

## 第7回バイオバンクオープンフォーラム

株式会社ステージン 統計解析事業部

上辻 茂男

Shigeo Kamitsuji, PhD, kamitsuji@stagen.co.jp

株式会社ステージンの上辻と申します。弊社は元理化学研究所センター長の鎌谷直之先生の指導の下、アカデミア、企業へゲノム研究支援サービスを提供する会社です。弊社は、これまでに6つのプロジェクトでBBJデータを活用させていただきました。

BBJデータの最大の魅力はサンプルサイズです。約16万人分の網羅的SNPデータ、臨床情報の利用が可能です。弊社案件の場合、希少疾患や、薬剤服用による副作用発症の遺伝的原因探索を目的としたゲノム研究が多く、ケース群のサンプルサイズが非常に小さいです。そこでBBJデータをコントロール群として利用することで、統計的検出力を高めることが可能です。利用申請前に、臨床情報の該当症例数を検索することができますので、検出力シミュレーションが可能です。また最近ではPost-GWAS解析のご依頼も増えてきました。その1つであるメンデル無作為化解析をBBJデータで行い、自社開発品のドラッグリポジショニングの可能性を検討するお客様もいらっしゃいます。

BBJは、申請からデータ取得までの期間が早いと皆様おっしゃいます。弊社案件ですと、ゲノムデータの取得に約3週間、臨床情報の取得に平均2か月です。また、データの保有は弊社だけで、弊社で解析を行い、お客様は結果だけを受領する、このような仕組みがBBJでは可能です。そのため申請は弊社が行い、BBJデータの利用者として弊社の名前のみが掲載されます。企業にとっては都合がよいとのこと。

もう1つお客様の声として多いことは、「BBJは疾患群で、コントロールデータとして利用することに問題はないのか」という質問です。

たとえば糖尿病患者への薬剤服用による副作用発症のゲノム研究を行う場合、糖尿病患者のみをBBJから選択しコントロールデータとします。ケース群、コントロールともに糖尿病患者ですので、疾患関連遺伝子が検出される可能性は低くなります。このときに「BBJに登録された糖尿病患者の中には薬剤を服用して副作用を発症する患者もいるのではないか？」とのご質問をいただきます。たしかに否定できませんが、発症率が低いのであれば、原因SNPのアレル頻度は低いことが予想されます。つまりBBJデータの中に、服用して副作用を発症とした症例がいたとしてもその影響はほとんどないと考えてよいでしょう。

次に、希少疾患患者をケース群に、BBJデータをコントロール群としたGWASを行うとします。ここで、東京大学医科学研究所、理化学研究所を中心に行われたBBJのGWAS結果が登録されているJENGERを利用しますと、それぞれの疾患における関連SNPはほとんど共通しておらず、またBBJのアレル頻度分布は、その他のデータベースとも差はないことから、アレルはランダムに割り振られていると考えることができ、統計的には一般集団と変わりません。なによりも、JENGERに登録された先駆的な研究論文が、BBJデータをコ

ントロール群として使うことの有用性を示しています。

BBJ データは進化し続けています。先ほどの森崎先生のご講演の中でも紹介されましたが「BBJ データパッケージ」の登場は、BBJ 事業の大きな変革点に感じます。これまで、すべての情報を利用するためには共同研究契約しかありませんでした。BBJ データパッケージは利用期間内で自由にゲノム研究が可能で、研究成果を自由に使えるとのこと。また疾患の有無や臨床検査値など症状に近い形質だけでなく、中流的なメタボローム・プロテオームデータが活用できることによって、メカニズムの解明や、薬剤のターゲットとなり得る分子の同定が期待されます。詳しくは、明日 8/2 15:10 からの「シンポジウム3「新たなつながりが創るデータ・バンク研究の未来」で、弊社会長の鎌谷直之先生がお話する予定ですので奮ってご参加ください。

日本が世界とスピード感を持ってゲノム研究を進めていくために、BBJ をはじめ、公開ゲノムコホートデータの活用は不可欠です。スタージェンはこれからも日本のゲノム研究の支援に尽力して参ります。ご清聴ありがとうございました。

以上