

8. 研究利用についての決定	42
誰がどんな決定をすべきか？	
監視及び諮問のしくみ	
多すぎる委員会	
監視の必要性	
9. データ共有のための保存管理	46
データの適正化と文書化	
ESRC/UKDA の事例	
医学・生物学データの保存管理	
登録と窓口	
10. 所見と結論	50
所見	
結論	
補遺1	54
研究事業協力者	
補遺2	56
著者	

## 序論

本書は、独立したコンサルタントから医学研究評議会 (Medical Research Council, MRC) 及びウエルカム財団への報告書であり、主にこの二つの組織が研究助成しているかまたは責任を負っている英国の集団を対象としたデータ・試料コレクションの研究利用の際の諸問題について報告しているものである。

本調査は英国において MRC 及び財団が単独または共同で支援したコレクションの全体あるいは全領域を利用する際の取り決めについて検討するため、研究助成機関がコンサルタントに依頼したもので、以下のことを決定する際の参考とすることを目的としている。

1. 現行の利用協定をどの程度、標準化できるか。
2. 倫理的、法的、実際的問題を考慮した場合、より広範囲で標準化の余地があるかどうか。
3. ひな形となるガバナンス構造が可能であるかどうか。もし可能なら複数のコレクションに通用する中心となる要素を有効に使うことができるようにとり入れることができるかどうか。
4. 研究助成したコレクション全体に適用できるガイドラインの作成が可能であるかどうか。

これらの質問に対する解答は、同意と守秘性に対する懸念、科学的課題、所有権に対する関心、決定権限と監視についての問題、及び実際的に考慮すべき事項により左右されることが検討開始早々に明らかになったため、MRC とウエルカム財団は付託検討事項の範囲を広げた。

本調査は多数の研究責任者、情報管理者、資料(情報・試料)管理者、研究助成機関代表者、監視または諮問委員会委員長との討議で進め、実施のための文書及び技術文献の検討で補完した。多忙にもかかわらず貴重な経験や意見や見識を提供いただいた皆様に心から感謝申し上げたい(ご協力くださった方々については、補遺1に記載)。

本調査は UK Data Archive の Dr Louise Corti、Ms Melanie Wright 両氏の作成した 2002 年度 MRC とウエルカム財団の報告書<sup>1</sup>を参考に構成したものである。

1 Corti L, Wright M. 'MRC Population Data Archiving And Access Project Consultants' Report: Developing an MRC policy for population data archiving and access' (「MRC 集団データ保管及び利用研究事業の検討報告のコンサルタントリポート: 集団データの保管と利用のための MRC 方針の開発」)(草案)。医学研究評議会; 2002 年 9 月  
[www.mrc.ac.uk/prn/index/strategy-strategy/strategy-science\\_strategy/strategy-strategy\\_implementation/strategy-other\\_initiatives/strategy-data\\_sharing/pdf-ukda\\_draft\\_report.pdf-link](http://www.mrc.ac.uk/prn/index/strategy-strategy/strategy-science_strategy/strategy-strategy_implementation/strategy-other_initiatives/strategy-data_sharing/pdf-ukda_draft_report.pdf-link) [2005 年 12 月 21 日アクセス]。

## 要約

コレクションの研究利用は改善可能であり、ほとんどの科学者がそのように考えている。しかし、研究利用を最適化するには、その障害を減らすだけでなく、研究利用の提供を積極的に奨励、指導し、資金援助し、その努力に報いる必要がある。

研究利用の活性化を求める意見には以下のものがある。

- ・ より高い情報価値と、投入した資金や努力に対してより多くの成果を引き出し、それによって社会の健康ベネフィットを増加させる。
- ・ 専門的または再現が困難か高価である研究資源への研究者の研究利用を拒否せずに、上記の良識や科学的自由の伝統に貢献できる。
- ・ 研究の検証や視点を変えた再現実験が可能となる。
- ・ 非生産的な努力が繰り返し行われていることを減らすことができる。
- ・ 患者や市民の研究参加や検体の提供、同意書提出の負担を最少にする。
- ・ データセットや試料の関連付け、集積、他のデータセットや試料との比較を促進する。
- ・ より多くの、より多様な研究者がデータ分析を行い、能力を高め、方法や手順を改善し、研究成果を発表・報告し、新しい研究方向を発見し、それによってコレクションの質が高められ、その内容が多様性をもって豊富になる。

適切で効率的な研究利用を増加させるには、以下のことに注意しなければならない。

- ・ 研究利用が提供された場合、同意、守秘、離脱の権利に関するものなど、参加者との最初の約束が完全に守られる。
- ・ 研究事業の一般社会での評価や患者との関係が損なわれない。
- ・ 研究事業の科学的評価が高い水準に保たれる。
- ・ 研究資源開発者や管理者の利益を保護し、その努力と善意に応える。
- ・ データの匿名化、文書化、情報の保管、研究利用申請の審査や研究利用協定の取り決め、データ及びまたは試料の準備、そして科学的解釈の援助など、研究利用が可能となるまでに生じた努力と経費に十分見合う。
- ・ 知的財産管理が適切になされる。

以下は、助成機関の質問に対する回答の要約である。

Q: 現行の研究利用に関する取り決めはどの程度まで標準化されていますか。

A: 共通点は多いですが、あまり標準化されていません。

Q: 倫理的、法的、実用的問題を考慮しても、さらに標準化を進める余地はありますか。

A: はい、あります。例えば研究利用、試料の提供協定の基本となる条項の標準化が求められています。資料への研究利用の基本項目の標準化は有用です。他にも考えられます。

