

Quality of life データの解析

- 欠測値への対応 -

【背景】

Patient reported outcome は患者本位の医療を実現するために重要であり、quality of life (QOL)はその代表格である。QOL は多くの臨床試験において副次評価項目であり、ベースライン、治療期間を経て追跡期間まで経時的測定が予定される。その目的は、特に予後の悪い疾患において、治療薬の QOL に対する影響を推定し「良く生きる」と「長く生きる」との利益を比較衡量するための情報を将来の患者に提供することに他ならない。しかし予後の悪い疾患を対象としているからこそ、患者の死亡や容体の悪化が頻発し、物理的、倫理的な理由で欠測データが生じる。すなわち QOL データにおいて欠測は本質であり、その解析は不完全データの枠組みの中で議論することができる。

【目的】

課題研究を見こし、QOL データの解析における問題点および対処の方法を先行研究をもとに総括する。

【方法論の紹介】

QOL 対象とした研究では 2 種の欠測が考えられる：(1)item non-response；(2)unit non-response。前者は調査票の質問項目のうち一部の回答を欠いている状況であり、後者は質問票そのものを欠いている状況である。(2)について次のように分類できる：(a)monotone missing；(b)non-monotone missing。前者は、ある対象者に対する最初の時点から時点 t までの QOL は観測されるが、時点 $t + 1$ 以降の QOL が観測されない状況である。後者は、前者以外のすべての状況を指す。実際のデータでは(1)および(2)(a)、(2)(b)が混在していると考えられる。このうち(2)(a)については定式化が可能であり、先行研究において欠測のメカニズムが議論されている。結果変数 Y_i の測定過程を表すパラメータを θ 、 Y_i で条件づけた欠測過程を表すパラメータを φ とし、脱落時点は D_i で示す。これらの同時分布が、

$$f(y_i, d_i; \theta, \varphi) = f(y_i; \theta) f(d_i | y_i; \varphi) = f(y_i; \theta) f(d_i | y_i^o, y_i^m; \varphi) = f(y_i; \theta) Pr(d_i | y_i^o, y_i^m; \varphi)$$

と分解できると仮定する。ただし完全なデータ Y_i に対し、実際に観測された部分を Y_i^o 、欠測している部分を Y_i^m とする。この分解に基づき欠測過程は次の 3 種に分類できる：missing completely at random (MCAR)；missing at random (MAR)；missing not at random (MNAR)。欠測過程が全てのデータと独立な場合、すなわち $Pr(D_i = d_i | Y_i^o, Y_i^m; \varphi) = Pr(D_i = d_i | \varphi)$ のとき、その欠測過程は MCAR と呼ばれる。欠測仮定が、脱落時点までに観測されたデータには依存するが、脱落時点以降の欠測データには依存しない場合、すなわち $Pr(D_i = d_i | Y_i^o, Y_i^m; \varphi) = Pr(D_i = d_i | Y_i^o; \varphi)$ のとき、その欠測過程は MAR と呼ばれる。欠測過程が、脱落時点以降の欠測データにも依存する場合、すなわち MCAR でも MAR でもないとき、その欠測過程は MNAR 呼ばれる。尤度に基づいた推定を行う場合、無視できる欠測メカニズムとは MCAR と MAR、無視できない欠測とは MNAR の場合を指すことになる。一方、経時データの解析によく用いられる一般化推定方程式など尤度に基づかない解析方法では、無視できない欠測とは MAR および MNAR、無視できる欠測とは欠測メカニズムが MCAR の場合のみを指す。

欠測を伴う QOL データの解析において最も懸念される問題は推定結果にバイアスが生じることである。推定値におけるバイアスの存否を議論するには、欠測理由の同定が重要となる。また欠測確率が性、年齢等の対象者特性に依存するのであれば、これらを調整することでバイアスの補正と推定精度の向上が期待される。以下では欠測のある QOL データの解析方法として提案されている方法を紹介する。

