

第5回A Iを活用したプログラム医療機器に関する専門 部会

日時 令和5年4月4日(火)
14:00～
場所 ウェブ会議

<開会>

○事務局（澁岡先端技術評価業務調整役） 第5回AIを活用したプログラム医療機器に関する専門部会を開催させていただきます。本日は、お忙しい中、お集まりいただきましてありがとうございます。

<出席状況及び配布資料確認等>

○事務局（澁岡先端技術評価業務調整役） 最初に、委員の出席状況について申し上げます。専門部会委員17名のうち、現在14名に御出席いただいております。全委員の過半数に達しておりますので、専門部会規程第7条の規定に基づき、本専門部会の成立を御報告いたします。欠席の委員は田中委員、今、遅れているのが武田委員と中岡委員です。また、森委員に関しては途中で退席されるとの連絡を頂いております。

次に、配布資料の確認をさせていただきます。議事次第・資料目録、資料取扱区分表、資料1から資料5、参考資料1から参考資料3です。事前にメールで送付させていただいております。資料5については昨日追加で送付させていただきました。もし、不足などがありましたら事務局までお申し付けください。

資料の取扱いについてです。資料取扱区分表を御覧ください。資料の内容に応じて取扱いを分けております。本日の配布資料のうち、資料2から資料5及び参考資料3については取扱注意とさせていただきます。厳重に保管していただき、コピー等の複製、第三者への開示は御遠慮ください。資料1及び参考資料1についてはその他ということになっておりますので、委員各自で適切に保管・管理・廃棄をお願いいたします。

今回もWeb会議ですので、発言に際しましては最初に名前を言っていただき、ミュートを解除して御発言ください。発言が終わりましたら、またミュートに戻すようお願いいたします。今回もWebの録音から文字おこしをして議事録を作成いたします。速記業者の録音ではありませんので、議事録確認の際に先生方に御協力いただく部分があるかと思っております。この点についてはあらかじめお詫びいたします。どうぞよろしくをお願いいたします。

以降の議事進行については佐久間部会長をお願いいたします。

<はじめに>

○佐久間部会長 本日はありがとうございます。
議題に移ります。令和5年3月2日に開催された第4回専門部

会WGの参加者と概要を資料1にまとめております。第4回専門部会WGでの議論を踏まえ、報告書案全体のまとめ方について、及び臨床データのデータベースに共通的な案件として何を出すかについて、副部会長の伊藤先生、中岡先生と確認させていただきましたが、これが資料2です。令和5年2月20日の第4回専門部会で各委員の先生方との議論を踏まえ、既に各担当項目について執筆していただいております報告書の更なる加筆とか改訂を依頼させていただいたところです。これについて、副部会長の伊藤先生、中岡先生と確認させていただいた内容については資料2です。それから、委員の先生方に加筆等の対応をいただきました各項目についての報告書案を取りまとめたものについては資料3をそれぞれ御参照ください。

それから、専門部会WGでは、最近先生方も御承知の対話型のAIといったものも今回のものに含めるのか。結論として今回は含めないわけですがけれども、これについて鎮西先生より情報提供していただいた資料を参考資料2として掲載しております。第4回専門部会において共有された報告書項目・素案についても改めて資料4に再掲しております。

まず、報告書のまとめ方について、私のほうから改めて御説明したいと思います。資料2を御覧ください。全体のまとめ方については2月20日に議論していただいて、今までの項目立てに少し改訂を加えました。

まずは、本報告の目的とScopeで、前回のAI部会報告との差分をある程度明確にしようということがあります。個人情報保護については、この専門部会とかWGでも結構議論になるのですがけれども、実は法律の案件で科学の案件ではないので、課題であることはそうなのですがけれども、本文に含めることはしないことが科学委員会の立場としては妥当であろうということです。ただし、この問題はかなり重要なので、参考あるいは付録のような形で言及することにしたいと考えております。

国内外の動向については、IMDRF、FDA、国内での活動について、その中で特に専門部会の先生方からGood Machine Learning Practiceと今回の議論がどのような関係になるのかといったようなこと、この辺りについても少し記載していただくといいだろうというようなことがあるということです。2-2のところを考察ということでその辺りを入れてもらうことを考えるということです。

3で機械学習におけるバイアスについて、4は、市販後学習にお

ける評価データの再利用に関する問題と問題解決に向けた研究の現状ということです。ここも、実はデータベースに関連することがあります。後ほど、清水先生からお話があると思いますが、ここで書かれている内容を一部データベースに持ってくるか、あるいはデータベースの議論の共通部分をここに持ってくるということについて、これは後で議論させていただきます。

データ augmentation の新たな手法ということについて、あとはシミュレーションです。それから、データベースについての展望ということです。

臨床データベースの共通的な案件としてどういうことがあるか。一般的なことについては列挙するとして、先日の議論を踏まえて、やはり患者に関する臨床背景情報をできるだけ連結可能な形で記録しておくことが重要だということがあるかと思います。それから、アノテーションの付け方なのですが、その根拠となる考え方です。例えば、論文とかガイドラインなどがあるのだろうということを明記するというのは、報告書に明記するのではなくて、データベースにそれが必要だろうと。これは、後でお話があると思います。実は、病理診断のつけ方そのものも結構頻繁に改訂されて、過去のアノテーションが今のある時点では正しくなくなっている、あるいは更に細分化されているといったようなことがある。その辺りのこともちゃんと書いておくことが必要だということですので。データベースだからといって、これで一回作ったから OK というものではなくて、やはりこれそのものが実は変わり得るものなのだということをおある程度議論しておくことが必要だろうということですので。

今のことで、時代とともに新たな臨床情報、あるいは新たなサイエンティフィックなバインディングが出てくると、それでアノテーションというのは変わるので、その辺りについてもちゃんと入ったことを記載するというので、使う側が、いつの時代の、どういう分類のものを使ったかということをお分かるようにしておくことが必要になる。それから、変更・追加の履歴をしっかりと書くと同時に、例えば、これを変更した結果、新しい標準的な考え方が変わるといったことがあります。それが、医薬品にどう影響するのかということが分かるようにしておくということ。エンジニアは、そういうことを全く理解していないと思いますので、そのところが分かるようにしておくことが重要だろうということですので。

それから、治療に関わるデータを記録している場合には、治療手技の変更とか、標準的な治療の変化というものがあるでしょうから、そういうことなどが書いてあることが必要だろうということです。もう1つ、データベースとしてそれがもし薬事に使われるとすると、信頼性調査に対応できるような形でのものが必要だろう、トレーサビリティがある程度必要だろうということを書きました。

そのようなことを書いたということです。大体こういう形で前回の議論を反映させるということを考えました。これについて、いやいや更にこういう点を入れたほうがいいのかということがあれば、御質問も含めて頂ければと思います。いかがでしょうか。鎮西先生どうぞ。

- 鎮西委員 今の最後のところですか。この手の評価データベースに対して、信頼性調査をやるかもしれないというのは、PMDAのこれまでの通知とか、そういう行政文書のどこかにそういう話が出ているのでしょうか。
- 佐久間部会長 この点についてPMDAのほうはどうですか。
- 鎮西委員 ここで初めて書いてしまうのはちょっとまずいです。
- 佐久間部会長 ここで、これだけ単独で読むと確かに強すぎるのです。実はこれには条件付けがあって、「申請データとして使う場合には」という言葉が出てくるのです。
- 鎮西委員 そうですね。
- 佐久間部会長 そうなのです。そこをどこまで書くかというのは、確かに議論になるところなのです。
- 鎮西委員 そもそも、申請データとしてそういうのはありなのかという話と、仮にあるとしたら、その信頼性調査を何かという話がここで出てしまうというのはどうなのかなという気がするのです。
- 佐久間部会長 PMDAのほうからはどうですか。率直に御意見を頂ければと思います。この辺りは非常に微妙な問題を含んでいると思うのです。一応は求めますよね、やはりそれがちゃんとできるようにしておくことが必要だというのが多分大原則なのですけれども、どこまでいくかという話にならないと。
- 鎮西委員 それに対して信頼性調査という言葉を使ってしまうと、どういうことをやるかということも含めて、一気にものすごく現実味が増してしまってますね。
- 高橋執行役員 PMDA執行役員の高橋です。御指摘ありがとうございます。確かに鎮西先生御指摘のとおり、いきなりここで信頼性調査の話がポ

コッと出てくるというのは、ちょっと違和感があるのかと思います。しかしながら、評価のデータとして、データには信頼性が必要だということは確かに事実だろうと思いますので、もしよろしければ、そのデータへの信頼性が必要なのだというようなトーンでまとめさせていただけると。この点については、我々のほうでもどういう書き方がいいのかということについて検討させていただきたいと思います。確かに、いきなり信頼性調査というと、これまでの流れの中で、ちょっと唐突な感じがしますので、その書き方については、また御相談させていただければと思います。以上です。

