための最新版ガイドライン. 薬理と治療 2010 ; 38 (11) : 939-47.

- 3) 津谷喜一郎. CONSORT 2010 訳者解説. 薬理と治療 2010 ; 38 (11) : 948-49.
- 4) Moher D, Jones A, Lepage I, for the CONSORT Group. Use of the CONSORT statement and quality reports of randomized trials: a comparative before-and-after evaluation. *JAMA* 2001; 285: 1992-5.
- 5) Plint AC, Moher D, Morrison A, Schulz K, Altman DG, Hill C, et al. Does the CONSORT checklist improve the quality of reports of randomised controlled trials? A systematic review. *Med J Aust* 2006; 185: 263-7.
- 6) Hopewell S, Dutton S, Yu L-M, Chan A-W, Altman DG. The quality of reports of randomized trials in 2000 and 2006: a comparative study of articles indexed by PubMed. *BMJ* 2010; 340: c723.
- 7) Zeng L, LA Drye, T Li. Characterizing current registration of phase 3 crossover trials on ClinicalTrials. gov. 2015. <http://jhir.library.jhu.edu/handle/1774.2/38151>.
- 8) Hui D, Zhukovsky DS, Bruera E. Which treatment is better? Ascertaining patient preferences with crossover randomized controlled trials. *J Pain Symptom Manage* 2015; 49: 625-31. doi: 10.1016/j.jpainsymman.2014.11.294
- 9) Cohlen BJ, te Velde ER, Looman CW, Eijckemans R, Habbema JD. Crossover or parallel design in infertility trials? The discussion continues. *Fertil Steril* 1998; 70: 40-5. doi: 10.1016/S0015-0282 (98) 00114-9
- 10) Makubate B, Senn S. Planning and analysis of cross-over trials in infertility. *Stat Med* 2010; 29: 3203-10. doi: 10.1002/sim.398115
- 11) Daya S. Differences between crossover and parallel study designs debate? *Fertil Steril* 1999; 71: 771-3.
- 12) Jones B, Kenward MG. Design and analysis of cross-over trials. 2nd edition ed. 2003, London: Chapman & Hall/CRC.
- 13) Senn S. Crossover-trials in clinical research. 2nd ed. Wiley, 2002. doi: 10.1002/0470854596.
- 14) Wellek S, Blettner M. On the proper use of the crossover design in clinical trials: part 18 of a series on evaluation of scientific publications. *Dtsch Arztebl Int* 2012; 109: 276-81.
- 15) Gart JJ. An exact test for comparing matched proportions in cross-over designs. *Biometrika* 1969; 56: 75-80. doi: 10.1093/biomet/56.1.75
- 16) Mainland D. Elementary Medical Statistics. W. B. Saunders, 1963.
- 17) Freeman PR. The performance of the two-stage analysis of two treatment, two-period crossover trials. *Stat Med* 1989; 8: 1421-32. doi: 10.1002/sim.4780081202
- 18) Chan AW, Altman DG. Epidemiology and reporting of randomized trials published in PubMed journals. *Lancet* 2005; 365: 1159-62. doi: 10.1016/S0140-6736 (05) 71879-1
- 19) Mills EJ, Chan AW, Wu P, Vail A, Guyatt GH, Altman DG. Design, analysis, and presentation of crossover trials. *Trials* 2009; 10: 27. doi: 10.1186/1745-6215-10-27
- 20) Yu LM, Chan AW, Hopewell S, Deeks JJ, Altman DG. Reporting on covariate adjustment in randomised controlled trials before and after revision of the 2001 CONSORT statement: a literature review. *Trials* 2010; 11: 59. doi: 10.1186/1745-6215-11-59
- 21) Inrig JK, Califf RM, Tasneem A, et al. The landscape of clinical trials in nephrology: a systematic review of ClinicalTrials. gov. *Am J Kidney Dis* 2014; 63: 771-80. doi: 10.1053/j.ajkd.2013.10.043
- 22) Odutayo A, Emdin CA, Hsiao AJ, et al. Association between trial registration and positive study findings: cross sectional study (Epidemiological Study of Randomized Trials-ESORT). *BMJ* 2017; 356: j917. doi: 10.1136/bmj.j917
- 23) Gewandter JS, McDermott MP, McKeown A, et al. Reporting of cross-over clinical trials of analgesic treatments for chronic pain: Analgesic, Anesthetic, and addiction clinical trial translations, innovations, opportunities, and networks systematic review and recommendations. *Pain* 2016; 157: 2544-51. doi: 10.1097/j.pain.0000000000000673
- 24) Straube S, Werny B, Friede T. A systematic review identifies shortcomings in the reporting of crossover trials in chronic painful conditions. *J Clin Epidemiol* 2015; 68: 1496-503. doi: 10.1016/j.jclinepi.2015.04.006
- 25) Lathyris DN, Trikalinos TA, Ioannidis JP. Evidence from crossover trials: empirical evaluation and comparison against parallel arm trials. *Int J Epidemiol* 2007; 36: 422-30. doi: 10.1093/ije/dym001
- 26) Li T, Yu T, Hawkins BS, Dickersin K. Design, analysis, and reporting of crossover trials for inclusion in a meta-analysis. *PLoS One* 2015; 10: e0133023. doi: 10.1371/journal.pone.0133023
- 27) Nolan SJ, Hambleton I, Dwan K. The use and reporting of the cross-over study design in clinical trials and systematic reviews: a systematic assessment. *PLoS One* 2016; 11: e0159014. doi: 10.1371/journal.pone.0159014
- 28) Ioannidis JP, Evans SJ, Gøtzsche PC, et al. CONSORT Group. Better reporting of harms in randomized trials: an extension of the CONSORT statement. *Ann Intern Med* 2004; 141: 781-8. doi: 10.7326/0003-4819-141-10-200411160-00009
- 29) Guo X, Kong X, Huang R, et al. Effect of Ginkgo biloba on visual field and contrast sensitivity in Chinese patients with normal tension glaucoma: a randomized, crossover clinical trial. *Invest Ophthalmol Vis Sci* 2014; 55: 110-6. doi: 10.1167/iops.13-13168
- 30) Johansson DP, Lee I, Risérus U, Langton M, Landberg R.

