

第9回データサイエンスラウンドテーブル会議

臨床開発でのデジタルバイオマーカー（dBM）の利活用

2025年7月3日

タイムテーブル

時間	所要時間	内容
10:00-10:15	15分	開会の挨拶
10:15-11:00	45分	情報提供
11:00-12:15	75分	第1部：グループディスカッション
12:15-13:15	60分	休憩
13:15-13:45	30分	第1部：各グループで議論した内容の発表（各グループ5分で発表） & 全体でのディスカッション（15分）
13:45-13:55	10分	情報提供
13:55-14:45	50分	第2部：グループディスカッション
14:45-15:15	30分	第2部：各グループで議論した内容の発表（各グループ5分で発表） & 全体でのディスカッション（15分）
15:15-15:30	15分	休憩
15:30-16:45	75分	各テーマで議論した内容の発表（25分×3テーマ）
16:45-17:00	15分	閉会の挨拶

本テーマのファシリテーター

第1部 & 第2部のディスカッションのファシリテーターを務めます

- 伊藤陽一（北海道大学病院）
- 横田勲（北海道大学）
- 青木琢巳（PMDA）
- 田中清士朗（PMDA）
- 吉田瑞樹（ファイザー）
- 柴田康晴（MSD）
- 川畑宣勲（中外）

情報提供

- 午前中
 - はじめに（吉田）
 - dBMの概要と活用意義，dBM開発の工程の概要（柴田）
 - 評価尺度の開発・PROとの対比（伊藤）
 - dBMと規制（青木）
 - 第1部 & 第2部の議論トピック（吉田）
 - 仮想事例の紹介（柴田，吉田，青木）
- 午後
 - dBM特有の統計的課題（横田）

第1部 & 第2部の議論トピック

本日議論する内容の概要

- 第1部：dBMの活用可能性
 - dBMの活用を含めた医薬品の開発戦略，dBM特有の問題やその解決策，規制上の課題や薬事上の留意点，そしてdBM活用において生物統計家が関与すべき点や注意すべき点について議論する
- 第2部：評価尺度の開発とバリデーション
 - PRO開発と対比させながら，信頼性，妥当性，反応性を評価し，意味のある変化の閾値を設定する試験デザインの立案や試験結果の解釈時の留意点について，dBMの活用場面ごとに統計的な論点も含めて議論する

第1部：dBM活用の利点と課題

- ご自身が担当している疾患領域でdBMを活用する利点は何か
- 想定される主な課題は何か
 - 統計的課題に加えて，技術面，運用面，規制面なども含む
- それらの課題への対処法は何か

第1部：dBMをピボタル試験の主要評価項目として用いるための開発戦略①

- どのようなdBM指標が主要評価項目となりうるか。仮想事例の評価項目は主要評価項目として適切か
 - 1時間あたりの平均咳嗽回数，1晩あたりの掻爬時間，Pain index
- dBM指標を主要評価項目として用いるためには何が必要か
 - 医薬品開発のどのタイミングでこういった試験を実施して，こういった特性を評価し，最終的にこういったエビデンスを揃える必要があるか
 - dBM指標を主要評価項目として使用する際に特有の課題（例えば安全性）はないか
 - 関心のある疾患領域で全く使用経験がない場合に加えて，他の疾患領域で主要評価項目として用いた経験がある場合，当該疾患領域で副次評価項目として用いた経験がある場合（副次評価項目として用いる場合）など，仮想的な状況に応じて何が必要か整理する
- 規制当局からどのような懸念や質問が予想されるか

3つの仮想事例

1. 慢性咳嗽患者を対象に、デジタル咳録音装置（VitaloJAK）を用いて24時間当たりの咳嗽回数を測定し、1時間当たりの平均咳嗽回数をピボタル試験の主要評価項目とする場合
2. アトピー性皮膚炎患者を対象に、腕時計型デバイス（アクチグラフィ）を用いて夜間睡眠時の掻爬行動を測定し、一晩あたりの掻爬時間をピボタル試験の主要評価項目とする場合
3. 子宮内膜症患者を対象に、ウェアラブルデバイスにより皮膚電位、心拍数及びその変動などの複数の生理学的パラメータを客観的かつ持続的に収集し、これらのデータにアルゴリズム処理を施して構築した、疼痛に相当するdBM指標であるPain Indexをピボタル試験の主要評価項目とする場合