○佐久間部会長 ありがとうございます。中田先生どうぞ。

○中田(典)委員 前回の会合で、自然言語処理については含めないと決まりました。しかし、この会議を開いている間にどんどん技術革新が進んで、例えば ChatGPT というのは裏で画像生成もやっています。今年の初めに放射線学会の雑誌に、正式に ChatGPT に対する放射線学会のアメリカの姿勢みたいなことを論文で出しています。なので、ChatGPT は、例えば自然言語処理はそれでやるのだけれども、それに画像の AI を融合させたようなものというのは近々日本でも出るだろうというのが、オープンデータから明らかになっています。なので、その部分については今後審議が必要であるみたいなことは入れておいたほうがいいと思います。要するに分離できないのです。

○鎮西委員 もう1年以内に出ると思います。

○佐久間部会長 そうでしょうね。

○中田(典)委員 要するに、対話型の AI が、あるところの画像 AI の結果ではこのようなものが出ていますみたいなのを、例えば個人情報保護法をスルーしながら、データを上げていってしまう。それから、政府が電子カルテのクラウド化を一部始めていると。そのクラウドを、例えばマイクロソフトが受注してしまうと、マイクロソフトは ChatGPT を使いますみたいなことを言い出すと思うのです。だから、そういうことは十分考えられます。これは、あまり科学的ではないのですけれども、何かちょっと留意が必要のような気がします。そのところの御議論をよろしくお願いします。

○鎮西委員 参考資料は出ますか。

○佐久間部会長 鎮西先生がくださった参考資料2ですよね。

○鎮西委員 参考資料2です。これは、もう1か月ぐらい前の話なので。

○佐久間部会長 これは、WGをやったときに出てきたものですよ。

- 鎮西委員 はい。これは、もうそんなにびっくりしないです。
- 佐久間部会長 これは、アーカイブだから、まだ査読が通っているわけではないと思うのです。
- 鎮西委員 はい。
- 佐久間部会長 そういう状況なので、いきなりこれまで含めると、この報告書はなかなか難しいところがあります。これは、後で議論しましょう。
- 鎮西委員 科学委員会にいずれ取り上げざるを得ないと思います。
- 佐久間部会長 今の2点、私が書きすぎたところがあります。信頼性調査のところは、少しPMDAサイドと適切な表現になるようにします。まだ、本文が出来てきているわけではないので、そこのところが入るような形にしたいと思います。
- 中田(典)委員 このChatGPTは自然言語処理のAIの技術を学んでいくとまだ序の口なのです。これから、もっともっと進むだろうということが非常に予測できるので、そういうところには留意すべきだと科学的には考えます。
- 佐久間部会長 中田先生、今のところを、その場で結構ですから、パサパサッと数行に書いていただいて、言葉にして残しておいていただけるとすごく有り難いです。
- 中田(典)委員 分かりました。
- 佐久間部会長 ありがとうございます。急をお願いしてしまって申し訳ないです。時間のこともありますので、各担当項目について、執筆していただきました内容について順番に委員の先生からお願いします。まず、資料3及び追加の資料5を見てください。最初に「はじめに」のところだけ説明させていただきます。

<機械学習におけるバイアスについて>

- 佐久間部会長 森先生のところから話をさせていただきます。資料を出していただけますか。資料3の私のところを出してください。最初の「はじめに」の所を出していただけますか。そこのところを1、2分、さっと説明してから森先生のところにいきたいと思います。

資料3です。私のところの「はじめに」は、とにかく今までの議論を少しまとめてということで書きました。2017年時点では、どちらかというと技術的側面について、あとは責任ということが書かれていると。基本的には技術的観点の議論から出ているのだけれども、ここではそういうことが現実化してきて、まだ市販後に変わるというものは限定的だし、ほとんどないです。こういう

ものが出てくるだろうということを視野に入れて、議論をまとめましょうということを行ったということが第2パラグラフです。

第3パラグラフでは、その頃に、はっきりしてこなかったいろいろ規制当局の状況が分かってきたということで、そこを言及しました。第4パラグラフでは、先ほどのChatGPTもありましたけれども、その技術についてはまだ特性うんぬんと書きました。この書き方は先ほどの議論を踏まえて直します。

現実になってくる課題についてありましたということで、各先生方のことを書いたということです。それから、次にGood Machine Learning Practiceとか、そういうものが出てきていますということの話をしました。各項目でどういうことがされているかということが整理されたというのを書くことを、私なりの推量でまとめて全体の概観ができる形にしてきたということです。

最終パラグラフのところでは、個人情報保護法については法的課題なので馴染まない課題なのだけれども、付録の形で記載することにしますよということを書いたということで、一応全体をまとめてみました。今日の話をもとに、また改訂しますが、こんな形の内容です。森先生が終わった後で、お話を頂ければと思います。

○森委員

私のファイルをそのまま共有します。私のほうから発表させていただきます。よろしくお願いいたします。

今回幾つか文章を足しました。校閲記録が残してあるといいのですが、ちょっとお待ちください。バイアスについて幾つか書き足しをしました。バイアスが增強されるバイアス増幅なども追加してまとめました。あとは、データ自身が持つバイアスということで、これは、これまでどおりの内容です。それぞれのところでの説明の追加を行っています。各バイアスについては、これまでよく知られているA社のCT、B社のCTでは全然うまくいかないとか、あとは、特定の病態とか年齢・人種・性別などからのデータを用いると、どうしてもバイアスがかかってしまうということをきちんと書きました。

利用手法についても書き足したところです。これについては少しだけ追加しています。それから、人によるバイアスについては、今後の大きな問題となってくる。ここは少し書き足しました。特に、機械学習を用いた画像分類等の教師データ作成に関するバイアスというのが、結構機械学習を使った医療機器において非常に大きな影響を及ぼす可能性がありますので、その部分を書きまし

た。特に、アノテーションバイアスというものが結構あって、そこから分類器が作られてきて、それがかなりのバイアスを持っているということもよくあると思います。それを、人によるバイアスという形でこの部分を追加しました。恐らく、人によるバイアスというものは、先ほどのようなデータのバイアスよりも、もっと大きな問題であるかと思うと、ここももうちょっと書き足しが必要かもしれません。今後、医療機器開発は人が行っていくところですので、この人によるバイアスについてはもう少し追加したいと思いますが、現在はこのような状況です。以上です。

○佐久間部会長 ありがとうございます。人によるバイアスの部分というのは、結局、データサンプリングしてくるときに、多分バイアスが生じてしまうとかいうようなことがあって、これも結局データベースとかそういうところにも絡んでくることなのかと理解しました。

先生方から御質問、追加はありませんか。中田先生が、前回だったか前々回に出されたプロスームの公共のデータベースを使ったらより良い結果が出てきてしまった。データクライムということを書かれていた論文を読みました。これも、結果として画像処理の中で、何というか、あるサンプリングをしてるところが、実はゼロパディングすることによってどこかに偏ってしまう。その結果としてそれが影響するという、なるほどと思って見ていました。結局、無意識のうちに、何というか処理そのものがバイアスを作っている可能性も実はある。

○森委員 はい。

○佐久間部会長 だから、そういうことというのが結構危険なことがあるのだということをおもいました。なるほどと思って、中田先生に出していただいたものを見ていました。他の先生方からバイアスに関して何かありますか。

例えば、何といいますかね、前回のプログラム調査会の中で少し議論になったことが1つありました。いわゆる、ここの中でも、例えば訓練データと他データとそれが同質なのかどうかという議論をするときに、平均値が同じなのか、分散が同じなのか。様々なものがあって、何を機能として実現するかということ意識した上で同じだと言わなければいけないのだけれども、その辺りが読者に意識として、こういうAIが機械学習ベースのものをやるときの基本的な危うさというのでしょうか、こここのところが分かるような形になるといいなと思って聞いておりました。

○森委員 COVIDのことをちょっと書かせてもらったのですが、画像を見た

ときにどの時点をゴールとするか、正解とするかということが、画像だけの判断なのか、それとも最終的な判断が求められるのかというのが、画像診断の結果と、それから最終診断の確定の診断というのは違うところもあります。そのずれというのが問題となってきます。よくあるのは、PCRの検査だけで画像と同じものを画像に求めたり、あるいはその逆をしたり、そういう作り手の考えの誤りというのがよくあったりするので、この点もきちんとした教育というのが大事なのかと思いました。これも書き足せば切りがないと思います。

- 佐久間部会長 そうですね。
- 森委員 十分な教育をするとか、分かりやすく説明している資料を開発会社の人にお渡しするということが一番大事なのではないかと思います。
- 佐久間部会長 その一環がこれだという話もあるのですけれども。
- 森委員 そうですね。
- 佐久間部会長 他にはありませんか。森先生、お忙しいところありがとうございました。村垣先生どうぞ。
- 村垣委員 人によるバイアスというのは非常に広い感じがします。森先生が出しておられた、例えばCTのセグメンテーションの領域だとか、認知バイアスが代表的であるのであれば、よく医学ではインター・オブザーバー・バリエーションといいたいまいしょうか、観察者間で異なるという意味で使いますが、（今回は）観察者バイアス等々がよい表現にも思いました。人によるバイアスというのは非常に広がってしまいますので。
- 森委員 広いですよ。
- 村垣委員 あるいは、どこかに、観察者バイアスという用語をいれていいのかなという気がしました。
- 森委員 ありがとうございます。人によるバイアスはもうちょっと区分分けしたほうがいいのですよね。
- 村垣委員 そうですね、それがよいかと思います。大項目としてはよいのですけれども、それが全てだとすると、いろいろな意味での捉え方が違ってくると思います。
- 森委員 ありがとうございます。
- 佐久間部会長 重要な指摘をありがとうございました。人による場合は非常に広い概念なので、確かに分類をする必要がありますね。