- Effects of unfermented and fermented whole grain rye crisp breads served as part of a standardized breakfast, on appetite and postprandial glucose and insulin responses: a randomized cross-over trial. *PLoS One* 2015; 10: e0122241. doi: 10.1371/journal.pone.0122241
- 31) Hopewell S, Clarke M, Moher D, et al. CONSORT for reporting randomised trials in journal and conference abstracts. *Lancet* 2008; 371: 281-3. doi: 10.1016/S0140-6736 (07) 61835-2
 - 32) Arruda-Olson AM, Mahoney DW, Nehra A, Leckel M, Pellikka PA. Cardiovascular effects of sildenafil during exercise in men with known or probable coronary artery disease: a randomized crossover trial. *JAMA* 2002; 287: 719-25. doi: 10.1001/jama.287.6.719
 - 33) Equi A, Balfour-Lynn IM, Bush A, Rosenthal M. Long term azithromycin in children with cystic fibrosis: a randomised, placebo controlled crossover trial. *Lancet* 2002; 360: 978-84. doi: 10.1016/S0140-6736 (02) 11081-6
 - 34) Rubio-Aurioles E, Porst H, Kim ED, et al. A randomized open-label trial with a crossover comparison of sexual self-confidence and other treatment outcomes following tadalafil once a day vs. tadalafil or sildenafil on-demand in men with erectile dysfunction. *J Sex Med* 2012; 9: 1418-29. doi: 10.1111/j.1743-6109.2012.02667.x
 - 35) Escudier B, Porta C, Bono P, et al. Randomized, controlled, double blind, cross-over trial assessing treatment preference for pazopanib versus sunitinib in patients with metastatic renal cell carcinoma: PISCES Study. *J Clin Oncol* 2014; 32: 1412-8. doi: 10.1200/JCO.2013.50.8267
 - 36) Konstas AG, Lake S, Economou AI, Kaltsos K, Jenkins JN, Stewart WC. 24-Hour control with a latanoprost-timolol fixed combination vs timolol alone. *Arch Ophthalmol* 2006; 124: 1553-7. doi: 10.1001/archophth.124.11.1553
 - 37) Julious SA, Campbell MJ, Altman DG. Estimating sample sizes for continuous, binary, and ordinal outcomes in paired comparisons: practical hints. *J Biopharm Stat* 1999; 9: 241-51. doi: 10.1081/BIP-100101174
 - 38) Gelderblom H, Wüstenberg T, McLean T, et al. Bupropion for the treatment of apathy in Huntington's disease: a multicenter, randomised, double-blind, placebo-controlled, prospective crossover trial. *PLoS One* 2017; 12: e0173872. doi: 10.1371/journal.pone.0173872
 - 39) Fogel RB, Rosario N, Aristizabal G, et al. Effect of montelukast or salmeterol added to inhaled fluticasone on exercise induced bronchoconstriction in children. *Ann Allergy Asthma Immunol* 2010; 104: 511-7. doi: 10.1016/j.anai.2009.12.011
 - 40) Freemantle N, Satram-Hoang S, Tang ET, et al, DAPS Investigators. Final results of the DAPS (Denosumab Adherence Preference Satisfaction) study: a 24-month, randomized, crossover comparison with alendronate in postmenopausal women. *Osteoporos Int* 2012; 23: 317-26. doi: 10.1007/s00198-011-1780-1
 - 41) Fleiss JL, Wallenstein S, Rosenfeld R. Adjusting for baseline measurements in the two-period crossover study: a cautionary note. *Control Clin Trials* 1985; 6: 192-7. doi: 10.1016/0197-2456 (85) 90002-9
 - 42) Willan AR, Pater JL. Using baseline measurements in the two period crossover clinical trial. *Control Clin Trials* 1986; 7: 282-9. doi: 10.1016/0197-2456 (86) 90036-X
 - 43) Carpenter JR, Kenward MG. Sensitivity Analysis with Multiple Imputation. In: Molenberghs G, et al. eds. *Handbook of Missing Data Methodology*. CRC Press New York, 2015: 446.
 - 44) Abell TL, Johnson WD, Kedar A, et al. A double-masked, randomized, placebo-controlled trial of temporary endoscopic mucosal gastric electrical stimulation for gastroparesis. *Gastrointest Endosc* 2011; 74: 496-503. e3.
 - 45) Hill J, Bird HA, Fenn GC, Lee CE, Woodward M, Wright V. A double-blind crossover study to compare lysine acetyl salicylate (Aspergesic) with ibuprofen in the treatment of rheumatoid arthritis. *J Clin Pharm Ther* 1990; 15: 205-11. doi: 10.1111/j.1365-2710.1990.tb00376.x
 - 46) Bonten TN, Snoep JD, Assendelft WJ, et al. Time-dependent effects of aspirin on blood pressure and morning platelet reactivity: a randomized cross-over trial. *Hypertension* 2015; 65: 743-50. doi: 10.1161/HYPERTENSIONAHA.114.04980
 - 47) Chen CY, Holbrook M, Duess MA, et al. Effect of almond consumption on vascular function in patients with coronary artery disease: a randomized, controlled, cross-over trial. *Nutr J* 2015; 14: 61. doi: 10.1186/s12937-015-0049-5
 - 48) Marchetti E, Mummolo S, Di Mattia J, et al. Efficacy of essential oil mouthwash with and without alcohol: a 3-day plaque accumulation model. *Trials* 2011; 12: 262. doi: 10.1186/1745-6215-12-262
 - 49) Markman JD, Frazer ME, Rast SA, et al. Double-blind, randomized, controlled, crossover trial of pregabalin for neurogenic claudication. *Neurology* 2015; 84: 265-72. doi: 10.1212/WNL.0000000000001168
 - 50) Graff GR, Maguiness K, McNamara J, et al. Efficacy and tolerability of a new formulation of pancrelipase delayed-release capsules in children aged 7 to 11 years with exocrine pancreatic insufficiency and cystic fibrosis: a multicenter, randomized, double-blind, placebo-controlled, two-period crossover, superiority study. *Clin Ther* 2010; 32: 89-103. doi: 10.1016/j.clinthera.2010.01.012
 - 51) Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995; 273: 408-12. doi: 10.1001/jama.1995.03520290060030