仮想事例1（グループ2）の議論トピック

- 慢性咳嗽患者を対象に、咳モニターによるdBMから構築した評価項目である「一時間当たりの咳の回数」を主要評価項目として用いるためには何が必要か
- 仮想的な状況に応じて何が必要か整理する
 - 評価項目として用いた経験が全くない場合
 - 副次評価項目として用いた経験がある場合（副次評価項目として用いる場合）
 - アトピー咳嗽患者を対象に主要評価項目として用いた経験がある場合

仮想事例2（グループ1）の議論トピック

- アトピー性皮膚炎患者を対象に、アクチグラフィによるdBMから構築した評価項目である「一晩あたりの掻爬時間」を主要評価項目として用いるためには何が必要か
- 仮想的な状況に応じて何が必要か整理する
 - 評価項目として用いた経験が全くない場合
 - 副次評価項目として用いた経験がある場合
（副次評価項目として用いる場合）
 - 乾癬患者を対象に主要評価項目として用いた経験がある場合

仮想事例3（グループ3）の議論トピック

- 子宮内膜症に対する開発において、“Pain-Index”を臨床試験の主要評価項目とするには何が必要か
- 仮想的な状況に応じて何が必要か整理する
 - 評価項目として用いた経験が全くない場合
 - 副次評価項目として用いた経験がある場合
（副次評価項目として用いる場合）
 - 類縁疾患（例：子宮腺筋症）で主要評価項目として使われている場合

第1部：dBMをピボタル試験の主要評価項目として用いるための開発戦略②

- dBM指標の主要評価項目としての活用を推進するためには、どのような取り組みが必要か
 - 産官学（+a）で必要な取り組みはあるか
 - どんなエビデンスを構築する必要があるか
 - どこ（製薬企業、アカデミア、デジタルデバイスの開発企業など）が主体となってエビデンスを構築すべきか

第1部：生物統計家が果たすべき役割

- 生物統計家はどの段階から関与すべきか
- 生物統計家が特に関与すべき事項は何か

第2部：各特性の評価方法と評価基準

- 各特性を評価するためにはどのような試験デザインが適切か
 - 仮想事例のdBm指標をピボタル試験の主要評価項目として用いる場合で考える。副次評価項目として用いる場合との違いについても考える
 - 各試験の対象集団，評価項目，比較対照となる評価尺度，症例数設計，解析手法
 - 評価基準
- ご自身が担当している疾患領域で特に重要な特性は何か。評価する上で何か課題はあるか

第2部：臨床的に意味のある差の設定

- 「臨床的に意味のある差」や「臨床的に意味のある変化の閾値」を設定するためにはどのような試験デザインが適切か
 - 仮想事例のdBM指標をピボタル試験の主要評価項目として用いる場合で考える。副次評価項目として用いる場合との違いについても考える
 - 適切なアンカーの選定，統計的手法の活用，臨床専門家や患者からのフィードバックの活用など，設定のアプローチとその妥当性を議論する

第2部：dBMを用いた試験計画・解析での懸念点や工夫すべき点の検討

- dBMを用いた試験を計画・解析する上での懸念点や工夫すべき点は何か
 - 症例数設計，膨大なデータの保持や品質管理，患者のデバイス装着のコンプライアンス不良，デバイス間ばらつき，時系列データ解析の難しさ
 - などの懸念点を挙げて，それらに対処するための方策を含めて議論する

第1部の議論結果

(グループごとに重点的に議論したトピックは異なります)

※本議論の内容は、参加者による意見交換をまとめたものであり、各参加者の所属団体の
公式見解を示すものではありません

【第1部：dBM活用の利点と課題】

ご自身が担当している疾患領域でのdBMを活用する利点，課題，対処法について

【利点】

- 従来は評価が困難だった症状などを評価できる
- 定期的かつ継続的に評価できるのが利点
 - 例：睡眠中の呼吸状態や、1晩あたりの掻く時間など
- 早期の段階で効果を検出しやすくなる

【課題】

- 欠測データやコンプライアンスの問題
 - 例：実際に寝ているかどうか分からない
- 測定値の算出アルゴリズムが不明瞭で、ブラックボックス化している
- 医薬品の開発戦略において、dBMのバリデーションをどのタイミングで行うかが課題となる
- 将来的な追加適応も見据え、開発初期段階からdBMの開発戦略を検討する必要がある