<国内外の関連動向の分析>

○佐久間部会長 それでは順番に戻って、2の「国内外の関連動向の分析」を中岡先生からお願いします。

○中岡副部会長 遅れて申し訳ありません。資料は、事務局で共有していただけますか。前回のものに関して、少しアップデートという形にさせていただきました。大まかなところは皆さんも結構読まれているところなので、ご存知かと思います。

各国の規制ということで、一番進んでいるのがFDAで、FDAの現状を紹介する形にしました。GMLPの導入に関する辺りは皆さんご存知のことなので、そこは飛ばして下のほうへスクロールしてください。ディスカッションペーパーに対する回答です。回答という形でアクションプランが出されています。論点が5つ示されています。今後、Machine Learningを利用したSaMDを実用化する上で、どういうところが重要になっているかというところをもう少し細かく説明させていただきました。GMLPに関しては、FDA単独ではなくて、カナダで英国と共同で発行されるということなので、そういう状況を紹介させていただきました。

下のほうへ行っていて、問題はこのPre-Certのところですね。Pre-Certは2017年にスタートしたのですが、2022年9月ですから、つい半年程前に成果物を出しました。その成果物というのは“The Software Pre-Cert Pilot Program: Tailored Total Product Lifecycle Approaches and Key Findings”という形で書いてあったのですが、実は、Key Findingsが何というのでしょうか、あまり具体的な対策が書いていないのです。留意点は書いてあり、要するに平たく言えば、Pre-Certをやったけれども、結局、Pre-Certという仕組みで製造販売業者が性能を変化させていくということ、勝手にそれを認めるというのは難しいということ、多分彼らは結論として言っていると思います。ですので、彼らにすれば、また更なる検討が必要であると。更なる検討は必要であるけれども、Pre-Certは一旦ここで終了ということで書かれているので、この後、Pre-Certの第2弾的なものを作るのかなという形のExecutive summaryが書かれているので、そういうことをここで紹介させていただきました。

FDAも、PMDAと同じようにSaMDの対応部局を作りました。そこに関しては、現状、私どもの調査では、別件のほうでもいろいろな調査はしているのですが、現状、そこがガイダンスとか課題の取りまとめをしているようなのですが、実際に審査に関与しているかどうかという情報がありません。これも委託調

査の結果ではあるのですけれども、今、600近い製品が、確かホームページでAI/ML利用SaMDのリストに載っているのです。これは結局、Machine Learningをちゃんと使って、しかも主機能で使っているかどうかということに関して判別するコードがありません。製品コード等の区別がないということなのです。実は、それから実際にMLを使っているかどうか、どのタイプのMLを使っているか。それが、ちゃんと主機能に使っているかということが判別できないという現状です。

その下は、FDAのホームページでAI/MLを利用したSaMDの承認リストを実際に今公開しているのですが、この公開に関しては、1年更改されないまま、いきなり2022年10月にボンと更新されました。そこで、1年ちょっとの間に155製品が新たに承認されています。ただ、それも先ほど言ったようにMLを使っているのがどれぐらいあるかというのが分からないということと、よく読むと、日本でいうところの一部変更承認申請に相当するものが混じり込んでいたり、単なるイメージデータの改善や他のシステムにデータを使用するためのフォーマット変更をするだけ。医療目的と言えなくもないのですけれども、そういうものも入っているので分からない。一部は私どもも調べに行きました。510(K)の中においても、いわゆるsummaryで文書が発行されていますけれども、そこにも使用アルゴリズムを明記していない製品が多々あります。それなので、やはりこれは単純にアメリカが、本当に承認品が進んでいるのかということ、この情報だけで判別するのはちょっと難しい状況であるというのが現状です。

その下の、FDAとHC、MHRAの話に関しては、先ほど言ったGMLPの共同文書の内容を少しここで記載しています。ここは、そのままちょっと。前回のものは飛ばさせていただきます。

それ以外に欧州の話だとか、韓国、中国の話を追記させていただきました。これも、弊部のほうで別途行っている調査・研究報告があります。一応今回の情報に関しては、令和3年度の調査報告書をベースにしています。近々令和4年度の調査報告書を公開しますので、そちらのほうの情報も可能なものは入れていきたいと思えます。この辺りは読んでいただければいいかと思えます。EUの状況、その次が韓国の話です。韓国もこのようにAIには力を入れていて、いろいろな取組をやっています。規制当局も含め、国策としてやっています。中国に関しても同じです。ガイダンスも9つぐらい出ています。ただ、言葉の問題でなかなか内容が把

握できないというところです。

下は日本です。これは、佐久間部会長から御要望があったもので、私が把握している限りでの取組です。次世代の評価指標作成をやったということ。私が代表者を務めさせていただいているAMEDのRS事業、RS研究事業でどういうことをやったかというところを書かせていただきました。あとは、令和4年6月7日に閣議決定された規制改革実施計画に基づいて、SaMDの認証制度に移行するための認証基準の策定及び改正が今進んでいるというところも一応紹介させていただいております。

その下のIMDRFに関しては特にアップデートというほどのものはなかったかと記憶しています。その後、国際標準化の動向のほうがどうも動き出しているようです。ここはフォローしきれていない部分があって、なかなか辛い部分があります。この内容は、ある程度御紹介させていただいたとおりです。このSNAIGのほうが、この間またレポートを確か出したはずですが、実際に、AI baseとかAI functionalの規格づくりのほうを始める方向であるという情報が載っていたと記憶しています。それを載せるのは間に合わなかったのですが、それに関しては、この後でまた追記させていただきたいと思っています。以上が、私のほうからの紹介になります。

○佐久間部会長 ありがとうございます。やはり、かなり早く動いている所があり追い付くのが大変です。今日、先生が少しお話になった、実はいろいろ議論しているけれども、まだ明確になることが難しそうだと、これまでやってみた結果として、どういうところに課題が出始めていそうだと。今日、先生がちょうど口頭で話されたと思うのです。その辺りを考察に簡単にまとめておいていただけると有り難いかと思います。

○中岡副部会長 はい。どういうふうに入れるかが、なかなか難しいところです。

○佐久間部会長 まだ少し時間がありますので、少し書いてみていただいて、あと、今日、全体で大分出てきましたから、ここでの議論が、例えば、この後のデータベースの話だとか、それから、データの再利用の話だとか、出てきたところの議論が正にこういう所でも議論されているということが関連性として読者に見えると、規制や標準化のところを見るときに、パッと見た人には、この議論は何なのだというところがなかなか分からないと思うのです。そこが少し道標になると、この文章はいいのかと思っているものですから、その辺りを少し御検討いただいて。私も少しやってみますけれど

も、ここはまた議論させてください。

○中岡副部会長 そうですね。先生が言われたとおり、実はこの部分は結構難しいところがあります。1つ面白かったのは、ずっと私どもの、この科学委員会専門部会でバイアスの話をしているかと思いますが、ガイダンスを見ていると、実はアメリカのほうは、Pre-cert を始める段階からバイアスの話をかなり強く懸念点として出しているのです。それが明確に記述してあり、そこはかなり意識して Pre-cert をやっていたと思うのです。

それにも関わらず、やはりリスクが高くなってしまふ。高リスクの医療機器というか、SaMD というふうな形になってしまふと、やはり、Pre-cert の仕組みではカバーできない。結局、彼らにしても GMP なのか QMS 的なものになるのか分かりませんが、ちゃんと査察しないと、変更が妥当、テストが妥当なのかということを保証できないということではないかと思っています。

○佐久間部会長 多分、この段階は「ないかな」ということでいいのではないかと思います。

○中岡副部会長 非常に個人的な考えですが、そういう形の考察であればかなりのことは書けるかと思います。

○佐久間部会長 私は、それでもいいのではないかと思います。やはり、規格をやっていてどうしてこの規格が出てくるのだというのが、多分、メンバーにとってはなかなか分かりにくいと思うのです。そのときにここでの議論などが入ってくると分かりやすいですし、それから、前回出てきたような Good Machine Learning Practice とどう関係しているのかという話になると思うので、そこはまた相談させてください。

○中岡副部会長 よろしくお願ひいたします。

○佐久間部会長 先生方から、ほかに御意見や御質問はございませんか。よろしいでしょうか。ありがとうございます。

<市販後学習における評価データの再利用に関する問題と問題解決に向けた研究の現状>

○佐久間部会長 それでは、次に行きたいと思います。4の「市販後学習における評価データの再利用に関する問題と問題解決に向けた研究の現状」について、清水先生、よろしくお願ひします。詳細に議論していただいてありがとうございます。

○清水委員 修正後の原稿を使って、どこをどう直したかという話をします。まず、1節の「はじめに」の所です。頂いたコメント、再利用のシ

ーンについては、開発時に評価データを再利用するケース、それから、市販後に新たに別の学習データを使って学習させ、再び同一の評価データを使うケースがあるという、2つの話を書いてくださいということでした。