Q: ひな形となるガバナンス構造を形成できる可能性はありますか。もしあるなら、複数のコレクションに通用する要素をとり入れることができますか。(前出序論3)

A: 最も多いガバナンスの形態は、監視委員会によるコレクションや情報の共有を監督するものです。多少の違いはありますが、これらの大半はよく似た権限や構造的特徴を持っています。現行の監督機関がもつ最良の特徴を組み入れたモデル構造の構築が可能でしょう。

この報告書では多数の監視または諮問委員会が連携して機能している状況事例を紹介しています。

Q: 助成を受けたコレクション全体に広く適用できるガイドラインを開発する余地はありますか。

A: 同意、守秘、匿名化については早急の明確化と指針の改定が必要です。徹底的に匿名化した「限定されたデータセット」の特徴に関する指針も有用でしょう。

ガイドラインが最善の手段であるかどうかは分かりませんが、データ提供者の権利と義務(情報利用申請者の資格審査、研究提案の利点、申請文書の質などに関して)、そして情報利用申請者の権利と義務(希望すれば共同研究を拒否できるのか、関連を示すことのできなかつた結果でも発表を求められるのかに関してなど)を明確にすることが今重要です。

複数のコレクションの研究利用について方針及び手続きを公表する際には非公式であっても上記のような問題に関するガイダンスが必要です。これは研究利用協定の要件についてのガイダンスと組み合わせて作ることも可能でしょう。

規準や指針は、データセットの科学的に確かな、そして自由な研究利用を支援するために必要な文書に基づいて作成するとよいでしょう。

つまり、多くのことができる可能性があります。

## 1. データ共有を推進するもの

最近、研究のための最適化された情報共有を求める声が強くなっている。それは、研究参加者の協力や、公的及び慈善資金による助成や、大学や医療機関による援助による社会的ベネフィットを最大に引き出そうとするものである。

データ共有には様々な課題があり、本書が‘研究利用’の問題として調査したもの（例えば保存・管理の問題と比較）は特に重要である。

### 1.1 MRC 及びウエルカム財団の方針

#### 1.1.1

MRC は‘Statement on Data Sharing and Preservation Policy’「データ共有化及び保存方針に関する声明」を採択し、その冒頭で次のように宣言している。

MRC が助成している研究から生まれた価値あるデータは、可能な限り制約をなくし、新しい研究が可能となるよう、科学者社会が入手できるものとするを MRC は希望している。またそのデータは時宜を得た責任ある行動で共有されなければならない。

データ共有化で可能となった新規の研究は、全 MRC 助成研究の科学的品質、倫理性、資金に見合った価値の高い標準に適ったものであり、かつ利用された元のデータセットに明確な価値を付加するものでなければならない。<sup>2</sup>

#### 1.1.2

また先に述べた声明にはデータの共有と保管に関する目標が明確に述べられているので、ここに広範囲に引用する必要があると考える。

データの共有化を可能にする研究者は、既存のデータを利用した二次的な研究の推進について、研究助成機関や研究が行われている学術機関、二次研究を始める新規利用者から、十分かつ適切な評価を当然受けるべきであると MRC は考える。

そのような研究は、新規利用者と元のデータ作成者または管理者などすべての関係者が共同作業を始める前に、責務と権利について合意に達しているときに最大の成果を上げることが多い。

一次研究のための明示された限定的期間、データが専有されることは妥当なことである。データの性質や価値、またデータが作られ利用される方法を考慮に入れた、様々な秩序のもとに様々

な取り組みが求められる。データセットの完成に貢献している進行中の研究を、時期尚早な、あるいはご都合主義的な共有化や分析で損なってはならない。共有化では常にデータの長期的価値を高めるように配慮しなければならない。

MRC の方針は、研究発表前の特許申請を阻むものではない、むしろ短期間発表を遅らせ出願書類の検討・作成の時間を取る必要のあることも時にはあると認識している。

個人データを扱う医学研究では、特に同意や守秘に関係した個人データに伴う特別な責務が生じる。従ってこの種のデータ共有化に当たっては事前に適切な管理許可(倫理、法律、制度上)を得ることが不可欠である。研究者、研究参加者、及び研究管理者は、法の定める規制要件の範囲内で、新規利用者の利用が不必要な制限なく最大になるように配慮しなければならない。また患者や一般人が享受可能な研究ベネフィットは、必ず認識されているリスクを上回らなければならない。個人情報漏えいなどのリスクは、そのリスクに見合う強固な方法で管理する必要がある。

有効な共有化を可能にするには、データをその生成から廃棄にわたって厳密に管理し、適切で高品質な管理データをと共に公開しなければならない。これは、財政的支援を受け一次データを作成あるいは収集した、通常個人か機関であるところのデータ所有者(owner)の責務である。

2006年1月1日から、MRCに研究助成金申請書を提出する申請者全員は、各人のデータ保管及び共有化についての戦略を説明する文書も提出しなければならないこととなった。受け取るMRCの研究助成金から生成されるデータを共有したくないと考えている申請者は、データを共有できない明確な理由を書いて提出しなければならない。<sup>3</sup>

### 1.1.3

ウエルカム財団はコレクションの利用全般にわたる方針を持っているわけではないが、その幹部や諮問機関はますます、研究課題ごとに、意義ある場合には広範囲な利用提供ができるように勤めている。助成する研究課題の多くに特定のデータ共有要件がある。財団ではヒトゲノム配列情報はパブリックドメインに置き無料で自由に利用可能にするべきだとの主張がなされ、財団の研究助成を受けた研究課題には公共のデータベースへの遺伝子型データの可及的迅速な提出を要求している。