【対処法】

- アルゴリズムの透明性を確保するため、dBM開発企業から情報提供を受けられるよう交渉する
- 欠測を防ぐ工夫（例：デバイスが欠測時にアラートを出す機能など）を導入する
- 欠測のメカニズムを考慮する

【事例】

- ディシェンヌ型筋ジストロフィー
 - 6分間歩行距離の代替としてdBMが主要評価項目に採用された例がある

【第1部：dBMをピボタル試験の主要評価項目として用いるための開発戦略①】

どのようなdBM指標が主要評価項目となりうるか

dBM指標を主要評価項目として用いるためには何が必要か

【検討事例：仮想事例2】**【どのようなdBM指標が主要評価項目となりうるか】**

- 夜間の搔爬時間が主要評価項目となり得るか？
 - 現時点では主要評価項目とするには難しいと考える
 - 「搔爬時間のみで臨床的な重症度を適切に評価できるか」が重要なポイント
 - 現行の評価指標の代替となり得るのであれば、主要評価項目としての可能性はある
 - ゴールドスタンダードの存在も重要だが、より適切な指標であれば主要評価項目として主張できる余地はある

【dBM指標を主要評価項目として用いるためには何が必要か】

- 主要評価項目、副次評価項目、探索的評価項目では、それぞれの性質の違いにより、dBMに求められる特性評価の重み付けが異なる
 - 探索的評価項目としての活用
 - 副次評価項目ほど厳密な要件は求められないと考える
 - まずは早期フェーズで探索的評価項目として設定し、dBMの特性を確認するためのデータを収集する
 - 副次評価項目としての活用
 - 主要評価項目と同様に、副次評価項目でも程度の差はあるものの、実施可能性、反応性、内容的妥当性、基準関連妥当性、構成概念妥当性といった特性について、一通りの確認が必要と考える

【第1部：dBM活用の利点と課題】

ご自身が担当している疾患領域でのdBMを活用する利点，課題，対処法について

【利点】

- 日常的な咳の回数など、従来は測定が困難だったデータを可視化できる

【課題】

- 患者への負担
- 場所の制限
- 装置装着による行動変化
- 参加集団への影響
 - 装置の装着を受け入れられる被験者のみが参加するため、選択バイアスが生じる可能性がある
- 解析への影響
 - 測定を継続できなかった場合、中間事象としてどのように扱うかを検討する必要がある。また、デバイスを外した理由も考慮する必要がある
- 外部対照試験への適用の難しさ
 - 装置の装着によって行動が変化する可能性がある場合、同一試験内での比較が必要になる可能性がある
- 測定データの安定性
 - 日内変動を正確に捉えられるかどうかは課題
- 費用の高さ
- バリデーションの必要性

【対処法】

- 患者の負担を軽減するため、装着時間をできるだけ短くする
- 少人数で試験的に導入し、長時間の装着が可能かどうかを検討する

【第1部：dBMをピボタル試験の主要評価項目として用いるための開発戦略①】

どのようなdBM指標が主要評価項目となりうるか

dBM指標を主要評価項目として用いるためには何が必要か

【検討事例：仮想事例1】**【dBM指標を主要評価項目として用いるためには何が必要か】**

- デバイスやデータ処理方法について、十分なバリデーションを行う
 - まずは副次評価項目として使用し、実用性や運用上の課題を確認する
 - QOLとの相関があるかどうかを検討する
- 臨床的に意義のある改善幅について、十分に検討する
 - 例：咳の回数がどの程度減少すれば「改善」と判断できるのか、といった基準の明確化

【第1部：dBMをピボタル試験の主要評価項目として用いるための開発戦略②】
dBM指標の主要評価項目としての活用を推進するためには、どのような取り組みが必要か

【活用推進のために必要な取り組み】

- コンソーシアムを開催し、産官学の連携による情報共有を促進する
 - 新たな装置の提案を行う
- 費用の削減
 - ただし、そもそもdBMの使用が推奨されるのかについては、現時点では議論の余地がある
- バイオマーカーの利点を広く周知する
 - 評価の精度が高まる
 - 症例数や試験期間の短縮が可能になる、など
- 複数の主要評価項目の一部として用いることを検討する

