

電子ラボノートを実装する前に何を検討すべきか

飯室 聡¹⁾

1) 国際医療福祉大学 未来研究支援センター

iimuro@iuhw.ac.jp

What Should We Consider Before Implementing Electronic Lab Notebooks

Satoshi Iimuro¹

1) Innovation and Research Support Center, International University of Health and Welfare

概要

実験ノートは、日々の研究、実験の記録であるとともに、研究の誠実さの根拠および不正疑惑に対する反論のエビデンスという側面を持つ。従来の紙の実験ノートから電子ラボノートへの切り替えが進みつつある現在、導入にあたって研究者とともに考えるべき課題が大きく2つある。一つはELNを「研究の誠実さの根拠および不正疑惑に対する反論のエビデンス」と捉えるために必要な研究データ管理の考え方である。そこで大切なのは、研究公正という視点である。ただし、本邦においては、研究倫理 Research ethics と研究公正 Research integrity という2つのコトバが明確な定義なく使われている印象がある。両者は似て非なるものである。もう一点は、ELN導入に際しては前段階での準備がかなり必要であるということである。多くの機関や研究者は「ELNを導入したら研究データ管理が楽になる」と考えがちであるが、実際には事前に多くの検討事項があり、その負荷は決して小さくない。事前にプロジェクト全体を把握し、それを管理できる体制を構築することは、プロセス管理の整備に他ならない。そしてプロセス管理された環境であるからこそ、その蓄積されたデータ/メタデータは研究の誠実さの根拠となるのである。

1 はじめに

近年の研究倫理教育の充実にも関わらず、いくつもの研究不正疑惑が耳目を集めている。これらを見ていて思うのは、それらの事象の検証になぜ時間がかかるのか、なぜすぐに報告書および有効な改善策の提示がないのか、ということである。たとえその事象が特定不正行為に該当していなかったとしても、少なからず疑惑を持たれるような行為を行ったことには間違いなく、そもそもなぜそのような事象を事前に防ぐことができなかつたのかという問題も存在する。

研究倫理教育教材やガイドラインでは「研究は誠実に行われるべきである」と強調される[1][2]。他人から言われずとも研究者であれば誰もが「誠実に研究を行う」ことは当たり前だと考えるはずである。ではどうすれば自分の研究について「誠実である」とエビデンスとともに示すことはできるだろうか。「誠実です」とコトバで言うだけではどうにもならない。上記の参考文献においては「誠実とは何か」の概念についての記載はあるが、では実際の研究においてどうすれば「誠実な研究」となるのかについての具体的な記載はない。

実験ノートは、日々の研究、実験の記録であるとともに、研究の誠実さの根拠および不正疑惑に対する反論のエビデンスという側面を持つ。この数年で従来の紙の実験ノート（以下、実験ノート）から電子ラボノート Electric Lab Notebook（以下、ELN）への切り替えが進みつつある。実験ノートとELNとは基本的に同じ目的を持つものではあるが、ELN導入に際してそれに過度の期待を寄せてしまい、「想像していたよりも使えない」とばかりにELNの利用をやめてしまう事例も散見される。実はこの状況は、20数年前の臨床研究における電子的データ収集システム Electric Data Capture System (EDC) 導入初期の状況とそっくりである。「導入してみたもののEDCは結局使いにくい、紙の調査票 Paper Case Report Form (pCRF) で十分なのではないか」という議論がつい数年前まで行われていた。もちろん、現在では特に規模の大きな臨床研究であればEDCを用いて研究を実施することが当たり前となっている。なぜ「当たり前」と言うか、そこには明確なメリットがあるからである。

本稿では、研究の公正性担保という観点からELN導入に際して研究者が事前に抑えておくべき

ポイントを2点にしぼって検討する。

一つはELNを「研究の誠実さの根拠および不正疑惑に対する反論のエビデンス」と捉えるために必要な研究データ管理の考え方である。端的には、ELN導入の目的は「研究データに誤りがないことを目指すため」ではなく、「研究データのある一定の水準で管理できる状態に置くこと」であると認識することである。そのためにはまず研究倫理と研究公正の区別と連関ということを理解しなくてはならない。

