

HGVDDB で既に公開済みのデータの取り扱いについて

(独)科学技術振興機構

バイオサイエンスデータベースセンター

ヒト由来データの統合および手続きの効率化を図るため、統合化推進プロジェクトの成果の一つであるヒトゲノムバリエーションデータベース(Human Genome Variation Database: HGVDDB)におけるデータ共有を、NBDC ヒトデータベースにおけるデータ共有と一本化することを検討している。

HGVDDB は NBDC の統合推進事業の前身にあたる「統合データベースプロジェクト」(平成 18-23 年)の実施課題の一つであり、当時よりヒトゲノム上のバリエーションデータの共有に必要なルールの策定および審査委員会の設置を進め、データ共有の仕組みを整えていた。しかしながら、NBDC が平成 25 年 4 月に策定した NBDC ヒトデータ共有ガイドラインと相違点があるため、一本化に際してすり合わせを行なう必要がある。

1. ガイドラインの相違

(1) データカテゴリ分類の違い

HGVDDB のカテゴリは以下の通りである。

レベル		1	2	3
データの種 類	GWAS データ	カテゴリ-A (頻度データ、統計解析結果)	カテゴリ-B (個体レベルでの CNV 情報)	カテゴリ-C (GWAS 遺伝子型データ) カテゴリ-D (GWAS 生データ)
	リシークエンシ ングデータ	カテゴリ-E (公知のリシークエンシングおよ び変異データ)	カテゴリ-F (リシークエンシングおよび変異 データ)	(該当なし)
データ使用時の手続き等		ウェブサイトにおいて、閲覧可 能。但しカテゴリ-A のデータを 大量取得する場合はレベル 2 と 同様の申請を要する。	氏名、職名、連絡先、使用目 的、e-mail アドレス(原則、所属 機関から発行されたアドレス)を 申請する。	データアクセス申請書を提出し、 許可を受ける。また、使用期間 に応じて、データ使用報告書の 提出を要する。

一方、NBDC ヒトデータベースのカテゴリは、データの種類を問わず、集計情報等の個別データではないデータをオープンデータとし、個別データを含むデータを制限公開データ(Type I / II)としている。ただし、インフォームドコンセントの説明・同意文書にカテゴリの指定がある場合は準ずるものとする。

(2) インフォームドコンセント等の確認状況

HGVDB では、データベース上で公開する集計・統計情報の公開は学会発表や学術雑誌等への結果の公開と同等と捉え、データベースへの公開に際して別途“データベース上での共有”についての記載は必要としないという判断をしてきた。

NBDC ヒトデータベースの共有ポリシーでは、オープンデータ・制限公開データにかかわらず、IC 説明・同意文書や研究計画書内への“データベースへのデータの提供および研究者間におけるデータの共有”に関する記載および倫理審査委員会による審議と承認が必須としている。

2. NBDC ヒトデータベースと HGVDB とのデータ共有化

1. の通り、NBDC ヒトデータベースと HGVDB におけるデータ共有ポリシーは異なるものの、スムーズなデータ移行を目指すため、以下の様な方針で一本化を実施し、NBDC ヒトデータベースの一部としてデータの共有化を進めることとしたい。

(1) HGVDB レベル1(カテゴリ A・E)データ

NBDC ヒトデータ共有ガイドライン策定以前の H25 年 4 月までに既に公開しているデータであることを踏まえて特例として扱うこととし、データベースでのデータ共有についての再同意や所属機関の倫理審査委員会による承認をデータ提供者に求めずに NBDC ヒトデータベースのオープンデータとして移行する。

ただし、カテゴリ A の疾患群のデータについては制限公開データ(Type I)とする。

(2) HGVDB レベル2(カテゴリ B・F)データ

より安全な取扱い方へ寄った形で、制限公開データ(Type I)として移行する。

(3) HGVDB レベル3(カテゴリ C・D)データ

制限公開データ(Type I)として移行する。

なお、HGVDB レベル1(カテゴリ A・E)データのうち、カテゴリ A の疾患群のデータおよび HGVDB レベル2(カテゴリ B・F)データについては、データ提供者にオープンデータでの提供で良いかを確認し、確認できたデータについてはオープンデータとして NBDC ヒトデータベースから公開していくこととする。

以上