## 1.2 NIH の方針

米国の国立衛生研究所(US National Institutes of Health, NIH)は、「参加者のプライバシーを守り、機密及び所有権データ安全管理措置を取った上で、可能な限り多くの人が無料でデータを入手できるようにすべきである」と考えている<sup>4</sup>。また、50万ドル以上の補助金を申請する者は申請書類に研究による成果データの共有化計画、または非共有化の明確で正当な理由を添えることを要求としている。

NIH の論理的根拠は以下の通りである。

データの共有化によって、自由な科学研究が促進され、多様な分析や意見が奨励され、新規研究が促され、新規またはもっと異なった仮説や分析方法を試すことが可能となり、データ収集方法や計測に関する研究が支援され、新しい研究者の教育が容易になり、最初の研究者が構想しなかった研究課題の探求が可能となり、そして複数のコレクションからのデータを統合した新規データセットの生成が可能となる<sup>5</sup>。

### 1.3 その他の推進策

#### 1.3.1

ウエルカム財団は 2003 年のフロリダでの会議で、発表前のデータ公開についての問題を話し合った<sup>6</sup>。特にこのときのゲノム配列データについての話し合いは、フォートローダーデール協定 (Fort Lauderdale Agreement) として知られるようになった。この協定では、公共データベースへの速やかな遺伝子配列データの公開が奨励され、そのための研究助成機関、データ作成者、データ利用者の役割の大枠が描かれている。生の配列データを迅速に公共データベースに掲載すること、(早い時には SNP マッピングされたその日のうちに掲載されることもある) は、ゲノム研究者社会の習慣となっている。

#### 1.3.2

フォートローダーデールで話し合った人々は、「人集団を対象とした研究事業として設立された大規模データ生産センターの他の種類のデータにも迅速な発表前公開の原則を適用すること」を推奨している。だが、この点についての発展はなかった。その理由はおそらくその人々がそのような発言する影響力をもつ人々ではなかったためであろう。

#### 1.3.3

多くの国々がこれまで述べたような理由でデータ共有化を推奨している。ウェブ上のデータベース、DNA バンク、公共の分子ライブラリー、臨床試験データの記録や公表の発展と、また、より自由な公共の科学成果の発表を目指す様々な動き、それらは広範囲な利用を高め合うものである。

#### 1.3.4

データ共有化を進めるもう一つの力について書くべきであろう。それはコホートや臨床試験に参加、検体を提供、自身の医療記録の分析に同意、あるいは別の方法による、データ生成に貢献した一般の人々が関心を寄せる問題である。自らの意思で研究に参加した人々は、データや試料の利用—普通に考えられる幅広い利用を意味するのだが—により、建設的で迅速な研究の進歩が生まれることを当然のこととして望んでいる。

- 2 'MRC Statement on Data Sharing and Preservation Policy' (「データ共有化と保存方針に関するMRC声明」) 医学研究評議会; 2005年9月. [www.mrc.ac.uk/index/strategy-strategy/strategy-science\\_strategy/strategy-strategy\\_implementation/strategy-other\\_initiatives/strategy-data\\_sharing/strategy-data\\_sharing\\_policy.htm](http://www.mrc.ac.uk/index/strategy-strategy/strategy-science_strategy/strategy-strategy_implementation/strategy-other_initiatives/strategy-data_sharing/strategy-data_sharing_policy.htm) [2005年12月21日アクセス].
- 3 前述
- 4 'Final NIH Statement on Sharing Research Data' (「研究データ共有化に関するNIH最終声明」). 米国国立衛生研究所; 2003. <http://grants.nih.gov/grants/guide/notice-files/NOT-OD-03-032.html> [2005年12月21日アクセス].
- 5 プレスリリース. 米国国立衛生研究所; 2002年3月1日.  
<http://grants.nih.gov/grants/guide/notice-files/NOT-OD-02-035.html>  
[2005年12月21日アクセス].
- 6 'Sharing Data from Large-scale Biological Research Projects: A system of tripartite responsibility' (「大規模な生物学的研究プロジェクトからのデータを共有: 三者間の責任システム」). ロンドン: ウェルカム財団; 2003.  
[www.wellcome.ac.uk/assets/wtd003207.pdf](http://www.wellcome.ac.uk/assets/wtd003207.pdf) [2005年12月21日アクセス].



## 2. コレクションと用語

### 2.1 いくつかのコレクションの概略

現在利用できる研究資源の多様性と豊かさの価値を知ることで、理に適った研究利用がなぜ重要であるのかが鮮明になる。以下は全体像の説明のために加えた少数を除き、MRC やウエルカム財団による助成を受けている英国の主要なコレクションの一部である。ほとんどについては本報告書において状況が述べられている。

1946 Birth Cohort(1946 年出生のコホート)(National Survey of Health and Development, NSHD)<sup>7</sup>. 1946 年 3 月の 1 週間に生まれた人々の健康と成長を追跡している進行中の研究。MRC による研究助成を受けている。University College London の Epidemiology and Public Health 学部、Michael Wadsworth 教授主導で実施されている。最も長期間追跡しているコホート研究の一つで、何回ものデータ収集を繰り返しながら徹底的に研究されている。研究利用は主に直接の共同研究により行われている。