もう一点は、ELN導入に際しては前段階での準備がかなり必要であるということである。その準備は主としてラボにおける研究活動の見直しと標準化ということになる。ELN導入を検討している多くの機関や研究者は「ELNを導入したら研究データ管理が楽になる」と考えがちである。もちろんその通りなのだが、そのためには導入前の準備が欠かせないのである。

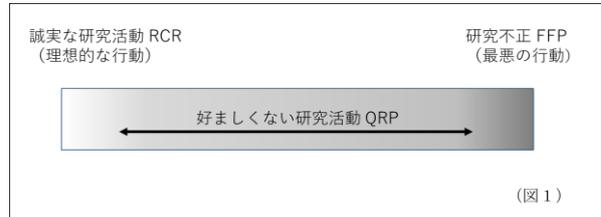
2 責任ある研究活動、研究の誠実さとは何か

2.1 研究倫理と研究公正の区別と連関

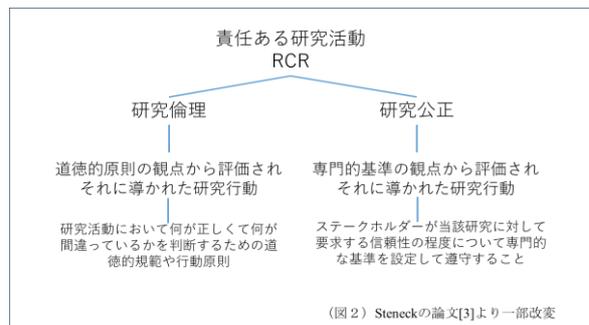
本邦においては、研究倫理 Research ethics と研究公正 Research integrity という2つのコトバが明確な定義なく使われている印象がある。自らの研究が誠実に実施されていること、責任ある研究活動を行っているということを示すためには、研究公正というコトバの意味を少し厳密に考える必要があると。

研究不正を広くとらえれば様々な行為が含まれるが、一般的には捏造 Fabrication、改ざん Falsification および盗用 Plagiarism、(それぞれの頭文字をとって FFP) の3つを特定不正行為と位置付けている[1]。研究倫理教育においては、これらの特定不正行為をしてはならないという規範が示される。一方で、研究活動の社会との関係性という観点から、研究活動は誠実な(責任ある)研究活動 Responsible Conduct of Research (RCR)、好ましくない研究活動 Questionable Research Practice (QRP)、特定不正行為 (FFP) の3つに分類され(図1) [3]、研究者は自らの研究活動がRCRでなければならないと教育される。ここから導かれることは、自らの研究がRCRであるためにはFFPさえ犯していなければよいというわけではなく、FFP/QRP含めて管理しておく必要がある

ということである。なお、QRPの中ではとくに不適切なオーサーシップや二重投稿などについても、それらの行為を所属している科学コミュニティ内でどのように定義すべきかという議論が残っているものの、研究倫理教育において「してはならない行為」という規範が示される。



Steneckは、RCRは2つの要素から構成されるとして「道徳的原則の観点から評価されそれに導かれた研究行動」と、「専門的基準の観点から評価されそれに導かれた研究行動」を挙げている。そして前者を研究倫理、後者を研究公正としている[3]。簡単に言えば、研究倫理とは「研究活動において何が正しくて何が間違っているかを判断するための道徳的なルールや行動原則」である。研究公正についてSteneckはもう少し詳しく説明しており、「専門機関、研究機関、関連する場合は政府および一般市民によって大枠が示されている高い道徳原則と専門的基準を把持し、それを厳密に遵守していること the quality of possessing and steadfastly adhering to high moral principles and professional standards, as outlined by professional organizations, research institutions and, when relevant, the government and public.」としている。この説明を踏まえて、自分は研究公正を「ステークホルダーが当該研究に対して要求する信頼性の程度について、専門的な基準を設定して遵守すること」と再定義した。