もちろん、2番については最初から入っているので、1番について、「なお」から始まる文章を付け加えました。「なお、市販前に同一のテストデータを繰り返し用いて開発する場合にも、同様のリスクが存在する。本質的な違いがなく、同じ注意点が存在することに注意されたい」というような書き方になっています。

それから、2つ目です。この辺り、現実の研究開発において医療データの取得が難しいこと、やむを得ない場合もあることなどに言及することの記述の補強ということでした。これは、既にここに書いている内容で該当するのかと思っておりますが、足りないということであれば、後でまたコメントいただければと思います。

それから、具体的なシーンとの関連で上記の2に関して一変の際に過去の申請で使用した評価データを再度使って評価すること自体に問題がある可能性。これも既に書いており、ここが該当するのではないかと思います。御覧いただいて、更に補強が必要であれば御指摘ください。

3つ目の同一データを評価に使うことが問題になるわけではないという説明で、どのようなリスクが顕在化する可能性があるかについても、ここを少し御確認いただいて、この点が分かりづらい等、教えていただければと思います。

それから、Bishopの本といえば、多分、工学系の方はよく御存じかと思いますが、その本から holdout の説明を持ってきました。それから、これもどのような場合に問題が起きるかの説明の追記。それから、過剰適合の検出の説明です。

それから、Dworkの少し分かりづらいアルゴリズムを示させていただきます。少し分かりづらいので説明を補強してほしい、直感的な説明を追記してほしいということなので、「差の絶対値が小さい場合にはユーザーのアクセスが training dataset に制限されて、差の絶対値が大きい場合には holdout dataset」と書いてある、同一テストデータへのアクセスが可能となるというような記述を追加しています。

それから、ずっと下のほうです。Kaggleからの教訓をまとめてくださいというような話など、この辺り、5節全体を少し見直してみました。ディスカッションについて厚くしたつもりですが、こ

ここで全部を説明するのは時間の都合上難しいと思いますので、後で御覧いただき、ここがよく分からない等々あれば御指摘いただければと思います。

今日、最初に佐久間先生から御指摘いただいた、ここには書き切れないので他に回したいというような件について3つほどありますので、最後に説明いたします。ここからが頂いた宿題の部分です。黒い文字の所は、既に説明した点です。

まず、「学習とテストの独立性」の話を書いてくださいということでした。入れようと思いろいろ考えてみたのですが、今、説明した話はテストデータの再利用におけるバイアスの話で、学習とテストの独立性の話を入れるとそこだけ浮かび上がってしまうので、是非、データセットの節で説明できればと考えています。

具体的にどういうことを説明するかというと、これも前回既に説明していることですが、簡単に申し上げると、独立性にはいろいろな独立性があります。時間的独立性、疾患独立性、臓器独立性、部位独立性と仮に呼んでいるのですが、そういう独立性のことを踏まえて、学習とテストデータの独立性について考える必要があるということデータベースのほうで入れたいと思います。それから、AUCについてコメントを頂きました。後で御覧いただき、多分、AUCで類似性の評価は難しいのではないかとことを書いております。

それから、もう1つ別の所に追記したいものは、「学習データへの摂動の追加」の話です。これも「テストデータの再利用におけるバイアス」とは独立した話ですので、データベース、あるいは、鎮西先生が書かれているシミュレーションの所に入れるか、別の所に書いたほうが収まりがいいのではないかと考えております。あとは、既に対応済みの件です。以上です。

○佐久間部会長 ありがとうございます。御質問はございますか。私も十分理解しておらずコメントを書きしてしまったので、対応と同時に、まず、後半の所ですが、確かに再利用の話とデータセットの独立性の話は違うので、そこを分けるのはよろしいのではないのでしょうか。他の先生方はいかがでしょう。

そこは、どちらかということデータベースやデータを作る所と共通の話ですよ。そういう御指摘だったと思うので、そこはしっかりしていったときに、ちょうどデータベースの所で、同じ人でも時間が違うなどということが出てきたので、その辺りを少し、どの段階で書くのが分かりやすいかは、実際のデータベースの例

が出てきて、こういうことがあるということをやった上で、それを総合的に見たほうがいいのか、あるいは、もう少し最初の段階で、こういう独立性があり、そういう目で見ると個別のデータベースはこういう意味合いがあるというようなことにつながるのか。それは書いてみてからだと思うのですが、そういう形になるのかと思いますけれども、この点はいかがでしょうか。

○清水委員 私もそのいずれかの方法がいいと思います。具体例があったほうがいいのかと思いますので、そういう意味では、いろいろな種類のデータベースがあるという話をする所で独立性の話を書きたいと思います。

それから、1点、注意しなければいけないと考えていることは、今回、学習データについてはそれほど強い縛りを与える必要がないというトーンで書き、評価データは、しっかり質を担保する必要があるということを押さえて書くことを大前提として、それでも学習データとテストデータの独立性はどうしても触れないといけない重要な概念だと思います。先ほどのように学習データに余り強い縛りはないのですが、極端な話、テストデータと全く同じ学習データを使うのはもちろん NG です。そういうところを押さえるような書き方にしたいと考えています。

○佐久間部会長 今の点、誤解がないようにしなければいけないのは、評価のときに現実、本質と違ったような応答をしないようにすることが必要で、だから、評価データは学習データと独立しなければいけないのです。

一方、先ほどの信頼性の話と関連するのですが、学習データについて何ら言っているわけではないわけで、評価データのところはちゃんとやるようにしてくださいという、その切り分けが、意外に開発者の特性にも依存していると思うのですが、例えば、あるデータセットがあり、学習と評価に分けて繰り返しシャッフルしながらやっていますから、それを普通の感覚でやってきているエンジニアリングの人にとってみると、非常に奇異に映る可能性があるかと思っています。今、先生がおっしゃったような形でその辺りがちゃんと言えるといいのかと、確かに思いました。

○清水委員 コメントありがとうございます。その辺りにも気を付けて書きたいと思いますので、また御意見いただければと。

○佐久間部会長 先生方から、ほかに何かありますか。前半は、非常に難しい話をしている所があります。私は、AUCのことを完全に誤解して先生

のレポートを見ておりました。過適合していることを検知するのはなかなか難しいのでしょうかけれども、その辺りについて、こういうことに気を付けるということがあると、完璧ではないのですが、今の科学的な議論の中ではこういうことがあるという紹介でいいと思うのです。

○清水委員 一応、書いています。これは十分説明しなかったかもしれませんが。過適合の検出の仕方としては、あるデータセットに過適合しているかどうかは、それによる評価値とそれとは異なるテストデータによる評価値との間の差が、過適合しているかどうかの検出のポイントになります。5節でも、このことに基づいて検出する例を示しています。これを御覧いただいて、またコメントを頂ければ。

○佐久間部会長 分かりました。これを読まれる読者層を見たときに、こういうことをやっている専門の人が見ることもあるでしょうし、それから、まず、ここで指摘している課題はどういうことなのかを大枠で捉えようとする人たちと、多分、幾つか想定しなければいけないと思っています。コンテンツはこれでいいと思うので、例えば、その所に誘導するというか、私も考えてみますけれども、そういうことができるかと思っております。余り変な誘導はできないので、科学的にいうとこれはこういうことだということが、簡単なサマリーみたいなものがあるといいのかもしれませんが。

○清水委員 文献を引用するなどではなくて、サマリー。

○佐久間部会長 そうです。先生が書かれているところの重要な概念というのでしょうか、そこがある程度示されて、それについてここに詳しく書かれているという感じの道標があると、先生の文章は読みやすくなると思います。

○清水委員 分かりました。

。

○佐久間部会長 ほかの先生方はどうでしょうか。大丈夫でしょうか。清水先生、ありがとうございます。

<データ augmentation の新たな手法>

○佐久間部会長 5、「データ augmentation の新たな手法」について、中田先生、お願いできますか。

○中田(典)委員 まだ出来ていないので。

○佐久間部会長 それでは、後にします。

<物理モデル・シミュレーションにより学習データ構築の現状と課題>

○佐久間部会長 鎮西先生、次、行きましょう。

○鎮西委員 資料の30ページです。このファイルについては、前回扱ったものの違いは最後の所です。それだけでは何をやったのか忘れてしまうので、ざっとおさらいいたします。数値計算、コンピュータシミュレーション・モデリングを使って出てくる結果をいかに機械学習と組み合わせるかという話で、2通りの方法があります。数値的に合成したデータを学習に使う、あるいは、data augmentationとして使う方法。2番目は、数値計算の結果をラベルとして学習するという事です。どちらもあり得ると考えています。

30ページの下の方では、メリットとしてデータが増える、それから、数値シミュレーションとしてはどちらかという順問題が多いので、数値計算屋としては楽だという感じがします。次のページです。あと、医療の場合では、患者さんを使っているとかえってパラメーターが偏る、あるいは、代表値しか得られないケースがあるけれども、数値シミュレーションでやや極端に振ったような値も得られるだろうという。

