

ベースラインハザードから思うこと

横田 勲 (北海道大学)

私は移動中に音楽を聴きながら、時折ベースラインハザード(関数)について考えることがあります。生存時間アウトカムの解析においてハザード比推定を行う際に標準的に用いられる Cox 比例ハザードモデルの右辺の最初に登場するアレです。ベースラインハザード関数にパラメトリックモデルを置かないことから、指数回帰や Weibull 回帰等に対し、Cox モデルはセミパラメトリックモデルと呼ばれます。ここでは、これまで Cox 回帰をはじめとした生存時間解析法を実データに適用する上でハザードについて感じたことを通して、計量生物学の未来に向けて思うことを記したいと思います。

まず、集団全員に共通するベースラインハザードをおくことへの違和感をもつことが少なくありません。現在、私は日本小児血液がんグループでの生物統計委員を務めており、中でも神経芽腫という疾患の臨床試験を計画・解析する機会がありました。試験は単群で行われますが、予後因子の解析では Cox 回帰が用いられます。神経芽腫は自然退縮(特に治療せずとも腫瘍が縮小したり消失したりすること)がみられることがあり、そのため病期の軽い場合は無治療経過観察が治療方針となることがあります。このような患児はベースラインハザードがゼロであるわけで、推定の問題はさておき、潜在的にイベントを起こしうるハザードが mixture した治療モデルをおくほうが自然です。別の機会に病理の臨床家が“腫瘍の顔”がよさそう、悪そうと仰ることを耳にしたことがあります。予後や治療反応性がよさそうな場合とそうでない場合を意味するのだそうです。このことが正しいならば、治療との交互作用として扱うこと、部分集団によって異なるベースラインハザードをおく層別の回帰モデル、対象者によってベースラインハザードが異なることをゆるす変量効果(frailty)モデルのような選択肢が想像できます。がん臨床試験では、time to treatment failure という、化学療法を中止せざるを得なくなるまでの期間が代替評価項目的に用いられることがありますが、これはすなわち、化学療法継続中ではハザードが小さく、中止後にハザードが大きくなる値をとるのでしょう。この結果としてみられる直線的に降下する生存曲線を目にすると、Weibull 分布が mixture するような表現だけでなく、多状態プロセスをおけば対象者間で異なる中止時点も考慮できるのかなあ等の想像が膨らみます。

(ベースライン)ハザードにおく分布を考えさせられることもあります。生理学的指標や臨床検査値のような連続変数がある閾値を超えた場合にイベント発生と定義した評価項目がしばしばみられます。この過程をギャンブラーの破滅(gambler's ruin)で知られるような first hitting time モデルとして捉え、閾値回帰を応用する手法を知ったときは見事だと思いました。

以上の内容を通して、私が普段関わることの多い医学データに、メカニズムや生成過程を反映するようなモデルを積極的に適用できる未来に近づけたらいいなと、今考えています。もちろん、上述したようなやや複雑なモデルは、実際の医学データに適用するとパラメータ推定が不安定になりやすいような困難があるかもしれません。また回帰係数を対数ハザード比として扱えるような簡単なモデルではないので、結果の提示方法にも多分のアイデアが必要となるでしょう。とはいえ日頃、疫学者や臨床家と共同研究する際に、研究者のメカニズムやストーリーを具現化したモデルにデータがうまくあてはまった時の快感は統計家だから感じられる楽しみであり、そのために方法論の整備をしていきたいと考えます。現状、他のことにかまけて、自身の目の前のデータを深追いすることさえ中々できない自戒の念を込めつつ、これからも計量生物学の発展に少しでも貢献できるよう頑張っていこうと思います。