つまり、Steneckが定義している研究公正の範疇から、高い倫理的原則 high moral principles を研究倫理の方に移し、研究の公正性を「ステークホルダーによって専門的な基準 professional standards を

設定しそれを遵守すること」とした。

このように再定義することで、RCR 実現のためには、研究倫理の観点から行動についての規範を遵守すること、および研究公正の観点から必要とされるレベルの信頼性を担保すること、その両面からの検討が必要なのだ、と理解できる (図2)。研究倫理についてはこれまでの研究倫理教育にあらためて付け加えることはない。ではもう一つの研究公正についてはどうであろうか。

2.2 研究の公正性とは何か

研究公正でまず参照すべき定義は「研究公正に関するシンガポール宣言」である[4]。この宣言では、研究の公正性は以下の4つの原則で表現される。すなわち 1)誠実性 Honesty、2)説明責任 Accountability、3)専門家としての礼儀および公平性 Professional courtesy and fairness、4)研究の適切な管理 Good stewardship of research、の4つである。この定義と、先ほどの「ステークホルダーが当該研究に対して要求する信頼性の程度について、専門的な基準を設定して遵守すること」はどのようにつながるのだろうか。

同宣言の「説明責任」および「研究の適切な管理」との整合性が最も分かりやすい。すなわち、研究者はステークホルダーに対して自らの研究が科学的、倫理的に適切な水準で管理されていること、そしてそれを説明できること、を担保することが公正性の土台になっている。もちろん「誠実性」および「専門家としての礼儀および公平」の実践についてもステークホルダーに対して自らの研究を説明できるということが土台にある。

2.3 データとメタデータ

では、「ステークホルダーが当該研究に対して要求する信頼性の程度について、専門的な基準を設定して遵守すること」とは具体的にどのような行動なのであろうか。ここではとくに基礎研究におけるデータの信頼性ということを具体例として検討する。

実験をするとデータが発生する。発生するのはデータだけではない。メタデータと言われる「データを説明するためのデータ」も無数に発生する。誰でも考えればすぐにいくつか思い浮かぶであろう。いつ、誰が、どのような環境下で、どの実験系を用いて実験をしたのか、試薬のロットナンバーや保存状態など実験ノートには沢山のメタデー

タが記載されている。そして、データは適切なメタデータとともに記録、保管する必要がある。

少し抽象的なレベルの例を挙げると、動物を用いた実験系のデータだとしても、それが探索段階でのデータなのか検証段階でのデータなのかによってそのデータの持つ意味は大きく変わってくる。探索段階のデータであれば、そのデータを用いてデータクリーニングの基準を設定するかもしれないが、検証段階のデータであればそのデータを見てからデータクリーニングの基準を変更するのは避けるべきである。つまり、探索か検証か、ということもメタデータとして重要な情報となる。もちろん、このデータクリーニングの過程自体も重要なメタデータである。

2.4 研究データ管理の三原則

適切なメタデータを選ぶ一番の目的は、その論文の results、具体的にはデータや図表についての説明責任を果たすことにある。論文に掲載されたデータの背後には採用されなかった沢山のデータが存在する。それらからいかなる過程を経て論文の図表になっているのか、そもそも元のデータはどのような実験経過で収集されたものなのか、どのような加工がされて results として提示されているのか、そのような情報がメタデータである。それは「データの素性を明らかにするためのデータ」と言ってもよい。

もちろん、メタデータはひとつの実験に限ったとしても無限に発生するものであり、必要なものに絞って収集する必要がある。どのメタデータを必要とするのかは、ステークホルダーとの協議で設定された信頼性の程度によって決まる。たとえば、新薬申請のための治験のデータであれば、「医薬品の臨床試験の実施の基準に関する省令」(平成9年3月27日 厚生省令第二十八号)にどのようなデータがどういうメタデータとともに記録され、提出されなければならないか、ということが定められている。文部科学省や経済産業省、国立研究開発法人日本医療研究開発機構などの公的研究費であれば、事前に Data Management Plan (以下、DMP)の提出が求められることが多い。DMPには、どのデータ/メタデータがどういう媒体で、どのように管理されるのか、ということをあらかじめ記載するが、DMPの内容は直接的には研究費配分機関と研究機関/研究者との協議によって決まる。もちろん、その背後には我々の所属する小社会がス