あと、数値計算をやると、どういうパラメーターの辺りが鈍感になるのか。多少のパラメーターの変動があっても結果はほとんど変わらないという解析もできるということで、開発の中で曖昧な部分を重点的に拾っていくことができる。今度は、その結果を機械学習の側でカバーすることも可能ということ。それから、赤字は新しく付け足した部分です。希少疾病や脆弱な被験者群からの同意取得など、いろいろな倫理的な問題とも整合性がいいだろうということでメリットを入れています。

ただし、留意事項として幾つか挙げております。この辺りは書き換えている部分がたくさんあります。これは要望に関して、少し見直したりしたものです。あと、data driven modelというのは、ASME V&V 40では範囲から外してありますと書いてあるので、それも書いておきました。この辺りは特にいじっておりません。

32ページです。ここからが、かなり書き足した部分です。数値シミュレーションと機械学習のすみ分け、それから、生成したデータを評価に使う場合はどうかということで、ここは先生方からいろいろ御意見を頂いた部分です。最終的に変えた部分としては、今見えている部分の少し下のほうへ行って、黄色く色を残した辺りからがポイントです。一番下の行です。「人為的に加工又は生成した評価データの位置付け」ということで、「数値シミュレー

ションの結果は、抽象化に伴う捨象、いろいろ省略したり簡単化した部分ということで曖昧さもあるので、特に医学では観測事実と並ぶものとして受け容れられていない。経験的には医療機器の開発と評価においては、位置付けと解釈が定まらない要素が増えるほど、全体的な評価プロセスが複雑化して難化する」、PMDA に突っ込まれることが増えてくる。機械学習の医療機器への応用も発展途上なので、そこに更に数値計算という曖昧なものを持ってくると、曖昧なもの2乗プロセスになってしまい、いろいろ突っ込まれるということだと思います。それは砕けた言い方になりますが、数値シミュレーションの結果を機械学習モデルの評価に用いることは、評価を更に難しくする可能性が高いので注意してくださいということ。やるなとまでは言わないけれどもという書き方にしました。「一方、工学では適切に得られたシミュレーションを、観察事実を補足するものとして利用することがあるから、FDAでも動物試験の結果を根拠なく援用するよりも、適切な数値シミュレーションのほうが良い場合もあり得るとしているように、現在はこれを受け容れる過渡期にあると言える」という言い方で逃げました。まとめ方はそうしてあり、最後に、これは新しく付けた部分なので、ここを先生方に見ていただければと思います。

確か前回の委員会のときに、具体例としてどういうものがあるかという話があったので、シナリオ例を付けてみました。体動への耐性を学習させたときに、体動の影響がどういうふうに画像に出てくるかを数値シミュレーション的にやってみたらどうかという。これは完全に創作なので、こういうものを載せることがいいのかどうかも含めて御意見を頂ければと思います。

考えたのは、こういうシナリオです。臨床ではモーションアーチファクトをわざわざやる形でデータを取らないし、逆に、被験者にそういうことをやらせてもらってデータを取ることは、特にCTの場合などは被ばくが増えるだけですから、なかなか正当化できないでしょう。

なので、それを数値的にシミュレートしてみたらどうでしょうか。Context of useとしては、肺のCT画像から病変部を検出するときに、体動、呼吸動などの影響を入れて、CTアルゴリズムに沿ってデータを集めたという想定でデータを再構成してみましょと。Verificationもある程度可能ですと。実験用CT装置はどこにあるのだと言われると、余り明確に言えないのですが、そういうことも可能です。

最終的に、Validation というのは、計算で出した画像がリアル
のモーションアーチファクトと似た画が出ているかどうかという
ことを先生に評価してもらおう。ただ、ここでは更にシナリオの一
部として、やってみたところ、どうも現実の CT 装置よりも過剰に
モーションアーチファクトが出てきました。最後の下から 6 行目
ぐらいです。原因としては、臨床用 CT 装置では、ちゃんと補正を
やっているのだけれども、このアルゴリズムにはそれが入ってい
ない。その補正方法は公開されていないから、シミュレーション
では加えていません。

結論としては、そういうシミュレーションをやるときには、過
剰なアーチファクトで学習させると、かえって他の重要な疾病に
関する特徴をスルーしてしまうアルゴリズム、ネットワークにな
ってしまうかもしれないので、ここは別途確認しましょう。これ
は本当に創作なので、こういうまとめ方にしてあります。私が追
加した部分は以上です。

○佐久間部会長 ありがとうございます。多分、後半の所が一番重要なポイント
になってくると思います。先生方、何か御質問はございますか。
今の最後の所で、先生が作られた数値シミュレーションの報告書
をしっかりと参照しながら、これを有効なデータとして使うときの
考え方はそこにあることは基本だから、そこで V&V が出てくると。

○鎮西委員 はい。

○佐久間部会長 一方、今のどういう使い方があるかという例は非常に分かりや
すいと思います。物理プロセスがはっきりしている。CT のアルゴ
リズム、体動の影響の見積もりなど理論的に明確なものもありま
す。ただ、実はブラックボックスもある。各企業固有の門外不出
のアルゴリズムが存在していて、そこで出てくる。だから、たと
え理論的には正しいとしてもちゃんと V&V をやってくださいとい
う話のところをちゃんとやった上で、適用を考えましょうとい
うことだと思いました。そこがすごくよく出ているように思いま
した。

ほかの先生方、多分、中田先生も data augmentation の所で、
人工知能だけではなくてこういう数値計算みたいな問題が出てく
るだろうという気がしていて、多分、学習データについて、こ
ういうことで少し増やしていくことはあり得て、一方、正しく増や
さなければいけないところがあり、そのときにこの共通的な考
え方が必要になってくるという感じだと思います。中田先生、何
かコメントはありますか。

- 中田(典)委員 これ、一応、ユースケースを書かれています、こういうユースケースは多少無理をしても書いたほうがいいと思うので、いいと思います。
- 佐久間部会長 すごく分かりやすいですね、これがあると。
- 中田(典)委員 これがないと困るのです。それから、無理をして作っているのは重々承知で、結局、実際に持ち込まれるものはこれとは違うものが出てくるとは思いますけれども、ないよりはいいと思いますので、御苦労されて作られたことには敬意を表します。以上です。
- 佐久間部会長 ありがとうございます。よろしいですか。この科学委員会専門部会の先回の報告書は非常に重要な位置付けを持っていて、だからイン・シリコ・ラーニングというのが結構出てきている。それがどこまでいけるかという話のときに、そこでもう1回、ここの話に戻ってというふうに誘導ができればいいと思うので、そういうことかと思えます。他の先生、いかがですか。利用データというのは取るのが難しいから、こういうところをやりたくなるのだからというのはよく分かって。
- 鎮西委員 あとは、また自然言語 AI の話に戻りますが、あれを見てみんなすごいと言って、シンギュラリティが起こっているという話を言い出している人もいるほどです。でも一方で、例えば数値計算のほうでこれだけ重箱の隅をつつくように評価をしてやっていますと。自然言語のほうは何となくフワッとした言い方で、何か合っているような気がするから、これでいいよねという言い方でこのままいくのはどうかなという気もしますから、数値計算側ではこれぐらいきちんと重箱の隅をつついて、いろいろなことを考えていますというのを残しておいてもいいのではないかと思えました。
- 佐久間部会長 そうですね。さっきのこういう課題がありますよということから、「はじめに」のことを書いていたときに、対話型 AI の回答が必ずしも正しくはない可能性がある。
- 鎮西委員 正しくないのと、今、出ているのは相当カンニングしている可能性があるのです。
- 佐久間部会長 そうそう。
- 鎮西委員 医師国家試験 5 年分全部合格しましたというのが、昨日ぐらいに出ていましたけれども、それは多分、過去問を見て、それを出してきているだけでしょうという気もするので、そういうことが起こっているのではないかという気がします。
- 佐久間部会長 だから、それが結局、今日の話の前半の学習データと評価デー

タの独立性みたいなところが実はあって、そういうところなのです。

- 鎮西委員 そうですね。
- 佐久間部会長 そこが評価のポイントなのだという事は一本通しておいて、そこに戻る形にすることを考えておく必要があるのかもしれませんが。一方で、非常に期待があるから期待先行でいくと、何でそんなことを言うのだということを勢いで言ってくる可能性があるときに、いやいや、科学的にはこういう課題があるのですよということをやちゃんと理解しながら、落ち着いて社会実装していくことだと思います。
- 鎮西委員 人間は言葉で言われてしまうと騙されやすいのです。
- 佐久間部会長 清水先生、どうぞ。
- 清水委員 1点、教えていただきたいことがあります。上のほうで「FDAが動物試験の結果を根拠なく援用するよりも、適切な数値シミュレーションのほうが良い場合もありうる」というのは、具体的にはどんな例でしょうか。
- 鎮西委員 これは、どこかでちゃんと引用したほうがいいですね。これはFDAがコンピューターシミュレーション・アンド・モデリングを推進するときに、いつも言っている常套句なのです。特定の文章で、ここにソースがあるというのを言えるかどうか分からないのですが、ただ、繰り返し使っているの、どこかに載っていると思うので、これはちゃんと引用を付けておこうと思います。
- 佐久間部会長 多分、今の御質問は、どんなケースなのかという先ほどの例だと思います。あれは拡大管の流体力学的効果か何かのところだったのではないでしたか。それがあったような気がしますね。
- 鎮西委員 はい。
- 佐久間部会長 いわゆる心血管領域の人工臓器のところの流体解析などの部分で、動物と、例えばそれによって体部構造が違ったりするから、その影響が出てくることを考えるときにということかもしれないですが、ちょっと調べてみなければいけないですね。
- 清水委員 評価全体を置き換えるという話なのか、それとも評価の一部にこれを加えるという話なのか、その辺は。
- 鎮西委員 FDAもかなり慎重なので、数値計算のほうがいい場合もあり得るというぐらいですね、まだ。
- 清水委員 では、一部分で数値計算の結果を利用するという事は、過渡期には十分考えられるという話。
- 鎮西委員 はい。