Complete case analysis : プロトコルにて予定された測定時点の QOL が全て観測された対象者のみを解析対象とする方法である。利点は極めて簡便であること、共分散分析など完全データの解析に用いる通常の方法が適用できることである。欠点は一般にサンプルサイズが激減するために検出力が低下してしまうこと、および欠測過程が MCAR でないかぎり推定結果がバイアスされることである。

Available case analysis : 観測された全てのデータを用いて解析に用いる方法である。本解析方法では脱落した対象者も解析に含めるため、complete case analysis に比べて効率は良い。また例えば混合モデルのような尤度に基づいた解析を行えば、欠測過程が MACR の場合に限らず MAR の場合でも推定値は一致性を持つ。なお一般に欠測過程が MAR であるか MNAR であるかを判別することは難しく、現実には MNAR を仮定したほうが自然な場合が多いことから、常に尤度に基づいた解析が妥当とは限らない。

Imputation method : 欠測値をある値を用いて補填する方法である。特に欠測値をある値で 1 回だけ置き換える補填方法は single imputation と呼ばれる。次の補填方法がある : 最後に観測された値で補填する last observation carried forward 法 ; 欠測時点以降に最初に観測された値で補填する next observation carried backward 法 ; 回帰式による推定値を用いて補填する regression imputation ; データを完備している対象者からランダムに選び出した値を用いて補填する hot deck imputation。Single imputation は理解しやすく実行が容易である。しかし欠測値の不確実性を考慮しておらず、推定値の標準誤差を過小評価することにつながる。この問題に対し、欠測値を異なる値で M (≥ 2) 回補填することで解決を図る方法が multiple imputation である。本方法では M 個の異なる完全なデータセットが作成される。それぞれのデータセットに関して推定値とその分散を求め、最後にそれらを併合して 1 つの結果とする。

Inverse probability censoring weighted method : アウトカムが観測された対象者を何らかの適当な値で重み付ける方法である。観測されたデータには選択バイアスがあり、IPCW 法ではその選択確率を考慮してバイアスを修正することを目的としている。ここで、選択確率とは、ある対象者がある時点で観測されるかどうかの確率である。選択確率はそれまでの結果変数の値や対象者の共変量のような、欠測に影響すると思われる要因から推定される。IPCW 法は、欠測メカニズムを考慮した解析の 1 つあり、欠測過程が MRA のときバイアスのない結果を導く。

Principal stratification method : 対象者を潜在的な中間結果変数によって層別し、特定の層内で興味のある結果変数の比較を行う方法である。それぞれの層は principle stratum と呼ばれる。割り当てられた治療を Z (1:試験薬, 2:対照薬)、ベースラインの共変量を X 、興味のある結果変数(QOL)を Y 、対象者の生存状態を S (1:生存, 2:死亡)と表す。 Y_z 、 S_z とは、それぞれ $Z = z$ のときの結果変数と生存状態である。死亡した対象者については結果変数は定義されないと想定する。このとき次式で定義される survivor average causal effect (SACE)を推定する : $SACE \equiv E(Y_1|S_1 = S_0 = 1) - E(Y_0|S_1 = S_0 = 1)$ 。

【まとめ】

Complete case analysis はデータが欠測している対象者を除外することで問題に対処する解析態度であり、一方、imputation 法は欠測値を人為的に補填することで問題に対処する解析態度である。Complete case analysis に比べ、一般に imputation 法は不完全データの解析に伴うバイアスの発生と効率の低下という問題に善処できると考えられている。しかし結果の解釈という観点からは、imputation 法は暗に死者の QOL の存在を仮定しこれを「欠測」したものとして補填することから、一般的な理解に馴染まない。本来、死者の QOL は科学という枠組みでは観念できず、定義されないと見なすのが妥当かもしれない。QOL の定義可能性に着目し、死亡のために QOL データが存在しない状況は特に“truncation by death”と呼ばれる。この点、Principal stratification method は、現に生存している対象者の中で試験薬が QOL に及ぼした効果、として解析結果を解釈できる。

課題研究では、それぞれの手法を比べ性能評価を行うとともに、結果がいかに解釈できるのかについても議論したい。