【第1部：生物統計家が果たすべき役割】

生物統計家はどの段階から関与すべきか。生物統計家が特に関与すべき事項は何か

【どの段階から関与すべきか】

- 機器の使用を検討する初期段階

【特に関与すべき事項は何か】

- 欠測データが発生した場合の対応
- 患者集団への影響評価
- 症例数や検出力の評価
- 審査官の視点を踏まえたインプットの提供
 - アルゴリズムがブラックボックス化していないかの確認
 - （例：咳モニターの場合）圧縮アルゴリズムに疑義がないか、など
- バリデーションの数学的側面の確認
 - カットオフ値の設定に関する検討
 - 必要に応じて、ベンダーとの協議も視野に入れる

【第1部：dBM活用の利点と課題】

ご自身が担当している疾患領域でのdBMを活用する利点，課題，対処法について

【担当領域】

- がん領域で、吐き気、全身状態、疼痛などを評価／測定できると有用

【利点】

- 評価の客観性の向上
- 生体学的な妥当性
- 新たな測定概念の創出

【課題及び対処法】

- 高齢者が対象となることが多いため、シンプルな（操作が簡単な）デバイスであることが望ましい
- dBMに関する教育やサポート体制を事前に準備しておくこと
- dBM指標の算出アルゴリズムなどがブラックボックスになっている場合、規制当局などに対してdBMの評価系としての妥当性を説明するのが難しい。そのため、dBMの開発会社から直接規制当局に説明できるようにしておくなどの準備をしておく

【第1部：dBMをピボタル試験の主要評価項目として用いるための開発戦略①】

どのようなdBM指標が主要評価項目となりうるか

dBM指標を主要評価項目として用いるためには何が必要か

【検討事例：仮想事例3】

【Pain Indexが主要評価項目となりうるか】

- ピボタル試験の主要評価項目とするのは難しい
- ピボタル試験よりも前の相の試験であれば、試験の目的によっては、主要評価項目とできる場合があるかもしれない
 - 後期相では、通常の生活を妨げるような痛みを軽減できるかが重要だが、早期相であれば、Pain Indexというスコア自体を評価することを試験の目的とする場合もありうる

【Pain Indexを主要評価項目として用いるために必要なこと】

- 客観性の確保
- 既存手法との相関（収束妥当性、判別妥当性）の確認
 - 連続データを用いることで、痛みの周期性に基づいた「痛みの軽減」の定義が可能となる可能性がある。その定義が既存手法とどの程度相関しているかを確認する必要がある
- 注目している痛みによって評価尺度が変化する一方、注目していない痛みによって変化しないか（例：子宮内膜症の痛みには反応するが、頭痛には反応しない）。変化してもその理由が妥当か
 - PROの課題：患者が過去に経験した最大の痛みによって、痛みのスケールが個人間で異なってしまう可能性がある

【第1部：dBMをピボタル試験の主要評価項目として用いるための開発戦略②】
dBM指標の主要評価項目としての活用を推進するためには、どのような取り組みが必要か

【活用推進のために必要な取り組み】

- 規制当局との早期の対話
 - dBM指標を主要評価項目として使用する判断をした時点で、できるだけ早期に規制当局と相談することが望ましい
- 膨大なデータ量への事前対応
 - 例：1人あたり1日で約30万行のデータが発生した事例がある
- 産官学の協力
 - PROの開発では、アカデミアが主導して基盤を構築した実績がある
 - dBMの場合はテック企業が中心となる可能性があるが、臨床バリデーションなどを単独で完結させるのは難しいため、産官学の連携は不可欠である

第2部の議論結果

(グループごとに重点的に議論したトピックは異なります)