- 清水委員 なるほど、分かりました。
- 鎮西委員 根拠なく使用するよりも、動物実験をやって、その結果がそのまま人に対して当てはめられるかどうかという点に関して、いろいろなケースに関してあまり根拠が明確でないことを、彼らは気にしているということです。
- 清水委員 なるほど。分かりました。
- 佐久間部会長 中田先生、どうぞ。
- 中田(典)委員 自然言語処理のことで一言、短く言いたいと思います。鎮西先生がカンニングと言われたのは正にそのとおりですが、最近、バイオGPTという、要するにPubMedのデータを全部把握しているGPTが現れて、GPTとは要するに文章の意味を自然言語処理的に理解しているわけです。その上でPubMedのデータを全部把握している。言わば、そこでカンニングしているというわけですから、これを医師が瞬時に使い出すというのは容易に想像が付くことなので、そういう使われ方、二次利用というものです。そういうことが恐らく行われるというのは間違いないだろうと思います。今までのは意味の理解というのができなかったの、あまり精度が良くなかったのですが、かなりの確に精度よく抄録の中からポイントとなるものを引き出してくることができるようになったので、そこが私は怖いなと思っています。以上です。
- 佐久間部会長 先生、今の点は、実はこのAIの医療応用が始まる時に、ゲノムの解析でゲノムの変異とある種の抗がん剤のマッチングを取る。そこで実は東大の医科研などがそれをずっとやって、今まで標準治療とされていないけれども、特に希少がんなどでこの薬が効きそうだと出してくれる。そういうところと結構似た話だなと思って聞いていました。あのときミヤノ先生がおっしゃっていたのは、これは医師の考えるツールだという言い方をされていて、多分、考えるツールの範囲のところと、意思決定支援に入ってきたとき、実は今回はそれを入れていなかったのです、そういうところがあって。だから、そういう分野のAIというもので、同じと言えば同じですけども、膨大な医療知識、PubMedならPubMedの知識を分かった上で、そこで適切なパターンマッチングをして持ってくるということです。
- 中田(典)委員 ミヤノ先生と私と、IBMのWatson Healthですか、その話は昔あったのですが、あの時代はまだ抄録の中の文章の言語学的な理解はされていないのです。今のは文法的に全部理解しているわけです。ですから、管理・応用が利いているので、あの時代のWatson

と同じだと考えないほうが良いと思います。

- 佐久間部会長 なるほど。
- 中田(典)委員 それから、コグニティブ・コンピューティングはもう消えてしまいました。実際、Watson Healthは解体してサービスをやめてしまいましたので、今、次元は違うと考えたほうが良いと思います。
- 佐久間部会長 分かりました。アプリケーションとしては似ている感じもしています。
- 中田(典)委員 ユーザーインターフェースは少し似ています。
- 佐久間部会長 なるほど。どういうふうに医療応用するかというところですね。
- 中田(典)委員 医療応用というより、電子カルテとつながると医師が勝手に使い始めるのが、今、恐らく問題になると思います。各国でいろいろな見識者が勝手に使い出すのを禁止するとか、制限するとか制限しないとか、そういうのが今は問題になっているのです。この科学委員会で議論する必要はないと思いますけれども、重大問題だと思います。以上です。
- 佐久間部会長 ありがとうございます。

<データベースについての展望>

- 佐久間部会長 次、データベースの話に移りたいと思います。順に伊藤先生からお願いできますか。
- 伊藤副部会長 現状、結局、この前書いた内容で、私の内容とほかの先生が書いている内容を見させていただいて、データベースの部分をどうやってまとめたらいいのか私なりの私見と言いますか、一応、データベースの現状の概要を、私も含めて4つの領域のデータベースについてここでまとめることになっていますので、それぞれについて、今、先生方が自由に記載されていると思います。ただ、大体、書くべき内容みたいなものをある程度決めて、データベースの概要として私は10個書きました。1つはデータベースの目的が何で、これを作り始めたのか。2番として収集されるデータの種類と規格はどんなもので、3番目、今、集めた収集データのデータベース規模がどれくらいか。4番目は紐付けられる医療情報としてどういったものが挙げられるか。5番目、異なる個人、地域、機関、国、時期におけるデータ収集というのは必要か否か。6番は人の手技に対する依存度の有無があるか否か。7番が撮影・記録にかかるデバイスの多様性、8番が使用するデバイスの多様性があるか否か。9番、これは結構重要だと思いますが、主なアノテーションの方法や教師データの作成方法がどのようなものなのか。想定される利

用方法、これはそれぞれ大分違うと思います。10番、これは除くかどうか議論がありました。同意・移植のプロセス、個人情報への配慮ということで、一応、いろいろほかの先生方はこちらに結構振れていた。この1番に書いた中で共通の課題とか問題というのがどういうものなのか。これは、ある程度書いていただいた中で私が抽出してもいいです。3番目、データベースごとの独自の課題、問題点は何か。この3つの柱でいくと意外にこのデータベースの全体像が分かりやすいというのが私の提案です。以上です。

- 佐久間部会長　まとめていただいて、ありがとうございました。一気にいきますけれども、ほかの先生方、どうですか。
- 伊藤副部会長　いろいろ言ってください。私が考えただけなので、もっといろいろな先生の御提案に従おうと思っています。
- 佐久間部会長　御意見があれば頂ければと思います。お願いします。
- 笹野委員　笹野です。伊藤先生、ありがとうございます。非常に明確にまとめていただいて、私が追加することはないかと思えます。先ほどちょっと出した時間的独立性と言いますか異なる時間、これは5番に含まれるのかもしれませんが、そのようなところが独立した項目として概要に1つあってもいいのかなと思いました。
- 伊藤副部会長　なるほど、分かりました。
- 笹野委員　あと、我々は自由に記載していますけれども、多分、結構共通してまとめられるところが多いと思うので。
- 伊藤副部会長　そうなのです。読んでいて正にそう思って、書いてある所が割と皆さん共通しているのですが、何となく自由記載なので、そういった意味では決めてやると共通点と違いが浮き彫りになるのかなと思って提案させていただきました。
- 笹野委員　ありがとうございます。多分、3番が案外少なくて、結構、きれいにまとまってくるだろうと、これを拝見して思いました。ありがとうございました。
- 佐久間部会長　佐々木先生、いかがですか。
- 佐々木委員　全く伊藤先生の御提案どおりで非常に整理されていると思いました。佐久間先生も最初におっしゃっていましたが、病理のデータの場合には診断名が変わることがありますので、どの時期のデータなのかということと、そのデータを使う側も必要に応じて付けられたアノテーション、診断名等を置換することが重要

だということを私の所には書き加えさせていただきましたが、時間的要素というものも必要ということで、今、お話にありましたように5番に入るのかなと思いました。非常によくまとめていただきましたので、この通り整理して、もう一度見直して書きたいと思います。ありがとうございます。

○佐久間部会長 鎮西先生、どうぞ。

○鎮西委員 この辺りの話ですが、例えば時代によっていろいろ変わってきたものを、今、我々人間がそれを見て、あっ、これはこの時代のだし、こういう診断が付いているけど、この画を見る限り、今だったらこっちの診断をするよねという判断は、現在、可能ですよね。だとすると、かなり近い将来に機械でもそれをやってくれるのではないかと。そうなってくると、その辺りの時代的なシフトといったバイアスについては、あるのを前提で逆に我々はあまり気にしなくて済む時代が来るのではないかという気がしますが、その辺りはいかがでしょうか。

例えば病理の写真が残っていて、それで見られるというのは割と簡単なケースですが、疫学調査になってくると、元のデータなどはとうになくなってしまっていて数値だけが残っているわけです。ただ、どういうところが変わってくるかということも逆に検証されていくはずなので、それさえあれば、例えば1970年代に得られたデータと2020年代に得られたデータで生じる問題点に関して、人間がああだこうだと頭を悩ませるより前に、機械のほうでちゃんと判断してくれることも想定されるのではないかという気がします。いかがでしょうか。

○伊藤副部会長 1点、いいですか。先生、取り扱うデータによって大分違うと思います。

○鎮西委員 なるほど。

○伊藤副部会長 例えば病理診断、この前おっしゃっていただいたように、あるときに診断が若干変わるような時代においては、今、全部、先生が見直して診断を付け直せばいいのかもしれませんが、病理診断としてはそういう問題点がある。手術においては使うデバイスが10年前と全然違うので、扱うデータ内容によって鎮西先生の議論がフィットする場所と、あまりフィットしないデータと2つあるような気がします。