Aberdeen 'Children of the 1950s' Cohort(アバディーンの「1950 年代出生小児」コホート)<sup>8</sup>. 再開された歴史的コホート。1950～1956 年にアバディーンで生まれ 1962 年の障害児学校教育研究に参加した 12000 名からなるコホートで、彼らの出生時の体重、妊娠期間、身長、体重、認知検査値、社会経済指標、その他のデータを収集した。1998 年から参加者大半について追跡調査し、新しいデータを収集、そしてスコットランド健康登録(Scottish Morbidity Records)システムのデータに関連付けされた。現在その研究資源は生涯の、及び世代を超えての健康への影響の研究に利用されている。

1958 Birth Cohort (1958 年出生のコホート)(National Child Development Study, NCDS)<sup>9</sup>. 1958 年 3 月の 1 週間に生まれた英国住民の身体的発育、及び教育、社会性、経済的発達についての進行中の研究。いくつかの社会科学的研究機関によって開発され、現在ロンドン大学 Institute of Education(教育研究所)、Centre for Longitudinal Studies(長期研究センター)によって実施されている。主に the Economic and Social Research Council(社会経済研究評議会)による研究費助成。MRC 研究助成によるコホート医学・生物学的調査は近年終了した。ウエルカム財団はヒト由来試料からの不死化細胞系の作成について研究助成している。一流の研究者が多数加わっている。そのデータセットを利用している生物・医学研究の一つに外部で行われているいくつかの症例対照研究がある。

1970 British Cohort Study(1970 年英国コホート研究)<sup>10</sup>. 1970 年 4 月の 1 週間に生まれた英国国民に関する進行中の研究。妊娠等、身体の状態についても取り扱うが、主として識字能力や雇用など社会科学的課題についての研究である。1998 年からロンドン大学 教育研究所の管轄となった。Jane Elliott 博士が研究責任者(Principal Investigator)である。MRC やウエルカム財団が関与してきたとは考えられず、1958 Birth Cohort のような医学・物学的研究要素はない。これも(事実上)一連の英国の出生コホート研究の一部として考えられるのでここで言及することにした。

ALSPAC(Avon Longitudinal Study of Parents and Children(親と子に関する Avon 縦断的研究))<sup>11</sup>. 「Avon Children of the 90s Study(90 年代研究の Avon の子供たち)」として知られている。1991～1992 年に登録した Bristol 地域の妊婦 14500 例を対象に小児の健康と成長に影響する要因を調べる進行中の研究。Bristol 大学が実施。MRC とウエルカム財団が主に研究助成しているが、同大学も出資。研究事業開始時からエグゼクティブ・ディレクターを務めてきた Jean Golding 教授が 2005 年夏退職し、指揮権は George Davey Smith 教授に移った。何億ものデータを作成し、胎盤から乳歯までの 50 万点にも上るヒト由来試料を有する ALSPAC の貴重なコレクションは多くの共同研究に利用されている。

Millennium Cohort Study(2000 年コホートの研究)<sup>12</sup>. 2000～2001 年に英国の 4 地域で生まれた小児 18000 名以上を対象に徹底した記録を行ったコホート研究。社会経済研究評議会及び政府の省庁連合による研究助成。ロンドン大学教育研究所の Centre for Longitudinal Studies(CLS=縦断的研究センター)により実施。Heather Joshi 教授が研究事業本部長を務める。多数の共同研究者が関与し、多くの課題について研究が行われている。

Southampton Women's Survey(サザンプトンの女性に関する調査研究)<sup>13</sup>. 数千名の妊婦を対象とした進行中の研究で、生まれて来た子供たちの後期健康に与える様々な胎児期及び出産前因子の影響を研究している。ある意味で ALSPAC を補完するもの。Hazel Inskip 博士の主導でサザンプトンの MRC Epidemiology Resource Centre(MRC 疫学研究資源センター)で実施されている。

Hertfordshire Cohort Study(ハートフォードシャーコホート研究)<sup>14</sup>. 1931～1939 年にハートフォードシャーで生まれ現在も住んでいる 3000 名を対象とする進行中の研究で、ゲノム因子や早期環境因子と後の慢性疾患発症との関連を研究。外部研究者多数と共同でサザンプトンの MRC 疫学研究資源センターで実施。

Arthritis Research Campaign (ARC) Epidemiology Unit(関節炎研究活動疫学ユニット)コレクション<sup>15</sup>. Alan Silman 教授が率いるマンチェスター大学の一部門であるこのユニットに、ARC が主に研究助成しているものだが、MRC、ウエルカム財団なども助成している。このユニットにはリウマチ障害や筋骨格障害に関して豊富なデータ及びヒト由来試料のコレクションがあり、多数の共同研究が行われている。

Whitehall and Whitehall II(ホワイトホール及びホワイトホール II)<sup>16</sup>. ホワイトホール研究事業は 1967 年に始まり、公務員男性の「社会環境の違い」に関連した健康格差について研究したもので、その成果として社会心理的因子は食事やその他の健康因子と同様に健康格差を説明する重要な因子でありうるという重要な示唆を生んでいる。1985 年以降、「仕事、ストレス、そして健康」研究としても知られているホワイトホール II がこれらの調査を引き継ぎ拡大している。非常に多様な研究助成で実施されている。University College London、疫学公衆衛生学部、Sir Michael Marmot 教授により主導。