※本議論の内容は、参加者による意見交換をまとめたものであり、各参加者の所属団体の
公式見解を示すものではありません

【第2部：各特性の評価方法と評価基準】

各特性を評価するためにはどのような試験デザインが適切か

【検討事例：仮想事例2】

- **信頼性**：同じ状態にある対象に対して、測定値が状態に依存せず一貫して得られること
 - － 開発の早期段階で評価すべき重要な指標である
 - － アトピー性皮膚炎の場合、Ph2試験での評価が適していると考える
 - － ゴールドスタンダード（ビデオ観察による真値）が存在する場合、それとの比較によりdBMの信頼性を評価できる
 - － ゴールドスタンダードが存在しない場合、Test-retest法（例：1週間後と2週間後の測定結果を比較）による評価が有効である
 - － Ph2試験で評価する場合、治療前のスクリーニング期間（無治療状態）に観察を行うことも一案である
- **反応性**：評価対象の状態が変化したときに、その変化を正確に捉えられること
 - － ゴールドスタンダードがある場合、その変化に応じてdBMの値も変化しているかを確認する
 - － dBMを含めた複数の指標間で効果サイズを比較することで、反応性を評価する。効果サイズが大きい場合、治療効果を捉えられている
 - － Ph2試験での評価が可能である
- **妥当性**：測定しようとしている事象をdBMが正しく捉えていること
 - － ビデオ観察などのゴールドスタンダードによる評価が困難な場合、基準関連妥当性に加えて構成概念妥当性の評価も検討すべきと考える
- **臨床的に意味のある差**
 - － 臨床的に意味のある差についても、dBMを主要評価項目として用いる際には明確に定義し、検討する必要があると考える

【第2部：dBMを用いた試験計画・解析での懸念点や工夫すべき点の検討】

dBMを用いた試験を計画・解析する上での懸念点や工夫すべき点は何か

【懸念点】

- 欠測の発生
 - 解析時点で評価可能な被験者のみに絞ると、対象集団の選択に偏りが生じる可能性があり、その妥当性に懸念が残る
 - デバイスが装着されていない場合、アラートを出すなどの仕組みが有用と考える
- データ転送のミスや機器の故障の頻度
 - 主要な臨床試験を開始する前に、他の試験や予備的な運用で試行し、転送エラーや故障の発生頻度を把握しておくことが望ましい
- デバイス由来の安全性に関する懸念事象への対応
 - 安全性に関する問題が発生した場合に備え、対応方針をあらかじめプロトコルに明記しておく必要がある
 - アトピー性皮膚炎の場合、デバイスを装着することによって生じる赤みや痒みなど
- アルゴリズムの開示
 - ゴールドスタンダードが存在する場合、アルゴリズムによって得られた結果がそのゴールドスタンダードと一致していれば、それだけで十分であり、アルゴリズムの開示は不要だという意見がある。一方で、それだけでは不十分であり、たとえ結果が一致していたとしても、アルゴリズムの中身を確認する必要があるという意見もある
 - アルゴリズムの開示が難しい可能性があるのであれば、デバイス選定の段階からその点を考慮する必要があると考える

【第2部：各特性の評価方法と評価基準】

各特性を評価するためにはどのような試験デザインが適切か

【検討事例：仮想事例1】

- 仮想事例：咳モニター
 - 医薬品開発にdBMを用いる場合でも、特別な試験デザインが必ずしも必要とは限らない可能性がある
- 評価項目をどう設定するか？
 - 回数の平均値による群間差の比較
 - 個人によって一般に咳の基準回数が異なるため、解釈が難しくなる可能性がある
 - 個人内のベースラインからの減少割合（%）の群間差の比較
 - 個人差を考慮した評価が可能と考える
 - 時間帯別の評価
 - 昼間の咳が減る、就寝中の咳が減ることで寝つきが良くなるなど、症状の変化が時間帯によって異なる可能性がある
 - 1日全体の咳回数ではなく、「昼間」「就寝中」など時間帯を分けて治療効果を評価することも検討に値すると考える
- 妥当性などを評価する試験では、変化が観察できる集団を対象とすることが必要
 - ほとんどの被験者が咳の回数ゼロのような状態では、反応性などの評価が困難になるため、適切な対象集団の選定が必要

【第2部：臨床的に意味のある差の設定】

「臨床的に意味のある差」や「臨床的に意味のある変化の閾値」を設定するためにはどのような試験デザインが適切か

【検討事例：仮想事例1】

- 必要なデータ
 - PROなどのアンカー指標とdBMを同時に測定したデータが必要である
 - 必ずしも介入試験である必要はなく、観察研究でも十分な場合があると考える
- 臨床的に意味のある差
 - 患者に対して「咳が何回減れば改善と感じるか」といったアンケートを実施することは有用
 - 対照治療（標準治療）における経時的な変化データを取得し、それに対して試験治療で期待される優越性（例：+〇回、〇倍）を上乗せして評価する
- 新しいdBM指標がゴールドスタンダードとなり得る場合
 - 例えば、これまで咳の回数を定量的に測定する手段がなかった場合、新たなdBM指標がそれ自体ゴールドスタンダードとして機能する可能性がある
 - このような場合、診断薬の評価と同様に、「咳の回数の真値」に近い出力が得られるかどうかを確認するだけでも、十分な評価となり得ると考える