○佐久間部会長 他、いかがでしょうか。多分、そこは、今、鎮西先生のお話を聞いていて確かにそういう可能性はあるけれども、それがあったとしても、それをどういうふうにデータとして記録するか的手法

は別として、そこで書かれた時間的情報が必要だということを言っているとすれば、今のまとめ方でも大丈夫だということですよ
ね。

○鎮西委員 多分、まだ少なくとも2、3年は過渡期なので、あまりそういったことを気にしなくてもいい。むしろ我々は、例えば時代が変わっても何かあまり変化しないものを今の機械で測って、そういったデータを取っておく。標準体を測ってその結果を記録しておくほうが大事なのかもしれませんね。AIの話とは離れてしまいますけれども。

○佐久間部会長 今のことで思ったのですが、結局、膨大になるから大変なので
すけれども、生データの重要性はますます増すような気がしてく
るのです。この辺り、データベースを運営する立場からいくとス
トレージの問題はかなり大きな問題だから、どうしてもその部
分はある程度フィルタリングを掛けて入れてくることになるのだ
と思います。村垣先生、どうぞ。

○村垣委員 伊藤先生、おまとめいただきありがとうございます。データ
ベースの共通と独自とある2
と3の違いがよく分からなくて、多分、1の所でも様々な課題など
が出てくると思いますので、シンプルに2、3を一緒にしてデータ
ベースにおける課題、問題点でもいいのかなどという気がいたし
ましたけれども、いかがでしょうか。

○伊藤副部会長 そういったまとめ方でもいいと思います。ただ、一番思ったの
が、皆さんのを見てみると、大体、結構共通課題が確かに多いの
です。

○村垣委員 はい。

○伊藤副部会長 なので、恐らく共通の所で多くはまとめられると思いますが、
ただ、データベースの種類ごとに独自の課題として残されている
ものがあって、今、たまたま4つのソースの中でデータベースが
論じられていますけれども、今でもこういうものがあるから、将
来、こういったものが想定されるということを記載してもいいの
かなと思って書きました。村垣先生がおっしゃるように、必要が
ないのであればシンプルなほうがいいかなと思っています。

○村垣委員 ありがとうございます。あと、今回、収集データの種類、規格
で大きく分けてスタティックなデータと、いわゆる連続データと
か時系列データという種類に分けられるかと思いますが、もしそ
のような視点がなければそのような観点も入れていただければと
思いました。以上です。

- 佐久間部会長 どうでしょうか。今、伊藤先生がおっしゃったとおり 4 つのケースですね。4 つのデータベースのケースである程度集約できると。ただし、それでカバーできないものもあるかもしれない。そこは今回の調査研究のリミテーションとしてはあるのですが、ここまでまとめていけるから、ちゃんとこういうことをやっていってねということについては、結構、道標になるだろうと思います。一方、これを見た人は全部これに当てはめなければいけないのかという可能性もあるので、そこは共通と独自で、独自に許容性もあったほうが、これを見て自分たちのものを考えるときに少し精神的バリアが低くすむというか、そこはあるかなという気もしますけれども、どうですか。
- 伊藤副部会長 私のイメージとしては、基本的に自由記載の中で 1 番の概要の内容をある程度含んでいただいて、かつ、独自の問題点に関しては 4 人の先生方に書いていただき、共通の課題に関しては私のほうで見て、ある程度まとめて皆さんにどうですかという形にすれば何となくまとまるのではないかと思っていました。皆さんの御意見、いかがでしょうか。
- 笹野委員 書き方としては、我々が 1 の項目について書いていき、伊藤先生に見ていただいて 2 と 3 に振り分けるみたいな形になるのですかね。
- 伊藤副部会長 はい。それで、ある程度まとめて皆さんにブラッシュアップしていただくのでも全然構わないと思います。
- 笹野委員 分かりました。独自のというところはどうしても残るかなと。特に心電図などはスタティックデータかバリアビリティかは結構大きな問題でもありますし、それはそれで項目としては残して、この 4 つだけでどこまで全てのデータベースに演繹できるのか分かりませんが、そういう項目はあったほうがいいかなと思います。以上です。
- 佐久間部会長 さっきのまとめ方の頭で、先生方に個々のいわゆる消化管の内視鏡というようなことと、内視鏡外科の映像と病理の映像、心電図というのがあって、そこで一通り書いていただいていますね。ここでそこをフィルタリングするというか、そこで出てきた情報を正にデータ圧縮して、それで 1 に落としていくとどういうことになるかということ、伊藤先生にやっていただくということですね。そのときの書きぶりとして、これは具体的なところで収集データの種類、規格というのが、例えば動画の場合もあるだろうし、静止画の場合もあるだろうしみたいなことが出てきます。伊

藤先生のイメージとしては、今回、全部まとめてしまうと、データの種類の病理ではこういうことです、これはこういうことですみたいな感じの書き方になるのか。その辺り、どんなイメージをお持ちですか。

○伊藤副部長 むしろ、今、先生が言われた元となるデータもそれぞれ違うので、それはもちろん独自のものだと思いますが、以前私が書いた中で、表でまとめたもの、まだあれも足りないなと思っていたので、例えばアノテーションの手法はどのようなものが共通なのか、この中で非常にクローズアップされる項目があると思います。なので、全てを共通の中で書き込む必要はないとされていて、その中で、本当にここで重要だとして議論されるべきものに絞り、ある程度最初に広めに書こうと思います。それで皆さんに必要なというのであればどんどん削っていただく形がいいかなと思います。

○佐久間部長 だから、各先生方をお願いするとしても、今までのスタイルで各分野についてまとめて、この1、2、3ぐらいの書きぶりで少し御自身の所の分野を書いていただき、それを伊藤先生が見て、そこから少し抽象化したようなものに落とし込んでいく。その二段構えみたいな項目でデータベースを書くと、読むほうも具体的な感じがあって分かりやすい。そこは重複があっても私はいいと思います。そんな感じでもいいのかと思います、どうですか。

○伊藤副部長 私のイメージはそういう感じなのと、私が一番聞きたかったのは、1番の織り込むべき内容の項目がこれで十分か否かというところがあって、先ほど言われた時期のデータのクロノロジカルなものに関しては独立させたほうがいいのか、ここについて御意見を頂いて、ある程度これを目安に書いていただけると書くほうも書きやすいのかなと思ったのです。

○佐久間部長 それと、今日、実は清水先生が言われたところで更にもうちょっと、臓器別独立性といった話は逆に医学サイドから来た話ですね。そこから後、もう1つ数学的な独立性みたいな話と、多分、清水先生、この場合、どういうふうに入れますか。

○清水委員 ありがとうございます。最初、2に入れようかなとも思っていたのですが、完全に独立して4にして、システム開発でデータベースをどう使うかという話なので、先生方から頂いたものは、一応、目を通した上で整合性がある形で4にしたほうが、他の先生方も書きやすいのではないかと考えています。

○伊藤副部長 そういうことですか。4としてどういう内容ですか。

- 清水委員 SaMD 設計のためのデータベースの利用法というような。
- 佐久間部会長 利用における留意点みたいな感じですか。利用法における、いろいろとその留意点みたいな感じですかね。
- 清水委員 そうですね。開発のための。
- 清水委員 データベースの利用法における注意点、ちょっと長くなりましたけど。タイトルは考えさせていただくとして、内容的にはそういうことで。
- 佐久間部会長 私のイメージから言うと、ここでこういうデータベースの目的があって、こういうものが出てきます。それをうまく使おうという人がいたときに、独立性というけれども、独立性という概念はそこでもいろいろ出ているわけで、そこに出ているように、単に例えば同じものを使ってはいけないとか、そういうことではないですよ。そのときにそういう概念があって、これは数学的にはこういうことですよということを、多分、清水先生が書かれて、そんなイメージですね。
- 清水委員 はい、正にそのことで書いておきたいと思います。できたら最初に前半部分を拝見させていただいた上で、全体にどのデータの場合にも共通で言える話と、この話はこのデータでしかないねということがあれば、その辺も区別しながら書いてみたい。
- 佐久間部会長 これは技術的な話、数学的な話と、どうリンクするかというところがポイントなのだろうと思います。
- 清水委員 はい。統計的には IID ですか、independent and identically distributed という学習とテストは、ある程度、独立性はないといけないというところから始まって、ただ、その独立性が、先ほど佐久間先生の話でもあったように細分化しています。できる限り開発する人が、自分の問題の場合にはどうしたらいいか、どういうことに注意したらいいかが分かるような書き方にしたいと思います。
- 伊藤副部会長 今の話だと、4 に関しては清水先生が後で加えていただける感じですね。
- 佐久間部会長 そうそう、そうです。だから、これはどちらかというとエンジニアからの視点でこれを見たときという話だと思います。一方で、1 から 3 というのは実際の臨床で作るときに、単にそこだけではなく、AI というものの特性を考えたときにデータをどう取るかということでもとめていくかという話だと思います。だから、両方ともある問題を違う方向から見ているようなところがあって、そこが分かるといいのかなという感じがして聞いていました。他