テークホルダーとして存在する。

さて、あなたの論文の図について疑義が寄せられたという状況を考えてみよう。それが単純なミスであって FFP ではないことをステークホルダーに対して説明するためには何が必要か。あなたはステークホルダーに対して、今回問題とされている事象を説明しなくてはならない。つまり、この図は実験で得られたデータから問題のない方法で作成したものでありこの図は正しい、あるいはその「方法」の一部に誤りがあってこの図になっているのでありその誤りを修正して得られる正しい図はこれである、という説明が必要となる。もう少し詳しく説明しよう。

まずその図から元データにまで遡ることが必要となる。これを追跡可能性 *traceability* と言う。次に元データから問題の図まで（もしあればミスも含めて）再現することが必要となる。これを再現可能性 *reproducibility* という。そして、そのミスを修正すると本来の正しい図はこうなる、ということの説明が必要がある。そのためのエビデンスの土台にあるのがメタデータである。どのレベルの詳細さ、どれくらいのメタデータが要求されるのかということは、ステークホルダーの協議によって決まる。なお、*traceability* と *reproducibility* を担保するためにはラボの活動がプロセス管理 *process control* されていること、つまり必要なメタデータが決まった手順で過不足なく収集、管理されていることが必要となる。この管理の手順を決める段階に必要なメタデータが同定されることになる。

前段落では論文の図について論じたが、データそのものについても同じことである。そのデータはどの実験で得られたのか、そこからどのような過程を経て今、目の前にあるのか。そういうことが説明できて初めてそのデータは信頼できる、誠実なデータであると言える。それを担保しているのはデータとともに収集、管理されているメタデータである。

なお、我々の研究グループではここで挙げた追跡可能性 *traceability*、再現可能性 *reproducibility* としてプロセス管理 *process control* を研究データ管理の三原則と位置付けている[5]。この三原則を踏まえて研究を行うことで、実験あるいは実験データの誠実さということがステークホルダーに要求されているレベルで保証されることになる。注意すべきは、この三原則によって FFP や QRP をゼロ

にできるわけではない、ということである。FFP/QRP の発生が、要求されているレベル以下に管理されていることを示すことができる、ということである。その意味において、研究管理あるいは研究データ管理という文脈においては、研究倫理のみならず研究公正という観点も合わせて必要となる、ということが明確になる。逆に言えば、研究倫理と研究公正を明確に区別しないままに論じてしまえば、下手をすると研究公正の概念と必要な行動が抜け落ちてしまうということになりかねないのである。

3 ELN 導入に抱く幻想

さて、本論文のもう一つのテーマについて検討していく。ELN の導入価格がかなりこなれてきたこともあり、その採用については多くの研究者が期待を寄せている。一方で、「ELN は使い勝手がよくない、期待は幻想だった」と ELN の使用を諦めてしまう研究者も散見される。

3.1 EDC 導入初期に何が起こったか

ここ数年の ELN の採用状況は、実は 20 数年前の臨床研究における電子的データ収集システム（EDC : Electronic Data Capture System）の採用状況に酷似している。当時、製薬企業の治験においては、データの品質保証の観点から EDC の採用は着実に進んでいた。一方でその導入や利用にかかる費用の高さ（ひとつの案件で年間数千万円）からアカデミアにおける採用は躊躇する大学が多かった。その後 10 年余の経過で価格破壊が起こり、現在はアカデミアの研究費でも治験対応レベルの EDC が利用できる価格（1 案件で年額数十万円）に落ち着いている。