【第2部：dBMを用いた試験計画・解析での懸念点や工夫すべき点の検討】
dBMを用いた試験を計画・解析する上での懸念点や工夫すべき点は何か

- 咳モニターでは、咳のない時間帯を除去してから咳の回数をカウントしている
 - 除去後のデータのみを保存すればよいのか、全測定時間のデータを保存すべきか、検討が必要である
 - ローデータは膨大ではあるものの、重要な情報が含まれている可能性があるため、基本的にローデータも保存することが望ましいと考える
- 試験実施中にコンプライアンス不良が発生した場合はどうするか？
 - 想像するだけでもリスクが大きいため、事前にパイロット試験などで実施可能性を確認しておくことが望ましい
- dBMの大きな特徴の1つは、連続的にデータを収集できる点である
 - 多時点での測定データを適切に扱える解析手法の検討が必要となる
- デバイスの測定が妥当かどうか担保されないまま、医師が研究目的で導入を希望するケースも想定される
 - 「妥当性が保証されていないから」と門前払いするのではなく、試験実施までにどのような準備が必要かを一緒に検討できる体制を整えることが重要である

【第2部：各特性の評価方法と評価基準】

各特性を評価するためにはどのような試験デザインが適切か

【検討事例：仮想事例3】

pEP=primary endpoint、sEP=secondary endpoint

(1) dBM開発の臨床試験として、どのような試験デザインが適切か（医薬品のピボタル試験で使うために）

- Verification（正確性）、Usability Validation（操作性）、Analytical Validation（信頼性、反応性）、Clinical Validation（内容妥当性、構成概念妥当性、基準関連妥当性）を検討する試験が必要
- 特定の1時点における従来指標（疼痛VAS等）との比較に加え、dBMは時系列データとして評価できる

(2) 医薬品のピボタル試験は、どのようなデザインが適切か（Pain indexが指標として確立されている前提で）

- ベースライン期間として2週間くらいとる（連続評価できるというdBMの強みを活かす）
- 割付因子にPain Indexを活用することで、疼痛VASを用いるよりもベースラインの偏りをより正確に避けられる可能性がある
- 試験期間や評価ポイント／タイムフレームは、dBMで測定される薬効発現が現れる期間を考慮する
- 臨床的に重要なのは1時点の疼痛ではなく、疼痛 × 持続時間（AUC）か（疾患ごとに患者の声やKOLの意見を聞くべき）
- いきなりpEPとして活用するのは難しそう。1つの医薬品開発よりもかなり長いスパンで置き換えていくのはどうか
 - ① pEPは疼痛VAS。Recallバイアス等を避けるため、患者がdBM（Pain Index）を参照して疼痛VASを記録する
 - ② pEPはdBMのスコアで補正した疼痛VASとする（例：VASでの「これまで経験した最も強い痛み」を上回ったことがdBMで確認されたときにVASのスケールの上限を補正する等）
 - ③ 「pEP：疼痛VAS、sEP：dBM」の試験経験を積み上げていくうちに、dBM（Pain IndexのAUC等）の方が妥当であるとのコンセンサスが得られてくる。そのうえで、pEPをdBMに置き換える

【第2部：dBMを用いた試験計画・解析での懸念点や工夫すべき点の検討】
dBMを用いた試験を計画・解析する上での懸念点や工夫すべき点は何か

- 欠測
 - デバイス外れによる欠測の懸念がある → 装着率とデータ取得率の関係等を事前に検討することにより、試験開始前に症例採否ルール（閾値）を設定できる
 - 欠測データの取扱い → Estimandを踏まえて検討する必要がある
- オペレーション
 - ウェアラブルデバイスによる測定ができるため、遠隔モニタリングのツールになりうる可能性
- デバイス、アルゴリズム
 - 採用したデバイスがモデルチェンジとなる懸念、経時的な精度劣化のリスク → 事前に必要量を調達し品質管理
 - デバイス内のソフトウェア更新により、センサーから取得した生データ処理（アルゴリズム）の挙動が変わってしまうリスク
→ 更新凍結 or 更新前後の挙動比較など事前に計画しておく
- 一般化可能性
 - きちんとデータが取れた集団（デジタルリテラシーが高く、継続装着できた患者）に偏ってしまう懸念
→ データが取れた集団とそうでない集団の患者背景を比較し、試験結果の一般化可能性の限界を消去法的に考察