- 52) Valentino LA, Rusen L, Elezovic I, Smith LM, Korth-Bradley JM, Rendo P. Multicentre, randomized, open-label study of on-demand treatment with two prophylaxis regimens of recombinant coagulation factor IX in haemophilia B subjects. *Haemophilia* 2014; 20: 398-406. doi: 10.1111/hae.12344
- 53) Elbourne DR, Altman DG, Higgins JP, Curtin F, Worthington HV, Vail A. Meta-analyses involving cross-over trials: methodological issues. *Int J Epidemiol* 2002; 31: 140-9. doi: 10.1093/ije/31.1.140
- 54) Stedman MR, Curtin F, Elbourne DR, Kesselheim AS, Brookhart MA. Meta-analyses involving cross-over trials: methodological issues. *Int J Epidemiol* 2011; 40: 1732-4. doi: 10.1093/ije/dyp345
- 55) Curtin F, Elbourne D, Altman DG. Meta-analysis combining parallel and cross-over clinical trials. II: binary outcomes. *Stat Med* 2002; 21: 2145-59. doi: 10.1002/sim.1206
- 56) O'Connor DW, Eppingstall B, Taffe J, van der Ploeg ES. A randomized, controlled cross-over trial of dermally-applied lavender (*Lavandula angustifolia*) oil as a treatment of agitated behaviour in dementia. *BMC Complement Altern Med* 2013; 13: 315. doi: 10.1186/1472-6882-13-315
- 57) Solomon DH, Garg R, Lu B, et al. Effect of hydroxychloroquine on insulin sensitivity and lipid parameters in rheumatoid arthritis patients without diabetes mellitus: a randomized, blinded crossover trial. *Arthritis Care Res (Hoboken)* 2014; 66: 1246-51. doi: 10.1002/acr.22285
- 58) Juszczak E, Altman DG, Hopewell S, Schulz K. Reporting of multi-arm parallel-group randomized trials: extension of the CONSORT 2010 statement. *JAMA* 2019; 321: 1610-20. doi: 10.1001/jama.2019.3087
- 59) Rietbergen C, Moerbeck M. The design of cluster randomized crossover trials. *J Educ Behav Stat* 2011; 36: 472-90. doi: 10.3102/1076998610379136.
- 60) Arnup SJ, Forbes AB, Kahan BC, Morgan KE, McKenzie JE. Appropriate statistical methods were infrequently used in cluster-randomized crossover trials. *J Clin Epidemiol* 2016; 74: 40-50. doi: 10.1016/j.jclinepi.2015.11.013
- 61) 折笠秀樹. クロスオーバー試験の計画および解析. *薬理と治療* 2016; 44: 1261-76.

* * *

Copyright © 2019 British Medical Journal Publishing Group Ltd. Reproduced with permission from BMJ Publishing Group Ltd.