価格面で解決されても、なお克服すべき壁があった。それは個別の臨床研究への実装である。EDC の選定や構築のスキルは教えてもらうことができるし、いったん EDC にデータを入力してしまえばそこから先は自動的に監査証跡がつく。それでも臨床試験の品質に疑義が突き付けられその疑義に満足に回答できなかった臨床研究がいくつもある。臨床研究法施行のきっかけとなったディオバン事件がその筆頭であろう。EDC を実装すれば臨床試験の品質が保証されるのではなく、EDC を使って臨床試験全体の品質を担保しつつ試験を遂行する方法を構築しなくてはならなかったのである。

実際に EDC を導入する際には、研究管理側、研究者側それぞれで事前に検討すべきことがある。研究管理者側であれば、その大学で主にどのような臨床研究が行われているか、大学内でデータ管理の実務まで実施するのかそれともその業務は外注するのか、モニタリングや監査についてはどうか、などの検討事項がある。その実態に照らして EDC の仕様を決め、最適な EDC を選定することになる。

研究者側では、まず研究仮説の明確化とその仮説を検証するのに必要なデータおよびメタデータの同定という作業が発生する。具体的には、その研究でどのようなデータを収集するのか、そのデータにどれくらいの精度を求めるのか、欠測を許容するのか否か、エンドポイントの定義とそのデータセット化はどうか、等々の検討を事前にしっかりと行うことが必要である。紙の調査票 (pCRF) を用いていた時には、これら検討を最後の症例の調査が終わって (LPO: last patient out) から実施することをおおむね許容していた。およそ研究に必要なデータが収集できるような pCRF さえ作成してしまえば、試験は割とすぐに開始できた。その代わりに、LPO からデータ固定までに多大の時間を要した。LPO の後にデータクリーニングの方針やイベント定義、解析方法を定めるような時代もあったのである。pCRF の余白に種々の「重要な」コメントなどの記載があり、それが別途データとして採用される、というようなこともあった。(このような自由度が良いことなのか否か、許容されるか否かは別問題である。)

EDC の場合は、事前に詳細を検討しているということを前提に LPO からデータ固定までの時間が大幅に短縮することが可能になった。それが、アカデミアにおける EDC 導入のモチベーションの一つだったのである。もし、満足な事前検討をすることなく pCRF と同じ感覚で EDC を使って臨床試験を実施した場合、せっかくセットアップした EDC にデータが収集されてもデータ固定から解析までの短縮にはつながらず、pCRF ほどの自由度もないということになる。だからこそ、EDC がアカデミアに導入されて 20 年近くなろうとしているのに一部で「EDC は不要だ」という議論がなされる。「不要」なのではなく、使いこなせていないだけなのである。

現実的には、実際の臨床試験への実装に当たっては研究者の自助努力のみでは困難であり、研究管理側のバックアップが必須であった。研究者を

支援しながら、研究仮説や収集データ、エンドポイント、解析計画など全体像を俯瞰し、どのステップにどのような作業が必要か、それを EDC でどのように実現するかを事前検討することで、初めて研究の質が上がる。もし研究終了後に疑義が生じたときには、事前に全体像が把握できているからこそ、どの段階に何が起きたのかを追跡することができ、それが不正に当たったのか否か、不正でないとなればしっかりと修正して再度論文として世に問うことが可能となる。これが、臨床研究支援として我々が 15 年に渡って経験してきた歴史であった。

3.2 ELN の置かれた状況

ELN も同じ状況である。数年前までは、見積りを依頼するとサーバの構築等の初期費用と年間のライセンス使用料で最低数百万円から数千万円という金額 (ラボの規模に依存) が普通に提示された。企業にて研究データの品質保証が絶対に必要という条件があれば採用の判断も可能であろうが、アカデミアでの実装はほぼ不可能であった。最近では、この数年のクラウド環境の整備とシステムそのものの価格低下もあり、1 ユーザーあたり年間数万円のライセンス使用料と初期費用としての PC あるいはタブレット端末の購入で導入が可能となってきている。ただし EDC と同様に、ラボへの実装に当たっては研究支援組織のバックアップが必要であろうと推測される。この研究支援組織が存外に大きな問題となることもある。