Caerphilly Prospective Study(ケアフィリー前向き研究)<sup>17</sup>. 1979~1999 年にかけて MRC 南ウェールズ疫学部門は、ウェールズの男性数千人を対象に心臓血管やその他健康面の調査を行った。1999 年その部門の閉鎖後、ブリストル大学 Social Medicine(社会医学)部の保存管理下で利用可能な状態で引き継がれたコレクションとして管理されている。運営(steering)グループが研究利用に関する決定を行っている。

EPIC(European Prospective Investigation of Cancer)(欧州前向き癌研究)- Norfolk<sup>18</sup>. 欧州の 10 カ国にわたる大規模研究事業で、生活様式(特に食事)、遺伝因子と、癌その他の慢性疾患との関連性を分析したもの。英国コホートは現在 Norwich 近郊に住む 24000 名から構成されている。Addenbrook's Hospital とケンブリッジ大学が提携。

MRC DNA コレクション<sup>19</sup>. 冠動脈疾患、遅発性アルツハイマー病、ぜんそく及びアトピー性皮膚炎、乳癌、結腸直腸癌、二型糖尿病、糸球体腎炎、高血圧、急性白血病、加齢性黄斑変性症、多発硬化症、パーキンソン病、単極性うつ病の疾患別に構成された 13 のコレクションであり、様々な大学に保管されている。

MRC DNA バンク<sup>20</sup>. 上記及びその他のコレクションの試料を保管。マンチェスターの Centre for Integrated Genomic Medical Research 及び Porton Down の European Collection of Cell Cultures から構成されている。

UK Biobank<sup>21</sup>. 幅広い研究利用に適合した供給ができる大規模かつ多目的の研究資源として開発されている。2006 年から 40~69 歳のボランティア 50 万名を募集し、アンケート調査をし、血液その他のヒト由来試料を集め、参加者の NHS 記録やその他のデータベースと関連付けをしている。データコレクションについては多数の大学や保健医療機関が参加している。MRC、ウエルカム財団、保健省、スコットランド政府による研究助成を受けている。マンチェスター大学を連携本部としている。UK Biobank は数十年コホートを追跡する計画を見込んでいる。

Generation Scotland<sup>22</sup>. 選ばれたいくつかの対象となる慢性疾患で苦しむ家族を含む多数の家族構成員(第1次では 15000 名、第2次では 35000 名)を対象にした、主要な疾患の危険因子や治療に関する研究事業。Scottish Higher Education Funding Council(スコットランド高等教育助成機関)、スコットランド政府、その他助成機関による支援。スコットランドの四つの大学医学部(アバディーン、ダンディー、エジンバラ、グラスゴー)、その他多数の機関が参加している。薬理遺伝学も研究対象としている。遺伝子抽出用試料の管理はエジンバラにあるウエルカム財団 Clinical Research Facility(臨床研究施設)、Genetics Core Laboratory(遺伝学中核研究所)で行われている。多様な共同研究を予定。

その他のコレクション. 英国だけでも他に何千という研究コレクションがある。特定の領域、年齢幅、性別、暴露、疾患、社会的地位または経験、医療処置、臨床試験、その他医療介入に従って組織化されてい

る<sup>23,24</sup>。研究用の仮想コレクションとして利用できる記録を関連付けた研究資源もある。また研究用と考えられる無数の病理組織及びその他の組織バンクもある。MRC 及びウエルカム財団はこれらの研究の多くに、過去・現在にわたり何らかの形で関与している。

## 2.2 本報告書における用語

「研究利用」とは、データまたはヒト由来試料の使用が許可されることを意味する。状況により、その使用には閲覧、複製、操作、解析、翻訳、符号化、関連付け、転送・移送、保存、保存管理、破棄が含まれることもある。研究利用はデータセットまたはその一部、ヒト由来試料、あるいは研究参加者か被験者について考えられる。

「二次研究利用」とは、最初の(一次)保存管理者でない研究者による研究利用を意味することを明確にしたいときに何カ所かで用いられている。とはいえ、本報告書中の「研究利用」は常に二次研究利用のことを意味している。

「データ対象者」とは、個人情報に属する人のことであり、UK 1988 Data Protection Act(データ保護法)で定義されている。

「匿名化」とは、個人データをその個人を容易に特定できない形に変換すること。識別可能性は多種多様であるため、匿名化とは相対的なものである。

「ヒト由来試料」とは、主にヒト組織またはヒト組織由来の DNA またはその他分子のことである。また、尿またはその他の収集物質を意味することもある。

「研究資源」とは、単にコレクションを意味することも時にはあるが、大抵は広く研究利用可能なコレクションを意味する。すべてのコレクションは研究資源なので、当然その区別は厳密なものではない。

「文書化」とは、選択、測定、確認、符号化、プログラム、サンプル処理など、経時的変化も含め、そのデータセットを初めて研究利用する人でも十分に独立した分析ができるよう、それらのことについて詳細に記述したすべてのものをこの分野では意味する。

「研究責任者」とは、研究チームの活動に責任を持つ上級科学者である。

「保存管理者」とは、通常1人以上の研究責任者のことを意味し、コレクションの保護及び使用に公的な責任を持つ人のことである。

「スチュワード」は、「保存管理者」の代わりに使われることが時にある。著者は保管物の保護・育成・増強、またその研究利用の賢明な提供を意味する上でそのより積極的な響きが気に入っている。(類例として、森林スチュワード、種子バンクスチュワード、教会や礼拝堂のスチュワード、貴重な原稿またはその他文化遺産のスチュワードなどがある)

7 [www.nshd.mrc.ac.uk](http://www.nshd.mrc.ac.uk).