の先生方、いかがですか。そしたら一旦、これでまとめてみましょう。あとで日程が来ますけど、まだWGがあって、もう1回ぐらいの時間があるので、そこの工程をまた後で考えたいと思います。ありがとうございます。あと10分ぐらい時間があります。佐々木先生、追加の点があれば、担当分の御説明を頂けますか。

○佐々木委員 ありがとうございます。佐久間先生に言っていたような先ほどの診断名が変わるということと、どのときに付けたアノテーションか、使う側もデータベースを作る側も留意しながら作ろうということを書き加えただけです、大きな改訂はございません。以上です。

○佐久間部会長 ありがとうございます。先生方、いかがでしょうか。質問はございますでしょうか。

○清水委員 清水です。森先生がチャットのほうで触れているのですが、ちょっと御覧いただいて。

○佐久間部会長 医学放射線学会などでデータベースを作っているの、加えていただけますか。

○伊藤副部会長 これ、いいと思いますけど。放射線のデータベース、誰に担当していただくかを決めれば。

○佐久間部会長 誰に担当いただいたらいいのか。

○清水委員 これは多分、御本人は会議中で出れないような気がします。

○佐久間部会長 これは後で確認しましょう。

○清水委員 多分、イメージをお持ちかもしれない。確認したほうが早いのかもかしない。

○佐久間部会長 確認しましょう。笹野先生の御担当分はいかがでしょう。

○笹野委員 私のほうは先ほどのお話で全て出ていまして、特に追加で申し上げることはございません。各データベースの独自の部分というのはしっかり書いて、そこは独立して、読んだ人が自分のケースに当てはめられるようにという記載は、もう少し充実させたほうがいいかなと思いましたが、それだけです。

<総括>

○佐久間部会長 残りの時間は全体のディスカッションに入りたいと思いますが、その前に今までの議論をもう1回整理します。「はじめに」については、今日の議論を反映して少しアップデートを入れること。それから、ChatGPTなどの自然言語処理系のAIというものも進んできているので、その取扱いについて適切な記述をする。この報告書でなかなかカバーしきれないのですが、重要な課題であるとい

った書き方はしましょう。それから、データベースの要件の所等で、私が記述メモの所で書き過ぎたところがあって、評価データということに関して、学習データと違い、よりトレーサビリティということが求められる可能性があることをどこかで適切に入れるということ。これはPMDAの方々と相談しながら適切に考えることをするという事です。中岡先生のところはかなりいろいろな動きが出てきていて、この動きに関して考察できる所を強化すると、各国の規制の動向、あるいは標準化の動向と今回の議論がどう関連しているかについて、少し考察の中に加える作業をするということ。これは中岡先生と佐久間ですね、担当することになるかと思います。

- 中岡副部長 佐久間先生、確認です。基本的には多分、ここの議論の中でということになると、各国の規制当局がいろいろトライアルをしている今の状況ですね。
- 佐久間部長 はい。
- 中岡副部長 その中で、日本として導入できそうなものがどういうものかという形の考察になるのかなと思いますけど。
- 佐久間部長 あと、各国で議論されていることはどういう科学的な、この後に出てくる議論と関連しているかというところがあった方がいい。何ができるかという話もありますが、結局、そことこれがどう関連しているかということで、その関連性を考えた上で何ができるかを考えなければいけない。先ほどの話で信頼性をやると言っても、できる部分とできない部分があって、日本の国情に合わせていくのですが、ただ、この議論の中でどういう点を押さえながらやらなければいけないか分かった上で、この制度の設計をしてもらわなければ困るわけですから、そういうことをやるための資料になるようにという考察ではないかと思います。
- 中岡副部長 分かりました。どこまで食い込めるか難しいところですけども。
- 佐久間部長 書いていただいて、これはもう1回行政チェックがありますし、書き過ぎのところは切りますから、そこは1回やってみることでいいと思います。
- 中岡副部長 分かりました。頑張ってみます。
- 佐久間部長 それから、清水先生のところは、そういう意味では、今言ったデータベース共通の部分を伊藤先生の項目の4に持ってくるという作業があって、それは全体を1回見ていただいてからということと、前半の所に関しては、私も1回読ませていただいています

けれども、数学的な話と意味的な話の所をより初学者というか、この分野に入ってきた人に、過度に怖がらせるのではなく、適切な懸念点を持って動いてもらえるような形の文章にするところなので、例えば同じデータを使ってはいけないと言っているわけではなく、同じデータを使ったときに、こういう危なさがあるから気をつけてねということが分かるような形の文章にしていって、それをまとめていっている作業だと思うので、ほとんどコンデンスは入っていると思いますから、あとは少し全体のあれかと思えますので、そこは先生と私で御相談させていただきたいと思えます。どちらかという先生は最後の所に注力していただくのが、この段階ではいいのかなという感じがしています。

○清水委員 お願いします。あと、データへの摂動を与えるという話、鎮西先生の所でということをし少し申し上げたのですが、もしよろしければ鎮西先生に書いていただいてもいいのかなと。

○佐久間部会長 そうですね。

○清水委員 私が違うトーンの記事で書いても思うのですが、鎮西先生、いかがですか。また後で御相談させていただければと思えます。

○鎮西委員 機械学習における数値シミュレーションの利用という節があるので、その中で使い方としてこういうのがありますということで書いておく手がある。

○清水委員 はい。また後で御相談させていただきたいと思えます。よろしくをお願いします。

○佐久間部会長 あとは、森先生の所は人によるバイアスを少し分類しましょうという話が出てきました。あと、データベースについては先ほど議論したとおり伊藤先生のまとめ方、これは後で伊藤先生、事務局に。

○伊藤副部会長 もう送りました。

○佐久間部会長 そうですね。皆さんでシェアしましょう。そうした上で、もう一度、データベースを先生方に、そんな形でリエディットしていただいて、それを1回持ち寄って伊藤先生に見ていただき、もう少し抽象化したような文章を作るという作業に次に入ると。まずは各データベースの部分を、伊藤先生の骨組みに従って再編集していただく作業になる。それで4番目の所に入るので、清水先生の所が今の1、2、3の議論にうまく合った形に、そこの内容をまたリエディットする。そんな感じの作業になるかなと思えます。大体、そんな事かと思いましたが、いかがでしょう。

それで最後、ChatGPTです。これはどうしますか。だから文法と

して言語構造も理解した上でやっているのので、かなり応用の範囲が出てくるような時代になってきたと。ただし、どういうふうに書きますかというところですが、これは宿題で考えましょうか。先ほど中田先生にそういう実績を書いていただくということでお願いしたので、そこを含めてですが、ChatGPTなどの言語の背景にある概念などは、コンピュータは理解できていません。この背後を捉えているわけではないので、ただ、あまりにも言語として真っ当な回答が返ってきていることを書いてあるので、それはまた継続審議しましょう。メール上でも議論ができればと思います。これは新しく、これを始めたときには想定していなかった事態なので。あとデータベースについて、医学放射線学会などについては森先生に確認しましょう。

以上が全体のまとめで、長くなっていますが、全体を通して先生方、何かございますか。よろしいですか。ありがとうございました。本日の議事は以上ですが、今日、まとめて最後に私がお願いした作業、事務局、最後のところだけメモでまとめていただいて、うまくしていただければと思います。

<その他>

○事務局（浜岡先端技術評価業務調整役） 本日はありがとうございました。

次回の専門部会ですが、6月19日(火)、14時から16時の開催予定で最終回の予定になっています。詳細については追って御連絡をいたします。また、その前に5月15日(月)、14時から16時に第5回目の専門部会のWGを予定しています。御参加いただく委員について申し上げます。佐久間部会長、伊藤副部会長、中岡副部会長、佐々木委員、笹野委員、清水委員、田中委員、鎮西委員、中田典生委員、森委員、横井委員でございます。よろしく願いいたします。

<今後の進め方について>

○佐久間部会長 お願いします。横井先生、いらっしゃいますか。

○横井委員 はい、おります。

○佐久間部会長 実は変なお願いですが、横井先生、WGに入っていていただいて、これである程度全体の流れができたと思いますので、それで一読していただいて、実は自分が書いているとチェックするときに思い込みが入ってしまうので、先生は第三者的な目で見ていただくことが可能なお立場かと思っておりますから、これから最終作業のところ

ではWGの中でそんな役割をしていただき、例えば先生が見て不思議だと思ふような所についてコメントいただくとか、そういうことをお願いできればと思います。よろしくお願いします。

○横井委員 はい、ありがとうございます。そうですね、やらせていただきます。

○佐久間部会長 15日まで、今言ったことで先生方にまとめていただいたものを連休明けぐらいのときにまとめて、15日にそれをガッチャンコして、もう1回議論し、それで最終的なものに持ち込んで修正をかけたものの原案を6月にかける。多分、15日に終わった後、そこでまとめて数回やり取りしながら、最終的にまとめていく作業になるかと思っています。よろしくお願いいたします。

<閉会>

○佐久間部会長 ということで、本日の専門部会は5分ほど延びてしまいましたが、ここまでとさせていただきます。本当に今日はありがとうございました。