簡単に言えば、機関 (大学) の中でどの部門が研究支援組織となるのか、ということである。候補としては以下である。

- URA (University Research Administrator)
- 図書館
- 情報システム部
- 臨床研究支援
- 研究倫理支援

URA とは、一般的に研究者および事務職員とともに、研究資源の導入促進、研究活動の企画・マネジメント、研究成果の活用促進を行って、研究者の研究活動の活性化や研究開発マネジメントの強化を支える業務に従事する人材 (あるいは部署) を意味する。ELN 導入を主導する部署としては適切であるが、各機関においてどのようなスキルを持った人材が確保できているかはまちまちであり、URA であればその任を担えるというわけではな

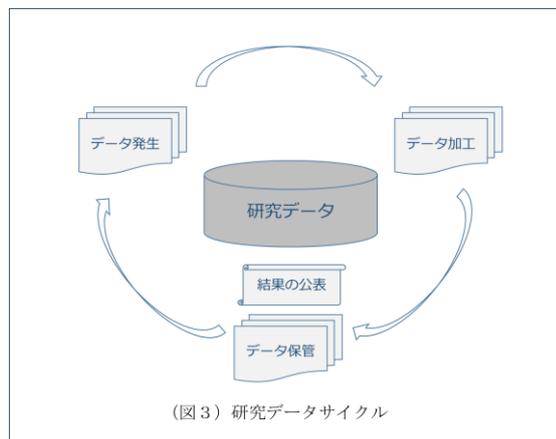
い。図書館は、機関リポジトリの管理をしていることが多く、その延長で研究データ管理もと期待すること大であるが、その可否はやはり図書館に勤務するスタッフの専門性に大きく依存する。残りの3つも、直接に研究データ管理、研究データの信頼性保証を専門としているわけではないという点において同様である。

つまり、ELN 実装に際しては、まずは機関内のステークホルダーを集め、どの部署が主導してそれを遂行するのかという検討から始める必要があるのである。その上で、以下の事項の検討を行うこととなる。

- ・ ELN 各機関における研究データポリシー、研究ガイドラインの確認
- ・ ステークホルダーの確認
- ・ 機関の研究室における研究活動の把握
- ・ ELN の要件、選定、費用
- ・ ELN 利用に関する研究者教育の整備と実施

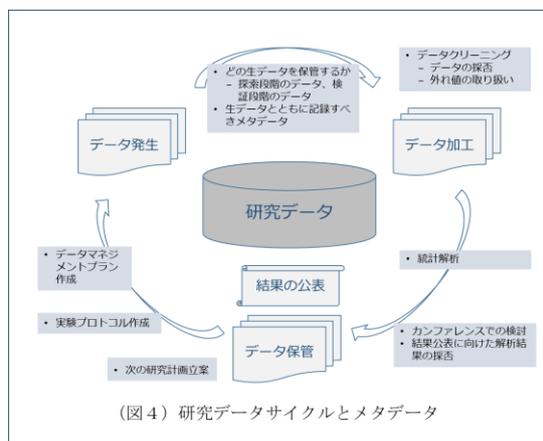
3.3 個々のラボでのELN実装において準備すること

ラボでは年間計画が設定され、大きな年間計画の下、日々の研究が実施される。そこで生成されるデータ/メタデータは次の研究、そして次の年次の研究計画へとつながっていく。これは「研究データサイクル」として表現されることも多い(図3)。



(図3) 研究データサイクル

この研究データサイクルは前章で述べたプロセス管理の土台に他ならない。データサイクルの中で発生するデータの種類とタイミング、発生したデータの管理、生データを加工して2次資料にする過程、次の研究計画の立案などが管理される。その管理に必要な内容がメタデータとなる(図4)。