8 参照 Batty GDら、1950年代のコホート研究のアバディーン小児：健康に及ぼす一生及び世代内影響の研究のための背景、方法、新規研究資源に関する追跡情報 *Paediatr Perinat Epidemiol* 2004;18:221-39.  
[www.epi.bris.ac.uk/staff/gdaveysmith/pdf/P427%20The%20Aberdeen%20Children.pdf](http://www.epi.bris.ac.uk/staff/gdaveysmith/pdf/P427%20The%20Aberdeen%20Children.pdf) 掲載論文 [accessed 21 December 2005].

9 参照 [www.clseros.ioe.ac.uk](http://www.clseros.ioe.ac.uk).

10 参照 [www.cls.ioe.ac.uk](http://www.cls.ioe.ac.uk).

11 [www.alspac.bris.ac.uk](http://www.alspac.bris.ac.uk).

12 参照 [www.cls.ioe.ac.uk](http://www.cls.ioe.ac.uk).

13 [www.swsurvey.soton.ac.uk](http://www.swsurvey.soton.ac.uk).

14 参照 [www.mrc.soton.ac.uk](http://www.mrc.soton.ac.uk).

15 参照 [www.medicine.manchester.ac.uk/arc/](http://www.medicine.manchester.ac.uk/arc/).

16 [www.ucl.ac.uk/whitehallIII](http://www.ucl.ac.uk/whitehallIII).

17 [www.epi.bris.ac.uk/caerphilly/caerphilly.htm](http://www.epi.bris.ac.uk/caerphilly/caerphilly.htm).

18 [www.srl.cam.ac.uk/epic](http://www.srl.cam.ac.uk/epic).

19 [www.postgenomeconsortium.com/cigmr](http://www.postgenomeconsortium.com/cigmr).

20 参照 [www.medicine.manchester.ac.uk/cigmr/](http://www.medicine.manchester.ac.uk/cigmr/) and [www.ecacc.org.uk](http://www.ecacc.org.uk).

21 [www.ukbiobank.ac.uk](http://www.ukbiobank.ac.uk).

22 [www.generationscotland.org](http://www.generationscotland.org).

23 Directory of Clinical Databases, [www.lshtm.ac.uk/docdat](http://www.lshtm.ac.uk/docdat).

24 UK Data Archive, [www.data-archive.ac.uk/findingdata/majorstudies.asp](http://www.data-archive.ac.uk/findingdata/majorstudies.asp).

### 3. データ及び研究利用の基本

#### 3.1 研究利用

##### 3.1.1

前述のように、「研究利用」とはデータまたはヒト由来試料の使用が許可されることを意味する。状況により、その使用には閲覧、複製、操作、解析、翻訳、符号化、関連付け、送信、保存、保存・管理、破棄が含まれることもある。

##### 3.1.2

一般的な研究利用は以下の方法で保存管理される。

- ・ データのコピー、またはヒト由来試料を送ってもらう。
- ・ データをオンラインでダウンロードする。
- ・ 直接センターに出向き、そこでデータを利用する。

##### 3.1.3

「研究利用」という言葉は多分事実上拡大解釈されている。そのような研究利用においては、データの保存管理者がデータまたはヒト由来試料の研究利用を直接許可（通常倫理的または法的限定のため）する以外の方法で提供される場合もある。研究利用申請時に保存管理者が特定の分析を行い、その分析結果を提供する場合もある。

##### 3.1.4

ある研究利用において研究参加者、試料ドナー、またはデータ対象者と接触や連絡の必要が生じることがある。コレクションのステュワードは倫理的な理由から外部研究者にそのような接触を許可すべきではない。しかし、例えば、追加・データ収集について、あるいは別の研究への参加可能性について保存管理者あるいは外部研究者が連絡することに、研究参加者が同意するかどうかを尋ねるなど、時には仲介者として連絡を行うこともある。（ちなみにあるコホート研究の報告によると、データまたはヒト由来試料についての研究利用よりも研究参加者にアクセスしたいとの要望が強かったとある）

##### 3.1.5

研究利用は以下のために求められている。

- ・ コレクションそのものを研究する。
- ・ 他のデータコレクションや分析の基礎となるプラットフォームとして一定のコレクションを使う。
- ・ あるデータを他のデータと関連付けてより大きな研究資源を作る。
- ・ 他の研究の参加候補者を特定するためコレクションに目を通す。
- ・ 他の研究の症例としてコレクションからデータを選択する。
- ・ メタ解析用に他のデータと一緒に保管する。

### 3.1.6

用語「研究利用」と「転送」に関する説明。UK 1988 Data Protection Act では、「データの研究利用」は「データ転送」の意味となっている。そして「ヒト由来試料移転」は単に「ヒト由来試料の研究利用」の意味で、それには例えば同じビルなどサンプルが近くに保存されていて実際には移転の必要のない場合も含む。これで混乱が生じることはほとんどない。問題は、データまたはヒト由来試料が提供されて研究活動ができるかどうかである。

## 3.2 データのライフステージ

### 3.2.1

データセットにはいくつかの段階がある。このことは研究利用に関しても意味を持つ。データのライフステージは、初めて測定され記録されるデータ、生データから始まる。生データは、品質管理、冗長部分の削除、転写・符号化の明らかな誤りの訂正、ファイリングや指標付けの改良によって、クリーンデータに変わる。さらに、多数の初期データ(臨床的測定や問診の検討から生成された疾病発症日など)から引き出された推論または「構築された」データ、あるいは、その研究資源に基づいた研究から得られた新しいデータ、あるいはそれらの研究に付加されたヒト由来試料を分析することから得られたデータ、例えばそのようなものを組み入れることでこのクリーンデータが検討されて高品位 (augmented) データになる。異なるデータにとっては異なる事柄を意味する、いわゆる成熟した (mature) データとして、データセットはデータベースに収載され、保存、記録管理される。匿名化されたデータが作成されることもある。

