

検討課題に関する意見

山口拓洋(東北大学医学統計学)

(1) データベースの必要な規模

以下の(2)でも述べますが、本事業においてはまずは内的妥当性(正確性)を確保することが重要と考えます。そのうえで、データの精度を考慮すべきです。データの精度を向上させるには、サンプルサイズを増やすこと、データの品質管理等が重要です。

ピオグリタゾン(アクトス)と膀胱がん発症との因果関係の証明に、米国では約 20 万人のコホート(Kaiser Permanente Northern California; KPNC データベースから、3200 万人加入)を利用、フランスでは約 150 万人のコホート(国民健康保険データベースから、6500 万人加入)を利用し、同薬により膀胱がん発生リスクが増加する傾向があることを証明しています。もちろん、ターゲットとするイベントの発症や薬剤への曝露の頻度(以上は経時的変化も考慮すべき)にも当然依存しますが、このような数千万規模のデータベースが利用できることのメリットは大きいと考えます。

(サンプルサイズの設定根拠の具体例は、例えば前回の参考資料5の12,13/41などが参考になるかと存じます。)

(2) データの代表性(一般化可能性)

まずはデータの内的妥当性及び集積可能性の担保を優先せざるをえないと思います。次には(1)の精度の問題が重要であり、代表性(一般化可能性)の議論はそのインフラがある程度整備されてからの議論でも遅くはないと思います。また、東大病院は別としても、他に加わっている大学病院が非常に特殊な患者層であるかと言えばそうとも言い切れない側面もありますし、徳洲会グループ、今後は NTT グループ等も入ってきますので、代表性云々で本事業の有用性が否定されるのは問題ではないでしょうか。

内的妥当性を保つ工夫としては、病院ベースのデータを集めるアメリカの研究の Case-control surveillance (CCS) というシステムにおいては(CCS 自体は 2009 年で終了)、病院の周囲 50 マイルの人たちに限ることによって比較可能性を確保しようとするアイデアが使われていました。

(3) 地域連携のあり方(実効性)を含むフォローアップの重要性

(特定個人が複数の医療機関で受診した結果を追いかけるようにするべき等)

東京大学薬剤疫学久保田教授が過去に実施された研究にて、東大病院のデータで糖尿病患者の透析への移行を見ようとされた事がありますが、東大病院には入院中の患者のための透析施設しかなく、透析が近づくと外来の患者は次々に大きな透析施設のある病院に移っていき、透析への移行を見ることはできないという結論になったと聞いております。医療情報データベースを用いた研究においても、ヘルシンキ宣言(30条)に定められている「結果の完全性と正確性の説明責任を負う」ためには、特定個人を見失わずに追跡し続けなければならない、結果の正確性を保てない研究は研究自体を行う価値がなく、倫理的ではないと個人的には考えます。例えば、久保田教授らのグループが以前やられた別の研究では、大学のデータだけを使って何とか比較的長期間のフォローをしたいと考えて、観察開始前の1年間を前期6か月と後期6か月に分けて両方の半年の期間に少なくとも1回ずつ来ている人をrepeat user、どちらか1回だけ来ている人をtemporal userに分けると、東大病院では外来患者はおおむねほぼ半分ずつに別れ、観察期間を3年としたとき、平均の観察期間はrepeat userで2.2年、temporal userは0.8年で、このうちrepeat usersのみを対象にして、比較的長期の観察が必要な事象発生をみたことがあるそうです。また、他の医療機関のデータを得ないと正確なイベント発生は不明であっても、ある程度把握でき、かつ把握できない程度が比較される二群で大きく異なるとは考えられないのなら、比較は許されるかもしれません。まずは、データを収集蓄積していくという観点が重要ではないでしょうか。また、脱落した人とそうでない人で観察開始時の背景因子が異なるか(どのような要因が脱落に影響を及ぼすのか)を検討するのは有用と考えます。

今後、本事業の医療情報データベースのみならず、他の様々なデータベースが必要に応じて互いに連結可能となり、正確性・信頼性の高い最終アウトカムを得ることが可能な臨床・疫学研究が盛んに実施され国民の生活向上に寄与できるような体制づくりを進めるべきであり、十分な議論が必要と考えます。

(4) 最後に

何のためのデータなのか？そのためにどのようなインフラを整備する必要があるのか？常に原点に立ち返ることが重要と個人的に思っています。

以上