(図4) 研究データサイクルとメタデータ

箇条書きにするとおおよそ以下のような内容となり、おそらく多くの研究にとってこれらはメタデータの最小限の組み合わせとなるであろう。

- ・ 研究プロトコル
 - ラボの大きな計画
 - 個々の実験の計画
- ・ データ管理計画
 - データ/メタデータの範囲を含む
- ・ 個々の実験におけるデータの採否や外れ値の取り扱い
- ・ データセットの固定の過程と固定されたデータセット
- ・ 統計解析
 - 解析用データセット作成の過程
 - 解析プログラム
- ・ カンファレンスでの検討内容
- ・ 公表に向けた解析結果の採否

注意すべきは、これらのデータ/メタデータのすべてを ELN 上に載せるわけではないということである。何を ELN 上で管理し、何は ELN 外で管理するのか、ということも事前にしっかりと検討しておかなくてはならない。

4 まとめ

第2章では ELN を「研究の誠実さの根拠および不正疑惑に対する反論のエビデンス」としてとらえるための基本的な考え方を紹介した。つまり、研究倫理の範囲でのみ考えるのではなく、研究公正の枠組みをしっかりと踏まえること、それは研究のプロセス管理を土台とするということであっ

た。

第3章ではELN導入の持つ問題点を少し具体的に紹介した。EDCとの類似性を見ながらELN導入に関する「最初の一步」を検討してみたが、両者に共通することは、実装に際しての「プロセス全体の理解の必要性」であった。臨床研究も基礎研究も、個々のプロセスを取り出してみれば、その取り組み自体は決して難しいことでもなく、これまでもやってきたことである。pCRFとEDC、あるいは実験ノートとELNの違いは、プロセスの理解と準備を研究の最後の段階でまとめてやるのか、プロジェクト(研究)が開始されるときに事前に実施するのかの違いである。

事前にプロジェクト全体を把握し、それを管理できる体制を構築することは、プロセス管理の整備に他ならない。そしてプロセス管理された環境であるからこそ、そこの蓄積されたデータ/メタデータは研究の誠実さの根底となるのである。

5 謝辞

本論文は以下の2つの研究成果を踏まえています。このような研究の機会を頂戴したことに深く感謝いたします。

・国立研究開発法人科学技術振興機構社会技術研究開発事業「科学技術イノベーション政策のための科学 研究開発プログラム」

・国立研究開発法人日本医療研究開発機構研究公正研究公正・業務推進部「研究データの質向上の指導者育成プログラム開発」

また、東大名誉教授の故大橋靖雄先生には研究者として駆け出しのころから親身にご指導いただきました。今、自分があるのはひとえに大橋靖雄先生の御指導のおかげであり、本論文の基本的な考え方についても先生から多くの御助言を賜りました。ここに深謝の意を表します。

参考文献

- [1] 日本学術振興会 「科学の健全な発展のために」丸善出版 p.3 - p.8、p.46
- [2] 文部科学省 研究活動における不正行為への対応等に関するガイドライン
chrome-extension://efaidnbmnribpcajpcglclefin
dmkaj/https://www.mext.go.jp/b_menu/houdou/2

6/08/_icsFiles/afieldfile/2014/08/26/1351568_02_1.pdf (最終閲覧 2023/9/13) p.4

- [3] Nicholas H. Steneck, Fostering Integrity in Research: Definitions, Current Knowledge, and Future Directions, Science and Engineering Ethics (2006) 12, 53-74
- [4] Marusic A. The Singapore statement on research integrity. Croat Med J. 2010;51(5):381-2.
- [5] 2018年度、2019年度 AMED 研究公正研究公正・業務推進部『研究データの質向上の指導者育成プログラム開発』
- [6] 大学ICT推進協議会 大学における研究データポリシー策定のためのガイドライン 2021年7月1日
- [7] 文部科学省 研究活動における不正行為への対応等に関するガイドライン 2014年8月26日
- [8] 文部科学省 研究機関における公的研究費の管理・監査のガイドライン 2007年2月15日 (2014年2月18日改正)