### 3.2.2

コレクションは別の展開の道をたどる。紙のファイルは電子化され、初期の電子データシステムはより近代的なものに更新される。より新しい分類コードに修正されることもある。健康の表現型あるいは社会的データコレクションは、それらに遺伝子型データを接続または関連付けることもある。データにヒト由来試料を関連付けることもある。過去に活動的だったコレクションは、寿命が来る、あるいは支援がなくなるなどにより、取り残された古いコレクションになっていることもある。その場合、手入れをするか破棄するか の決定が下される。保存管理者または研究助成機関とコレクションとの関係は常に変化することが考えられる。

### 3.2.3

二次研究利用が適正であるかどうかは多くの事柄の中で特に、かかるデータセットがいかに成熟しているか、いかに徹底的に文書化されているか、どれほど検索可能であるかということによる。生データはそれを集めているユニット以外の人々が使用することはめったにない。対象者をどのように選んだか、どのような理由でどのような問診を行ったか、どのような測定技術を用いたか、どのようにサンプルを処理したか、どのような分類コードを使ったかなどについて、最初のデータ収集者はよく分かっている、有能だが「データ収集に関係していなかった」研究利用者でもそのデータを適正に分析できるほど、(かなり

の)労を惜しまずデータを詳述しているとは限らない。さらに現在の保存管理者でさえ、早期の段階から関与しているとは限らないため、すべてのあやふやな領域 (soft area) や落とし穴 (pitfall) までは熟知していない場合がある。コレクションやその変化が十分に文書化されていなければ、そのデータセットは二次利用する研究者にとって効率的に利用可能なものでなく、その意味では一次利用する研究者にとっても利用可能なものでなくなる。研究利用の決定にはこれらすべてことを考慮に入れなければならない。

#### 3.2.4

研究資源を構築した者が保存管理の直接の権限範囲を超えて第三者保存管理機関にデータを転送したり、保存センターにサンプルを預けたりするときは、様々な安全保護、文書化、そして研究利用の問題に注意しなければならない。

### 3.3 研究利用可能性の程度

#### 3.3.1

コレクションには幅広い研究利用に提供されないものから、最初から多様な利用や多数の利用者を前提に設計されたものまで様々ある。それらの中間には、保存管理者の研究上の利益を重視しながら、限られた条件の許で他者も利用できるコレクションも多数ある。

#### 3.3.2

研究利用申請者は、保存管理者の親しい研究仲間かまたは同じ研究機関の職員かもしれない。あるいは保存管理者が知っている尊敬している外部研究者であるかもしれない。あるいは保存管理者とは全く関係のない遠方の研究者かもしれない。当然ながら当事者の相互作用ではその当事者間の関係は経時的に変化する。

#### 3.3.3

研究利用はまさに共同で行う努力であり、共生的協力といえるだろう。保存管理者チーム以外の研究者が共同研究事業に補完的研究資源やその能力を提供したり、保存管理者がデータやヒト由来試料を提供するだけでなく、解釈やその他の知的な貢献を行ったりする。また別の極端な例として、データ保存管理者と全く関係のない研究に利用されることもある。これらの極端な例の間に多種多様な協力関係が存在するのである。状況によって協力関係は研究利用提供者や研究利用申請者にとって有利なものあるいは不利なものいずれかと考えられる。

#### 3.3.4

従って、二次研究利用が適切かどうかは、データのライフステージについての考慮すべき事項以外にも、研究利用提供に対する保存管理者の心構えや、その研究利用がどの程度保存管理者の協力を必要とするかにもよる。



### 3.3.5

企業やその他商業用の研究利用の申請については知的財産権や社会的イメージを考慮して、特別に扱われる傾向がある。

### 3.4 「パブリック(公共)データ」と「自由な研究利用」

これらの用語については注意が必要である。政府機関が収集したデータか、あるいは政府、慈善団体が研究助成した研究事業による収集データであることを意味するときは時折「パブリック(公共)データ」と呼ばれる。厄介なのは、そのデータがパブリックドメインに置かれていない(この意味では British Library のカタログはパブリックドメインにある)ことがあり、この場合条件付でデータの一部が切り離されて研究利用される。また、ワールド・ワイド・ウェブに置かれたゲノム配列のように「自由な研究利用」が本来の意味を発揮していることもあるが、大抵の場合そのコレクションの研究利用を申請できることを示している(申請できても拒否されることもある)。この二つの表現には注意しなければならない。あまりに無頓着にこれらの言葉を使用すると国民に誤解されることになる。

## 4. 研究利用の条件

この章では研究利用協定及びその要件について簡潔に述べ、多くの課題について触れてみたい。その後、より重要な課題について述べることにする。

### 4.1 協定

#### 4.1.1

パブリックドメインにないデータの研究利用は、協定、要件を特定した契約の締結、方針と法律の照会により取り決められる。

#### 4.1.2

研究利用のために提供する者の側での協定は、研究助成機関、研究責任者、その他の監視団体、大学、その他コレクションを管理する研究機関、またはこれら団体の連合によって締結される。

#### 4.1.3

研究利用をする者側の協定は、職員や学生を含むすべての者が必ず条件を満たし遵守することについて責任を負う研究責任者かその他指導者、大学またはその研究機関によって締結される。

#### 4.1.4

法律、規制、管理法に適合することは協議されるものではない。しかし、協定には NHS 研究倫理審査に関する詳細や時期、NHS Research Governance、MRC ガイダンス、人組織法 (Human Tissue Act)、個人情報保護法 (Data Protection Act) やその他の法律、the Council for Organization of Medical Sciences (CIOMS) や Human Genome Organization (HUGO) などの国際機関の倫理ガイダンスの遵守について書き込むことができる。

#### 4.1.5

協定は監視機関やその他管理機構の意思決定権限について言及することができる。

### 4.2 協定の要件

#### 4.2.1

研究利用協定はその置かれている状況によって異なり、驚くほど標準化されていない。あるものは他に比べ形式が整っている。とはいえ、協定はすべて契約から成り立っている。おそらく以下の点をすべて網羅している協定はなく、ここに挙げた一覧表はデータの研究利用やヒト由来試料移転協定に最も多く見られる基本項目で、方針や協定の作成やその改訂の際に参考にすることができる。

#### 4.2.2

##### 専門職としての適格性の確認

申請者はデータベースまたはゲノム研究の経験または関心のある健康問題に関する論文発表の経験についての証拠提出を求められる。かかる証拠提出の依頼は許容できない不適切な利用から研究資源の評判を保護し、研究資源スチュワードの努力を無駄にせず、以下に挙げる事柄のいくつかを行うためである。

ケンブリッジにあるthe Juvenile Diabetes Research Foundation/Wellcome Trust Diabetes and Inflammation Laboratory(JDRF/WT DIL)が採用している利用資格規準の一例を挙げる。

「誠実な研究者」は次のいずれかとして定義される。JDRF/WT DILのみによって管理されているデータを研究利用する場合は、PubMedに載っている適切な査読論文を書いた者で、現在も同分野で研究して者のことである。JDRF/WT DILのデータ対象者に関するデータが幅広い共同研究連携の一部となっている場合は、誠実な研究者であるかどうかの判断は、その共同研究連携のデータ研究利用委員会の責任で行われる<sup>25</sup>。

データや試料利用申請者の適格性についての評価と承認に関する保存管理者の責務については検討すべきである。

(これらの科学的課題については第6章を参照)

#### 4.2.3

##### 科学的利点及び適格性についての審査

研究利用協定において、あるいは従来からの慣習で、研究提案の科学的妥当性(独創性、分析力、適格性など)を審査する権利を保有しているコレクションもあるし、また発表物を草稿の段階で綿密に審査する権利を主張するコレクションもある。またその他に、かかる審査は保存管理者の権限ではないと考えているコレクションもある。

#### 4.2.4

##### 提供されるものの明細

データ研究利用協定には通常、そのデータ(そのデータセットの関連データコード、コンピュータープログラム、全般にわたる記録を伴うこともある)をどのようにいつ提供するかが明記されている。専門的なあるいはITについての仕様書について補足事項が付いている場合もある。

ヒト由来試料移転協定には通常、どのようにして、いつ提供するか、輸送のときどのように保護するかについても記載されている。取り扱いや保管の方法についての補足事項が付いている場合もある。

#### 4.2.5

##### 同意

これは個人識別可能なデータまたはヒト由来試料が関与するときはいかなるときも考慮しなければならない基本的事柄であり、常に協定で注意を払わなければならないことである。この言葉が最初の同意を意味したり、その後の同意または特別な同意、あるいは同意要件の放棄を意味することもある。

(第5章の同意についての議論を参照)

#### 4.2.6

##### 目的の制限

これは同意に関係することで、同意が関与するときはほとんどの場合、ある健康状態か疾患のみ研究できることと明記される。目的の制限は従来からの制約事項で個人情報保護法やその他の法律、指針に盛り込まれており、データ対象者やヒト由来試料提供者の関心事項であるだろう。

データまたはヒト由来試料を、他のデータと比較する症例対照研究に使うかどうかという目的に関する論点が適宜議論されなければならない。他には、一つのコレクションを使ってそれとは別の臨床試験または他の研究事業のためのデータ対象候補者を特定したり連絡を取ったりするかどうかという問題もある。

多くの協定において、商業目的を追求する、または利潤につながる可能性があるとの認識の下で行う科学研究は、そのことについての明確な承諾を得て、しかも知的財産契約に署名した場合を除いて許されない。

研究利用とは別の項目として、地域の専門教育用にデータまたはヒト由来試料を使用することについて触れている協定もある。

(多様な要因の中で研究をするこの時代に狭い目的制限を奨励できるかどうかの疑問はある。また一旦データ・ヒト由来試料が移転させてしまったら、目的制限の調査やその遵守の強制は難しいことを注意すべきである。)

#### 4.2.7

##### 守秘性

Data Protection Act(個人情報保護法)、the Human Tissue Act(人組織法)、MRC指針、その他法律または指針の要件、また医学上の守秘についてのコモンロー上の義務について厳重な注意喚起をしておく。

匿名化や開示に際しての個人情報保護の諸側面については論議されている。これらの論議は時折技術的に複雑になる。匿名化が可逆性の場合(MRC用語で「linked anonymised(関連付けた匿名化)」)、どのように識別子を保管するか、誰が識別子を使用して再識別する権限を持つか、再識別を管理する規